Yaglom limit via Holley inequality

Pablo A. Ferrari and Leonardo T. Rolla

Universidad de Buenos Aires

Abstract. Let S be a countable set provided with a partial order and a minimal element. Consider a Markov chain on $S \cup \{0\}$ absorbed at 0 with a quasi-stationary distribution. We use Holley inequality to obtain sufficient conditions under which the following hold. The trajectory of the chain starting from the minimal state is stochastically dominated by the trajectory of the chain starting from any probability on S, when both are conditioned to non-absorption until a certain time. Moreover, the Yaglom limit corresponding to this deterministic initial condition is the unique minimal quasi-stationary distribution in the sense of stochastic order. As an application, we provide new proofs to classical results in the field.

1 Introduction

Let S be a countable set, 0 be an element outside S and consider a Markov chain $X_n \in S \cup \{0\}$ with transition matrix $Q(x, y)$, $x, y \in S \cup \{0\}$. We assume that the matrix Q yields a chain which is *absorbed at* 0, meaning that $Q(0, 0) = 1$. We assume also $Q(x, 0) < 1$ for all $x \in S$.

For a probability measure ν on S, we define νT_n as the conditional distribution at time n of the chain started with law ν given that it is not absorbed until time n. More precisely,

$$\nu T_n(y) := \frac{\nu Q^n(y)}{1 - \nu Q^n(0)}, \qquad y \in S. \tag{1.1}$$

A probability measure ν is a *quasi-stationary distribution* (or simply *q.s.d.*) if $\nu = \nu T_1$ (and thus $\nu T_n = \nu$ for all $n \geq 1$). Rewriting (1.1), a measure ν is a q.s.d. if and only if

$$\nu(y) = \sum_{x \in S} \nu(x)\big[Q(x, y) + Q(x, 0)\nu(y)\big], \qquad y \in S. \tag{1.2}$$

The *Yaglom limit* of ν is $\lim_n \nu T_n$, if the limit exists and is a probability measure. In this case, the limit is called a *quasi-limiting distribution*.

Assume that S has a partial order denoted $\leq$ and a minimal state called 1. Let $\nu \preceq \nu'$ denote the stochastic order of measures on S induced by $\leq$.

We say that a probability measure μ on S^n is *irreducible* if the set $\{(x_1, \ldots, x_n) \in S^n : \mu(x_1, \ldots, x_n) > 0\}$ is connected in the sense that any element of S^n with positive μ-probability can be reached from any other via successive coordinate changes without passing through elements with zero μ-probability. We say that a Markov chain on $S \cup \{0\}$ with transition matrix Q and initial distribution ν on S has *irreducible trajectories in* S if for each $n \geq 1$ the measure μ on S^{n+1} defined by $\mu(x_0, \ldots, x_n) := \nu(x_0) Q(x_0, x_1) \cdots Q(x_{n-1}, x_n)$ is irreducible.

Let δ_x be the probability distribution on S concentrated on the state $x \in S$.

Theorem 1. *Let S be a partially ordered countable space with a minimal element called* 1. *Let Q be the transition matrix of a Markov chain on $S \cup \{0\}$ absorbed at* 0. *Assume that the chain with initial distribution δ_1 has irreducible trajectories in S. If, for all $x, x', z, z' \in S$ with $x \leq x'$, $z \leq z'$, whenever the denominators are positive,*

$$\frac{Q(x, \cdot) Q(\cdot, z)}{Q^2(x, z)} \preceq \frac{Q(x', \cdot) Q(\cdot, z')}{Q^2(x', z')}, \tag{1.3}$$

$$\frac{Q(x, \cdot)}{1 - Q(x, 0)} \preceq \frac{Q(x', \cdot)}{1 - Q(x', 0)}, \tag{1.4}$$

as probability measures on S, then the following hold:

   (i) *The sequence $(\delta_1 T_n)_{n \geq 1}$ is monotone*: $\delta_1 T_n \preceq \delta_1 T_{n+1}$, *for all $n \geq 0$.*

  (ii) *For any probability ν on S, $\delta_1 T_n \preceq \nu T_n$.*

 (iii) *In particular, if ν is a q.s.d., then $\delta_1 T_n \preceq \nu$, for all $n \geq 0$.*

 (iv) *If there is a q.s.d. for Q, then the Yaglom limit of δ_1 converges. The limit distribution $\underline{\nu} := \lim_n \delta_1 T_n$ is a q.s.d. and satisfies $\underline{\nu} \preceq \nu$ for any other q.s.d. ν.*

The proof of Theorem 1 is an application of Holley inequality in the space of finite-length trajectories of the chain. Roughly speaking, Holley inequality says that the local dominations (1.3) and (1.4) imply that the conditional law of a length-n trajectory of the chain starting with δ_1 given nonabsorption by time n is stochastically dominated by the conditional law of a length-n trajectory of the chain starting from any other measure ν.

In Section 2, we state and prove Holley inequality, and use it to prove Theorem 1. We then give a sufficient condition for $\underline{\nu}$ to be the minimal q.s.d. in the sense of absorption time rather than stochastic domination.

Convergence of the Yaglom limit has been studied for birth-and-death chains and one-dimensional random walks in both continuous and discrete time. Theorem 1 gives an alternative proof to many of these classical results.

References to previous works and details of our approach to the aperiodic cases are discussed in detail in Section 3. Periodic chains are discussed in Section 4, after condition (1.3) is relaxed so as to include this case.

There is a large literature on q.s.d.'s compiled and periodically updated by Pollett (2014). We quote the recent book of Collet, Martínez and San Martín (2013) and the work of Kesten (1995) on Yaglom limits of discrete-time Markov chains.

2 Yaglom limit via Holley inequality

2.1 Trajectory distribution

For integers $n < m$ let $\mathcal{X}_n^m := \{(x_n, \ldots, x_m) : x_k \in S \text{ for } n \leq k \leq m\}$ be the set of possible trajectories of the chain with transition matrix Q in the time interval $[n, m]$ which are not absorbed by 0 in that time interval.

Let ν be a probability measure on S and define the measure $\mu_n^m(\nu, Q)$ on $\mathcal{X}_n^m$ by

$$\mu_n^m(\nu, Q)(\mathbf{x}_n^m) := \frac{\nu(x_n) Q(x_n, x_{n+1}) \cdots Q(x_{m-1}, x_m)}{1 - \nu Q^n(0)}, \tag{2.1}$$

where $\mathbf{x}_n^m = (x_n, \ldots, x_m)$. The measure $\mu_n^m(\nu, Q)$ is the conditional distribution of the chain $X_n^m = (X_n, \ldots, X_m)$ with initial distribution ν at time n and transition probabilities Q, given that the chain is not absorbed during the time interval $[n, m]$.

Due to the conditioning, the first marginal of $\mu_n^m(\nu, Q)$ is not ν in general, but its last marginal is νT_{m-n}. Indeed, by (1.1) and (2.1),

$$\nu T_{m-n}(y) = \sum_{(x_n, \ldots, x_{m-1})} \mu_n^m(\nu, Q)(x_n, \ldots, x_{m-1}, y). \tag{2.2}$$

2.2 Holley inequality

Let Ω be a set endowed with a partial order $\leq$. Let μ, μ' be probability measures on Ω. The stochastic domination $\mu \preceq \mu'$ is equivalent to the existence of a measure $\tilde{\mu}$ on $\Omega \times \Omega$ with marginals μ and μ' such that $\tilde{\mu}((\omega, \omega') : \omega \leq \omega') = 1$, see, for instance, Lindvall (1999). In this case, we say that $\tilde{\mu}$ is a *monotone coupling* of μ and μ'.

Let us endow $\mathcal{X}_n^m$ with the partial order given by the coordinate-wise order of trajectories: $\mathbf{x}_n^m \leq \mathbf{y}_n^m$ if $x_k \leq y_k$ for all $k \in [n, m]$.

A local domination condition for global domination of measures is provided by Holley Inequality (Holley (1974)). Here is a version suited to our context.

Proposition 2 (Holley inequality). *Let S be a partially ordered countable space. Let $\nu \preceq \nu'$ be probabilities on S and let Q, Q' be transition matrices on $S \cup \{0\}$ absorbed at 0. Denote the conditional laws of trajectories by $\mu = \mu_n^m(\nu, Q)$ and $\mu' = \mu_n^m(\nu', Q')$, respectively. Assume that μ is an irreducible probability on the space of trajectories $\mathcal{X}_n^m$. If, for all $x, x', z, z' \in S$ with $x \leq x', z \leq z'$, whenever*

the denominators are positive,

$$\frac{\nu(\cdot)Q(\cdot,z)}{\nu Q(z)} \preceq \frac{\nu'(\cdot)Q'(\cdot,z')}{\nu'Q'(z')}, \tag{a}$$

$$\frac{Q(x,\cdot)Q(\cdot,z)}{Q^2(x,z)} \preceq \frac{Q'(x',\cdot)Q'(\cdot,z')}{Q'^2(x',z')}, \tag{b}$$

$$\frac{Q(x,\cdot)}{1-Q(x,0)} \preceq \frac{Q'(x',\cdot)}{1-Q'(x',0)} \tag{c}$$

as measures on S, then $\mu \preceq \mu'$.

Holley inequality was proved in Georgii, Häggström and Maes (2001) for finite state space S. We use the Markovian structure of μ and condition (c) to get around this assumption.

Proof of Proposition 2. Let $(\eta_t : t \geq 0)$ be the Gibbs sampler for μ, a Markov jump process on $\mathcal{X}_n^m$ with the following evolution: the value at each site $k \in [n,m]$ is updated at rate 1 to a new value using the conditional distribution of μ given the configuration at the sites $[n,m] \setminus \{k\}$. Different sites are never updated simultaneously since they use independent Poisson clocks. This amounts to use the measures in the left-hand side of (a), (b) and (c) to update sites n, $[n+1, m-1]$ and m, respectively. The measure μ is reversible for η_t. Analogously, let $(\eta_t' : t \geq 0)$ be the Gibbs sampler for μ' for which the updating is done with the measures in the right-hand side of (a), (b) and (c), respectively. The measure μ' is reversible for η_t'.

We will use the stochastic inequalities (a), (b), (c) to construct a monotone coupling $((\eta_t, \eta_t') : t \geq 0)$ of both Gibbs sampler processes. In this coupling, at rate 1 the value at each site in $[n,m]$ is simultaneously updated for both marginal trajectories using a monotone coupling of the measures in (a), (b) and (c) to update sites n, $[n+1, m-1]$ and m, respectively. If the trajectories are ordered at time 0, then they will remain ordered at future times, that is, if $\eta_0 \leq \eta_0'$ then $\eta_t \leq \eta_t'$ for all $t \geq 0$. We thus need to find $\eta_0 \leq \eta_0'$.

We claim that, given any trajectory $\mathbf{z}_n^m$ with positive μ'-probability, there exists a trajectory $\mathbf{x}_n^m \leq \mathbf{z}_n^m$ with positive μ-probability. We prove this by constructing $\mathbf{x}_n^m$ as follows. Since $\nu'(z_n) > 0$ and $\nu \preceq \nu'$, it is possible to choose $x_n \leq z_n$ such that $\nu(x_n) > 0$. Suppose that $x_n \leq z_n, \ldots, x_k \leq z_k$ have been chosen. Since $Q'(z_k, z_{k+1}) > 0$, from condition (c) it is possible to choose $x_{k+1} \leq z_{k+1}$ such that $Q(x_k, x_{k+1}) > 0$. This proves the claim.

The coupled process starts from (η_0, η_0'), where η_0' is distributed with the reversible measure μ' and, given η_0', a trajectory $\eta_0 \leq \eta_0'$ with positive μ-probability is chosen according to the previous claim. Since $\eta_0 \leq \eta_0'$, by the coupling we have $\eta_t \leq \eta_t'$ for all $t \geq 0$, and thus the law of η_t is stochastically dominated by that of η_t'. Irreducibility implies that η_t converges in distribution to its unique invariant

measure μ. On the other hand, the distribution of η_t' is μ' for all t. Letting $t \to \infty$, we get $\mu \preceq \mu'$. $\qquad\square$

The proof of Holley inequality works also for the nonhomogeneous case. Consider a family o transition matrices $\mathcal{Q} = (Q_k, k \in \mathbb{Z})$ and let (X_k) be a (nonhomogeneous) Markov chain satisfying $P(X_{k+1} = y | X_k = x) = Q_k(x, y)$, that is, the transition matrix Q_k is used at time k. Then we have the following corollary of the proof of Proposition 2.

Corollary 3. *Proposition 2 holds for nonhomogeneous families of transition matrices $\mathcal{Q}$ and $\mathcal{Q}'$ such that Q_n and Q_n' satisfy* (a), *Q_{m-1} and Q_{m-1}' satisfy* (c), *and that, for $k = n+1, \ldots, m-1$,*

$$\frac{Q_{k-1}(x, \cdot) Q_k(\cdot, z)}{Q_{k-1} Q_k(x, z)} \preceq \frac{Q_{k-1}'(x, \cdot) Q_k'(\cdot, z)}{Q_{k-1}' Q_k'(x, z)}. \tag{b$'$}$$

2.3 Monotonicity and Yaglom limit

Proof of Theorem 1. We first prove (ii) using Holley inequality with $Q' = Q$ and $\nu = \delta_1$. The probability measure on the left-hand side of condition (a) is δ_1, and, since 1 is minimal in S, this condition is satisfied for any ν' on S. Conditions (b) and (c) are being assumed in (1.3) and (1.4). Irreducibility of $\mu_0^n(\delta_1, Q)$ has also been explicitly assumed. By Holley inequality, $\mu_0^n(\delta_1, Q) \preceq \mu_0^n(\nu', Q)$, which by (2.2) implies $\delta_1 T_n \preceq \nu' T_n$, concluding the proof of (ii).

If ν' is a q.s.d., then $\nu' = \nu' T_n$. Together with (ii), this implies (iii).

To prove (i), we introduce a nonhomogeneous chain $\mathcal{Q}$ forced to make the first jump into state 1 while the rest of the jumps are governed by Q. Let $\mathcal{Q} = (Q_k, -1 \leq k \leq n)$ be given by $Q_{-1}(x, 1) = 1$ for all $x \in S$ and $Q_k = Q$ for $k = 0, \ldots, n$. By definition of Q_{-1}, the projection of $\mu_{-1}^n(\delta_1, \mathcal{Q})$ onto $\mathcal{X}_0^n$ is $\mu_0^n(\delta_1, Q)$. Hence, the time-n marginal of $\mu_{-1}^n(\delta_1, \mathcal{Q})$ is $\delta_1 T_n$, the same as the time-n marginal of $\mu_{-n}^0(\delta_1, Q)$. Let $\mathcal{Q}' = (Q_k', -1 \leq k \leq n)$ be given by $Q_k' = Q$ for $k = -1, \ldots, n$ (i.e., the homogeneous chain). The time-n marginal of $\mu_{-1}^n(\delta_1, \mathcal{Q}')$ is $\delta_1 T_{n+1}$, the same as the time-n marginal of $\mu_{-1}^n(\delta_1, \mathcal{Q})$. Again, condition (c) has been assumed in (1.4). Writing $\nu = \nu' = \delta_1$, condition (a) holds trivially. Condition (b$'$) is trivial for $k = 0$, and for $k = 1, 2, \ldots, n-1$ it is assumed in (1.3). Also, irreducibility of $\mu_0^n(\delta_1, Q)$, and thus of $\mu_{-1}^n(\delta_1, \mathcal{Q})$, has been explicitly assumed. Using Corollary 3, we get $\mu_{-1}^n(\delta_1, \mathcal{Q}) \preceq \mu_{-1}^n(\delta_1, \mathcal{Q}')$, and thus $\delta_1 T_n \preceq \delta_1 T_{n+1}$, proving (i).

To show (iv), let ν' be a q.s.d. By (i), $\delta_1 T_n$ is an increasing sequence of measures and by (iii) all elements of the sequence are dominated by ν'. Hence, there is a limit $\underline{\nu} := \lim_n \delta_1 T_n \preceq \nu'$. To check that $\underline{\nu}$ is a q.s.d., use that T_n is a semigroup and that T_1 is continuous to get

$$\underline{\nu} = \lim_n \delta_1 T_{n+1} = \lim_n \delta_1 T_n T_1 = \underline{\nu} T_1. \qquad\square$$

2.4 Yaglom limit and minimal q.s.d.

For a measure ν on S, denote

$$a(\nu) := 1 - \nu Q(0), \qquad (2.3)$$

the mass staying at S after one step for the chain starting with ν. If ν is a q.s.d., then ν is a left eigenvector for $Q_{|S}$ with eigenvalue $a(\nu)$:

$$\nu Q_{|S} = a(\nu)\nu.$$

Let $a_* := \inf\{a(\nu) : \nu \text{ is a q.s.d.}\}$. If there exists a q.s.d. ν with $a(\nu) = a_*$, then it is called *minimal* and denoted $\nu_{\min}$.

The following lemma gives sufficient conditions in terms of Q so that the measure $\underline{\nu}$ given by Theorem 1 coincides with $\nu_{\min}$.

Lemma 4. *If Q is such that $Q(x, 0) \geq Q(x', 0)$ for all $x \leq x' \in S$, and $\underline{\nu}$ is a q.s.d. such that $\underline{\nu} \preceq \nu$ for any other q.s.d. ν, then $\underline{\nu} = \nu_{\min}$.*

Proof. The function $f : S \to \mathbb{R}^+$ given by $f(y) = Q(y, 0)$ is nonincreasing, whence $\underline{\nu} f \geq \nu f$ for any q.s.d. ν. Thus, $a(\underline{\nu}) \leq a(\nu)$, and taking the infimum over ν we get $a(\underline{\nu}) = a_*$, which proves the lemma. $\qquad\qquad\square$

3 The birth-and-death chain

In this section, we consider $S = \mathbb{N}$ with the usual order and birth-and-death processes. The transition matrix Q is defined by:

$$p_x, r_x, q_x > 0, \qquad q_x + r_x + p_x = 1 \qquad \text{for all } x \geq 1;$$

$$Q(x, x-1) = q_x, \qquad Q(x, x) = r_x, \qquad Q(x, x+1) = p_x, \qquad (3.1)$$

$$Q(x, y) = 0 \qquad \text{if } |x - y| > 1 \quad \text{and} \quad Q(0, 0) = 1.$$

In this case, there exist a q.s.d. if the absorption time of the chain starting from a fixed state has an exponential moment; see, for instance, van Doorn and Schrijner [(1995a), Corollary 4.1], Ferrari, Martínez and Picco [(1992), Theorem 6.1] and Ferrari et al. (1995). Under these conditions, Cavender (1978) shows that there is a critical value $\gamma > 0$ such that there is a one-parameter family of q.s.d.'s $\{\nu : \nu(1) \in (0, \gamma]\}$ indexed by $\nu(1)$. Cavender fixes $\nu(1) \leq \gamma$ and computes explicitly the other values using the equation (1.2) and the nearest-neighbor structure (this procedure does not yield a probability if $\nu(1) > \gamma$). Cavender also shows that any pair of q.s.d.'s ν, ν' satisfy a monotone likelihood ratio: $\nu'(1) > \nu(1)$ implies $\frac{\nu(1)}{\nu'(1)} < \frac{\nu(2)}{\nu'(2)} \leq \frac{\nu(3)}{\nu'(3)} \leq \cdots$, which in turn implies the domination $\nu' \preceq \nu$.

van Doorn and Schrijner (1995b) use the Karlin and McGregor polynomial representation of the chain to give a sufficient condition for the Yaglom limit to converge to an explicit limit. Ferrari, Martinez and Picco (1991) describe the domain of attraction of q.s.d.'s and show in particular that the Yaglom limit of δ_x converges to the minimal q.s.d., for any initial state x. Daley (1969) and Iglehart (1974) showed the Yaglom limit for random walks with negative drift and finite variance, respectively for discrete and continuous space.

In the sequel we develop Theorem 1's conditions and make them explicit for the case of birth-and-death chains. Corollary 7 is about space-homogeneous discrete-time random walks with delay.

Item (iv) of Corollary 8 gives the Yaglom limit for continuous-time walks. It was originally proven by Seneta (1966) using direct computation. Our proof uses monotonicity of the trajectories instead.

Corollary 10, presented in the next section, gives the Yaglom limit for the discrete-time periodic chain. It provides an alternative proof to that of Seneta and Vere-Jones (1966).

The conditions of Theorem 1

Since the state space $S = \mathbb{N}$ is totally ordered and the transitions are only to nearest neighbors, we can obtain conditions (1.3) and (1.4) in explicit terms of p_k, r_k and q_k. Take Q as defined in (3.1). Define for positive integers x, z, y:

$$b((x, z), y) := \sum_{w \geq y} \frac{Q(x, w) Q(w, z)}{Q^2(x, z)}; \tag{3.2}$$

$$c(x, y) := \sum_{w \geq y} \frac{Q(x, w)}{1 - Q(x, 0)}. \tag{3.3}$$

Conditions (1.3) and (1.4) are equivalent to

$$b((x, z), y) \leq b((x', z'), y) \qquad \text{for } z \leq z', x \leq x'; \tag{3.4}$$

$$c(x, y) \leq c(x', y) \qquad \text{for } x \leq x', \tag{3.5}$$

for all $y \geq 1$, whenever the denominators of both sides are positive.

Inequalities (3.4) hold trivially when $y = 1$ or $\{x, x', z, z'\} \not\subset \{y - 1, y\}$. The remaining cases are the following. For $y \geq 2$, the conditions (3.4) are equivalent to the following conditions:

$$\begin{aligned} b((y - 1, y - 1), y) &\leq b((y, y - 1), y) \leq b((y, y), y), \\ b((y - 1, y - 1), y) &\leq b((y - 1, y), y) \leq b((y, y), y). \end{aligned} \tag{3.6}$$

Using the convention $p_0 = 0$, conditions (3.6) for $y \geq 2$ read

$$
\begin{aligned}
\frac{p_{y-1}q_y}{r_{y-1}^2 + p_{y-1}q_y + q_{y-1}p_{y-2}} &\leq \frac{r_y q_y}{q_y r_{y-1} + r_y q_y} \\
&\leq \frac{r_y^2 + p_y q_{y+1}}{r_y^2 + p_y q_{y+1} + q_y p_{y-1}}, \\
\frac{p_{y-1}q_y}{r_{y-1}^2 + p_{y-1}q_y + q_{y-1}p_{y-2}} &\leq \frac{p_{y-1}r_y}{p_{y-1}r_y + r_{y-1}p_{y-1}} \\
&\leq \frac{r_y^2 + p_y q_{y+1}}{r_y^2 + p_y q_{y+1} + q_y p_{y-1}}.
\end{aligned}
\tag{3.7}
$$

Analogously, conditions (3.5) on $c(x, y)$ hold trivially when $y = 1$ or $(x, x') \neq (y-1, y)$. Hence, (3.5) is equivalent to

$$
c(y-1, y) \leq c(y, y), \qquad y \geq 2,
$$

which in the case $y = 2$ and $y \geq 3$ read, respectively,

$$
\frac{p_1}{p_1 + r_1} \leq r_2 + p_2, \qquad p_{y-1} \leq p_y + r_y \qquad \text{for } y \geq 3.
\tag{3.8}
$$

We summarize these computations as a lemma.

Lemma 5. *Let Q be the transition matrix for the birth-and-death chain defined in (3.1). Then conditions (1.3) and (1.4) are equivalent to (3.7) and (3.8).*

We are ready to state the result in this case.

Corollary 6. *Assume that the birth-and-death chain absorbed at zero defined in (3.1) has at least one q.s.d. and satisfies conditions (3.7) and (3.8). Then (i), (ii), (iii), (iv) of Theorem 1 hold. The Yaglom limit of δ_1 coincides with $\nu_{\min}$, the minimal q.s.d. in the sense of absorption time. Furthermore, for any $x \in \mathbb{N}$, the Yaglom limit of δ_x also converges to $\nu_{\min}$.*

Proof. Since for $x \geq 1$ the probability of transitions from x to x and to nearest neighbors of x are positive, the birth-and-death chain starting with δ_1 has irreducible trajectories in $\mathbb{N}$. By Lemma 5, the conditions of Theorem 1 are equivalent to the present conditions, hence (i), (ii), (iii), (iv) of Theorem 1 hold and the Yaglom limit of δ_1 converges to $\underline{\nu}$. Since $Q(x, 0) = 0$ for all $x > 1$, Lemma 4 applies and $\underline{\nu} = \nu_{\min}$. By Ferrari, Martinez and Picco [(1991), Theorem 3.1], if the Yaglom limit of δ_1 exists, then it coincides with the Yaglom limit of δ_x for any x, concluding the proof. $\qquad\square$

3.1 Random walk with delay

The absorbed delayed random walk is a particular case of birth-and-death chain on $\mathbb{N} \cup \{0\}$ defined in (3.1) with constant transition probabilities along $\mathbb{N}$:

$$p_x \equiv p, \qquad q_x \equiv q, \qquad r_x \equiv r,$$
$$p, q, r > 0, \qquad p + q + r = 1, \qquad p < q. \tag{3.9}$$

This walk has a drift towards 0 and it is absorbed at 0. A probability ν on $\mathbb{N}$ is a q.s.d. if and only if it satisfies the equations (1.2), which in this case are

$$\nu(x+1)q + \nu(x-1)p + \big(q\nu(1) - (p+q)\big)\nu(x) = 0, \qquad x \geq 1, \tag{3.10}$$

with the convention $\nu(0) = 0$. Cavender (1978) proved that the set of q.s.d.'s is a family indexed by $\nu(1)$ with $\nu(1) \in (0, (1 - \sqrt{\lambda})^2]$, where $\lambda = p/q$. Since the absorption probability of a q.s.d. ν is $\nu Q(0) = q\nu(1)$, the q.s.d. with maximal $\nu(1)$ is the minimal q.s.d. $\nu_{\min}$, a negative binomial with parameters 2 and $\sqrt{\lambda}$:

$$\nu_{\min}(x) = (1 - \sqrt{\lambda})^2 x (\sqrt{\lambda})^{x-1}, \qquad x \geq 1. \tag{3.11}$$

The remaining q.s.d. are given in function of $\nu(1) \in (0, (1 - \sqrt{\lambda})^2)$ by

$$\nu(x) = \frac{\nu(1)}{c}\left[\left(\frac{\lambda + 1 - \nu(1) + c}{2}\right)^x - \left(\frac{\lambda + 1 - \nu(1) - c}{2}\right)^x\right], \tag{3.12}$$

where $c = [(\nu(1) - \lambda - 1)^2 - 4\lambda]^{1/2}$, Cavender (1978), p. 585.

Corollary 7. *Consider the random walk with delay absorbed at zero defined in* (3.1) *with constant rates* (3.9). *If $pq \leq r^2$, then the conclusions* (i), (ii), (iii), (iv) *of Theorem* 1 *hold with $\underline{\nu} = \nu_{\min}$ given by* (3.11). *Furthermore, for any $x \in \mathbb{N}$ the Yaglom limit of δ_x converges to $\underline{\nu}$.*

Proof. In the present context the worst case of (3.7) is when $y = 2$, which reduces to:

$$\frac{pq}{pq + r^2} \leq \frac{1}{2} \leq \frac{r^2 + pq}{r^2 + 2pq}. \tag{3.13}$$

On the other hand (3.8) reads

$$\frac{p}{p+r} \leq r + p, \qquad p \leq p + r. \tag{3.14}$$

Condition $pq \leq r^2$ implies both (3.13) and (3.14). The result thus follows from Corollary 6. $\qquad \square$

If $r < \sqrt{pq}$, then trajectory domination is not true. Although the Yaglom limit of δ_1 is known to hold in this case (Daley (1969)), it does not seem to follow from the arguments presented here, except for the periodic case $r = 0$ discussed in Section 4.

3.2 The continuous-time random walk

Take positive $p < q$ with $p + q = 1$ and consider a family of random walks with delay (X_n^r), indexed by $r \in [0, 1)$, with transition probabilities

$$Q_r(x, x-1) = q(1-r), \qquad Q_r(x, x) = r, \qquad Q_r(x, x+1) = p(1-r),$$

$$Q_r(x, y) = 0, \qquad \text{otherwise,} \qquad x \geq 1; \qquad Q_r(0, 0) = 1.$$

Define the rescaled process

$$Y_t^r := X_{[t/(1-r)]}^r.$$

As $r \to 1$, the process (Y_t^r) converges in finite time-intervals to the process $(\hat{Y}_t)$, a continuous-time random walk with rates p, q to jump one unit forward or backwards, respectively, and absorbed at 0. Call $\hat{U}_t$ the corresponding semigroup:

$$\hat{U}_t(x, y) := P(\hat{Y}_t = y | \hat{Y}_0 = x).$$

Define νT_t^r as the probability given by

$$\nu T_t^r(y) := \frac{\nu Q_r^{[t/(1-r)]}(y)}{1 - \nu Q_r^{[t/(1-r)]}(0)},$$

that is, νT_t^r is the distribution at time t of the walk Y_t^r starting with ν, conditioned to nonabsorption. This distribution converges as $r \to 1$ to the distribution at time t of the continuous-time walk $\hat{Y}_t$ under the same condition:

$$\lim_{r \to 1} \nu T_t^r(y) = \nu \hat{T}_t(y) := \frac{\nu \hat{U}_t(y)}{1 - \nu \hat{U}_t(0)}. \tag{3.15}$$

The resulting operator $\hat{T}_t$ is a semigroup. For any $r \in [0, 1)$, the q.s.d.'s for Y_t^r satisfy equations (3.10) because the factors $(1 - r)$ cancel out. Moreover, the q.s.d.'s for the continuous-time walk $\hat{Y}_t$ also satisfy the same equations. Indeed, $\nu = \nu \hat{T}_t$ if and only if $\nu(\hat{U}_t - I) + \nu \hat{U}_t(0) \cdot \nu = 0$; dividing by t and letting $t \to 0$ yields (3.10). As a consequence, the minimal q.s.d. for both Y_t^r and $\hat{Y}_t$ is given by (3.11) while the remaining q.s.d. are given by (3.12). In the continuous-time case, p and q may be any positive real numbers satisfying $p < q$; the definitions (3.11) and (3.12) depend on p and q only through the ratio $\lambda = p/q$.

Corollary 8. *The continuous-time random walk with rates p, q absorbed at zero satisfies*:

(i) *The sequence $(\delta_1 \hat{T}_t, t \geq 0)$ is monotone: $\delta_1 \hat{T}_s \preceq \delta_1 \hat{T}_t$ for $0 < s \leq t < \infty$.*
(ii) *If ν is a probability measure on $\mathbb{N}$, then $\delta_1 \hat{T}_t \preceq \nu \hat{T}_t$ for all $t \geq 0$.*
(iii) *In particular, if ν is a q.s.d., then $\delta_1 \hat{T}_t \preceq \nu$.*
(iv) *The Yaglom limit of δ_1 converges to $\nu_{\min}$ given by (3.11).*

Proof. Take r sufficiently close to one so that $pq(1 - r)^2 \leq r^2$, to be under the conditions of Corollary 7.

To show (i), we use Corollary 7(i) to get $\delta_1 T_t^r \preceq \delta_1 T_{t+s}^r$, for all $t, s \geq 0$, and then use (3.15) to conclude. To prove (ii), (iii), we use Corollary 7(ii), (iii) to get $\delta_1 T_t^r \preceq \nu T_t^r$ (which equals ν if it is a q.s.d.), and again use (3.15) to conclude.

Let us show (iv). As discussed above, $\nu_{\min}$ given by (3.11) is a q.s.d., the other q.s.d.'s are given by (3.12), and in particular $\nu_{\min}$ is minimal also in the sense of stochastic ordering. By (i), (iii), there is $\underline{\nu} = \lim_t \delta_1 \hat{T}_t$. As in the proof of Theorem 1, using the semigroup property of $\hat{T}_t$, the limit $\underline{\nu}$ is a q.s.d. It follows from (iii) that $\underline{\nu} \preceq \nu_{\min}$, and therefore $\underline{\nu} = \nu_{\min}$.   □

4 The periodic case

Assume that the matrix Q is irreducible in S and that Q restricted to S has period $d \geq 2$. Let $S_1, \ldots, S_d \subset S$ be the cyclic subclasses, that is, the equivalence classes induced by the equivalence relation $\sim$ defined by $x \sim y$ if and only if $Q^{d\ell}(x, y) > 0$ for some $\ell \geq 1$. Assume that the classes are labeled so that $x \in S_j$, $Q(x, y) > 0$ implies $y \in S_{j+1} \cup \{0\}$ (with the convention $S_{d+1} = S_1$).

Theorem 9. *Let S be a partially ordered countable set with a minimal element called 1 and let X_n be a Markov chain on $S \cup \{0\}$, absorbed at 0, irreducible in S and with period d when restricted to S. Let $S_1, \ldots, S_d \subset S$ denote the cyclic subclasses of the chain restricted to S, choosing $S_1 \ni 1$. Assume that the chain with initial state 1 has irreducible trajectories.*

If, for all $x, x', z, z' \in S$ with $x \leq x'$, $z \leq z'$, x and x' in the same class, the stochastic inequalities (1.3) and (1.4) are satisfied whenever the denominators are positive, then the following hold.

(i) *Monotonicity: $\delta_1 T_n \preceq \delta_1 T_{n+d}$ for any $n \geq 0$.*
(ii) *For any probability ν on S_1, one has $\delta_1 T_n \preceq \nu T_n$.*
(iii) *If ν is a q.s.d., then $\delta_1 T_{dk+j-1} \preceq \nu(\cdot|S_j)$ for any $k \geq 0$.*
(iv) *If the chain has a q.s.d., then there is a q.s.d. $\nu_\star$ such that the Yaglom limit of δ_1 along d-periodic subsequences is given by*

$$\lim_k \delta_1 T_{dk+j-1} = \nu_\star(\cdot|S_j).$$

Moreover, for any other q.s.d. ν, one has $\nu_\star(\cdot|S_j) \preceq \nu(\cdot|S_j)$ for all j.
(v) *If moreover Q is such that $\nu \preceq \nu'$ implies $a(\nu) \leq a(\nu')$, then $\nu_\star = \nu_{\min}$.*

Before giving the proof, we discuss the particular case of the p-q random walk. As an application of the above theorem, we prove convergence of the Yaglom limit to the minimal q.s.d. based on monotonicity of trajectories.

The *p-q discrete-time random walk* is defined as follows. Consider the periodic random walk with transition probabilities

$$Q(0, 0) = 1, \qquad Q(x, x - 1) = q, \qquad Q(x, x + 1) = p \qquad \text{for } x \geq 1,$$

$$Q(x, y) = 0, \qquad \text{otherwise}; \qquad p + q = 1, \qquad p < q. \tag{4.1}$$

The chain has period 2 and, starting from δ_1, the walk visits odd sites at even times and vice-versa. The q.s.d.'s for this random walk satisfy (3.10) as before. The minimal q.s.d. $\nu_{\min}$ is given by (3.11), and the remaining q.s.d.'s are given by (3.12). The cyclic subclasses are $S_1 = 2\mathbb{N} - 1$ and $S_2 = 2\mathbb{N}$.

Corollary 10. *Let X_n be the discrete-time p-q random walk with transition probabilities* (4.1). *The Yaglom limit of δ_1 converges along even and odd times to projections of $\nu_{\min}$ given by* (3.11). *That is, for both $j = 1, 2$,*

$$\lim_n \delta_1 T_{2n+j-1} = \nu_{\min}(\cdot | S_j).$$

Moreover, $\nu_{\min}(\cdot | S_j) \preceq \nu(\cdot | S_j)$ for any other q.s.d. ν.

Proof. By Theorem 9(iv) there is $\nu_\star$ with the above properties, and the Yaglom limit converges to projections of $\nu_\star$ along even or odd subsequences. By Theorem 9(v) we have $\nu_\star = \nu_{\min}$, concluding the proof. □

In order to prove Theorem 9, we start with some basic properties of q.s.d.'s for periodic chains. For a probability ν on S, write $\nu = \sum_j m_j \nu_j$, where $\nu_j := \nu(\cdot | S_j)$ and $\sum_j m_j = 1$. For shortness, let S_j, m_j and ν_j be indexed by $j \in \mathbb{Z}_d$, so that $S_{d+1} = S_1$, etc. Recall that $a(\nu)$ is defined in (2.3).

Lemma 11. *Let Q be the transition matrix for a d-periodic chain in S absorbed at 0. If ν is a q.s.d., then for each class j, $m_j a(\nu_j) = a(\nu) m_{j+1}$, and $\nu_j T_n = \nu_{j+n}$, for all $n \geq 0$. In particular, $(a(\nu))^d = a(\nu_1) \cdots a(\nu_d)$.*

Proof. For any measure ν on S, $\nu_j Q$ is supported on $S_{j+1} \cup \{0\}$, and thus $\nu_j Q_{|S} = a(\nu_j) \nu_j T_1$. Hence, $\nu Q_{|S} = \sum_j m_j \nu_j Q_{|S} = \sum_j m_j a(\nu_j) \nu_j T_1$. On the other hand, if ν is a q.s.d., $\nu Q_{|S} = a(\nu) \nu = \sum_j a(\nu) m_j \nu_j$, and thus $\sum_j m_j a(\nu_j) \nu_j T_1 = \sum_j a(\nu) m_{j+1} \nu_{j+1}$. Now notice that, for each class j, the measures $\nu_j T_1$ and ν_{j+1} are probabilities supported on S_{j+1}, and these sets are disjoint. Therefore, $\nu_j T_1 = \nu_{j+1}$, and $m_j a(\nu_j) = a(\nu) m_{j+1}$. Iterating the former identity, we get $\nu_j T_n = \nu_j T_1^n = \nu_{j+n}$, and taking the product over j of both sides of the latter, we get $(a(\nu))^d = a(\nu_1) \cdots a(\nu_d)$. □

Proof of Theorem 9. Under the present assumptions on the transition matrix, Holley inequality holds for any pair of measures supported on the subspace of

trajectories that start in a given cyclic subclass. Therefore, parts (i) and (ii) can be proved just as in the proof of Theorem 1.

To prove (iii), let ν be a q.s.d. By (ii) and Lemma 11, $\delta_1 T_n \preceq \nu_1 T_n = \nu_{1+n}$. Claim (iii) follows by taking $n = dk + j - 1$.

Proof of (iv). As in the proof of Theorem 1(iv), by (i), (iii) the limits

$$\underline{\nu}_j := \lim_k \delta_1 T_{dk+j-1}$$

exist and satisfy

$$\underline{\nu}_j T_1 = \underline{\nu}_{j+1},$$

and moreover $\underline{\nu}_j \preceq \nu_j$ for any q.s.d. ν. It remains to find the right constants m_j and show that $\nu_\star$ given by $\nu_\star = \sum_j m_j \underline{\nu}_j$ is a q.s.d., that is, that there exists an $\alpha \in (0, 1)$ such that $\nu_\star Q_{|S} = \alpha \nu_\star$. Since $\underline{\nu}_j Q_{|S} = a(\underline{\nu}_j) \underline{\nu}_{j+1}$, the problem is equivalent to find $\alpha, m_1, \ldots, m_d$ solving the system of d linear equations given by $m_j a(\underline{\nu}_j) = \alpha m_{j+1}$. The system has a nonzero solution if and only if $\alpha^d = a(\underline{\nu}_1) \cdots a(\underline{\nu}_d)$. Choosing the positive α that satisfies this identity, the space of solutions is one-dimensional and its elements have coordinates which agree in sign. Choosing m to be the unique solution to satisfy $\sum_j m_j = 1$, we have that $\nu_\star$ is a probability and moreover it is a q.s.d. with $a(\nu_\star) = \alpha$, concluding the proof of (iv).

Proof of (v). It suffices to prove that $a(\nu_\star) \leq a(\nu)$ for all q.s.d. ν. By Lemma 11,

$$\left(a(\nu_\star)\right)^d = a(\underline{\nu}_1) \cdots a(\underline{\nu}_d) \leq a(\nu_1) \cdots a(\nu_d) = \left(a(\nu)\right)^d,$$

where the inequality comes from (iv) and the hypothesis of (v). □

Acknowledgments

References

Cavender, J. A. (1978). Quasi-stationary distributions of birth-and-death processes. *Adv. in Appl. Probab.* **10**, 570–586. MR0501388

Collet, P., Martínez, S. and San Martín, J. (2013). *Markov Chains, Diffusions and Dynamical Systems. Probability and Its Applications (New York)*. Heidelberg: Springer. MR2986807

Daley, D. J. (1969). Quasi-stationary behaviour of a left-continuous random walk. *Ann. Math. Statist.* **40**, 532–539. MR0246374

van Doorn, E. A. and Schrijner, P. (1995a). Geometric ergodicity and quasi-stationarity in discrete-time birth–death processes. *J. Austral. Math. Soc. Ser. B* **37**, 121–144. MR1359177

[mr] van Doorn, E. A. and Schrijner, P. (1995b). Ratio limits and limiting conditional distributions for discrete-time birth–death processes. *J. Math. Anal. Appl.* **190**, 263–284. MR1314118

Ferrari, P. A., Kesten, H., Martinez, S. and Picco, P. (1995). Existence of quasi-stationary distributions. A renewal dynamical approach. *Ann. Probab.* **23**, 501–521. MR1334159

Ferrari, P. A., Martinez, S. and Picco, P. (1991). Some properties of quasi-stationary distributions in the birth and death chains: A dynamical approach. In *Instabilities and Nonequilibrium Structures, III*, *Valparaíso*, *1989*. *Math. Appl.* **64**, 177–187. Dordrecht: Kluwer Acad. Publ. MR1177850

Ferrari, P. A., Martínez, S. and Picco, P. (1992). Existence of nontrivial quasi-stationary distributions in the birth–death chain. *Adv. in Appl. Probab.* **24**, 795–813. MR1188953

Georgii, H.-O., Häggström, O. and Maes, C. (2001). The random geometry of equilibrium phases. In *Phase Transitions and Critical Phenomena, Vol. 18*. 1–142. San Diego, CA: Academic Press. MR2014387

Holley, R. (1974). Remarks on the FKG inequalities. *Comm. Math. Phys.* **36**, 227–231. MR0341552

Iglehart, D. L. (1974). Random walks with negative drift conditioned to stay positive. *J. Appl. Probab.* **11**, 742–751. MR0368168

Kesten, H. (1995). A ratio limit theorem for (sub) Markov chains on {1, 2, . . .} with bounded jumps. *Adv. in Appl. Probab.* **27**, 652–691. MR1341881

Lindvall, T. (1999). On Strassen's theorem on stochastic domination. *Electron. Commun. Probab.* **4**, 51–59 (electronic). MR1711599

Pollett, P. K. (2014). Quasi-stationary distributions: A bibliography. Available at http://www.maths.uq.edu.au/~pkp/papers/qsds/qsds.pdf.

Seneta, E. (1966). Quasi-stationary behaviour in the random walk with continuous time. *Austral. J. Statist.* **8**, 92–98. MR0208688

Seneta, E. and Vere-Jones, D. (1966). On quasi-stationary distributions in discrete-time Markov chains with a denumerable infinity of states. *J. Appl. Probab.* **3**, 403–434. MR0207047

Departamento de Matemática
Facultad de Ciencias Exactas y Naturales
Universidad de Buenos Aires
Pabellón 1, Ciudad Universitaria
1428 Ciudad de Buenos Aires
Argentina
E-mail: pferrari@dm.uba.ar
          leorolla@dm.uba.ar
URL: http://mate.dm.uba.ar/~pferrari/
        http://mate.dm.uba.ar/~leorolla/