Published in *Result. Math.* **37** (2000) 246–273

Balanced partitions for Markov chains

María Teresa Gallegos and Gunter Ritter

In memoriam Ed Hewitt

Abstract: We call a finite partition of the state space of a (discrete–time) Markov chain *balanced* if the flows in both directions between any two of its classes are equal in equilibrium. If a Markov chain is reversible then any finite partition is balanced. We use this notion in order to gain insight into the structure of the stationary distributions of not necessarily reversible transition kernels. We illustrate our theory with an asymptotic analysis of a non–reversible Markovian star network with loss.[1]

1 Introduction

1.1 Background and outline Markov chains are standard tools for modelling and analyzing stochastic algorithms and real–world dynamical systems containing randomness; examples are stochastic optimization algorithms, learning schemes, and queueing and communication networks. Many of their performance measures flow from their *stationary distributions*, since these describe the long term behavior of the chains. Computation of a stationary distribution amounts to computing a left eigenvector of the eigenvalue 1 of the transition kernel, a function on the square of the state space. Unfortunately, state spaces are often astronomically large. This circumstance may make these distributions inaccessible and determining the fine structure of a stationary distribution is often a discouraging task.

If the chain is *reversible* then access to the stationary distribution is easier but even then the normalizing constant may pose a problem. *Lumpabilities* [3, 10] are concepts that allow to aggregate the state space thus cutting down its size. However, these conditions are restrictive and often not met. Therefore, additional approaches are necessary. The present paper deals with such an approach, *balanced partitions*, cf. Section 2.2. In the case of a finite state space, balanced partitions were introduced in the first author's [4] doctoral dissertation under the name $\mathcal{Z}$*–reversibility*; the method was shown to be strong enough to access the thermodynamic limit of a certain learning scheme, reinforcement learning, for an arbitrary finite number of coins. However, the range of applications of this notion goes far beyond. The purpose of this communication is twofold: first, we extend the notion to an arbitrary state space and, second, we apply it in order to analyze a non–reversible Markovian star network with loss.

The outline of the paper is as follows. In the remainder of this section, we describe some prerequisites on Markov chains necessary for the sequel. In Section 2, we introduce and discuss balancedness of a (finite) partition of the state space of a Markov chain and give

[1]**AMS 1991 Subject Classification**: Primary 60J05; secondary 60J10, 60J20
Key words: Markov chains on arbitrary state spaces, structure of stationary distributions, balanced partitions, Markovian analysis of queueing networks

necessary and sufficient conditions for its validity. In Section 3, we show how balanced partitions induce estimates of the stationary distributions of certain events. In Section 4, we finally demonstrate the applicability of the approach by an analysis of a multiclass service system with loss.

1.2 Notation and preliminaries The symbols $\mathbb{N} = \{0, 1, 2, ...\}$, $\mathbb{N}_> = \{1, 2, 3, ...\}$, and $\mathbb{R}$ stand for the sets of natural, strictly positive natural, and real numbers, respectively. We denote the interval of natural numbers between m and n by $m..n$. Given a set F and $n \in \mathbb{N}$, $\binom{F}{n}$ denotes the system of all subsets of F with n elements.

The symbol S denotes an arbitrary measurable space, i.e., a set (which we also denote by S) endowed with a σ–algebra $\mathcal{B}(S)$ on it. Measures on a measurable space are always assumed to be σ–finite and will be denoted by lower case Greek letters. The point mass at a point a is δ_a and the image measure of a measure ρ on S with respect to a measurable function $f : S \to \mathbb{R}$ is denoted by ρ_f. The symbol $\preceq$ denotes the stochastic ordering on the set of probability distributions on $\mathbb{R}$ or $\mathbb{Z}$; i.e., for two such probability distributions μ, ν, we have $\mu \preceq \nu$ if their tail distributions satisfy the relation $\mu([t, \infty[) \le \nu([t, \infty[)$ for all $t \in \mathbb{R}$. For general information and more details on measures we refer the reader to Bauer [2] or Hewitt and Stromberg [5].

A measurable *kernel* L from a measurable space $(S, \mathcal{B}(S))$ to another measurable space $(T, \mathcal{B}(T))$ is a mapping

$$L : S \times \mathcal{B}(T) \to \mathbb{R}$$

such that $L(\cdot, B)$ is measurable for each $B \in \mathcal{B}(S)$ and $L(x, \cdot)$ is a measure for each $x \in S$. Associated with a measure μ on S and the kernel L there are two more measures, the *tensor product* $\mu \otimes L(dx, dy) = \mu(dx)L(x, dy)$ on $\mathcal{B}(S) \otimes \mathcal{B}(T)$ and the *composition* $\mu L(dy) = \mu \otimes L(S, dy) = \int_S \mu(dx)L(x, dy)$ on $\mathcal{B}(T)$. The kernel L is *Markovian* if $L(x, \cdot)$ is even a probability measure for each $x \in S$. If $f : T \to [0, \infty]$ is measurable then the assignment $x \mapsto \int_T L(x, dy)f(y)$ defines a measurable function on S. The same is true if f is measurable, real–valued, and bounded and if L is Markovian. If $A \in \mathcal{B}(S)$ and $B \in \mathcal{B}(T)$ then L induces in a natural way a kernel from A to B, the restriction $L_{A,B}$ of L to $A \times B$.

In what follows, K denotes a Markov kernel from S to S. Together with any *initial distribution* μ on S, K generates an essentially unique *Markov chain* $X_t : (\Omega, \mathcal{F}, P) \to S$, $t \in \mathbb{N}$, on S; here $(\Omega, \mathcal{F}, P)$ is some suitable probability space. This means that $X = (X_t)_t$ is a stochastic sequence in S such that $X_0 \sim \mu$ and, for all $t \ge 1$ and all $B \in \mathcal{B}(S)$, the conditional probabilities $P[X_{t+1} \in B \mid X_0 = x_0, \cdots, X_t = x_t]$ do not depend on $x_0, \cdots, x_{t-1}$ for $P_{(X_0,\ldots,X_t)}$–a.a $(x_0, \ldots, x_t) \in S^{t+1}$; they are, in fact, all equal to $K(x_t, B)$. This Markov chain X is *homogeneous* since the transition from t to $t + 1$ is controlled by the same kernel K no matter what the index t is. In order to indicate the initial distribution μ, we will use the notations P^μ and E^μ for probabilities of events and expectations of functionals of the process.

A Markov chain X is called *stationary*, if the joint distribution of $X_n, \cdots, X_{t+n}$ does not depend on $n \in \mathbb{N}$ for all $t \in \mathbb{N}$. A distribution γ on S is called *stationary* with respect to K if $\gamma K = \gamma$. If the initial distribution of the Markov chain X is stationary then the whole chain is stationary and vice versa. The stationary distributions are crucial to the study of

K and X since they control the long term behavior of X through various *ergodic theorems*. A distribution γ on S is called *reversible* (with respect to K) if the tensor product $\gamma \otimes K$ is symmetric on the product $S \times S$ with respect to swapping the two factors. Any reversible distribution is plainly stationary.

There exists a refined theory of existence and uniqueness of stationary distributions on discrete as well as on arbitrary state spaces S and we refer the interested reader to Meyn and Tweedie's [9] treatise. Let us briefly sketch the main ideas as far as they are related to this paper. For a measurable subset $B \subseteq S$, let $T_B = \inf\{t \geq 1 / X_t \in B\}$ be the *first return time* of the chain X to B and define the *return probabilities* $L(x, B) = P^{\delta_x}[T_B < \infty]$, $x \in S$. The kernel K is called *irreducible* if there exists a nontrivial measure φ on S such that $L(x, B) > 0$ for all $x \in S$ whenever $\varphi(B) > 0$. In this case, there exists a (σ–finite or probability) measure ψ on S which is *maximal* with respect to absolute continuity with this property. Moreover, this maximal "irreducibility measure" ψ is unique up to equivalence. The system $\mathcal{B}^+(S)$ of measurable sets on S with strictly positive ψ–measure plays a central rôle in the theory.

If K is irreducible then it possesses *at most one* stationary distribution. If there is one (and K is irreducible) then K is called *positive recurrent* and the stationary distribution is equivalent to ψ. The positive recurrent case is of particular interest since the long term behavior of the chain is then independent of its initial distribution. If K is positive recurrent then it is also *recurrent*, i.e., $\sum K^n(x, B) = \infty$ for all $x \in S$ and all $B \in \mathcal{B}^+(S)$. It follows that $L(x, B) = 1$ for all such x and B.

For establishing a practical condition for the *existence* of stationary distributions a topological structure on S is of great help. Thus, let S be endowed with a locally compact and separable metric and denote its set of *bounded, continuous* real–valued functions by $\mathcal{C}_b(S)$. The kernel K has the *Feller property*, if $K\mathcal{C}_b(S) \subseteq \mathcal{C}_b(S)$. The *taboo kernel* ${}_CK$ of K with respect to a nonempty subset $C \subseteq S$ is defined by ${}_CK(x, B) = K(x, B \setminus C)$, $x \in S$, $B \in \mathcal{B}(S)$. There is the following classical theorem.

Lyapunov–Foster Criterion. Assume that

(α) K has the Feller property and

(β) there exists a nonempty, compact subset $C \subseteq S$ and a measurable function $L : S \to [0, \infty]$ such that

(i) ${}_CKL$ is bounded on C, and
(ii) ${}_CKL \leq L - 1$ off C.

Then K has a stationary distribution.

In the sequel, we assume that K is positive recurrent with (unique) stationary distribution γ. If not specified otherwise, essential infima ess inf and essential suprema ess sup are meant with respect to a maximal irreducibility measure ψ (or γ).

The great scientist to whom this paper is dedicated had a keen interest in measure theory. Since probability theory is one of its major domains of application we hope that he would have found something of interest in it.

2 Balanced partitions

Let $\mathcal{Z} = \{S_1, \cdots, S_N\}$ be partition of S consisting of sets $S_k \in \mathcal{B}^+(S)$ for all $k \in 1..N$ (so that $\gamma(S_k) > 0$). We will often identify $\mathcal{Z}$ with the interval $1..N$ and we will occasionally need the canonical projection $\pi : S \to \mathcal{Z}$. We denote conditional distributions with respect to S_k by γ_{S_k}, $\gamma_{S_k}(B) = \frac{\gamma(B\cap S_k)}{\gamma(S_k)}$. The stochastic matrix $K_{\mathcal{Z}} \in R^{\mathcal{Z}\times\mathcal{Z}}$ defined by

$$K_{\mathcal{Z}}(k,l) = \frac{1}{\gamma(S_k)} \int_{S_k} \gamma(dx) K(x, S_l), \quad k, l \in \mathcal{Z},$$

is called the *ideal aggregation* of K with respect to $\mathcal{Z}$ (and γ), cf. [8], p. 140, [10], and [3]. By the law of total probability, this is equivalent to $K_{\mathcal{Z}}(k,l) = P^{\gamma}[X_1 \in S_l / X_0 \in S_k]$. Although $K_{\mathcal{Z}}$ is a transition matrix for some Markov chain, it is not true in general that this Markov chain is equivalent to the aggregated process $\hat{X}_n = \pi(X_n)$; $\hat{X}_n$ is not even necessarily Markovian. It is well known that $K_{\mathcal{Z}}$ is the transition matrix of the aggregated process if the original chain has some lumpability property [3, 10]. Apart from this application, ideal aggregation has not received much attention yet. The following proposition shows that $K_{\mathcal{Z}}$ possesses the unique stationary distribution γ_π; for finite S, it was proved in [3], Satz 1.

2.1 Proposition Positive recurrence of K implies positive recurrence of $K_{\mathcal{Z}}$. Moreover, $\gamma_\pi = (\gamma(S_1), \cdots, \gamma(S_N))$ is the stationary distribution of $K_{\mathcal{Z}}$.

Proof. Since

$$\sum_k \gamma(S_k) K_{\mathcal{Z}}(k,l) = \int_S \gamma(dx) K(x, S_l) = \gamma(S_l),$$

the stochastic matrix $K_{\mathcal{Z}}$ has the stationary distribution γ_π. In order to show that $K_{\mathcal{Z}}$ is irreducible, let $k, l \in \mathcal{Z}$. Irreducibility of K and strict positivity of $\gamma(S_l)$ imply $L(x, S_l) > 0$ for all $x \in S$. It follows that, for all $x \in S$, there exists an $n \in \mathbb{N}$ such that $K^n(x, S_l) > 0$. In particular,

$$\bigcup_{n\geq 1} \{x \in S_k / K^n(x, S_l) > 0\} = S_k$$

and from $\gamma(S_k) > 0$ we infer that there exists $n \in \mathbb{N}$ such that $\gamma\{x \in S_k / K^n(x, S_l) > 0\} > 0$. This implies

$$\begin{aligned} & \int_{S_k} \gamma(dx_0) \int_E K(x_0, dx_1) \int_E K(x_1, dx_2) \cdots \int_E K(x_{n-2}, dx_{n-1}) K(x_{n-1}, S_l) \\ = & \int_{S_k} \gamma(dx) K^n(x, S_l) > 0. \end{aligned}$$

Since $\mathcal{Z}$ is a partition of S there is a finite sequence $S_k = S^0, S^1, \ldots, S^n = S_l$, $S^j \in \mathcal{Z}$, such that

$$\int_{S^0} \gamma(dx_0) \int_{S^1} K(x_0, dx_1) \int_{S^2} K(x_1, dx_2) \cdots \int_{S^{n-1}} K(x_{n-2}, dx_{n-1}) K(x_{n-1}, S^n) > 0.$$

But this relation together with the fact that γ is stationary implies

$$\begin{aligned} & \gamma(S^j) K_{\mathcal{Z}}(S^j, S^{j+1}) \\ = & \int_{S^j} \gamma(dx_j) K(x_j, S^{j+1}) \\ = & \int_S \gamma(dx_0) \int_S K(x_0, dx_1) \cdots \int_{S^j} K(x_{j-1}, dx_j) \int_{S^{j+1}} K(x_j, dx_{j+1}) \cdots \int_S K(x_{n-2}, dx_{n-1}) K(x_{n-1}, S) \\ > & \ 0 \end{aligned}$$

for all $j \in 0..(n-1)$. Hence, there is a path from S_k to S_l with respect to $K_{\mathcal{Z}}$; this finishes the proof. □

However, this knowledge does not help to gain information on γ in a direct way since the definition of $K_{\mathcal{Z}}$ needs the stationary distribution γ. The following notion appears in [4] in the case of a finite Markov chain; it is called $\mathcal{Z}$–reversibility there.

2.2 Definition Let $\mathcal{Z} = \{S_1, \cdots, S_N\}$ be a partition of S into sets $S_k \in \mathcal{B}^+(S)$ (so that $\gamma(S_k) > 0$). We call $\mathcal{Z}$ *balanced* (with respect to K) if the stationary distribution γ_π of the ideal aggregation $K_{\mathcal{Z}}$ is reversible.

Thus, balancedness of $\mathcal{Z}$ means

$$\gamma(S_k) K_{\mathcal{Z}}(k, l) = \gamma(S_l) K_{\mathcal{Z}}(l, k),$$

or, equivalently,

$$\int_{S_k} \gamma(dx) K(x, S_l) = \int_{S_l} \gamma(dx) K(x, S_k), \tag{1}$$

or also

$$P^\gamma[X_0 \in S_k, X_1 \in S_l] = P^\gamma[X_0 \in S_l, X_1 \in S_k]$$

for all $k, l \in 1..N$. The last two equalities can be interpreted as "balance of flow" between S_k and S_l in equilibrium. Finally, if $K_{\mathcal{Z}}(k, l) > 0$ or, equivalently, $K_{\mathcal{Z}}(l, k) > 0$ then balancedness of $\mathcal{Z}$ implies

$$\frac{\gamma(S_k)}{\gamma(S_l)} = \frac{\int_{S_l} \gamma_{S_l}(dx) K(x, S_k)}{\int_{S_k} \gamma_{S_k}(dx) K(x, S_l)}. \tag{2}$$

2.3 Remarks (a) The stationary distribution γ of K is reversible if and only if any finite partition of S is balanced. If S is finite then this is of course the case if and only if the finest partition $\mathcal{Z} \simeq S$ is balanced.

(b) Any splitting into *two* measurable subsets is balanced. (This case appears implicitly already in [7], Lemma 1.4.) It follows that balancedness of a partition does not imply that the aggregated process is Markovian. It is sufficient to consider a 3×3–permutation matrix and any partition into two classes.

(c) A notion related to balancedness is the recently introduced partition reversibility [1] of an ergodic Markov *jump* processes. Transferred to a positive recurrent Markov matrix K on a discrete state space, *partition reversibility* over $\mathcal{Z} = \{S_1, \cdots, S_N\}$ would mean

$$\gamma(x)K(x, S_l) = \sum_{y \in S_l} \gamma(y)K(y, x)$$

for all $x \in S_k$, $k, l \in 1..N$. Summing up over $x \in S_k$, we see that this notion of partition reversibility implies balancedness of $\mathcal{Z}$. The converse is not true: A counterexample is furnished by $S = 1..3$, $\mathcal{Z} = \{\{1, 2\}, \{3\}\}$, and

$$K = \begin{pmatrix} 1/4 & 0 & 3/4 \\ 1/2 & 1/2 & 0 \\ 1/3 & 2/3 & 0 \end{pmatrix}.$$

The only stationary distribution is $\gamma = \frac{1}{11}(4, 4, 3)$. As noted in (b), $\mathcal{Z}$ is balanced. Since $\gamma(1)K(1, S_2) = \gamma(1)K(1, 3) \neq \gamma(3)K(3, 1)$, K is not partition reversible over $\mathcal{Z}$.

The following Propositions 2.4 and 2.5 state conditions equivalent to balancedness.

2.4 Proposition For $\mathcal{Z} = \{S_1, \cdots, S_N\}$ to be balanced it is sufficient (and necessary) that

$$(3) \quad \int_{S_k} \gamma(dx)K(x, S_l) \leq \int_{S_l} \gamma(dx)K(x, S_k)$$

for all $k, l \in \mathcal{Z}$ such that $k < l$.

Proof. Let us show by induction on k that (1) holds for all $l \geq k$. We first have

$$\sum_l \int_{S_0} \gamma(dx)K(x, S_l) = \gamma(S_0) = (\gamma K)(S_0) = \sum_l \int_{S_l} \gamma(dx)K(x, S_0).$$

A comparison of this equality with (3) shows the claim for $k = 0$. Assume now that the claim has been proved up to some $k < N$. Similarly as above, we have

$$\begin{aligned} & \sum_{l \leq k} \int_{S_{k+1}} \gamma(dx)K(x, S_l) + \sum_{l > k} \int_{S_{k+1}} \gamma(dx)K(x, S_l) \\ = \; & \gamma(S_{k+1}) \\ = \; & (\gamma K)(S_{k+1}) \\ = \; & \sum_{l \leq k} \int_{S_l} \gamma(dx)K(x, S_{k+1}) + \sum_{l > k} \int_{S_l} \gamma(dx)K(x, S_{k+1}). \end{aligned}$$

Since the inductive assumption implies for $l \leq k$

$$\int_{S_l} \gamma(dx)K(x, S_{k+1}) = \int_{S_{k+1}} \gamma(dx)K(x, S_l),$$

we have from (4)

$$\sum_{l>k} \int_{S_{k+1}} \gamma(dx)K(x, S_l) = \sum_{l>k} \int_{S_l} \gamma(dx)K(x, S_{k+1})$$

and the claim follows from (3). □

A *coarsening* of $\mathcal{Z}$ is a partition consisting of unions of classes of $\mathcal{Z}$. We call a coarsening $\{C_1, \ldots, C_M\}$ of $\{S_1, \ldots, S_N\}$ *increasing* if $i < j$, $S_k \subseteq C_i$, and $S_l \subseteq C_j$ imply $k < l$. We deal next with heredity with respect to coarsenings of $\mathcal{Z}$.

2.5 Proposition (a) Any coarsening of a balanced partition is balanced.

(b) (**3-Lemma**) A partition $\mathcal{Z}$ is balanced if and only if any increasing coarsening of $\mathcal{Z}$ consisting of three classes is balanced.

Proof. (a) Let $\{N_1, \ldots, N_M\}$ be a partition of $1..N \simeq \mathcal{Z}$, let $C_i = \bigcup_{k \in N_i} S_k$, $i \in 1..M$, and let $\mathcal{Y} = \{C_1, \ldots, C_M\}$ a coarsening of $\mathcal{Z}$. Then, we have

$$\int_{C_i} \gamma(dx)K(x, C_j) = \sum_{k \in N_i} \int_{S_k} \gamma(dx)K(x, \bigcup_{l \in N_j} S_l) = \sum_{k \in N_i} \sum_{l \in N_j} \int_{S_k} \gamma(dx)K(x, S_l).$$

By balancedness of $\mathcal{Z}$, the last expression is symmetric with respect to i and j.

(b) We have to show (1) for all $k, l \in 1..N$. Without loss of generality, let $1 \leq k < l \leq N$ and consider first the partitions

$$\mathcal{Y}_1 := \{\bigcup_{m=1}^{k} S_m, \bigcup_{m=k+1}^{l} S_m, \bigcup_{m=l+1}^{N} S_m\} \quad \text{and} \quad \mathcal{Y}_2 := \{\bigcup_{m=1}^{k} S_m, \bigcup_{m=k+1}^{l-1} S_m, \bigcup_{m=l}^{N} S_m\}.$$

(a union over the empty index set is omitted). Plainly, both $\mathcal{Y}_1$ and $\mathcal{Y}_1$ are increasing coarsenings of $\mathcal{Z}$ consisting of at most three classes. Now, the hypothesis together with the fact that any splitting into two subsets in $\mathcal{B}^+(S)$ is balanced shows that both $\mathcal{Y}_1$ and $\mathcal{Y}_2$ are balanced. Hence,

$$\int_{\bigcup_{m=1}^{k} S_m} \gamma(dx)K(x, \bigcup_{m=k+1}^{l} S_m) = \int_{\bigcup_{m=k+1}^{l} S_m} \gamma(dx)K(x, \bigcup_{m=1}^{k} S_m).$$

and

$$\int_{\bigcup_{m=1}^{k} S_m} \gamma(dx)K(x, \bigcup_{m=k+1}^{l-1} S_m) = \int_{\bigcup_{m=k+1}^{l-1} S_m} \gamma(dx)K(x, \bigcup_{m=1}^{k} S_m).$$

Substracting corresponding sides, we obtain

$$(4) \quad \int_{\bigcup_{m=1}^{k} S_m} \gamma(dx)K(x, S_l) = \int_{S_l} \gamma(dx)K(x, \bigcup_{m=1}^{k} S_m).$$

Consider next the partitions

$$\mathcal{Y}_3 := \{\bigcup_{m=1}^{k-1} S_m, \bigcup_{m=k}^{l} S_m, \bigcup_{m=l+1}^{N} S_m\} \quad \text{and} \quad \mathcal{Y}_4 := \{\bigcup_{m=1}^{k-1} S_m, \bigcup_{m=k}^{l-1} S_m, \bigcup_{m=l}^{N} S_m\}.$$

The same arguments as above yield

$$\int_{\bigcup_{m=1}^{k-1} S_m} \gamma(dx)K(x, \bigcup_{m=k}^{l} S_m) = \int_{\bigcup_{m=k}^{l} S_m} \gamma(dx)K(x, \bigcup_{m=1}^{k-1} S_m),$$

$$\int_{\bigcup_{m=1}^{k-1} S_m} \gamma(dx)K(x, \bigcup_{m=k}^{l-1} S_m) = \int_{\bigcup_{m=k}^{l-1} S_m} \gamma(dx)K(x, \bigcup_{m=1}^{k-1} S_m),$$

and

$$(5) \quad \int_{\bigcup_{m=1}^{k-1} S_m} \gamma(dx)K(x, S_l) = \int_{S_l} \gamma(dx)K(x, \bigcup_{m=1}^{k-1} S_m).$$

Subtracting finally (5) from (4) we obtain (1). □

One obtains the following immediate corollary.

2.6 Corollary The stationary distribution γ is reversible if and only if any partition of S into three subsets is balanced.

In order to verify balancedness of a partition $\mathcal{Z}$ via Definition 2.2 or via Propositions 2.4, 2.5, one needs γ. We deal next with conditions that are formulated solely in terms of K and a maximal irreducibility measure of K, cf. 1.2. We first show that Kolmogorov's [7], p. 117, famous loop criterion for reversibility of a Markov matrix can be carried over to balancedness. Contrary to the former case, our present condition is only *sufficient* in the latter.

2.7 Proposition (generalized loop condition) Assume

$$\begin{aligned} & K(x_0, S^1)K(x_1, S^2)\ldots K(x_{n-1}, S^n)K(x_n, S^0) \\ (6) \qquad = \; & K(x_0, S^n)K(x_n, S^{n-1})\ldots K(x_2, S^1)K(x_1, S^0) \end{aligned}$$

for all $n \in \mathbf{N}$, all choices $S^0, \ldots, S^n \in \mathcal{Z}$, and for $\psi^{\otimes(n+1)}$–a.a. $(x_0, \ldots, x_n) \in S^0 \times \ldots \times S^n$. Then $\mathcal{Z}$ is balanced.

Proof. We show that the assumption implies the loop criterion for the ideal aggregation, i.e.,

$$K_{\mathcal{Z}}(k_0,k_1)K_{\mathcal{Z}}(k_1,k_2)\dots K_{\mathcal{Z}}(k_n,k_0)=K_{\mathcal{Z}}(k_n,k_{n-1})K_{\mathcal{Z}}(k_{n-1},k_{n-2})\dots K_{\mathcal{Z}}(k_0,k_n)$$

for all $n\in\mathbf{N}$ and all $k_0,...,k_n\in 1..N$. Putting $S^i:=\pi^{-1}(k_i)$, $i\in 0..n$, and $S_0^n=S^0\times...\times S^n$, we have

$$\begin{aligned}
&K_{\mathcal{Z}}(k_0,k_1)K_{\mathcal{Z}}(k_1,k_2)\dots K_{\mathcal{Z}}(k_{n-1},k_n)K_{\mathcal{Z}}(k_n,k_0)\\
=\;&\frac{1}{\prod_i\gamma(S^i)}\int_{S^0}\gamma(dx_0)K(x_0,S^1)\int_{S^1}\gamma(dx_1)K(x_1,S^2)\dots\\
&\qquad\int_{S^{n-1}}\gamma(dx_{n-1})K(x_{n-1},S^n)\int_{S^n}\gamma(dx_n)K(x_n,S^0)\\
=\;&\frac{1}{\prod_i\gamma(S^i)}\int_{S_0^n}\gamma(dx_0)\gamma(dx_1)\dots\gamma(dx_{n-1})\gamma(dx_n)K(x_0,S^1)K(x_1,S^2)\dots K(x_{n-1},S^n)K(x_n,S^0)\\
=\;&\frac{1}{\prod_i\gamma(S^i)}\int_{S_0^n}\gamma(dx_1)\gamma(dx_2)\dots\gamma(dx_n)\gamma(dx_0)K(x_1,S^0)K(x_2,S^1)\dots K(x_n,S^{n-1})K(x_0,S^n)\\
=\;&\frac{1}{\prod_i\gamma(S^i)}\int_{S^1}\gamma(dx_1)K(x_1,S^0)\int_{S^2}\gamma(dx_2)K(x_2,S^1)\dots\\
&\qquad\int_{S^n}\gamma(dx_n)K(x_n,S^{n-1})\int_{S^0}\gamma(dx_0)K(x_0,S^n)\\
=\;&K_{\mathcal{Z}}(k_1,k_0)K_{\mathcal{Z}}(k_2,k_1)\dots K_{\mathcal{Z}}(k_n,k_{n-1})K_{\mathcal{Z}}(k_0,k_n),
\end{aligned}$$

where the third equality follows from the generalized loop condition (6) and equivalence of ψ and γ. This is the claim. □

2.8 Notation Let us define a graph $G_K^{\mathcal{Z}}=(1..N,E_K^{\mathcal{Z}})$ on $1..N$ associated with K and $\mathcal{Z}$ by putting

$$E_K^{\mathcal{Z}}:=\{(k,l)\in 1..N|\exists_{x\in S_k}[K(x,S_l)>0]\text{ or }\exists_{x\in S_l}[K(x,S_k)>0]\}.$$

2.9 Corollary If $G_K^{\mathcal{Z}}$ is a tree then $\mathcal{Z}$ is balanced.

Proof. Since $G_K^{\mathcal{Z}}$ has no loops, the only way to have a chain $x_0,S^1,x_1,S^2,\dots,x_{n-1},S^n,x_n,S^0$ as required on the left side of the generalized loop condition (6) is to balance each of its links x_k,S^l with its counterpart x_l,S^k. But then, the assumption of 2.7 is plainly satisfied since both products consist of the same factors. □

The following immediate consequence of Corollary 2.9 concerns stochastic *block-band* matrices.

2.10 Corollary With the notation $S_0=S_{N+1}=\emptyset$ assume that $K(x,S_k)=0$ for all $x\notin S_{k-1}\cup S_k\cup S_{k+1}$, $1\le k\le N$. Then $\mathcal{Z}$ is balanced.

The generalized loop condition is sufficient for balancedness. The following proposition deals with its necessity. Let us call K *lumpable* for $\mathcal{Z}$ if, for all $k,l\in\mathcal{Z}$ and for $\psi\otimes\psi$–a.a. $(x,y)\in S_k\times S_k$, we have

(7) $\quad K(x,S_l)=K(y,S_l).$

2.11 Proposition If K is lumpable for $\mathcal{Z}$ then the following statements are equivalent.

(a) $\mathcal{Z}$ is balanced;

(b) the generalized loop condition (6) is satisfied for all $n \in \mathbf{N}$, all choices $S^0, \dots, S^n \in \mathcal{Z}$, and for $(x_0, \dots, x_n)$ in some subset of $S^0 \times \dots \times S^n$ of strictly positive $\psi^{\otimes(n+1)}$–measure;

(c) the generalized loop condition (6) is satisfied for all $n \in \mathbf{N}$, all choices $S^0, \dots, S^n \in \mathcal{Z}$ and for $\psi^{\otimes(n+1)}$–a.a. $(x_0, \dots, x_n) \in S^0 \times \dots \times S^n$.

Proof. Integrating (7) with respect to γ_{S_k} over $y \in S_k$ we obtain first

$$(8) \quad K(x, S_l) = \int_{S_k} \gamma_{S_k}(dy) K(y, S_l) = K_{\mathcal{Z}}(k, l)$$

for ψ–a.a. $x \in S_k$.

We now show that (a) implies (b). Given sets $S^i \in \mathcal{Z}$, $i \in 0..n$, balancedness and (8) imply

$$\begin{aligned}
& \prod_k \gamma(S^k) \cdot K(x_0, S^1) K(x_1, S^2) \dots K(x_{n-1}, S^n) K(x_n, S^0) \\
= & \prod_k \gamma(S^k) \cdot K_{\mathcal{Z}}(S^0, S^1) K_{\mathcal{Z}}(S^1, S^2) \dots K_{\mathcal{Z}}(S^{n-1}, S^n) K_{\mathcal{Z}}(S^n, S^0) \\
= & \prod_k \gamma(S^k) \cdot K_{\mathcal{Z}}(S^0, S^n) K_{\mathcal{Z}}(S^n, S^{n-1}) \dots K_{\mathcal{Z}}(S^2, S^1) K_{\mathcal{Z}}(S^1, S^0) \\
= & \prod_k \gamma(S^k) \cdot K(x_0, S^n) K(x_n, S^{n-1}) \dots K(x_2, S^1) K(x_1, S^0)
\end{aligned}$$

and, hence, (6) for $\psi^{\otimes(n+1)}$–a.a. $(x_0, \dots, x_n) \in S^0 \times \dots \times S^n$.

If (b) holds then (8) implies (6) $\psi^{\otimes(n+1)}$–almost everywhere, i.e., (c). The fact that (c) implies (a) was proved in Proposition 2.7.

3 Estimation of probabilities of events

Our main interest in balanced partitions is their ability to furnish lower and upper estimates of the stationary distribution γ. While ideal aggregation cuts down the size of the state space and would thus be useful for efficiently analyzing the stationary distribution, the crux with it is that it needs this very distribution for its definition. If we want to gain lower and upper estimates of the *fractions* $\frac{\gamma(S_k)}{\gamma(S_l)}$, $k, l \in 1..N, k \neq l$, without knowing the stationary distribution γ then the equality $\gamma K = \gamma$ always yields

$$(9) \quad \int_{S_k} \gamma(dx) K(x, S_l) + \int_{S_l} \gamma(dx) K(x, S_l) \leq \gamma(S_l)$$

or, equivalently,

$$\int_{S_k} \gamma(dx) K(x, S_l) \leq \int_{S_l} \gamma(dx)(1 - K(x, S_l)).$$

Passing to the conditional distributions γ_{S_k} and γ_{S_l}, elementary algebraic operations show

$$\frac{\gamma(S_k)}{\gamma(S_l)} \le \frac{\int_{S_l} \gamma_{S_l}(dx)(1-K(x,S_l))}{\int_{S_k} \gamma_{S_k}(dx)K(x,S_l)} \le \frac{1-\text{ess inf}_{x\in S_l} K(x,S_l)}{\text{ess inf}_{x\in S_k} K(x,S_l)} = \frac{\text{ess sup}_{x\in S_l} K(x,\complement S_l)}{\text{ess inf}_{x\in S_k} K(x,S_l)}$$

and an analogous lower estimate of $\gamma(S_k)/\gamma(S_l)$. The following statements show how balancedness allows to improve these estimates. The idea behind them is simple but the method turns out to be very useful for the analysis of stationary distributions of Markov chains. We deal with upper bounds of $\gamma(S_k)/\gamma(S_l)$, only, since lower bounds follow easily from them by swapping k and l. Let us abbreviate $S_{k,l} = \{x \in S_k / K(x,S_l) > 0\}$, $k,l \in \mathcal{Z}$.

3.1 Proposition Let $\mathcal{Z}$ be balanced.

(a) If $k,l \in \mathcal{Z}$ are such that $K_{\mathcal{Z}}(k,l) > 0$ or $K_{\mathcal{Z}}(l,k) > 0$ then we have $\gamma(S_{k,l}) > 0$ and $\gamma(S_{l,k}) > 0$. Moreover,

$$(10)\quad \frac{\gamma(S_{k,l})}{\gamma(S_{l,k})} = \frac{\int_{S_{l,k}} \gamma_{S_{l,k}}(dx)K(x,S_k)}{\int_{S_{k,l}} \gamma_{S_{k,l}}(dx)K(x,S_l)}.$$

(b) If μ and ν are two distributions on $\mathbb{R}$ such that $(\gamma_{S_{l,k}})_{K(\cdot,S_k)} \preceq \mu$ and $\nu \preceq (\gamma_{S_{k,l}})_{K(\cdot,S_l)}$ then $E\,\mu > 0$ and we have the estimate

$$(11)\quad \frac{\gamma(S_{k,l})}{\gamma(S_{l,k})} \le \frac{E\,\mu}{E\,\nu} \qquad \text{(the right side may be infinite).}$$

Proof. (a) Strict positivity of $K_{\mathcal{Z}}(k,l)$, balancedness of $\mathcal{Z}$, and the standing assumption $\gamma(S_k) > 0$ together imply $K_{\mathcal{Z}}(l,k) > 0$, i.e., we have $K_{\mathcal{Z}}(l,k) > 0$ in any case. Now, the estimate

$$(12)\quad 0 < K_{\mathcal{Z}}(l,k) = \frac{1}{\gamma(S_l)} \int_{S_l} \gamma(dx)K(x,S_k) = \frac{1}{\gamma(S_l)} \int_{S_{l,k}} \gamma(dx)K(x,S_k).$$

shows that $\gamma(S_{l,k}) > 0$. By symmetry, we also have $\gamma(S_{k,l}) > 0$. In order to prove (10) note that, by balancedness (1) of $\mathcal{Z}$, we have

$$1 = \frac{\int_{S_l} \gamma(dx)K(x,S_k)}{\int_{S_k} \gamma(dx)K(x,S_l)} = \frac{\int_{S_{l,k}} \gamma(dx)K(x,S_k)}{\int_{S_{k,l}} \gamma(dx)K(x,S_l)} = \frac{\gamma(S_{l,k}) \int_{S_{l,k}} \gamma_{S_{l,k}}(dx)K(x,S_k)}{\gamma(S_{k,l}) \int_{S_{k,l}} \gamma_{S_{k,l}}(dx)K(x,S_l)}.$$

This is the claim.

(b) The claim on strict positivity in (b) follows from (12) and the assumption $(\gamma_{S_{l,k}})_{K(\cdot,S_k)} \preceq \mu$. For proving (11), use (10) to estimate

$$\frac{\gamma(S_{k,l})}{\gamma(S_{l,k})} = \frac{\int_{S_{l,k}} \gamma_{S_{l,k}}(dx)K(x,S_k)}{\int_{S_{k,l}} \gamma_{S_{k,l}}(dx)K(x,S_l)} = \frac{E(\gamma_{S_{l,k}})_{K(\cdot,S_k)}}{E(\gamma_{S_{k,l}})_{K(\cdot,S_l)}} \le \frac{E\mu}{E\nu}. \qquad \square$$

Note that, in (11), we have nothing lost yet: by (2), there is equality in (11) for $\mu = (\gamma_{S_{l,k}})_{K(\cdot,S_k)}$ and $\nu = (\gamma_{S_{k,l}})_{K(\cdot,S_l)}$. However, these measures depend on the stationary distribution and it remains the problem to find suitable distributions ν and μ that allow to control their expectations. Our following propositions exploit two general choices.

3.2 Proposition Let $\mathcal{Z}$ be balanced. If $k, l \in \mathcal{Z}$ are such that $K_{\mathcal{Z}}(l,k) > 0$ then the number $\operatorname{ess\,sup}_{x \in S_{l,k}} K(x, S_k)$ is strictly positive and we have the estimate

$$\frac{\gamma(S_{k,l})}{\gamma(S_{l,k})} \leq \frac{\operatorname{ess\,sup}_{x \in S_{l,k}} K(x,S_k)}{\operatorname{ess\,inf}_{x \in S_{k,l}} K(x,S_l)} = \frac{\operatorname{ess\,sup}_{x \in S_l} K(x,S_k)}{\operatorname{ess\,inf}_{x \in S_{k,l}} K(x,S_l)}$$

(the bound may again be infinite).

Proof. Strict positivity follows from (12). For the first bound, put $\mu = \delta_{\operatorname{ess\,sup}_{x \in S_{l,k}} K(x,S_k)}$ and $\nu = \delta_{\operatorname{ess\,inf}_{x \in S_{k,l}} K(x,S_l)}$ and use Proposition 3.1(b). For the second bound, note that $\operatorname{ess\,sup}_{x \in S_{l,k}} K(x,S_k) = \operatorname{ess\,sup}_{x \in S_l} K(x,S_k)$. □

Our next proposition needs a few preliminaries.

3.3 Lemma Let ρ and σ be two finite measures on the real line such that $\rho \leq \sigma$, i.e., $\int d\rho f \leq \int d\sigma f$ for all *positive*, measurable functions f on $\mathbb{R}$.
(a) If the interval $[s, \infty[$ is a support of σ then

$$\rho + (\sigma(\mathbb{R}) - \rho(\mathbb{R}))\delta_s \preceq \sigma.$$

(b) If the interval $]-\infty, s]$ is a support of σ then

$$\sigma \preceq \rho + (\sigma(\mathbb{R}) - \rho(\mathbb{R}))\delta_s.$$

□

3.4 Notation and explanation Let K be recurrent and let $B \in \mathcal{B}^+(S)$. The first return distribution of the chain X into B starting from $x \in B$ is the Markovian kernel

$$K_B(x,F) = \sum_{t \geq 1} P^{\delta_x}[X_t \in F, T_B = t], \qquad F \subseteq B.$$

This kernel describes the motion of the chain X in B. Its stationary distribution is thus γ_B, $\gamma_B K_B = \gamma_B$. Now, for $t \geq 2$ and $F \subseteq B$, we have

$$\begin{aligned}
& P^{\delta_x}[X_t \in F, T_B = t] \\
= \ & P^{\delta_x}[X_1 \in \complement B, \ldots, X_{t-1} \in \complement B, X_t \in F] \\
= \ & \int_{\complement B \times \ldots \times \complement B} P^{\delta_x}[X_1 \in dy_1, \ldots, X_{t-1} \in dy_{t-1}, X_t \in F] \\
= \ & \int_{\complement B \times \ldots \times \complement B} K(x,dy_1)K(y_1,dy_2)\ldots K(y_{t-2},dy_{t-1})K(y_{t-1},F) \\
= \ & \int_{\complement B \times \complement B} K(x,dy_1)K^{t-2}_{\complement B,\complement B}(y_1,dy_{t-1})K(y_{t-1},F) \\
= \ & K_{B,\complement B}K^{t-2}_{\complement B,\complement B}K_{\complement B,B}(x,F).
\end{aligned}$$

Hence, the first return distribution to B has the analytic representation

$$K_B = K_{B,B} + K_{B,\complement B}(\sum_{t\geq 0}(K_{\complement B,\complement B})^t)K_{\complement B,B}.$$

For $n \geq 0$, let us abbreviate

$$K_{B;n} = K_{B,B} + K_{B,\complement B}(\sum_{t<n}(K_{\complement B,\complement B})^t)K_{\complement B,B}.$$

$K_{B;n}$ is a sub–Markovian kernel on B and, as n tends to ∞, the sequence $K_{B;n}$ increases to K_B. The following proposition uses the kernels $K_{S_{k,l};n}$ for estimating γ.

3.5 Proposition Let $\mathcal{Z}$ be balanced and let $k, l \in \mathcal{Z}$ such that $K_{\mathcal{Z}}(l,k) > 0$. For all $u \geq 0$, the numerator on the following right side is strictly positive and we have for all $u, v \geq 0$

$$\begin{aligned} &\frac{\gamma(S_{k,l})}{\gamma(S_{l,k})} \\ \leq\ &\frac{\text{ess sup}_{x\in S_{l,k}}(K_{S_{l,k};u}K_{S_{l,k},S_k})(x,S_k) + (1 - \text{ess inf}_{x\in S_{l,k}} K_{S_{l,k};u}(x,S_{l,k}))\text{ess sup}_{x\in S_{l,k}} K(x,S_k)}{\text{ess inf}_{x\in S_{k,l}}(K_{S_{k,l};v}K_{S_{k,l},S_l})(x,S_l) + (1 - \text{ess sup}_{x\in S_{k,l}} K_{S_{k,l};v}(x,S_{k,l}))\text{ess inf}_{x\in S_{k,l}} K(x,S_l)}. \end{aligned}$$

Proof. In view of Proposition 3.1(b), define

$$\mu = \left(\gamma_{S_{l,k}}K_{S_{l,k};u}\right)_{K(\cdot,S_k)} + \left(1 - (\gamma_{S_{l,k}}K_{S_{l,k};u})(S_{l,k})\right)\delta_{\text{ess sup}_{x\in S_{l,k}} K(x,S_k)} \quad \text{and}$$
$$\nu = \left(\gamma_{S_{k,l}}K_{S_{k,l};v}\right)_{K(\cdot,S_l)} + \left(1 - (\gamma_{S_{k,l}}K_{S_{k,l};v})(S_{k,l})\right)\delta_{\text{ess inf}_{x\in S_{k,l}} K(x,S_l)}.$$

By 3.4, $\gamma_{S_{l,k}}K_{S_{l,k};u} \leq \gamma_{S_{l,k}}K_{S_{l,k}} = \gamma_{S_{l,k}}$ and, hence, $(\gamma_{S_{l,k}}K_{S_{l,k};u})_{K(\cdot,S_k)} \leq (\gamma_{S_{l,k}})_{K(\cdot,S_k)}$. Since the interval $\left[\text{ess inf}_{x\in S_{l,k}} K(x,S_k), \text{ess sup}_{x\in S_{l,k}} K(x,S_k)\right]$ is a support of $(\gamma_{S_{l,k}})_{K(\cdot,S_k)}$, Lemma 3.3(b) implies $(\gamma_{S_{l,k}})_{K(\cdot,S_k)} \preceq \mu$. Similarly, $(\gamma_{S_{k,l}}K_{S_{k,l};v})_{K(\cdot,S_l)} \leq (\gamma_{S_{k,l}})_{K(\cdot,S_l)}$ and $\left[\text{ess inf}_{x\in S_{k,l}} K(x,S_l), \text{ess sup}_{x\in S_{k,l}} K(x,S_l)\right]$ is a support of $(\gamma_{S_{k,l}})_{K(\cdot,S_l)}$. Hence, Lemma 3.3(a) implies $\nu \preceq (\gamma_{S_{k,l}})_{K(\cdot,S_l)}$. Furthermore, we have

$$\begin{aligned} E\mu &= \int(\gamma_{S_{l,k}}K_{S_{l,k};u})(dx)K(x,S_k) + (1 - (\gamma_{S_{l,k}}K_{S_{l,k};u})(S_{l,k}))\text{ess}\sup_{x\in S_{l,k}} K(x,S_k) \\ &= \int\gamma_{S_{l,k}}(dx)(K_{S_{l,k};u}K_{S_{l,k},S_k})(x,S_k) + \left(1 - \int\gamma_{S_{l,k}}(dx)K_{S_{l,k};u}(x,S_{l,k})\right)\text{ess}\sup_{x\in S_{l,k}} K(x,S_k) \\ &\leq \text{ess}\sup_{x\in S_{l,k}}(K_{S_{l,k};u}K_{S_{l,k},S_k})(x,S_k) + (1 - \text{ess}\inf_{x\in S_{l,k}} K_{S_{l,k};u}(x,S_{l,k}))\text{ess}\sup_{x\in S_{l,k}} K(x,S_k) \end{aligned}$$

and

$$\begin{aligned} E\nu &= \int(\gamma_{S_{k,l}}K_{S_{k,l};v})(dx)K(x,S_l) + (1 - (\gamma_{S_{k,l}}K_{S_{k,l};v})(S_{k,l}))\text{ess}\inf_{x\in S_k} K(x,S_l) \\ &= \int\gamma_{S_{k,l}}(dx)(K_{S_{k,l};v}K_{S_{k,l},S_l})(x,S_l) + \left(1 - \int\gamma_{S_{k,l}}(dx)K_{S_{k,l};v})(x,S_{k,l})\right)\text{ess}\inf_{x\in S_{k,l}} K(x,S_l) \\ &\geq \text{ess}\inf_{x\in S_{k,l}}(K_{S_{k,l};v}K_{S_{k,l},S_l})(x,S_l) + (1 - \text{ess}\sup_{x\in S_{k,l}} K_{S_{k,l};v}(x,S_{k,l}))\text{ess}\inf_{x\in S_{k,l}} K(x,S_l). \end{aligned}$$

The claim now follows from Proposition 3.1(b). □

3.6 Use of shortest paths The bounds given in Propositions 3.2 and 3.5 are meaningless if the denominators vanish. There is sometimes a way to efficiently improve given estimates of quotients and to create additional bounds. This allows to draw a finer picture of γ. The improvement uses an algorithm for the *shortest path problem* in a directed graph and may, in particular, be applied to parameter–free transition kernels. Let $\frac{\gamma(B)}{\gamma(C)} \le U(B,C) \le \infty$, $B, C \in \mathcal{S}$ be the given upper bounds. (If Proposition 3.1, 3.2, or 3.5 is used for the initial bounds then $\mathcal{S} = \{S_{k,l} \in \mathcal{B}^+(S)/1 \le k, l \le N, k \ne l\}$. If there is no finite bound then $U(B,C)$ is defined as ∞.) Lower bounds are again treated by symmetry. Consider the weighted, directed graph $\mathcal{G}$ defined by the set $\mathcal{S}$ of vertices, the links (B,C) if there is a *finite* estimate $U(B,C)$, and the corresponding weights $\log U(B,C)$. If $T_{k_1}, \cdots, T_{k_n}, T_{k_1}$ is any (directed) cycle in $\mathcal{G}$ then its length (= sum of weights) is ≥ 0 since $\log U(T_{k_1}, T_{k_2}) + \cdots + \log U(T_{k_n}, T_{k_1}) \ge \log \frac{\gamma(T_{k_1})}{\gamma(T_{k_2})} + \cdots + \log \frac{\gamma(T_{k_n})}{\gamma(T_{k_1})} = 0$. In this situation, if there is a (directed) path from B to C in $\mathcal{G}$ then there is a shortest one, cf. [6]. The length of this shortest path may be smaller than $\log U(B,C)$, no matter whether (B,C) is a link or not. Its exponential is an upper bound of the quotient $\frac{\gamma(B)}{\gamma(C)}$.
By way of example, consider $S = 1..4$, the irreducible transition matrix

$$K = \frac{1}{4} \begin{pmatrix} 1 & 3 & 0 & 0 \\ 1 & 1 & 1 & 1 \\ 0 & 2 & 1 & 1 \\ 0 & 1 & 1 & 2 \end{pmatrix},$$

and $S_1 = \{1\}$, $S_2 = \{2\}$, $S_3 = \{3,4\}$. The stationary distribution is $\gamma = \frac{1}{41}(5, 15, 9, 12)$. The generalized loop condition shows that $\mathcal{Z}$ is balanced and Proposition 3.2 yields the bounds

$$\frac{1}{3} \le \frac{\gamma(1)}{\gamma(2)} \le \frac{1}{3} \quad \text{and} \quad \frac{1}{2} \le \frac{\gamma(2)}{\gamma(3,4)} \le 1.$$

Applying shortest paths, we obtain the additional bounds

$$\frac{1}{6} \le \frac{\gamma(1)}{\gamma(3,4)} \le \frac{1}{3}.$$

4 An application to systems analysis

In order to illustrate the versatility of the previous approach, we apply it to derive a fairness property of a non–reversible Markovian system discussed by communication and queueing theorists.

4.1 A multiclass service system with loss Alexopoulos, El-Tannir, and Serfozo [1] analyze a continuous-time service system with customer classes and class–dependent loss (they call it “blocking” but the quality of the system operation is rather that of a loss system). We deal here with a similar discrete–time system. It has the advantage of allowing (independent) arrival streams with arbitrary distributions. Our system operates as follows. It serves $r \ge 2$ *classes* or *types* of customers. At each time $t \in \mathbb{N}_>$ there arrive a random number $A_{k,t} : (\Omega, P) \to \mathbb{N}$ of customers of class $k \in 1..r$ at the system. We assume that the $A_{k,t}$'s are integrable and that the r processes $(A_{k,t})_{t\in\mathbb{N}}$ are all i.i.d., independent of one another, and independent of the initial number of customers, X_0, on the system. Moreover, we assume that the system can hold only one class of customers at any time. When customers of one class are served, any arrivals of other classes are rejected. Moreover, when the system is empty and customers of different classes arrive simultaneously then a random and uniformly distributed "coin" decides which class enters the system. The system has m servers: if there are $x \in \mathbb{N}$ customers on the system, then $x \wedge m$ of them are served at the same time. The service times of all customers are independent and geometrically distributed with rate q_k. In order to describe the departure process, let us introduce the independent and binomial$_{x,m,q_k}$–distributed random variables $D_{k,t}^{(x)} : (\Omega, P) \to 0..(x \wedge m)$, $x, t \in \mathbb{N}$, $k \in 1..r$, independent of the $A_{k,t}$'s. If, at time $t \in \mathbb{N}$, there are x customers of class k on the system then the number of customers leaving the system at time t is $D_{k,t}^{(x\wedge m)}$.

Denoting the kth unit vector in $\mathbb{N}^r$ by e_k, $k \in 1..r$, we define the state space $S = \bigcup_{k=1}^r \mathbb{N}e_k$, where $\mathbb{N}e_k = \{0, e_k, 2e_k, \cdots\}$. From the description above it is clear that the customer class and number of customers on the system at time $t \in \mathbb{N}$ may be represented by an r–dimensional random variable $X_t : (\Omega, P) \to S$. By the assumptions of independence, the process $(X_t)_{t\geq 0}$ is a Markov chain. It cannot transfer from a state xe_k to a state ye_l, $x, y > 0$, $l \neq k$, unless it passes through 0. In particular, with the notation $R_k = \mathbb{N}_{>}e_k$, the state space S is the star $\{0\} \cup \bigcup_{k=1}^r R_k$ with center 0 and rays R_k, $k \in 1..r$. The transition graph consists of forward arrows of possibly arbitrary lengths (depending on the distribution of $A_{k,0}$) and backward arrows of lengths at most m in each set $\{0\} \cup R_k$.

Let us define the blocks

$$B_0(k) = \{0\}, \quad B_a(k) = \{((a-1)m+1)\, e_k, \ldots, am\, e_k\},$$

$k \in 1..r$, $a > 0$, and the intervals

$$B_b^c(k) = \bigcup_{j=b}^{c} B_j(k), \quad B_b^\infty(k) = \bigcup_{j\geq b} B_j(k),$$

$0 \leq b \leq c$. If it is clear from the context which k is referred to we often omit it. Some entries of the transition matrix are

$$K(0,0) = \prod_{k=1}^{r} P[A_{k,0} = 0] \quad \text{and}$$

(13) $$K(xe_k, ye_k) = P[A_{k,0} - D_{k,0}^{(x\wedge m)} = y - x], \qquad y \geq 0,$$

for all $x \geq 1$ and all $k \in 1..r$. Moreover, for these x, k and all $a \geq 0$, we have

(14) $$K(xe_k, B_a(k)) = P[A_{k,0} - D_{k,0}^{(x\wedge m)} \in B_a(k) - x]$$

and

(15) $$K(xe_k, B_a^\infty(k)) = P[A_{k,0} - D_{k,0}^{(x\wedge m)} \in B_a^\infty - x].$$

Finally, if $a \geq 1$ then

$$(16) \quad K(0, B_a^\infty(k)) = P[A_{k,0} \in B_a^\infty] \left[\sum_{l=0}^{r-1} \frac{1}{l+1} \left(\sum_{J \in \binom{1..r \setminus \{k\}}{l}} \prod_{j \in J} P[A_{j,0} > 0] \prod_{j \notin J \cup \{k\}} P[A_{j,0} = 0] \right) \right]$$

As usual, we write $\rho_k = EA_{k,0}/ED_{k,0}^{(m)} = EA_{k,0}/(mq_k)$, $k \in 1..r$. If

$$(17) \quad 0 < P[A_{k,0} = 0] < 1 \quad \text{and} \quad 0 < q_k < 1 \quad \text{for all } k \in 1..r$$

then K is *irreducible*. We assert that it is *positive recurrent* if, in addition,

$$(18) \quad \rho_k < 1, \quad k \in 1..r.$$

Let us prove this assertion by means of the Lyapunov–Foster Criterion. Let $C = B_0 \cup \bigcup_{k=1}^r B_1(k)$ and let $L : S \to \mathbb{R}_+$ be defined by $L(xe_k) = x$ for all $x \in \mathbb{N}$, $k \in 1..r$. First, note that

$$({}_C KL)(0) = \sum_{k=1}^r \sum_{x>0} {}_C K(0, xe_k) L(xe_k) \leq \sum_{k=1}^r \sum_{x>0} P[A_{k,0} = x]x = \sum_{k=1}^r EA_{k,0} < \infty.$$

Moreover, simple algebraic manipulations based on (13) show

$$\begin{aligned}
& ({}_C KL)(xe_k) \\
= \; & \sum_{y>m} K(xe_k, ye_k) L(ye_k) \\
= \; & \sum_{y>m} \sum_{j=0}^{x \wedge m} P[A_{k,0} = y - x + j] P[D_{k,0}^{(x \wedge m)} = j] y \\
= \; & \sum_{j=0}^{x \wedge m} P[D_{k,0}^{(x \wedge m)} = j] \sum_{z > m - x + j} P[A_{k,0} = z](z + x - j) \\
\leq \; & \sum_{j=0}^{x \wedge m} P[D_{k,0}^{(x \wedge m)} = j](EA_{k,0} + x - j) \\
= \; & EA_{k,0} + x - ED_{k,0}^{(x \wedge m)} \\
= \; & L(xe_k) + EA_{k,0} - ED_{k,0}^{(x \wedge m)}
\end{aligned}$$

if $x \geq 1$. Hence, if $1 \leq x \leq m$ then we have $({}_C KL)(xe_k) \leq L(xe_k) + EA_{k,0} - ED_{k,0}^{(x)} < \infty$ and if $x > m$ then $({}_C KL)(xe_k) - L(xe_k) \leq EA_{k,0} - ED_{k,0}^{(m)} < 0$ for all $k \in 1..r$ by assumption (18). Thus, both Conditions (i),(ii) of the Lyapunov–Foster Criterion are satisfied. Therefore, under (17) and (18), the system has a unique stationary distribution γ.

Kolmogorov's loop criterion shows that γ is not reversible in general. However, Corollary 2.9 implies that the partitions

$$\mathscr{Z}_a(k) = \{R_1, \ldots, R_{k-1}, B_0^a(k), B_{a+1}^\infty(k), R_{k+1}, \ldots, R_r\},$$

are balanced for all $a \in \mathbb{N}$ and all $k \in 1..r$. We use *all* these partitions in combination with a recursion over the blocks in order to prove the asymptotic result that, as r increases to infinity, each of the rays R_k receives a fair share of the equilibrium probability, viz., under some mild conditions on the expectations of arrival and service times we have $\inf_{r\geq 1} r \inf_{k\in 1..r} \gamma(R_k) > 0$, cf. Theorem 4.9.

We subdivide the proof of this main theorem in a series of lemmas, some of them of separate interest on their own. In order to apply Section 3, we first compute some minima and maxima of the transition probabilities.

4.2 Lemma Let $a \geq 1$.

(a) We have

$$(19)\quad \min_{x\in B_{a+1}} K(x, B_a) = \min_{x\in B_1} K(x, B_0) = P[A_{k,0} - D_{k,0}^{(m)} = -m] = P[A_{k,0} = 0]q_k^m;$$

$$(20)\quad \max_{x\in B_1} K(x, B_0) = P[A_{k,0} - D_{k,0}^{(1)} = -1] = P[A_{k,0} = 0]q_k;$$

$$(21)\quad \max_{x\in B_{a+1}} K(x, B_a) = P[A_{k,0} - D_{k,0}^{(m)} \in (-m)..(-1)].$$

(b) For all $1 \leq a < b$ we have

$$(22)\quad \min_{x\in B_a} K(x, B_b^\infty) = \begin{cases} P[A_{k,0} - D_{k,0}^{(1)} \geq (b-1)m], & a = 1, \\ P[A_{k,0} - D_{k,0}^{(m)} \geq (b-a)m], & a \geq 2; \end{cases}$$

$$(23)\quad \max_{x\in B_a} K(x, B_b^\infty) = P[A_{k,0} - D_{k,0}^{(m)} \geq (b-a-1)m+1].$$

(c) The sequences

$$\left(K(0, B_a^\infty), \max_{x\in B_1} K(x, B_a^\infty), \cdots, \max_{x\in B_{a-1}} K(x, B_a^\infty)\right) \text{ and}$$

$$\left(K(0, B_a^\infty), \min_{x\in B_1} K(x, B_a^\infty), \cdots, \min_{x\in B_{a-1}} K(x, B_a^\infty)\right)$$

are increasing.

Proof. The claims follow from (13), (14), (15) and the assumption of independence between the various random variables if one takes into account the relations

$$(24)\quad D_{k,0}^{(1)} \preceq \ldots \preceq D_{k,0}^{(m)} \text{ and } D_{k,0}^{(1)} - 1 \succeq \ldots \succeq D_{k,0}^{(m)} - m.$$

Specifically, the minima (19) are assumed at the states $x = (a+1)m$ and $x = m$, respectively, the maximum (20) at the state $x = 1$, that in (21) at $x = am + 1$, the minimum (22) at $x = (a-1)m + 1$, and the maximum (23) at $x = am$.

As a sample proof, let us derive (22). By (15),

$$\min_{x \in B_a} K(x, B_b^\infty) = \min_{x \in B_a} P[A_{k,0} - (D_{k,0}^{x \wedge m} - x) \in B_b^\infty].$$

Now, since the random variables $D_{k,0}^{x\wedge m} - x$ are stochastically decreasing as $x \in B_a$ increases the sequence $(A_{k,0} - (D_{k,0}^{x\wedge m} - x))_{x \in B_a}$ is stochastically increasing; this proves the claim. □

The normalizations of the following measures μ_k^L and μ_k^U on $\mathbb{N}$ will turn out to be lower and upper stochastic bounds of the coarsening with respect to the blocks of the restriction of γ to R_k. We define these measures recursively.

4.3 Definitions Let $k \in 1..r$.

(a) Let $\mu_k^L(0) = 1$. If $\mu_k^L(0), \cdots, \mu_k^L(a)$ are defined, let

$$(25)\quad \mu_k^L(a+1) = \frac{\sum_{b=0}^{a} \mu_k^L(b) \min_{x \in B_b} K(x, B_{a+1}^\infty)}{\max_{x \in B_{a+1}} K(x, B_a)}.$$

(b) Let $\mu_k^U(0) = 1$. If $\mu_k^U(0), \cdots, \mu_k^U(a)$ are defined, let

$$(26)\quad \mu_k^U(a+1) = \frac{\sum_{b=0}^{a} \mu_k^U(b) \max_{x \in B_b} K(x, B_{a+1}^\infty)}{\min_{x \in B_{a+1}} K(x, B_a)}.$$

4.4 Lemma For all $a \in \mathbb{N}$ we have

$$(27)\quad \frac{\mu_k^L(a+1)}{\mu_k^L(0..a)} \le \frac{\gamma(B_{a+1})}{\gamma(B_0^a)} \le \frac{\mu_k^U(a+1)}{\mu_k^U(0..a)},$$

$$(28)\quad \frac{\mu_k^L(a)}{\mu_k^L(0..a)} \le \frac{\gamma(B_a)}{\gamma(B_0^a)} \le \frac{\mu_k^U(a)}{\mu_k^U(0..a)},$$

$$(29)\quad \left(\frac{\mu_k^L(b)}{\mu_k^L(0..a)}\right)_{b \in 0..a} \preceq \left(\frac{\gamma(B_b)}{\gamma(B_0^a)}\right)_{b \in 0..a} \preceq \left(\frac{\mu_k^U(b)}{\mu_k^U(0..a)}\right)_{b \in 0..a},$$

and

$$(30)\quad \mu_k^L(a) \le \frac{\gamma(B_a)}{\gamma(B_0)} \le \mu_k^U(a).$$

Proof. By symmetry, it is sufficient to prove the right hand sides of the estimates, only. We first prove (28), (29), and (27) simultaneously using mathematical induction. The estimates (28) and (29) are trivial for $a = 0$. In order to prove (27) for $a = 0$, we apply Proposition 3.2 with $S_k := B_1^\infty$ and $S_l := B_0$. By (13), $S_{k,l} = B_1$ and $S_{l,k} = B_0$. Therefore,

$$\frac{\gamma(B_1)}{\gamma(B_0)} \leq \frac{K(0, B_1^\infty)}{\min_{x\in B_1} K(x,0)} = \frac{\mu_k^U(1)}{\mu_k^U(0)},$$

which is (27) for $a = 0$. Suppose now that the three estimates have already been proved for all $b < a$. Claim (28) for a follows immediately from (27) for $a-1$. In order to show (29), we estimate

$$\begin{aligned}
&\left(\frac{\gamma(B_b)}{\gamma(B_0^a)}\right)_{b\in 0..a} \\
= &\left(\frac{\gamma(B_0^{a-1})}{\gamma(B_0^a)}\left(\frac{\gamma(B_b)}{\gamma(B_0^{a-1})}\right)_{b<a}, \frac{\gamma(B_a)}{\gamma(B_0^a)}\right) \\
\preceq &\left(\frac{\mu_k^U(0..(a-1))}{\mu_k^U(0..a)}\left(\frac{\mu_k^U(b)}{\mu_k^U(0..(a-1))}\right)_{b<a}, \frac{\mu_k^U(a)}{\mu_k^U(0..a)}\right) \\
= &\left(\frac{\mu_k^U(b)}{\mu_k^U(0..a)}\right)_{b\in 0..a};
\end{aligned}$$

here, the "$\preceq$" relation follows from the induction hypothesis together with the just proved estimate (28).

It remains to prove (27). Let us apply Proposition 3.1(a) to the balanced partition $\mathcal{Z} = \mathcal{Z}_a$ and to $S_k := B_{a+1}^\infty$ and $S_l := B_0^a$. First, (13) implies $S_{k,l} = B_{a+1}$ and $S_{l,k} = B_0^a$. Since we have $K(x, B_0^a) = K(x, B_a)$ for all $x \in B_{a+1}$, Proposition 3.1(a), (29), and Lemma 4.2(c) yield

$$\begin{aligned}
\frac{\gamma(B_{a+1})}{\gamma(B_0^a)} &= \frac{\sum_{x\in B_0^a}\gamma_{B_0^a}(x)K(x,B_{a+1}^\infty)}{\sum_{x\in B_{a+1}}\gamma_{B_{a+1}}(x)K(x,B_0^a)} \leq \frac{\sum_{b=0}^a \frac{\gamma(B_b)}{\gamma(B_0^a)}\max_{x\in B_b} K(x, B_{a+1}^\infty)}{\min_{x\in B_{a+1}} K(x,B_a)} \\
&\leq \frac{\sum_{b=0}^a \frac{\mu_k^U(b)}{\mu_k^U(0..a)}\max_{x\in B_b} K(x, B_{a+1}^\infty)}{\min_{x\in B_{a+1}} K(x,B_a)}.
\end{aligned}$$

By (26), this is the right hand side of (27) and the induction is finished.

The claim (30) is trivial for $a = 0$ and for $a > 0$ we use (27) to estimate

$$\begin{aligned}
&\frac{\gamma(B_a)}{\gamma(B_0)} = \frac{\gamma(B_a)}{\gamma(B_0^{a-1})}\prod_{b=1}^{a-1}\frac{\gamma(B_0^b)}{\gamma(B_0^{b-1})} = \frac{\gamma(B_a)}{\gamma(B_0^{a-1})}\prod_{b=1}^{a-1}(1+\frac{\gamma(B_b)}{\gamma(B_0^{b-1})}) \\
\leq &\frac{\mu_k^U(a)}{\mu_k^U(0..(a-1))}\prod_{b=1}^{a-1}(1+\frac{\mu_k^U(b)}{\mu_k^U(0..(b-1))}) = \frac{\mu_k^U(a)}{\mu_k^U(0..(a-1))}\prod_{b=1}^{a-1}\frac{\mu_k^U(0..b)}{\mu_k^U(0..(b-1))} = \frac{\mu_k^U(a)}{\mu_k^U(0)}. \quad \square
\end{aligned}$$

Had we tried to use Proposition 3.2 instead of 3.1(a), we would have had to define the measures

$$\tilde{\mu}_k^U(a+1) = \frac{\max_{x\in B_a} K(x, B_{a+1}^\infty)}{\min_{x\in B_{a+1}} K(x,B_a)} = \frac{P[A_{k,0} - D_{k,0}^{(m)} \geq 1]}{P[A_{k,0} - D_{k,0}^{(1)} = -m]}.$$

But these measures do not provide insight into the structure of γ since they are infinite. We next show that the measures μ_k^U are finite. We introduce first some notation.

4.5 Notation Let $k \in 1..r$.

(a) $$F_k(a) = \begin{cases} K(0, B_a^\infty(k)), & a \geq 1, \\ 0, & a \leq 0. \end{cases}$$

(b) $$G_k^L(a) = \begin{cases} & \min_{x\in B_1(k)} K(x, B_{a+1}^\infty(k)), & a \geq 1, \\ - & \max_{x\in B_1(k)} K(x, B_0(k)), & a = 0, \\ & 0, & a < 0; \end{cases}$$

$$H_k^L(a) = \begin{cases} & \min_{x\in B_2(k)} K(x, B_{a+2}^\infty(k)), & a \geq 1, \\ - & \max_{x\in B_2(k)} K(x, B_1(k)), & a = 0, \\ & 0, & a < 0. \end{cases}$$

(c) $$H_k^U(a) = \begin{cases} & \max_{x\in B_1(k)} K(x, B_{a+1}^\infty(k)), & a \geq 1, \\ - & \min_{x\in B_1(k)} K(x, B_0(k)), & a = 0, \\ & 0, & a < 0. \end{cases}$$

(d) For all $j \leq m$, let $A_{k,0}^{(j)} : (\Omega, P) \to \mathbb{N}$ be the random variable defined by

$$A_{k,0}^{(j)} = \begin{cases} 0, & \text{if } A_{k,0} \in 0..(j-1), \\ a, & \text{if } A_{k,0} \in B_a(k) + j - 1, \quad a \geq 1. \end{cases}$$

Plainly, we have $A_{k,0}^{(j)} > a$ if and only if $A_{k,0} \geq am + j$ and, hence,

$$EA_{k,0}^{(j)} = \sum_{a\geq 0} P[A_{k,0} \geq am + j]. \tag{31}$$

Moreover, since $(A_{k,0}^{(j)} - 1)m + j \leq A_{k,0} \leq A_{k,0}^{(j)}m + j - 1$, we have for $j \leq m$

$$\frac{EA_{k,0} - j + 1}{m} \leq EA_{k,0}^{(j)} \leq \frac{EA_{k,0} - j + m}{m}. \tag{32}$$

For abbreviation, let us denote the factor in brackets on the right hand side of (16) by $p(r, k)$, $r \in \mathbb{N}_>$, $k \in 1..r$. It is the conditional probability for arriving customers of class k to enter the system.

4.6 Lemma We have

$$p(r,k)\frac{EA_{k,0}}{m} \leq \sum_a F_k(a) \leq p(r,k)\frac{EA_{k,0} + m - 1}{m}, \tag{33}$$

$$\sum_a G_k^L(a) = -1 + \sum_{a\geq 0} P[A_{k,0} - D_{k,0}^{(1)} \geq am], \tag{34}$$

$$\frac{EA_{k,0} - ED_{k,0}^{(m)} + 1 - m}{m} \leq \sum_a H_k^L(a) = -1 + \sum_{a\geq 0} P[A_{k,0} - D_{k,0}^{(m)} \geq am] \leq \frac{EA_{k,0} - ED_{k,0}^{(m)}}{m}, \tag{35}$$

$$\sum_a H_k^U(a) \leq \frac{EA_{k,0} - ED_{k,0}^{(m)} - 1 + m}{m}. \tag{36}$$

Proof. By (16) and (31), we have

$$\sum_{a\geq 0} F_k(a) = \sum_{a\geq 1} K(0, B_a^\infty) = p(r,k)\sum_{a\geq 0} P[A_{k,0} > am] = p(r,k)EA_{k,0}^{(0)}.$$

and the two estimates on F_k follow from (32). Moreover, by (20) and (22) we have

$$\begin{aligned}
&\sum_{a\geq 0} G_k^L(a) \\
= \quad & -P[A_{k,0} - D_{k,0}^{(1)} = -1] + \sum_{a=1}^{\infty} P[A_{k,0} - D_{k,0}^{(1)} \geq am] \\
= \quad & -1 + \sum_{a\geq 0} P[A_{k,0} - D_{k,0}^{(1)} \geq am],
\end{aligned}$$

i.e., (34).

Turning to (35), use (21), (22), and (31) to compute

$$\begin{aligned}
&H_k^L(0) + \sum_{a=1}^{\infty} H_k^L(a) \\
= \quad & -\max_{x\in B_2(k)} K(x, B_1(k)) + \sum_{a=1}^{\infty} \min_{x\in B_2(k)} K(x, B_{a+2}^\infty(k)) \\
= \quad & -P[A_{k,0} - D_{k,0}^{(m)} \in (-m)..(-1)] + \sum_{a=1}^{\infty} P[A_{k,0} - D_{k,0}^{(m)} \geq am] \\
= \quad & -1 + \sum_{a\geq 0} P[A_{k,0} - D_{k,0}^{(m)} \geq am] \\
= \quad & -1 + \sum_{j\leq m} P[D_{k,0}^{(m)} = j] \sum_{a\geq 0} P[A_{k,0} \geq am + j] \\
= \quad & -1 + \sum_{j\leq m} P[D_{k,0}^{(m)} = j] EA_{k,0}^{(j)}.
\end{aligned}$$

The claims now follow from (32).

We, finally, use (19), (23), (31), and (32) to estimate

$$\begin{aligned}
&H_k^U(0) + \sum_{a=1}^{\infty} H_k^U(a) \\
= \quad & -\min_{x\in B_1(k)} K(x, B_0(k)) + \sum_{a=1}^{\infty} \max_{x\in B_1(k)} K(x, B_{a+1}^\infty(k)) \\
= \quad & -P[A_{k,0} - D_{k,0}^{(m)} = -m] + \sum_{a=0}^{\infty} P[A_{k,0} - D_{k,0}^{(m)} \geq am + 1]
\end{aligned}$$

$$
\begin{aligned}
&= -P[D_{k,0}^{(m)} = m]P[A_{k,0} = 0] + P[D_{k,0}^{(m)} = m]\sum_{a=0}^{\infty} P[A_{k,0} \geq (a+1)m + 1] \\
&\quad + \sum_{j=0}^{m-1} P[D_{k,0}^{(m)} = j]\sum_{a=0}^{\infty} P[A_{k,0} \geq am + j + 1] \\
&= -P[D_{k,0}^{(m)} = m]P[A_{k,0} = 0] - P[D_{k,0}^{(m)} = m]P[A_{k,0} > 0] + P[D_{k,0}^{(m)} = m]EA_{k,0}^{(1)} \\
&\quad + \sum_{j=0}^{m-1} P[D_{k,0}^{(m)} = j]EA_{k,0}^{(j+1)} \\
&\leq -P[D_{k,0}^{(m)} = m] + P[D_{k,0}^{(m)} = m]\frac{EA_{k,0} - 1 + m}{m} + \sum_{j=0}^{m-1} P[D_{k,0}^{(m)} = j]\frac{EA_{k,0} - j - 1 + m}{m} \\
&= \sum_{j=0}^{m} P[D_{k,0}^{(m)} = j]\frac{EA_{k,0} - j - 1 + m}{m} \\
&= \frac{EA_{k,0} - ED_{k,0}^{(m)} - 1 + m}{m}. \qquad \square
\end{aligned}
$$

4.7 Lemma

(a) If $ED_{k,0}^{(m)} - EA_{k,0} > 0$ then $\quad p(r,k)\frac{EA_{k,0}}{ED_{k,0}^{(m)} - EA_{k,0} - 1 + m} \leq \sum_{a>0} \mu_k^L(a) < \infty.$

(b) If $ED_{k,0}^{(m)} - EA_{k,0} > m - 1$ then $\quad \sum_{a>0} \mu_k^U(a) \leq p(r,k)\frac{EA_{k,0} - 1 + m}{ED_{k,0}^{(m)} - EA_{k,0} + 1 - m}.$

Proof. (a) The hypothesis implies that γ is finite; therefore, finiteness of μ_k^L follows immediately from (30). Next, it is straightforward to verify using 4.5 that (25) can be rewritten as the convolution equation

$$
(F_k(a) - H_k^L(a)) + \mu_k^L(1)(G_k^L(a-1) - H_k^L(a-1)) + (\mu_k^L \star H_k^L)(a) = 0, \qquad a \geq 0.
$$

Let the measure ν_k^L be equal to μ_k^L with the 1 at the origin replaced with a 0. This measure satisfies the convolution equation

$$
(\nu_k^L \star H_k^L)(a) = \nu_k^L(1)\left(H_k^L(a-1) - G_k^L(a-1)\right) - F_k(a), \qquad a \geq 0.
$$

By Lemma 4.6, all sequences appearing here have finite sums; summation yields

$$
\sum_a H_k^L(a) \sum_a \nu_k^L(a) = \nu_k^L(1)\left(\sum_a H_k^L(a) - \sum_a G_k^L(a)\right) - \sum_a F_k(a).
$$

Now, from Lemma 4.6 and (24) we have $\sum_a H_k^L(a) \leq \sum_a G_k^L(a)$ and, hence,

$$
\sum_{a\geq 0} H_k^L(a) \sum_{a>0} \mu_k^L(a) \leq -\sum_{a\geq 0} F_k(a).
$$

The lower estimate in Part (a) now follows from (33) and (35); note that, by assumption, $\sum_a H_k^L(a)$ is negative.

In order to prove Part (b), first use Lemma 4.2 in order to rewrite the recursion (26) in the form

$$\mu_k^U(a+1)H_k^U(0) + \mu_k^U(0)F_k(a+1) + \sum_{b=1}^{a} \mu_k^U(b)H_k^U(a+1-b) = 0$$

or, equivalently,

$$F_k(a+1) + \sum_{b=1}^{a+1} \mu_k^U(b)H_k^U(a+1-b) = 0, \qquad a \geq 0.$$

Using 4.5(a), we extend this to the convolution equation

$$(37) \quad (\nu_k^U \star H_k^U)(a) = -F_k(a), \qquad a \geq 0,$$

where the measure ν_k^U is equal to μ_k^U except at the origin where it is zero. Since, by assumption and by (36), $\sum_{a\geq 0} H_k^U(a)$ is strictly negative we obtain, summing up (37),

$$(38) \quad -\sum_{a\geq 0} F_k(a) = \sum_{a\geq 0} (\nu_k^U \star H_k^U)(a) = \sum_{a\geq 0} H_k^U(a) \sum_{b>0} \mu_k^U(b).$$

Since $\sum_{a\geq 0} F_k(a)$ is finite by (33) this relation shows that both sums $\sum_{a\geq 0} H_k^U(a)$ and $\sum_{b>0} \mu_k^U(b)$ are finite. Claim (b) finally follows from (38) and Lemma 4.6. □

4.8 Lemma (a) For all $r \in \mathbb{N}_>$ and all $k \leq r$, we have $r\,p(r,k) \geq 1$.
(b) For all $r \in \mathbb{N}_>$ and all $h, k \leq r$, we have $p(r,h) \leq 2\,p(r,k)$.

Proof. (a) By the definition of $p(r,k)$ we have

$$\begin{aligned} r\,p(r,k) &= \sum_{l=0}^{r-1} \frac{r}{l+1} \left(\sum_{J\in\binom{1..r\setminus\{k\}}{l}} \prod_{j\in J} P[A_{j,0} > 0] \prod_{j\notin J\cup\{k\}} P[A_{j,0} = 0] \right) \\ &\geq \sum_{l=0}^{r-1} \sum_{J\in\binom{1..r\setminus\{k\}}{l}} \prod_{j\in J} P[A_{j,0} > 0] \prod_{j\notin J\cup\{k\}} P[A_{j,0} = 0] \\ &= 1. \end{aligned}$$

(b) Let $j, k \in 1..r$, $j \neq k$. The same argument leading to formula (16) yields the equalities

$$\begin{aligned} K_1 &:= P[X_1 \in R_k / X_0 = 0, A_{h,0} > 0, A_{k,0} > 0] \\ &= \sum_{l=0}^{r-2} \frac{1}{l+2} \left(\sum_{I\in\binom{1..r\setminus\{h,k\}}{l}} \prod_{i\in I} P[A_{i,0} > 0] \prod_{i\notin I\cup\{h,k\}} P[A_{i,0} = 0] \right) \\ &= \sum_{l=0}^{r-2} \frac{1}{l+2} q(r,l) \end{aligned}$$

and

$$\begin{aligned} K_2 &:= P[X_1 \in R_k / X_0 = 0, A_{k,0} > 0, A_{j,0} = 0] \\ &= \sum_{l=0}^{r-2} \frac{1}{l+1} \left(\sum_{I \in \binom{1..r \setminus \{h,k\}}{l}} \prod_{i \in I} P[A_{i,0} > 0] \prod_{i \notin I \cup \{h,k\}} P[A_{i,0} = 0] \right) \\ &= \sum_{l=0}^{r-2} \frac{1}{l+1} q(r,l), \end{aligned}$$

where $q(r,l)$ stands for the sum in parantheses. Plainly, $K_1 \leq K_2 \leq 2\,K_1$. Now, the formula of total probability together with the independence assumptions between the various random variables specifying the system shows

$$\begin{aligned} p(r,j) &= P[X_1 \in R_j / X_0 = 0, A_{j,0} > 0] \\ &= P[X_1 \in R_j / X_0 = 0, A_{j,0} > 0, A_{k,0} > 0] P[A_{k,0} > 0] \\ &\quad + P[X_1 \in R_j / X_0 = 0, A_{j,0} > 0, A_{k,0} = 0] P[A_{k,0} = 0] \\ &= K_1 P[A_{k,0} > 0] + K_2 P[A_{k,0} = 0] \end{aligned}$$

and, similarly,

$$p(r,k) = K_1 P[A_{j,0} > 0] + K_2 P[A_{j,0} = 0]$$

and the claim follows from

$$K_1 P[A_{k,0} > 0] + K_2 P[A_{k,0} = 0] \leq K_2 \leq 2K_1 P[A_{j,0} > 0] + 2K_2 P[A_{j,0} = 0]. \qquad \square$$

We are now prepared to prove the main asymptotic result of this section.

4.9 Theorem (Fairness) If

(i) $\inf_{k \geq 1} \rho_k > 0$ and

(ii) $\inf_{k \geq 1} \left(ED_{k,0}^{(m)} - EA_{k,0} \right) > m - 1$

then we have

$$\inf_{r \geq 1} r \inf_{k \in 1..r} \gamma(R_k) > 0.$$

Proof. We start with (30) obtaining

(39) $\mu_k^L(\mathbb{N}_>) \leq \gamma(R_k)/\gamma(\{0\}) \leq \mu_k^U(\mathbb{N}_>)$.

Summing the upper estimate in (39) over all $k \le r$ and adding 1, we find that $\gamma(\{0\})$ is bounded below by $1/(1 + \sum_{k\le r} \mu_k^U(\mathbb{N}_>))$. Thus, by the lower estimate in (39) and by Lemma 4.7, we have

$$\gamma(R_k) \ge \mu_k^L(\mathbb{N}_>)/(1 + \sum_{j\le r} \mu_j^U(\mathbb{N}_>)) \ge \frac{p(r,k)\frac{EA_{k,0}}{ED_{k,0}^{(m)} - EA_{k,0} - 1 + m}}{1 + \sum_{j\le r} p(r,j)\frac{EA_{j,0} - 1 + m}{ED_{j,0}^{(m)} - EA_{j,0} + 1 - m}}.$$

Now, let $\rho := \inf_{k\ge 1} \rho_k$ and $\delta := \inf_{k\ge 1}\left(ED_{k,0}^{(m)} - EA_{k,0}\right) - m + 1$; by the assumptions (i) and (ii), both constants are strictly positive. Moreover, $\delta < 1$,

$$\frac{EA_{k,0}}{ED_{k,0}^{(m)} - EA_{k,0} - 1 + m} = \frac{\rho_k ED_{k,0}^{(m)}}{ED_{k,0}^{(m)} - EA_{k,0} - 1 + m} \ge \frac{\rho_k}{2(1-\rho_k)} \ge \frac{\rho}{2(1-\rho)},$$

and

$$\frac{EA_{j,0} - 1 + m}{ED_{j,0}^{(m)} - EA_{j,0} + 1 - m} \le \frac{m}{\delta} - 1.$$

Collecting the last three assertions and applying Lemma 4.8, we finally obtain

$$r\ \gamma(R_k) \ge \frac{r\ p(r,k)\frac{\rho}{2(1-\rho)}}{1 + \left(\frac{m}{\delta} - 1\right)\sum_{j\le r} p(r,j)} \ge \frac{1}{2}\frac{\rho}{1-\rho}\frac{1}{1 + 2\left(\frac{m}{\delta} - 1\right)}.$$

This proves the theorem. □

4.10 Remark The system is non–reversible even in the single–server case, $m = 1$. Nevertheless, in this case, the estimates in Lemmas 4.4, 4.6, and 4.7 become equalities since all blocks are one–point sets. In particular, we have

$$\mu_k^L(a) = \frac{\gamma(B_a)}{\gamma(B_0)} = \mu_k^U(a)$$

and, after summation, Lemma 4.7 shows

$$\frac{\gamma(R_k)}{\gamma(\{0\})} = \mu_k^L(\mathbb{N}_>) = p(m,k)\frac{\rho_k}{1-\rho_k}.$$

This formula leads to a representation of the equilibrium probabilities of the idle state and the rays analogous to the estimates in the proof of Theorem 4.9.

References

[1] C. Alexopoulos, A.A. El-Tannir, and R.F. Serfozo. Partition-reversible Markov processes. *Operations Research*, 47:125–130, 1999.

[2] Heinz Bauer. *Ma– und Integrationstheorie*. de Gruyter, Berlin–New York, 1990. Zu Kapitel 2.

[3] P. Buchholz. Exact and ordinary lumpability in finite Markov chains. *J. Appl. Prob.*, 31:59–75, 1994.

[4] María Teresa Gallegos-Wendel. *$\mathcal{Z}$–Reversibilität von Markov–Ketten und diskretes Verstärkungslernen*. PhD thesis, Fakultät für Mathematik und Informatik, Universität Passau, Shaker Verlag, Aachen, 1997.

[5] Edwin Hewitt and Karl Stromberg. *Real and Abstract Analysis*. Springer, Berlin–Heidelberg–New York, 1969.

[6] D. Jungnickel. *Graphen, Netzwerke und Algorithmen*. Bibl. Institut, 1987.

[7] Frank P. Kelly. *Reversibility and Stochastic Networks*. Wiley, Chichester, New York, Brisbane, Toronto, 1979.

[8] J. Kemeny and L. Snell. *Finite Markov Chains*. Springer, New York–Berlin–Heidelberg–Tokyo, 1976.

[9] S.P. Meyn and R.L. Tweedie. *Markov Chains and Stochastic Stability*. Springer, London–Berlin–Heidelberg–New York etc., 1993.

[10] G. Rubino and B. Sericola. On weak lumpability in Markov chains. *J. Appl. Prob.*, 26:744–756, 1989.

María Teresa Gallegos
Gunter Ritter

Universität Passau
Fakultät für Mathematik und Informatik
94 030 Passau, Germany