Drawing and Recolouring

Bart Jacobs
iHub, Radboud University, Nijmegen, The Netherlands

Márk Széles
iHub, Radboud University, Nijmegen, The Netherlands

Abstract

Drawing a ball from an urn filled with balls of different colours is one of the basic models in probability theory. The probability of drawing a ball of a particular colour is determined by the proportion / fraction of balls of that colour. This paper introduces a new stochastic model for such urns: draw a ball, recolour (repaint) it, and put it back into the urn. One can distinguish four modes of drawing-and-recolouring, namely whether done proportionally or uniformly (both for drawing and recolouring). These modes can be reformulated in financial terms as redistribution of wealth or in biological terms as evolutionary drift. The resulting four operations form a coalgebra for the distribution monad, on the set of multisets of a fixed size. In fact they form a Markov chain and even a hidden Markov model, in combination with the frequentist learning map as emission channel. This paper identifies fixed points, capturing stable situations, for these four modes of drawing-and-recolouring.

2012 ACM Subject Classification Mathematics of computing → Markov processes

Keywords and phrases Markov-chain, Urn model, Multiset

Digital Object Identifier 10.4230/LIPIcs.CALCO.2025.8

1 Introduction

Imagine an urn filled with various coloured balls. One typically assumes that the urn is well-shaken so that drawing a ball from the urn happens according to a probability distribution that is proportional to the numbers of balls of each colour. We shall describe such an urn as a multiset, written with ket notation, for instance as: $3|R\rangle + 2|G\rangle + 1|B\rangle$. It represents an urn with three red (R), two green (G), and one blue (B) ball. The corresponding probability distribution is $\frac{1}{2}|R\rangle + \frac{1}{3}|G\rangle + \frac{1}{6}|B\rangle$. It is obtained via normalisation, an operation that will appear as "frequentist learning" below: learning a distribution by counting.

Commonly, three modes of drawing are distinguished, namely multinomial, hypergeometric and Pólya, corresponding to draw-replace, draw-delete, and draw-duplicate. In multinomial mode one draws a ball, inspects its colour, and returns it to the urn, so that the urn is the same for a next draw. In hypergeometric mode one draws, inspects, and keeps the drawn ball outside the urn, so that the urn shrinks. In Pólya mode one draws, inspects, and returns the drawn ball together with an extra ball (of the same colour) to the urn. In this mode the urn grows in size – and displays clustering behaviour. For a systematic description of these three modes, see [4, 6]. For an overview of urn models we refer to [7, 9, 10].

This paper introduces a new stochastic urn model, in terms of drawing and recolouring. One "somhow" draws a ball, "somhow" repaints it, and then returns it to the urn – so that the number of balls in the urn remains the same. We consider in total four different modes for what we just called "somhow". In all these modes the recolouring does not depend on the colour of the drawn ball.

We describe the four modes informally, with labels PP, PU, UP and UU, where P stands for proportional and U for uniform. The first letter is for the mode of drawing and the second one for the mode of recolouring. Technical details in terms of multisets will appear

11th Conference on Algebra and Coalgebra in Computer Science (CALCO 2025).
Editors: Corina Cîrstea and Alexander Knapp; Article No. 8; pp. 8:1–8:15
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

later on. The draw-and-recolour operations work on a fixed urn, of a particular size $K \in \mathbb{N}$ with a particular number $N \in \mathbb{N}$ of colours – not all of which have to be present in the urn. The size of the urn is the total number of balls in it.

PP means that the draw is proportional and the recolouring too. This means that the fraction of balls of each colour in the urn determines which colour is drawn, but also which colour is used for the repainting.

PU means that the drawing is proportional and the recolouring is uniform. The latter means that the probability of each colour (of the returned ball) is $\frac{1}{N}$, where N is the number of colours.

UP means that the draw is uniform and the recolouring happens proportionally to the colours in the (original) urn. What it means for a draw to be uniform is a bit subtle: We look at the number of colours that occur in the urn, say $M \in \mathbb{N}$, with $M \leq N$, and draw uniformly with probability $\frac{1}{M}$. In this way we ignore the numbers of occurrences of each colour that occurs in the urn.

UU means that both the drawing and the recolouring happens uniformly.

This drawing-and-recolouring may be repeated. We are interested in fixed points of such repetitions. It is not immediately clear what they are. We give away one outcome: in the *PU* case one obtains a uniform distribution, since colours average out.

One can interpret the situation at hand in economic / financial terms. The colours then correspond to people and the number of balls of each colour represents the wealth of these people, as natural number. The drawing-and-recolouring can then be seen as a wealth distribution mechanism, for instance via taxing. In the *PU* case people are taxed proportional to their wealth, and the tax money is distributed uniformly. Our fixed point claim says that this will lead to a uniform situation, with a uniform distribution. Intuitively this may be clear. The reader is invited to try and find what the stable situations are in the remaining three cases.

The four modes *PP*, *PU*, *UP*, *UU* described above each lead to a function of the form $\mathcal{M}[K](X) \to \mathcal{D}\big(\mathcal{M}[K](X)\big)$. We write $\mathcal{D}$ for the distribution monad and $\mathcal{M}[K](X)$ for the set of multisets (urns) of size K, over a finite set of colours X. These maps form coalgebras, of the distribution monad $\mathcal{D}$, with $\mathcal{M}[K](X)$ as (finite) state space. In fact, such a coalgebra forms a Markov chain. Fixed points of Markov chains, also called stationary distributions, are a well-studied topic (see *e.g.* [8]).

We have said that *PU* leads to a uniform distribution. In order to make this precise we need to extend our Markov chains with an emission channel, so that they become hidden Markov models. The emission channel is, in each of the four cases, the frequentist learning map $\mathit{Flrn}\colon \mathcal{M}[K](X) \to \mathcal{D}(X)$, as informally mentioned in the beginning of this introduction. We can now say more precisely that *PU* leads to a stationary distribution of the Markov chain, which gives the uniform distribution when the emission channel *Flrn* is applied. The composition of the transition channel of the Markov chain and the application of the emission channel all involve, as usual, channel composition. In categorical terms it is Kleisli composition.

This paper is structured as follows. Section 2 recalls some basic theory of multisets, of discrete probability distributions, and of discrete-time Markov-chains with a finite, discrete state space. Section 3 presents the definitions of the four hidden Markov models for drawing and recolouring. Finally, Section 4 analyses the long-term behaviour of all four models. This involves computing fixed-points and proving convergence properties of four Markov-chains, and of the associated hidden Markov models.

Related work

Models similar to our urn models have appeared in the literature, albeit not as Markov chains on multisets. The *UU* mode has been used to model a situation in which agents exchange money. In each step, two agents are selected to interact. The loser transfers a unit of money to the winner. In [1], this scenario is investigated using tools from statistical mechanics. Under the physical interpretation, the agents correspond to particles randomly exchanging energy. Conservation of the total amount of money in the system then corresponds to conservation of total energy.

The *PP* mode occurs in evolutionary biology. In this case, the colours correspond to competing genetic variations. The process captured by *PP* is called neutral drift, which is a special case of a model called the Moran process [11, 12]. During neutral drift, eventually, all but one of the genetic variations are expected to disappear. In our situation this corresponds to stationary distributions with multisets containing only one element (multiple times), see Section 4.2.

2 Background

2.1 Multisets

A **multiset** is like a subset, except that its elements may occur multiple times. In the beginning of the introduction we mentioned an urn $3|R\rangle + 2|G\rangle + 1|B\rangle$ with coloured balls as example. For an arbitrary set X we shall write $\mathcal{M}(X)$ for the set of (finite) multisets with elements from X. A multiset $\varphi \in \mathcal{M}(X)$ may be described in two equivalent ways, namely via kets as a formal sum $\varphi = n_1|x_1\rangle + \cdots + n_K|x_K\rangle = \sum_i n_i|x_i\rangle$, with elements $x_i \in X$ and associated multiplicities $n_i \in \mathbb{N}$. Alternatively, we may describe a multiset $\varphi \in \mathcal{M}(X)$ as a function $\varphi\colon X \to \mathbb{N}$ with finite support, where $\mathrm{supp}(\varphi) = \{x \in X \mid \varphi(x) \neq 0\}$. We can combine these notations and write $\varphi = \sum_x \varphi(x)|x\rangle$, where the summands with multiplicity zero are omitted.

The size $\|\varphi\| \in \mathbb{N}$ of a multiset $\varphi \in \mathcal{M}(X)$ is the total number of its elements, including multiplicities. Thus, $\|\varphi\| = \sum_x \varphi(x)$. There is one multiset in $\mathcal{M}(X)$ of size zero, namely the empty multiset, written as $\mathbf{0} \in \mathcal{M}(X)$. As a function, it satisfies $\mathbf{0}(x) = 0$, for each $x \in X$, giving x multiplicity 0. We shall write $\mathcal{M}[K](X) \subseteq \mathcal{M}(X)$ for the subset of multisets of size $K \in \mathbb{N}$.

For each set X and number K there is an "accumulation" map $acc\colon X^K \to \mathcal{M}[K](X)$ that turns lists into multisets via $acc\big(x_1, \ldots, x_K\big) \coloneqq 1|x_1\rangle + \cdots + 1|x_K\rangle$. For instance $acc\big(c, b, a, c, a, c\big) = 2|a\rangle + 1|b\rangle + 3|c\rangle$. This accumulation map is the coequaliser of all transposition maps $X^K \xrightarrow{\cong} X^K$ induced by permutations of K, see [3, 6].

The next result collects some basic combinatorial results about multisets. Proofs may be found in [4, 6].

▶ **Lemma 1.** *Fix a non-empty finite set X with $N \coloneqq |X| \geq 1$ elements and a number $K \in \mathbb{N}$.*

1. *The set $\mathcal{M}[K](X)$ of multisets of size K over X has $\left(\!\binom{N}{K}\!\right)$ many elements, where $\left(\!\binom{N}{K}\!\right) \coloneqq \binom{N+K-1}{K} = \frac{(N+K-1)!}{K!\cdot(N-1)!}$ is the multi-choose coefficient.*
2. *This multi-choose coefficient satisfies the following equations.*

$$\left(\!\binom{N+1}{K}\!\right) + \left(\!\binom{N}{K+1}\!\right) = \left(\!\binom{N+1}{K+1}\!\right) \qquad \sum_{0\leq m\leq K} \left(\!\binom{N}{m}\!\right) = \left(\!\binom{N+1}{K}\!\right). \tag{1}$$

3. *For* $\varphi \in \mathcal{M}[K](X)$ *there are* $(\varphi) \in \mathbb{N}$ *many lists that accumulate to* φ*, where:*

$$(\varphi) \coloneqq \frac{\|\varphi\|!}{\varphi!!} \qquad \textit{with} \qquad \varphi!! \coloneqq \prod_{x\in X} \varphi(x)!.$$

4. *These multinomial coefficients* (φ) *satisfy*

$$\sum_{\varphi\in\mathcal{M}[K](X)} (\varphi) = N^K.$$

5. *The multinomial coefficients* (φ) *satisfy the following recurrence equation:*

$$\sum_{x\in\mathrm{supp}(\phi)} (\varphi - 1|x\rangle) = (\varphi).$$

2.2 Distributions

A (finite discrete probability) **distribution** $r_1|x_1\rangle + \cdots + r_K|x_K\rangle$ is written as a multiset, except that the multiplicities r_i are now probabilities from the unit interval $[0,1]$ that add up to one: $\sum_i r_i = 1$. We shall write $\mathcal{D}(X)$ for the set of distributions with elements from X. A distribution $\omega \in \mathcal{D}(X)$ may be written in ket form, as we just did, but also as a function $\omega\colon X \to [0,1]$ with finite support $\mathrm{supp}(\omega) = \{x \in X \mid \omega(x) \neq 0\}$ and with $\sum_x \omega(x) = 1$.

Each non-empty multiset $\varphi \in \mathcal{M}(X)$ can be turned into a distribution $\mathit{Flrn}(\varphi) \in \mathcal{D}(X)$, where *Flrn* stands for frequentist learning. This involves learning by counting:

$$\mathit{Flrn}(\varphi) \coloneqq \sum_{x\in\mathrm{supp}(\varphi)} \frac{\varphi(x)}{\|\varphi\|} \,|x\rangle. \tag{2}$$

For instance, for the urn $\upsilon = 3|R\rangle + 2|G\rangle + 5|B\rangle$ we get $\mathit{Flrn}(\upsilon) = \frac{3}{10}|R\rangle + \frac{1}{5}|G\rangle + \frac{1}{2}|B\rangle$. The distribution $\mathit{Flrn}(\upsilon)$ captures the probabilities of drawing a ball of a particular colour from the urn υ.

A **channel** is a function of the form $c\colon X \to \mathcal{D}(Y)$. It maps each element $x \in X$ to a distribution $c(x)$ on Y. As such it occurs often as conditional probability, written as $P(y|x)$. The function notation is more convenient and flexible, for instance since it supports sequential (and parallel) composition of channels, introduced below as $\circ$. This emphasises that channels are well-behaved probabilistic computations. In fact we simplify the notation and use a special arrow $\rightsquigarrow$ for channels. Thus we simply write $c\colon X \rightsquigarrow Y$ for the more cumbersome $c\colon X \to \mathcal{D}(Y)$.

A basic operation of channels is **pushforward**. Given a channel $c\colon X \rightsquigarrow Y$ and a distribution $\omega \in \mathcal{D}(X)$ on its domain X, we can form a new distribution $c_*(\omega) \in \mathcal{D}(Y)$ on its codomain Y, namely:

$$c_*(\omega) \coloneqq \sum_{x\in X} \omega(x)\cdot c(x) = \sum_{y\in Y}\left(\sum_{x\in X} \omega(x)\cdot c(x)(y)\right)|y\rangle. \tag{3}$$

Given another channel $d\colon Y \rightsquigarrow Z$ we can define channel composition $d \circ c\colon X \rightsquigarrow Z$ as: $\big(d \circ c\big)(x) \coloneqq d_*(c(x)) \in \mathcal{D}(Z)$. This composition $\circ$ is associative and has a unit given by the “Dirac” identity channel id with $\mathrm{id}(x) = 1|x\rangle$. This composition structure arises because $\mathcal{D}$ is a monad.

Given two distributions $\omega, \omega' \in \mathcal{D}(X)$, one can define the **total variation distance** *tvd* as

$$tvd(\omega,\omega') = \tfrac{1}{2}\cdot \sum_{x\in X} \big|\omega(x) - \omega'(x)\big|.$$

This makes $\mathcal{D}(X)$ a metric space, see *e.g.* [5]. When discussing convergence of sequences of distributions, we will assume that the set of distributions $\mathcal{D}(X)$ is equipped with this total variation distance.

We will need the following result about the interaction of state transformation and limits of sequences of distributions.

▶ **Lemma 2.** *Let X, Y be finite sets, $c : X \rightsquigarrow Y$ a channel, and $(\omega_i)_{i\in\mathbb{N}}$ a sequence of distributions in $\mathcal{D}(X)$ such that the limit $\omega = \lim_n \omega_n$ exists. Then the limit $\lim_n c_*(\omega_n)$ also exists, and it is equal to $c_*(\omega)$.*

Proof. The state transformation operation c_* defines a continuous mapping $c_* : \mathcal{D}(X) \to \mathcal{D}(Y)$. Therefore, it preserves limits of convergent sequences. ◀

2.3 Markov chains

A (discrete-time) **Markov chain** is a stochastic process in which the next state only depends on the actual state. In this paper, we will focus on Markov-chains that have a finite, discrete state space. Given a finite set X, a Markov chain with state space X can be conveniently represented as a channel $c : X \rightsquigarrow X$. Every such Markov chain c has a corresponding stochastic matrix of size $X \times X$, called the transition matrix of c.

We now recall some of the basic theory of finite-space Markov chains. We refer to [8] for a detailed introduction to the topic.

▶ **Definition 3.** *A Markov-chain $c : X \rightsquigarrow X$ is called* irreducible *if every state can reach every other state with a non-zero probability. Formally, c is irreducible if for all pairs of states $x, y \in X$ there exists an $n \in \mathbb{N}$ such that $c^n(x)(y) > 0$, where c^n is the n-fold Kleisli composition $c \circ \cdots \circ c \colon X \to \mathcal{D}(X)$.*

▶ **Definition 4.** *Let $c : X \rightsquigarrow X$ be a Markov-chain.*

1. *The* period *of a state $x \in X$ is the natural number*

$$\gcd \left\{ n > 0 \;\middle|\; c^n(x)(x) > 0 \right\}.$$

That is, the period is the greatest common divisor of all possible numbers of transitions in which the state x can be reached from itself.

2. *The Markov-chain c is called* aperiodic *if all of its states have period one.*

▶ **Definition 5.** *Let $c : X \rightsquigarrow X$ be a Markov chain.*

1. *A state $x \in X$ is called* absorbing *if it is not possible to transition out if it. That is, if $c(x) = 1|x\rangle$.*
2. *The Markov-chain c is* absorbing *if it has at least one absorbing state, and each state x can reach an absorbing state. In this situation, the non-absorbing states are called* transient.

▶ **Definition 6.** *Let $c : X \rightsquigarrow X$ be a Markov chain. A distribution $\omega \in \mathcal{D}(X)$ is called a* stationary distribution *of c if it satisfies $c_*(\omega) = \omega$.*

Such a stationary distribution represents a possible equilibrium of the Markov-chain. It is well-known that every Markov-chain has at least one stationary distribution. One is often interested in whether a Markov-chain eventually stabilises, by converging to such a stationary distribution. We present two standard results, without proof.

▶ **Proposition 7.** *Let $c : X \mathrel{\circ\!\!\rightarrow} X$ be an irrudicible, aperiodic Markov-chain.*

1. *There is a unique stationary distribution $\omega \in \mathcal{D}(X)$ for c.*
2. *Moreover, c converges to this stationary distribution from any starting state. That is, the limit:*

$$c_\infty \;=\; \lim_{n\to\infty} c^n : X \mathrel{\circ\!\!\rightarrow} X$$

exists, and satisfies $c_\infty(x) = \omega$ for all states $x \in X$.

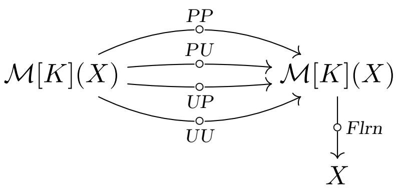
▶ **Proposition 8.** *Let $c : X \mathrel{\circ\!\!\rightarrow} X$ be an absorbing Markov-chain with the set $Y \subseteq X$ of its absorbing elements.*

1. *A state $\omega \in \mathcal{D}(X)$ is stationary for c if and only if $\mathrm{supp}(\omega) \subseteq Y$. That is, if it is of the form $\omega = \sum_{y\in Y} r_y|y\rangle$, for suitable probabilities $r_y \in [0,1]$ summing up to 1.*
2. *Let $x \in X$ be a state. Then the limit distribution $c_\infty(x) = \lim_n c^n(x)$ exists, and is a stationary distribution for c.*

There is a known formula to compute c_∞. We will not present it here, as it will not be needed in our developments. We refer the interested reader to [8, Chapter III]. An example calculation for the *UU* model will be shown in Subsection 4.3.

3 Hidden Markov models for drawing and recolouring

We consider four different hidden Markov models acting on multisets of a fixed size $K \in \mathbb{N}$, over a finite set X. In each of these cases the emission channel is frequentist learning $\mathit{Flrn}\colon \mathcal{M}[K](X) \mathrel{\circ\!\!\rightarrow} X$. We will introduce four different transition channels in a situation:



(4)

The transition channels are defined on K-sized multisets $\varphi \in \mathcal{M}[K](X)$ as described below, where $N = |X|$ is the number of elements in the finite set X. In each case an element $x \in \mathrm{supp}(\varphi)$ is drawn and $y \in X$ is added as recolouring. In case of proportional drawing, the probability of drawing a ball of colour $x \in \mathrm{supp}(\varphi)$ is $\frac{\varphi(x)}{K}$. Similarly, when recolouring propositionally, each added colour y is selected with probability $\frac{\varphi(y)}{K}$. If the draw is uniform, the probability of drawing $x \in \mathrm{supp}(\varphi)$ is $\frac{1}{|\mathrm{supp}(\varphi)|}$. Uniform recolouring selects each added colour y with probability $\frac{1}{N}$.

$$\begin{aligned}
\mathit{PP}(\varphi) &\coloneqq \sum_{x\in\mathrm{supp}(\varphi)} \sum_{y\in X} \frac{\varphi(x)\cdot\varphi(y)}{K^2}\, \Big|\, \varphi - 1|x\rangle + 1|y\rangle \Big\rangle \\
\mathit{PU}(\varphi) &\coloneqq \sum_{x\in\mathrm{supp}(\varphi)} \sum_{y\in X} \frac{\varphi(x)}{K\cdot N}\, \Big|\, \varphi - 1|x\rangle + 1|y\rangle \Big\rangle \\
\mathit{UP}(\varphi) &\coloneqq \sum_{x\in\mathrm{supp}(\varphi)} \sum_{y\in X} \frac{\varphi(y)}{|\mathrm{supp}(\varphi)|\cdot K}\, \Big|\, \varphi - 1|x\rangle + 1|y\rangle \Big\rangle \\
\mathit{UU}(\varphi) &\coloneqq \sum_{x\in\mathrm{supp}(\varphi)} \sum_{y\in X} \frac{1}{|\mathrm{supp}(\varphi)|\cdot N}\, \Big|\, \varphi - 1|x\rangle + 1|y\rangle \Big\rangle.
\end{aligned} \tag{5}$$

$$PP = \begin{pmatrix} 1 & 4/25 & 0 & 0 & 0 & 0 \\ 0 & 17/25 & 6/25 & 0 & 0 & 0 \\ 0 & 4/25 & 13/25 & 6/25 & 0 & 0 \\ 0 & 0 & 6/25 & 13/25 & 4/25 & 0 \\ 0 & 0 & 0 & 6/25 & 17/25 & 0 \\ 0 & 0 & 0 & 0 & 4/25 & 1 \end{pmatrix} \qquad PU = \begin{pmatrix} 1/2 & 1/10 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 1/5 & 0 & 0 & 0 \\ 0 & 2/5 & 1/2 & 3/10 & 0 & 0 \\ 0 & 0 & 3/10 & 1/2 & 2/5 & 0 \\ 0 & 0 & 0 & 1/5 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/10 & 1/2 \end{pmatrix}$$

$$UP = \begin{pmatrix} 1 & 2/5 & 0 & 0 & 0 & 0 \\ 0 & 1/2 & 3/10 & 0 & 0 & 0 \\ 0 & 1/10 & 1/2 & 1/5 & 0 & 0 \\ 0 & 0 & 1/5 & 1/2 & 1/10 & 0 \\ 0 & 0 & 0 & 3/10 & 1/2 & 0 \\ 0 & 0 & 0 & 0 & 2/5 & 1 \end{pmatrix} \qquad UU = \begin{pmatrix} 1/2 & 1/4 & 0 & 0 & 0 & 0 \\ 1/2 & 1/2 & 1/4 & 0 & 0 & 0 \\ 0 & 1/4 & 1/2 & 1/4 & 0 & 0 \\ 0 & 0 & 1/4 & 1/2 & 1/4 & 0 \\ 0 & 0 & 0 & 1/4 & 1/2 & 1/2 \\ 0 & 0 & 0 & 0 & 1/4 & 1/2 \end{pmatrix}$$

Figure 1 Examples of transition matrices for the four draw-and-recolour operations (5).

When we take $X = \{a, b\}$ and $K = 5$ we get a set $\mathcal{M}[5]\big(\{a,b\}\big)$ of multisets with $\big(\!\binom{2}{5}\!\big) = 6$ elements, namely $\mathcal{M}[5]\big(\{a,b\}\big) = \big\{\varphi_1, \ldots, \varphi_6\big\}$ where:

$$\begin{array}{lll} \varphi_1 = 5|a\rangle & \varphi_2 = 4|a\rangle + 1|b\rangle & \varphi_3 = 3|a\rangle + 2|b\rangle \\ \varphi_4 = 2|a\rangle + 3|b\rangle & \varphi_5 = 1|a\rangle + 4|b\rangle & \varphi_6 = 5|b\rangle. \end{array} \tag{6}$$

The stochastic 6×6 transition matrices associated with the four channels in (5) are described in Figure 1, using this order of multisets. We see a similar "tridiagonal" structure, with zeroes on the upper-right and lower-left, but with different probabilities inbetween. Composition of channels $PP \circ PP$ can equivalently be expressed via product of the corresponding matrices. We prefer to use channel composition since it gives an explicit handle on the inputs and outputs.

4 Long-term behaviour

In this section, we prove properties about fixed points and long-term behaviour of the models introduced in Section 3. Table 1 summarises some of these properties. The stationary distributions in the first two cases are of the same absorbing kind. The stationary distributions in the last two cases are new distributions. They were not found as eigenvectors via the solution of a linear equations. Their form was recognised in experiments with our own implementation. Below, proofs are provided that they are stationary indeed.

4.1 The uniform-proportional (UP) case

For this transition channel UP absorbing states play an important role.

▶ **Theorem 9.** *Consider for a number $K \in \mathbb{N}$ and a non-empty finite set X the uniform-proportional transition channel $UP\colon \mathcal{M}[K](X) \to \mathcal{M}[K](X)$ as Markov chain.*

1. *This UP chain has all K-point multisets of the form $K|x\rangle \in \mathcal{M}[K](X)$, for $x \in X$, as absorbing states.*
2. *Each state (multiset) $\varphi \in \mathcal{M}[K](X)$ can reach such an absorbing state.*
3. *A distribution $\omega \in \mathcal{D}\big(\mathcal{M}[K](X)\big)$ is stationary for UP if and only if it is of the form $\omega = \sum_x r_x \,\big|K|x\rangle\big\rangle$, for suitable probabilities $r_x \in [0,1]$ summing up to 1.*

Table 1 Summary of the properties of the Markov chains introduced in Section 3.

Markov-chain	Irreducible?	Aperiodic?	Absorbing?	Stationary distribution(s)
PP	No	Yes	Yes	$\sum_{x\in X} r_x \,\big\lvert K\lvert x\rangle\big\rangle$
UP	No	Yes	Yes	$\sum_{x\in X} r_x \,\big\lvert K\lvert x\rangle\big\rangle$
UU	Yes	Yes	No	$\sum_{\varphi\in\mathcal{M}[K](X)} \frac{\big\lvert \mathrm{supp}(\varphi)\big\rvert}{\mathcal{S}(N,K)} \big\lvert \varphi\big\rangle$
PU	Yes	Yes	No	$\sum_{\varphi\in\mathcal{M}[K](X)} \frac{(\varphi)}{N^K} \big\lvert \varphi\big\rangle$

Proof.

1. It is immediate that $UP\big(K|x\rangle\big) = 1\big|K|x\rangle\big\rangle$ since $\mathrm{supp}(K|x\rangle) = \{x\}$.
2. The fact that these absorbing states can be reached is "obvious" when one inspects the definition of the channel UP in (5), since the multisets become more "spread out". We illustrate this dynamics for the multisets $\varphi_i \in \mathcal{M}[5]\big(\{a,b\}\big)$ in (6), where $\varphi_1 = 5|a\rangle$ and $\varphi_6 = 5|b\rangle$ are absorbing. For $\varphi_3 = 3|a\rangle + 2|b\rangle$ one has:
$$\begin{aligned} UP(\varphi_3) &= \tfrac{3}{10}\big|\varphi_2\big\rangle + \tfrac{1}{2}\big|\varphi_3\big\rangle + \tfrac{1}{5}\big|\varphi_4\big\rangle \\ \big(UP \circ UP\big)(\varphi_3) &= \tfrac{3}{25}\big|\varphi_1\big\rangle + \tfrac{3}{10}\big|\varphi_2\big\rangle + \tfrac{8}{25}\big|\varphi_3\big\rangle + \tfrac{1}{5}\big|\varphi_4\big\rangle + \tfrac{3}{50}\big|\varphi_5\big\rangle. \end{aligned}$$
So after two steps one reaches the absorbing state φ_1. After another step also φ_6 is reached. From then on the probabilities for φ_1 and φ_6 increase, while the probabilities for the other, non-absorbing states/multisets decrease.
3. This is a consequence of the previous item, and Proposition 8. Clearly, if $\omega = \sum_x r_x \,\big|K|x\rangle\big\rangle$ then:
$$UP_*(\omega) = \sum_{x\in X} r_x \cdot UP\big(K|x\rangle\big) = \sum_{x\in X} r_x \cdot 1\big|K|x\rangle\big\rangle = \sum_{x\in X} r_x\big|K|x\rangle\big\rangle = \omega.$$
In the other direction, in iterations $UP^n(\varphi)$, for arbitrary $\varphi \in \mathcal{M}[K](X)$, probabilities accumulate at the absorbing states $K|x\rangle$, as illustrated in the previous point. ◀

An intriguing question is if we can determine the probabilities r_x in the stationary distribution $\sum_x r_x\big|K|x\rangle\big\rangle$ that is approached by an arbitrary $\varphi \in \mathcal{M}[K](X)$. We have no closed formula, but we do have a method for determining these probabilities. We first illustrate the method for the multisets $\varphi_i \in \mathcal{M}[5]\big(\{a,b\}\big)$ from (6). Subsequently, we give a stepwise description of the general approach.

Let us write $r_i \in [0,1]$ for $\lim_n UP^n(\varphi_i) = r_i\big|5|a\rangle\big\rangle + (1-r_i)\big|5|b\rangle\big\rangle$. Then $r_1 = 1$ and $r_6 = 0$, since $\varphi_1 = 5|a\rangle$ and $\varphi_6 = 5|b\rangle$. For the four other cases $\varphi_1, \varphi_2, \varphi_3, \varphi_4$ we can first consider the transition equations, corresponding to the lower-left matrix in Figure 1.

$$\begin{aligned} UP(\varphi_2) &= \tfrac{2}{5}\big|\varphi_1\big\rangle + \tfrac{1}{2}\big|\varphi_2\big\rangle + \tfrac{1}{10}\big|\varphi_3\big\rangle & \qquad UP(\varphi_4) &= \tfrac{1}{5}\big|\varphi_3\big\rangle + \tfrac{1}{2}\big|\varphi_4\big\rangle + \tfrac{3}{10}\big|\varphi_5\big\rangle. \\ UP(\varphi_3) &= \tfrac{3}{10}\big|\varphi_2\big\rangle + \tfrac{1}{2}\big|\varphi_3\big\rangle + \tfrac{1}{5}\big|\varphi_4\big\rangle. & \qquad UP(\varphi_5) &= \tfrac{1}{10}\big|\varphi_4\big\rangle + \tfrac{1}{2}\big|\varphi_5\big\rangle + \tfrac{2}{5}\big|\varphi_6\big\rangle. \end{aligned}$$

From these equations we can deduce equations for the probabilities r_1, r_2, r_3, r_4 in the stationary distributions for $\varphi_1, \varphi_2, \varphi_3, \varphi_4$, namely:

$$\begin{aligned} r_2 &= \tfrac{2}{5}\cdot 1 + \tfrac{1}{2}\cdot r_2 + \tfrac{1}{10}\cdot r_3 & \qquad r_4 &= \tfrac{1}{5}\cdot r_3 + \tfrac{1}{2}\cdot r_4 + \tfrac{3}{10}\cdot r_5. \\ r_3 &= \tfrac{3}{10}\cdot r_2 + \tfrac{1}{2}\cdot r_3 + \tfrac{1}{5}\cdot r_4. & \qquad r_5 &= \tfrac{1}{10}\cdot r_4 + \tfrac{1}{2}\cdot r_5 + \tfrac{2}{5}\cdot 0. \end{aligned}$$

Solving these equations yields: $r_2 = \frac{15}{16}$, $r_3 = \frac{11}{16}$, $r_4 = \frac{5}{16}$, $r_5 = \frac{1}{16}$.

There is a general approach behind this, which we briefly sketch. For details, see [8, Chapter III]. We assume a set X with $N := |X| > 0$ elements and a state space $\mathcal{M}[K](X)$ of multisets of size $K \in \mathbb{N}$. We assume that these multisets in $\mathcal{M}[K](X)$ are somehow available as list $(\varphi_1, \ldots, \varphi_M)$ for multichoose coefficient $M = \left(\!\binom{N}{K}\!\right)$, corresponding to the entries – see Lemma 1 (1).

- Trim the matrix of UP, as in Figure 1, to UP^- by removing all rows and columns corresponding to absorbing states $K|x\rangle$. The matrix UP is $M \times M$ and UP^- is $(M - N) \times (M - N)$.
- For each $x \in X$ form the matrix equation $UP^- \cdot V + C_x = V$, where $V = (v_1, \ldots v_{M-N})$ is a vector of variables and C_x is the (deleted, trimmed) column in UP corresponding to the absorbing multiset $K|x\rangle$.
- Solve these equations, via matrix inversion, with solutions $V_x = (UP^- - I)^{-1}(-C_x)$ for each $x \in X$, where I is the diagonal (identity) matrix of size $(M - N) \times (M - N)$.
- Augment these lists V_x of length $M - N$ by re-introducing entries for absorbing multisets: in V_x only at the position of $K|x\rangle$ a 1 is added, and a 0 for all other absorbing multisets. The resulting vectors of length M are called W_x.
- By construction, $\lim_n UP^n(\varphi_i) = \sum_x W_{x,i} \big| K|x\rangle \big\rangle$.

We have implemented this procedure, in Python. It yields, for instance,

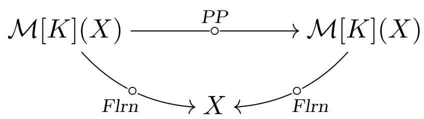
$$\lim_{n\to\infty} UP^n\Big(1|a\rangle + 2|b\rangle + 3|c\rangle\Big) \;=\; \tfrac{17}{216}\big|6|a\rangle\big\rangle + \tfrac{41}{144}\big|6|b\rangle\big\rangle + \tfrac{275}{432}\big|6|c\rangle\big\rangle.$$

The outcomes do not yield distributions that we recognise.

4.2 The proportional-proportional (PP) case

We begin with a result that distinguishes the PP transition channel from the other three.

▶ **Lemma 10.** *The following diagram of channels commutes.*



Proof. For a multiset $\varphi \in \mathcal{M}[K](X)$ we get:

$$\begin{aligned}
&\mathit{Flrn}_*(PP(\varphi)) \\
&= \sum_{z\in X} \sum_{x\in \mathrm{supp}(\varphi)} \sum_{y\in X} \frac{\varphi(x)\cdot\varphi(y)}{K^2} \cdot \frac{(\varphi - 1|x\rangle + 1|y\rangle)(z)}{K} \big| z \big\rangle \\
&= \frac{1}{K^3} \cdot \sum_{z\in X} \sum_{x\in \mathrm{supp}(\varphi)} \sum_{y\in X} \Big(\varphi(x)\cdot\varphi(y)\cdot(\varphi - 1|x\rangle)(z) \;+\; \varphi(x)\cdot\varphi(y)\cdot 1|y\rangle(z)\Big) \big| z \big\rangle \\
&= \frac{1}{K^3} \cdot \sum_{z\in X} K \cdot \left(\sum_{x\in \mathrm{supp}(\varphi)} \varphi(x)\cdot(\varphi - 1|x\rangle)(z) \right) \;+\; K\cdot\varphi(z) \big| z \big\rangle \\
&= \frac{1}{K^2} \cdot \sum_{z\in X} \left(\sum_{x\in \mathrm{supp}(\varphi),\, x\neq z} \varphi(x)\cdot\varphi(z) \right) + \varphi(z)\cdot(\varphi(z) - 1) \;+\; \varphi(z) \big| z \big\rangle \\
&= \frac{1}{K^2} \cdot \sum_{z\in X} (K - \varphi(z))\cdot\varphi(z) + \varphi(z)^2 \big| z \big\rangle \;=\; \frac{1}{K} \cdot \sum_{z\in X} \varphi(z) \big| z \big\rangle \;=\; \mathit{Flrn}(\varphi).
\end{aligned}$$ ◀

We are now ready to describe the long-term behaviour of the *PP* process. Observe that once all balls are of the same colour in the urn, the content of the urn cannot change anymore. From this, one can easily see that K-point multisets are absorbing states, and the Markov chain *PP* is absorbing. Therefore, from each inital urn, an absorbing state is reached with probability one in an infinite run. In economic terms, this means that one person is guaranteed to accumulate all wealth in the long term. The situation is summarised in the following theorem. The first three points are exactly as in Theorem 9.

▶ **Theorem 11.** *Let $K \in \mathbb{N}$ be a number, and X a non-empty, finite set. Consider the proportional-proportional transition channel $PP\colon \mathcal{M}[K](X) \rightsquigarrow \mathcal{M}[K](X)$ as Markov chain.*

1. *This PP chain has N absorbing states, namely $K|x\rangle$ for $x \in X$.*
2. *Each state (multiset) $\varphi \in \mathcal{M}[K](X)$ can reach such an absorbing state.*
3. *A distribution $\omega \in \mathcal{D}\big(\mathcal{M}[K](X)\big)$ is stationary for PP if and only if it is of the form $\omega = \sum_x r_x \,\big|K|x\rangle\big\rangle$, for suitable probabilities $r_x \in [0,1]$ summing up to 1.*
4. *Let $\varphi \in \mathcal{M}[K](X)$. Then:*
$$\lim_{n\to\infty} \mathit{Flrn}_*\Big(PP^n(\varphi)\Big) \;=\; \mathit{Flrn}(\varphi).$$
5. *Let $\varphi \in \mathcal{M}[K](X)$. Then:*
$$\lim_{n\to\infty} PP^n(\varphi) \;=\; \sum_{x\in X} \frac{\varphi(x)}{K}\,\big|K|x\rangle\big\rangle \;=\; \sum_{x\in X} \mathit{Flrn}(\varphi)(x)\,\big|K|x\rangle\big\rangle.$$
That is, the absorbing probabilities are proportional to the number of elements in the initial multiset.

Proof.

1. It is immediate that $PP\big(K|x\rangle\big) = 1\big|K|x\rangle\big\rangle$ since $\mathrm{supp}(K|x\rangle) = \{x\}$.
2. Immediate from the definition of *PP*. The reader may draw up a similar illustration as we did for the *UP* case.
3. This is a consequence of the previous point, and Proposition 8.
4. By Lemma 10.
5. Let $\varphi \in \mathcal{M}[K](X)$. The limit $\lim_n PP^n(\varphi)$ exists, because *PP* is an absorbing Markov chain, see Proposition 8. We also know that the limiting distribution is supported by absorbing states. Therefore, it is of the form
$$\lim_{n\to\infty} PP^n(\varphi) \;=\; \sum_{x\in X} r_x\,\big|K|x\rangle\big\rangle.$$
We calculate as follows.
$$\mathit{Flrn}_*\Big(\lim_{n\to\infty} PP^n(\varphi)\Big) \;=\; \lim_{n\to\infty} \mathit{Flrn}_*\Big(PP^n(\varphi)\Big) \;=\; \mathit{Flrn}(\varphi)$$
The first equality is an instance of Lemma 2. The second equality is Item 4.
We then have:
$$\mathit{Flrn}(\varphi) \;=\; \mathit{Flrn}_*\left(\sum_{x\in X} r_x\,\big|K|x\rangle\big\rangle\right) \;=\; \sum_{x\in X} r_x\,\big|\mathit{Flrn}\big(K|x\rangle\big)\big\rangle \;=\; \sum_{x\in X} r_x|x\rangle.$$ ◀

To finish this subsection, we briefly relate the *PP* chain to the Moran process, a well-known model in evolutionary biology [11, 12]. The Moran process describes a situation, in which two alleles (genetic variations) compete with each other. Let $X = \{A, B\}$ be the set of the two alleles. For a natural number K, the genetic variation in a K-sized population can be encoded as a K-sized multiset $\varphi \in \mathcal{M}[K](X)$. In each step of the Moran process, a random individual is selected for death, and at the same time, a random individual reproduces itself by cloning. The behaviour of the process may be further influenced by a so-called fitness parameter. The *PP* process for $N = 2$ is a special case of this Moran process, where the the fitness is independent of φ, and none of the alleles have a fitness advantage. This case is known as **neutral drift** in evolutionary biology.

Moran processes have been generalised to three or more competing genetic variations [2]. The *PP* process for $N > 2$ is again an instance of this generalised Moran process.

4.3 The uniform-uniform (UU) case

We start with a new result, characterising the sum of the sizes of the supports of multisets.

▶ **Proposition 12.** *Let a number $K \in \mathbb{N}$ and a finite set X be given, of size $N := |X| > 0$.*

1. *One can describe the sum of support sizes via a function $\mathcal{S}$, as:*

$$\sum_{\varphi \in \mathcal{M}[K](X)} \big|\mathrm{supp}(\varphi)\big| = \mathcal{S}(N, K) := \begin{cases} 0 & \text{if } K = 0 \\ N \cdot \left(\!\!\binom{N}{K-1}\!\!\right) & \text{otherwise.} \end{cases} \tag{7}$$

2. *Further, for each element $x \in X$,*

$$\sum_{\varphi \in \mathcal{M}[K](X)} \big|\mathrm{supp}(\varphi)\big| \cdot \varphi(x) = \frac{K}{N} \cdot \mathcal{S}(N, K).$$

Proof.

1. The equation obviously holds for $K = 0$, so we may assume $K \geq 1$. We use induction on $N \geq 1$. When $N = 1$, the left-hand-side is equal to 1, and so is the right-hand-side. We now reason as follows. Let $X = \{x_1, \ldots, x_N\}$ and let $y \notin X$. Then:

$$\begin{aligned}
&\sum_{\varphi \in \mathcal{M}[K](X \cup \{y\})} \big|\mathrm{supp}(\varphi)\big| \\
&= \sum_{0 \leq m \leq K} \; \sum_{\varphi \in \mathcal{M}[K-m](X)} \big|\mathrm{supp}(\varphi + m|y\rangle)\big| \\
&= \sum_{\varphi \in \mathcal{M}[K](X)} \big|\mathrm{supp}(\varphi)\big| + \sum_{1 \leq m \leq K} \; \sum_{\varphi \in \mathcal{M}[K-m](X)} 1 + \big|\mathrm{supp}(\varphi)\big| \\
&\overset{\text{(IH)}}{=} N \cdot \left(\!\!\binom{N}{K-1}\!\!\right) + \sum_{1 \leq m \leq K} \left(\!\!\binom{N}{K-m}\!\!\right) + \sum_{\varphi \in \mathcal{M}[K-m](X)} \big|\mathrm{supp}(\varphi)\big| \\
&\overset{\text{(IH)}}{=} N \cdot \left(\!\!\binom{N}{K-1}\!\!\right) + \sum_{1 \leq m \leq K} \left(\!\!\binom{N}{K-m}\!\!\right) + \sum_{1 \leq m \leq K-1} N \cdot \left(\!\!\binom{N}{K-m-1}\!\!\right) \\
&= N \cdot \left(\!\!\binom{N}{K-1}\!\!\right) + \sum_{0 \leq m \leq K-1} \left(\!\!\binom{N}{(K-1)-m}\!\!\right) + N \cdot \sum_{0 \leq m \leq K-2} \left(\!\!\binom{N}{(K-2)-m}\!\!\right) \\
&\overset{(1)}{=} N \cdot \left(\!\!\binom{N}{K-1}\!\!\right) + \left(\!\!\binom{N+1}{(K-1)-m}\!\!\right) + N \cdot \left(\!\!\binom{N+1}{K-2}\!\!\right) \\
&= N \cdot \left[\left(\!\!\binom{N}{K-1}\!\!\right) + \left(\!\!\binom{N+1}{K-2}\!\!\right) \right] + \left(\!\!\binom{N+1}{K-1}\!\!\right) \\
&\overset{(1)}{=} N \cdot \left(\!\!\binom{N+1}{K-1}\!\!\right) + \left(\!\!\binom{N+1}{K-1}\!\!\right) \\
&= (N+1) \cdot \left(\!\!\binom{N+1}{K-1}\!\!\right).
\end{aligned}$$

2. The sum on the left-hand-side of the equation is the same for each element x, so that by the previous point:

$$\begin{aligned}\sum_{\varphi\in\mathcal{M}[K](X)} \big|\mathrm{supp}(\varphi)\big| \cdot \varphi(x) &= \frac{1}{N}\cdot\sum_{x\in X}\sum_{\varphi\in\mathcal{M}[K](X)} \big|\mathrm{supp}(\varphi)\big|\cdot\varphi(x) \\ &= \frac{1}{N}\cdot\sum_{\varphi\in\mathcal{M}[K](X)} \big|\mathrm{supp}(\varphi)\big|\cdot\sum_{x\in X}\varphi(x) \\ &= \frac{1}{N}\cdot\sum_{\varphi\in\mathcal{M}[K](X)} \big|\mathrm{supp}(\varphi)\big|\cdot K \;=\; \frac{K}{N}\cdot\mathcal{S}(N,K).\end{aligned}$$ ◀

▶ Remark 13. The support size function $\mathcal{S}$ defined in (7) can be described more explicitly, for $K \geq 1$, as:

$$\begin{aligned}\mathcal{S}(N,K) \;=\; N\cdot\left(\!\!\binom{N}{K-1}\!\!\right) &= N\cdot\frac{(N+K-2)!}{(K-1)!\cdot(N-1)!} \\ &= \frac{N^2}{(K-1)!}\cdot(N+1)\cdot\ldots\cdot(N+K-2) \\ &= \frac{N^2}{(K-1)!}\cdot\prod_{1\leq i<K-1} N+i.\end{aligned}$$

For $N = 3$ this yields the sequence $\mathcal{S}(3,-)$ of so-called triangular matchstick numbers[1] 0, 3, 9, 18, 30, 45, 63, 84, 108, 135, 165, 198, 234, 273, 315, ... For $N > 3$ the sequence $\mathcal{S}(N,-)$ does not seem to be known.

▶ **Theorem 14.** *Let $K \in \mathbb{N}$ and let X be non-empty finite set of size N. Define the distribution $\mathrm{st}_{UU} \in \mathcal{D}\big(\mathcal{M}[K](X)\big)$ as:*

$$\mathrm{st}_{UU} \;\coloneqq\; \sum_{\varphi\in\mathcal{M}[K](X)} \frac{|\,\mathrm{supp}(\varphi)\,|}{\mathcal{S}(N,K)}\,\big|\varphi\big\rangle.$$

By Proposition 12 (1) this st_{UU} is indeed a distribution.

1. *st_{UU} is a stationary distribution for the channel UU.*
2. *The process UU is an aperiodic, irreducible Markov chain. Therefore, st_{UU} is the unique stationary distribution of UU. Moreover, UU converges to this stationary distribution. That is, for every initial multiset $\varphi \in \mathcal{M}[K](X)$*

$$\lim_{n\to\infty} UU^n(\varphi) \;=\; \mathrm{st}_{UU}.$$

3. *$Flrn_*(\mathrm{st}_{UU})$ is the uniform distribution $\sum_{x\in X}\frac{1}{N}|x\rangle$ on X.*

Proof.

1. We calculate:

$$\begin{aligned}UU_*(\mathrm{st}_{UU}) &= \sum_{\varphi\in\mathcal{M}[K](X)} \mathrm{st}_{UU}(\varphi)\cdot UU(\varphi) \\ &= \sum_{\varphi\in\mathcal{M}[K](X)}\sum_{x\in\mathrm{supp}(\varphi)}\sum_{y\in X} \frac{|\,\mathrm{supp}(\varphi)\,|}{\mathcal{S}(N,K)}\cdot\frac{1}{|\,\mathrm{supp}(\varphi)\,|\cdot N}\,\Big|\varphi - 1|x\rangle + 1|y\rangle\Big\rangle \\ &= \sum_{\varphi\in\mathcal{M}[K](X)} \frac{1}{\mathcal{S}(N,K)}\cdot\sum_{x\in\mathrm{supp}(\varphi)}\sum_{y\in X}\frac{1}{N}\,\Big|\varphi - 1|x\rangle + 1|y\rangle\Big\rangle \\ &\overset{(*)}{=} \sum_{\psi\in\mathcal{M}[K](X)} \frac{|\,\mathrm{supp}(\psi)\,|}{\mathcal{S}(N,K)}\,\big|\psi\big\rangle \;=\; \mathrm{st}_{UU}.\end{aligned}$$

[1] Also called sequence A045943, see oeis.org/A045943.

For the marked equation on the last line, notice that one can write a multiset $\psi \in \mathcal{M}[K](X)$ as $\psi = \varphi - 1|x\rangle + 1|y\rangle$ in $N \cdot |\text{supp}(\psi)|$-many ways. This is because for y can make $|\text{supp}(\psi)|$-many choices and N-many for $x \in X$.

2. It is clear from the definition of UU that any state can reach any other state. Aperiodicity follows from the fact that every state can transition to itself. For the convergence result, we apply Proposition 7.
3. Via Proposition 12 (2) we obtain the uniform distribution on X.

$$\begin{aligned} \mathit{Flrn}_*(\text{st}_{UU}) &= \sum_{x\in X} \sum_{\varphi\in\mathcal{M}[K](X)} \text{st}_{UU}(\varphi)\cdot \mathit{Flrn}(\varphi)(x)\,\big|\,x\,\big\rangle \\ &= \sum_{x\in X} \sum_{\varphi\in\mathcal{M}[K](X)} \frac{|\text{supp}(\varphi)|}{\mathcal{S}(N,K)}\cdot\frac{\varphi(x)}{K}\,\big|\,x\,\big\rangle \\ &= \sum_{x\in X} \frac{1}{\mathcal{S}(N,K)\cdot K}\cdot \sum_{\varphi\in\mathcal{M}[K](X)} \big|\text{supp}(\varphi)\big|\cdot\varphi(x)\,\big|\,x\,\big\rangle \\ &= \sum_{x\in X} \frac{1}{\mathcal{S}(N,K)\cdot K}\cdot\frac{K}{N}\cdot\mathcal{S}(N,K)\,\big|\,x\,\big\rangle \;=\; \sum_{x\in X}\frac{1}{N}\,|\,x\,\rangle. \end{aligned}$$ ◀

4.4 The proportional-uniform (PU) case

Similarly to the previous case, the hidden Markov model associated to PU converges to a uniform distribution. The next theorem is analogous to Theorem 14.

▶ **Theorem 15.** *Let $K \in \mathbb{N}$ be a number, and let X be non-empty finite set of size N. Define the distribution $\text{st}_{PU} \in \mathcal{D}\big(\mathcal{M}[K](X)\big)$ as:*

$$\text{st}_{PU} \;\coloneqq\; \sum_{\varphi\in\mathcal{M}[K](X)} \frac{(\varphi)}{N^K}\,\big|\,\varphi\,\big\rangle.$$

By Item (4) of Lemma 1, this is well-defined. Then:

1. st_{PU} *is a stationary distribution for the channel PU.*
2. *The Markov-chain PU is aperiodic and irreducible, so that st_{PU} is the unique stationary distribution of PU. Moreover, PU converges to its stationary distribution. That is, for every initial multiset $\varphi \in \mathcal{M}[K](X)$*

$$\lim_{n\to\infty} PU^n(\varphi) \;=\; \text{st}_{PU}.$$

3. $\mathit{Flrn}_*(\text{st}_{PU})$ *is the uniform distribution* $\sum_{x\in X}\frac{1}{N}|x\rangle$ *on X.*

Proof.

1. We calculate:

$$
\begin{aligned}
PU_*(\mathrm{st}_{PU}) &= \sum_{\varphi\in\mathcal{M}[K](X)} \mathrm{st}_{PU}(\varphi)\cdot PU(\varphi) \\
&= \sum_{\varphi\in\mathcal{M}[K](X)}\ \sum_{x\in\mathrm{supp}(\varphi)}\ \sum_{y\in X} \frac{\big(\varphi\big)}{N^K}\cdot\frac{\varphi(x)}{K\cdot N}\Big|\,\varphi-1|x\rangle+1|y\rangle\Big\rangle \\
&= \sum_{\varphi\in\mathcal{M}[K](X)}\ \sum_{x\in\mathrm{supp}(\varphi)}\ \sum_{y\in X} \frac{K!}{\varphi!!\cdot N^K}\cdot\frac{\varphi(x)}{K\cdot N}\Big|\,\varphi-1|x\rangle+1|y\rangle\Big\rangle \\
&= \sum_{\varphi\in\mathcal{M}[K](X)}\ \sum_{x\in\mathrm{supp}(\varphi)}\ \sum_{y\in X} \frac{(K-1)!}{(\varphi-1|x\rangle)!!}\cdot\frac{1}{N^{K+1}}\Big|\,\varphi-1|x\rangle+1|y\rangle\Big\rangle \\
&= \sum_{\varphi\in\mathcal{M}[K](X)}\ \sum_{x\in\mathrm{supp}(\varphi)}\ \sum_{y\in X} \big(\varphi-1|x\rangle\big)\cdot\frac{1}{N^{K+1}}\Big|\,\varphi-1|x\rangle+1|y\rangle\Big\rangle \\
&= \sum_{\varphi\in\mathcal{M}[K-1](X)}\ \sum_{y\in X} N\cdot\big(\varphi\big)\cdot\frac{1}{N^{K+1}}\Big|\,\varphi+1|y\rangle\Big\rangle \\
&= \sum_{\varphi\in\mathcal{M}[K-1](X)}\ \sum_{y\in X} \frac{\big(\varphi\big)}{N^K}\Big|\,\varphi+1|y\rangle\Big\rangle \\
&= \sum_{\varphi\in\mathcal{M}[K](X)}\ \sum_{x\in\mathrm{supp}(\varphi)} \frac{\big(\varphi-1|x\rangle\big)}{N^K}\Big|\,\varphi\Big\rangle \\
&= \sum_{\varphi\in\mathcal{M}[K](X)} \frac{\big(\varphi\big)}{N^K}\big|\,\varphi\big\rangle \qquad \text{by Lemma 1 (5)} \\
&= \mathrm{st}_{PU}
\end{aligned}
$$

2. It is clear from the definition of PU that any state can reach any other state. Aperiodicity follows from the fact that every state can transition to itself. For the convergence result, we apply Proposition 7.

3. We reason as follows.

$$
\begin{aligned}
\mathit{Flrn}_*(\mathrm{st}_{PU}) &= \sum_{x\in X}\ \sum_{\varphi\in\mathcal{M}[K](X)} \mathrm{st}_{PU}(\varphi)\cdot \mathit{Flrn}(\varphi)(x)\,\big|\,x\big\rangle \\
&= \sum_{x\in X}\ \sum_{\varphi\in\mathcal{M}[K](X)} \frac{\big(\varphi\big)}{N^K}\cdot\frac{\varphi(x)}{K}\,\big|\,x\big\rangle \\
&= \sum_{x\in X}\ \sum_{\varphi\in\mathcal{M}[K](X)} \frac{1}{N^K}\cdot\frac{K!}{\varphi!!}\cdot\frac{\varphi(x)}{K}\,\big|\,x\big\rangle \\
&= \sum_{x\in X} \frac{1}{N^K}\cdot \sum_{\varphi\in\mathcal{M}[K](X),\,\varphi(x)>0} \frac{(K-1)!}{(\varphi-1|x\rangle)!!}\,\big|\,x\big\rangle \\
&= \sum_{x\in X} \frac{1}{N^K}\cdot \sum_{\varphi\in\mathcal{M}[K](X),\,\varphi(x)>0} \big(\varphi-1|x\rangle\big)\,\big|\,x\big\rangle \\
&= \sum_{x\in X} \frac{1}{N^K}\cdot \sum_{\varphi\in\mathcal{M}[K-1](X)} \big(\varphi\big)\,\big|\,x\big\rangle \\
&= \sum_{x\in X} \frac{1}{N^K}\cdot N^{K-1}\,\big|\,x\big\rangle \;=\; \sum_{x\in X}\frac{1}{N}|\,x\rangle.
\end{aligned}
$$

◀

5 Concluding remarks

This paper has used multisets as states in a systematic description of four hidden Markov models. One of these models is known from evolutionary biology, where it is usually not described in terms of multisets. The long-term behaviour of these four models has been identified, in terms of (absorbing) stationary distributions.

These results, about long-term behaviour, can be interpreted in economic terms. The *PP* and *UP* processes, corresponding to models where the state's income from taxes is redistributed proportionally to the existing wealth of individuals, lead to a monopoly with probability one. On the other hand, the *PU* and *UU* processes are expected to result in uniform distribution of wealth. While the models we presented are fairly simple, one could extend them to reason about the effects of more complicated taxation and wealth-redistribution mechanisms. We leave such possible extensions to future work.

References

1. A. Dragulescu and V.M. Yakovenko. Statistical mechanics of money. *The European Physical Journal B-Condensed Matter and Complex Systems*, 17:723–729, 2000.
2. E. Ferreira and A. Neves. Fixation probabilities for the Moran process with three or more strategies: general and coupling results. *Journal of Mathematical Biology*, 81(1):277–314, 2020. `doi:10.1007/s00285-020-01510-0`.
3. B. Jacobs. Basic combinatorics and sufficient statistics for mutations on multiple datatypes. Researchers.One, `https://researchers.one/articles/22.11.00003`, 2022.
4. B. Jacobs. Urns & tubes. *Compositionality*, 4(4), 2022. `doi:10.32408/compositionality-4-4`.
5. B. Jacobs. Drawing with distance. *Logical Methods in Comp. Sci.*, 21(2), 2025. `doi:10.46298/lmcs-21(2:11)2025`.
6. B. Jacobs. Structured probabilitistic reasoning. Book, in preparation, see `http://www.cs.ru.nl/B.Jacobs/PAPERS/ProbabilisticReasoning.pdf`, 2025.
7. N. Johnson and S. Kotz. *Urn Models and Their Application: An Approach to Modern Discrete Probability Theory*. Wiley, 1977.
8. J. Kemeny and J. Snell. *Finite Markov chains*. van Nostrand Princeton, NJ, 1969.
9. S. Kotz and N. Balakrishnan. *Advances in Urn Models during the Past Two Decades*, pages 203–257. Birkhäuser Boston, Boston, MA, 1997. `doi:10.1007/978-1-4612-4140-9_14`.
10. H. Mahmoud. *Pólya urn models*. Chapman and Hall/CRC, 2008. `doi:10.1201/9781420059847`.
11. P. Moran. Random processes in genetics. *Mathematical proceedings of the Cambridge philosophical society*, 54(1):60–71, 1958. `doi:10.1017/S0305004100033193`.
12. M. Nowak. *Evolutionary dynamics: exploring the equations of life*. Harvard Univ. Press, 2006. `doi:10.2307/j.ctvjghw98`.