A Successive Lumping Procedure for a Class of Markov Chains

DRAFT 2012-April-23.

Please do not cite or distribute without permission, comments welcome

Michael N. Katehakis
Department of Management Science and Information Systems
Rutgers Business School, Newark and New Brunswick
Newark NJ 07210, USA
Email: mnk@rutgers.edu

Laurens C. Smit
Mathematisch Instituut,
Universiteit Leiden
Niels Bohrweg 1, 2333 CA, The Netherlands
Email: lsmit@math.leidenuniv.nl

Abstract

A class of Markov chains we call successively lumpable is specified for which it is shown that the stationary probabilities can be obtained by successively computing the stationary probabilities of a propitiously constructed sequence of Markov chains. Each of the latter chains has a(typically much) smaller state space and this yields significant computational improvements. We discuss how the results for discrete time Markov chains extend to semi-Markov processes and continuous time Markov processes. Finally we will study applications of successively lumpable Markov chains to classical reliability and queueing models.

Keywords: Markov Chains; Lumping; Reliability; Queueing

1. INTRODUCTION

In this paper we identify a class of Markov chains that we call successively lumpable, for which it is shown that the stationary probabilities can be obtained by successively computing the stationary probabilities of a propitiously constructed sequence of Markov chains. Each of the latter chains has a, typically much, smaller state space and a successive method of solution becomes possible with significant computational improvements. Lumping of states was first discussed in Kemeny and Snell [8] in 1960. Methods and benefits of aggregation/disaggregation are thoroughly described in Schweitzer, Puterman and Kindle [16], Miranker and Pan [12] and Yap [18]. In our construction the new key idea is to identify conditions (cf., Definition 2) on the transition matrix of the Markov chain under which it is *successively lumpable*. A necessary condition for a chain to be successively lumpable is the existence of "entrance states" cf., Definition 2. These states are called "input states" by Feinberg and Chui [2] and they are a special case of the "mandatory states" which have been studied in Kim and Smith [9] and Kim and Smith [10].

The paper is organized as follows. In Section 2.1 after some preliminaries we provide the basic framework for the first lumping stage. The successively lumpable class of Markov chains is defined in Section 2.2 and their main properties are given in Theorems 2 and 3. These theorems are the main results of the paper. In Section 3 another class of Markov chains is introduced for which, using our results of Section 2, we construct a multiple successive lumping procedure. In Section 4, we discuss the ramifications of the work in Sections 2 and 3 to the case of semi-Markov processes and continuous time Markov processes. In Section 5 we study applications of successively lumpable Markov chains to certain classical reliability/queueing problems. Versions of these reliability/queueing models have been studied before in Derman, Lieberman and Ross [1], Frostig [3], Hooghiemstra and Koole [4], Katehakis and Derman [6] as well as in Katehakis and Melolidakis [7], Righter [14], Zhang, Mi, Riska and Smirni [19] and references therein.

2. SUCCESSIVELY LUMPABLE MARKOV CHAINS

Let $X(t)$ denote an irreducible and positive recurrent Markov chain on a finite or countable state space $\mathcal{X}$. Clearly $\mathcal{X}$ can be partitioned into a (possibly infinite) sequence of mutually *exclusive* and *exhaustive* sets $\mathcal{D} := \{D_0, D_1, \ldots, D_M\}$, with $M \leqslant \infty$, $\cup_{m=0}^{M} D_m = \mathcal{X}$, and $D_m \cap D_{m'} = \emptyset$, when $m \neq m'$. For notational convenience, the elements of each set D_m will be denoted (relabelled) as $\{(m,1),(m,2),\ldots,(m,\ell_m)\}$, for some fixed constants $\ell_m \leqslant \infty$. The transition matrix of $X(t)$ will be denoted by $\underline{\underline{P}} = [p(m,j \,|\, m',j')]$, where its $((m',j'),(m,j))$−element is

$$p(m,j \,|\, m',j') = \mathbf{Pr}[X(t+1) = (m,j) \,|\, X(t) = (m',j')].$$

In the sequel we will denote the stationary probabilities with $\pi(m,j) = \lim_{t\to\infty} \mathbf{Pr}[X(t) = (m,j)]$. These probabilities exist, because the Markov chain $X(t)$ is irreducible and positive recurrent. We will use the notation

$$\underline{\pi} = \big(\pi(0,1),\ldots,\pi(0,\ell_0),\pi(1,1),\ldots,\pi(1,\ell_1),\ldots,\pi(M,1),\ldots,\pi(M,\ell_M)\big).$$

In the sequel, to avoid trivial cases we assume that $M \geqslant 2$, i.e., the partition $\mathcal{D}$ has at least two subsets. Note also, that we will use the symbol $\underline{\underline{A}}$ to denote a matrix where $a(i,j)$ will denote its (i,j)-th element and $\underline{a}(i)$ (respectively $\underline{a}'(j)$) will denote its i-th row (respectively j-th column) vector.

2.1. DEFINITIONS AND PROOFS FOR THE FIRST LUMPING STAGE

We start with the definition of the **entrance state** of a subset D_m of a partition $\mathcal{D}$ of the state space $\mathcal{X}$.

Definition 1 *A subset D_m of $\mathcal{D}$ has an* ***entrance state*** $(m, \boldsymbol{\varepsilon}_m(\mathcal{D})) \in D_m$ *if and only if*

$$p(m,j \,|\, m',j') = 0, \textit{ for all } m' \neq m \textit{ with } j \neq \boldsymbol{\varepsilon}_m(\mathcal{D}), \textit{ and all } j' \in D_{m'}.$$

Remark 1.

i) Note that from the positive recurrence assumption it follows that if $D_m \in \mathcal{D}$ has an entrance state, there exists some $(m',j') \in D_{m'}$ with $m' \neq m$ such that

$$p(m, \boldsymbol{\varepsilon}_m(\mathcal{D}) \,|\, m',j') > 0.$$

ii) An entrance state of a set D_m is the *only* state via which the set D_m can be entered by the chain $X(t)$ from a state in $\mathcal{X}\backslash D_m$, where given two sets A and B, $A\backslash B$ denotes the elements of A that do not belong to B.

iii) Note also that in the familiar one dimensional notation for the states, a subset D of $\mathcal{X}$ has an **entrance state** $\boldsymbol{\varepsilon} \in D$ if

$$p(j|j') = 0 \text{ for all } j \neq \boldsymbol{\varepsilon},\ j \in D \text{ and all } j' \notin D.$$

Given a partition $\mathcal{D}$ with an entrance state $(0, \varepsilon_0(\mathcal{D})) \in D_0$ we construct the following Markov chains.

a) A Markov chain $Z_0(t)$ on state space D_0 with transition matrix $\underline{\underline{\mathrm{U}}}_{D_0}$ which elements are as follows:

$$u_{D_0}(0,j\,|\,0,i) = \begin{cases} p(0,\varepsilon_0(\mathcal{D}))\,|\,0,i) + \displaystyle\sum_{(k,j')\notin D_0} p(k,j'\,|\,0,i), & \text{if } j = \varepsilon_0(\mathcal{D}), \\ p(0,j\,|\,0,i), & \text{otherwise.} \end{cases} \tag{1}$$

b) A Markov chain $X_1(t)$ with state space $\mathcal{X}_1 = \{(1,0)\} \cup D_1 \cup \ldots \cup D_M$ and transition matrix $\underline{\underline{\mathrm{P}}}_1$ where its $((k,j),(k',j'))$-th element is defined by Eq. (2) below if $(k,j) = (k',j') = (1,0)$ and by Eq. (3), otherwise.

$$p_1(1,0\,|\,1,0) = \sum_{(0,i'),\,(0,i)\in D_0} p(0,i'\,|\,0,i)\upsilon_{D_0}(0,i), \tag{2}$$

$$p_1(k',j'\,|\,k,j) = \begin{cases} \displaystyle\sum_{(0,i)\in D_0} p(k',j'\,|\,0,i)\upsilon_{D_0}(0,i), & \text{if } (k,j) = (1,0), \\ \displaystyle\sum_{(0,i)\in D_0} p(0,i\,|\,k,j), & \text{if } (k',j') = (1,0), \\ p(k',j'\,|\,k,j), & \text{otherwise.} \end{cases} \tag{3}$$

It is easy to see that both chains $Z_0(t)$ and $X_1(t)$ are irreducible and positive recurrent, because $X(t)$ has these properties as well. The steady state probabilities of Markov chain $Z_0(t)$ will be denoted by $\upsilon_{D_0}(0,i) = \lim_{t\to\infty} \mathbf{Pr}[Z_0(t) = (0,i)]$. The vector of the steady state probabilities of the Markov chain $X_1(t)$, will be denoted by:

$$\underline{\pi}_1 = \big(\pi_1(1,0); \pi_1(1,1), \ldots, \pi_1(1,\ell_1), \ldots, \pi_1(M,1), \ldots, \pi_1(M,\ell_M)\big).$$

Note that in the above construction of the new process $X_1(t)$, we have introduced an artificial state we denote as $(1,0)$. This state $(1,0)$ essentially represents the "lumped states" of the set D_0 of the initial process $X(t)$; we have used a semicolon in the above notation for $\underline{\pi}_1$ to emphasize this fact.

We will use the notation $\underline{\underline{\mathrm{U}}}_{D_0} = [\underline{u}'_{D_0}(0,1), \ldots, \underline{u}'_{D_0}(0,\ell_0)]$, where $\underline{u}'_{D_0}(0,j)$ denotes the j-th column of the transition matrix $\underline{\underline{\mathrm{U}}}_{D_0}$. Similarly,

$$\underline{\underline{\mathrm{P}}} = [\underline{p}'(0,1), \ldots, \underline{p}'(0,\ell_0), \ldots, \underline{p}'(M,1),\ \ldots, \underline{p}'(M,\ell_M)].$$

It is well known that $\underline{\pi}$ is the solution to the following system of equations: $\underline{\pi}\,\underline{\underline{\mathrm{P}}} = \underline{\pi}$ and $\underline{\pi}\,\underline{1}' = 1$. Here $\underline{1}$ will always denote a vector of ones of the same dimension as $\underline{\pi}$.

We will next state and prove the following proposition and theorem.

Proposition 1 *If D_0 has an entrance state $(0, \boldsymbol{\varepsilon}_0(\mathcal{D}))$, then the following is true for all $(0, i) \in D_0$:*

$$\upsilon_{D_0}(0, i) = \frac{\pi(0, i)}{\sum_{(0,j)\in D_0} \pi(0, j)}. \tag{4}$$

Proof. Let $\underline{\upsilon}_{D_0} = \big(\upsilon_{D_0}(0, 1), \ldots, \upsilon_{D_0}(0, \ell_0)\big)$. It is clear that for $\underline{\upsilon}_{D_0}$, defined by Eq. (4), the statement $\underline{\upsilon}_{D_0} \underline{1}' = 1$ holds. To prove that this choice of $\underline{\upsilon}_{D_0}$ is the solution, we will show that it also satisfies:

$$\underline{\upsilon}_{D_0} \underline{\underline{\mathrm{U}}}_{D_0} = \underline{\upsilon}_{D_0}. \tag{5}$$

By uniqueness of solutions to Eq. (5) (with $\underline{\upsilon}_{D_0} \underline{1}' = 1$) it then follows that $\underline{\upsilon}_{D_0}$ is indeed the steady state vector.

To show that Eq. (5) holds, we distinguish two cases: the entrance state $(0, \boldsymbol{\varepsilon}_0(\mathcal{D}))$ or any of the other states.

We will use the following derivation:

$$\begin{aligned}
\sum_{(0,j)\in D_0} p(0, \boldsymbol{\varepsilon}_0(\mathcal{D}) \,|\, 0, j)\pi(0, j) =& \pi(0, \boldsymbol{\varepsilon}_0(\mathcal{D})) - \sum_{(k,i')\notin D_0} p(0, \boldsymbol{\varepsilon}_0(\mathcal{D}) \,|\, k, i')\pi(k, i') \\
=& \pi(0, \boldsymbol{\varepsilon}_0(\mathcal{D})) - \sum_{(k,i')\notin D_0} \big(1 - \sum_{(k',i'')\notin D_0} p(k', i'' \,|\, k, i')\big)\pi(k, i') \\
=& \pi(0, \boldsymbol{\varepsilon}_0(\mathcal{D})) - \sum_{(k,i')\notin D_0} \pi(k, i') \\
& + \sum_{(k',i''),(k,i')\notin D_0} p(k', i'' \,|\, k, i')\pi(k, i') \\
=& \pi(0, \boldsymbol{\varepsilon}_0(\mathcal{D})) - \sum_{(k,i')\notin D_0} \pi(k, i') \\
& + \sum_{(k',i'')\notin D_0} (\, \pi(k', i'') - \sum_{(0,j)\in D_0} p(k', i'' \,|\, 0, j)\pi(0, j)) \\
=& \pi(0, \boldsymbol{\varepsilon}_0(\mathcal{D})) - \sum_{(0,j)\in D_0} \sum_{(k,i')\notin D_0} p(k, i' \,|\, 0, j)\pi(0, j).
\end{aligned}$$

Now, using this equality we obtain for $(0,i) = (0, \boldsymbol{\varepsilon}_0(\mathcal{D}))$:

$$
\begin{aligned}
&\underline{\upsilon}_{D_0}\, \underline{u}'_{D_0}(0, \boldsymbol{\varepsilon}_0(\mathcal{D})) = \\
&= \sum_{(0,j)\in D_0} \upsilon_{D_0}(0,j) u_{D_0}(0, \boldsymbol{\varepsilon}_0(\mathcal{D}) \,|\, 0, j) \\
&= \frac{\sum_{(0,j)\in D_0} \pi(0,j) \left(p(0, \boldsymbol{\varepsilon}_0(\mathcal{D}) \,|\, 0,j) + \sum_{(k,i')\notin D_0} p(k,i' \,|\, 0,i) \right)}{\sum_{(0,i')\in D_0} \pi(0,i')} \\
&= \frac{\pi(0, \boldsymbol{\varepsilon}_0(\mathcal{D})) - \sum_{(0,j)\in D_0} \pi(0,j) \sum_{(k,i')\notin D_0} (p(k,i' \,|\, 0,j) - p(k,i' \,|\, 0,j))}{\sum_{(0,i')\in D_0} \pi(0,i')} \\
&= \frac{\pi(0, \boldsymbol{\varepsilon}_0(\mathcal{D}))}{\sum_{(0,j)\in D_0} \pi(0,j)} \\
&= \upsilon_{D_0}(\epsilon_0(\mathcal{D})).
\end{aligned}
$$

Similarly, for $(0,i) \neq (0, \boldsymbol{\varepsilon}_0(\mathcal{D}))$:

$$
\begin{aligned}
\underline{\upsilon}_{D_0}\, \underline{u}'_{D_0}(0,i) &= \sum_{(0,j)\in D_0} \upsilon_{D_0}(0,j) u_{D_0}(0,i \,|\, 0,j) \\
&= \frac{1}{\sum_{(0,i')\in D_0} \pi(0,i')} \sum_{(0,j)\in D_0} \pi(0,j) p(0,i \,|\, 0,j) \\
&= \frac{\pi(0,i)}{\sum_{(0,i')\in D_0} \pi(0,i')} \\
&= \upsilon_{D_0}(0,i).
\end{aligned}
$$

Thus, $\underline{\upsilon}_{D_0}\, \underline{u}'_{D_0}(0,i) = \upsilon_{D_0}(0,i)$ for all $(0,i) \in D_0$ and the proof is complete. $\square$

For the chain $X_1(t)$ we have the following main result concerning the steady state distribution.

Theorem 1 *If D_0 has an entrance state $(0, \boldsymbol{\varepsilon}_0(\mathcal{D}))$, then the following are true regarding the Markov chains $X(t)$ and $X_1(t)$.*

i) If $(k,j) \neq (1,0)$, then

$$\pi_1(k,j) = \pi(k,j),$$

ii) If $(k,j) = (1,0)$, then

$$\pi_1(k,j) = \sum_{(0,i)\in D_0} \pi(0,i).$$

Proof. We need to show that the above choice of $\underline{\pi}_1$ satisfies the steady state equations of the $X_1(t)$ process, i.e., it is the unique solution of the linear system

$$\underline{\pi}_1\, \underline{\underline{\mathrm{P}}}_1 = \underline{\pi}_1,$$

together with $\underline{\pi}_1 \; \underline{1}' = 1$.

The latter equality is easy to see.

Next, for each state $(k,j) \neq (1,0)$ we have:

$$
\begin{aligned}
\underline{\pi}_1 \underline{p}'_1(k,j) = & \sum_{(k',j') \in \mathcal{X}_1} p_1(k,j \,|\, k',j') \pi_1(k',j') \\
= & \sum_{(k',j') \in \mathcal{X}_1 \setminus \{(1,0)\}} p(k,j \,|\, k',j') \pi(k',j') \\
& + \sum_{(0,i) \in D_0} p(k,j \,|\, 0,i) \upsilon_{D_0}(0,i) \sum_{(0,i') \in D_0} \pi(0,i') \\
= & \sum_{(k',j') \in \mathcal{X}_1 \setminus \{(1,0)\}} \pi(k',j') p(k,j \,|\, k',j') + \sum_{(0,i) \in D_0} p(k,j \,|\, 0,i) \pi(0,i) \\
= & \pi(k,j) = \pi_1(k,j),
\end{aligned}
$$

and thus $\pi_1(k,j) = \pi(k,j)$ satisfies the steady state equations of the $X_1(t)$ process for all $(k,j) \neq (1,0)$.

Finally, for $(k,j) = (1,0)$, we have:

$$
\begin{aligned}
\underline{\pi}_1 \, \underline{p}'_1(1,0) = & \sum_{(k',j') \in \mathcal{X}_1} p_1(1,0 \,|\, k',j') \pi_1(k',j') \\
= & \sum_{(k',j') \in \mathcal{X}_1 \setminus \{(1,0)\}} \sum_{(0,i) \in D_0} p(0,i \,|\, k',j') \pi(k',j') \\
& + \sum_{(0,i),(0,i') \in D_0} p(0,i \,|\, 0,i') \upsilon_{D_0}(0,i') \sum_{(0,i'') \in D_0} \pi(0,i'') \\
= & \sum_{(0,i) \in D_0} \sum_{(k',j') \in \mathcal{X}_1 \setminus \{(1,0)\}} p(0,i \,|\, k',j') \pi(k',j') + \sum_{(0,i),(0,i') \in D_0} p(0,i \,|\, 0,i') \pi(0,i') \\
= & \sum_{(0,i) \in D_0} \sum_{(k',j') \in \mathcal{X}} p(0,i \,|\, k',j') \pi(k',j') \\
= & \sum_{(0,i) \in D_0} \pi(0,i) = \pi_1(1,0).
\end{aligned}
$$

So indeed the choice of $\pi_1(1,0) = \sum_{(0,i) \in D_0} \pi(0,i)$ satisfies the steady state equation of $X_1(t)$, that corresponds to state $(1,0)$.

The proof is now complete. $\square$

In section 2.2 below we provide conditions under which it is possible to successively (or sequentially) use the lumping procedure of Section 2.1 over the sets $D_1, \; D_2, \ldots, D_m$. In section 2.3 we present the algorithm and a computational example.

2.2. DEFINITIONS AND PROOFS FOR SUCCESSIVE LUMPING

We start with the following extended notation and definitions. For a Markov chain $X(t)$, with state space $\mathcal{X}$, transition matrix $\underline{\underline{P}}$ and a partition $\mathcal{D} = \{D_0, \ldots, D_M\}$, we define $\Delta_0 = D_0$, $\Delta_m = \{(m,0)\} \cup D_m$, with $m = 1, 2, \ldots M$, where $(m,0)$ is an artificial state, representing the lumped states: $\bigcup_{k=0}^{m-1} D_k$.

We further define the partitions $\mathcal{D}_m = \{\Delta_m, D_{m+1}, \ldots, D_M\}$ and the state spaces: $\mathcal{X}_m = \Delta_m \cup D_{m+1} \cup \ldots \cup D_M$, for $m = 0, \ldots, M$.

For notational consistency, we will use the notation: $X_0(t) = X(t)$, $\mathcal{X}_0 = \mathcal{X}$, $\mathcal{D}_0 = \mathcal{D}$, $\underline{\underline{P}}_0 = \underline{\underline{P}}$, and $\underline{\pi}_0 = \underline{\pi}$. Furthermore without loss of generality we will denote the entrance state of Δ_m in $\mathcal{D}_m$ with $(m, \epsilon_m(\mathcal{D}_m))$, if it exists.

We next state the following definition.

Definition 2 *A Markov chain $X(t)$ is called **successively lumpable** with respect to partition $\mathcal{D} = \{D_0, \ldots, D_M\}$ if and only if the set $D_0 \cup D_1 \cup \ldots \cup D_i$ has an entrance state for all $i = 0, \ldots, M$.*

The above definition means that there exists only one state in $\{D_0 \cup \ldots \cup D_{m'}\}$ that can be entered from a state in D_m when $m > m' > 0$. Note also that the definition implies that transitions out of states in $D_{m'}$ can only lead to states in D_m with $m \geqslant m' \geqslant 0$ or to the entrance state of the set $\Delta_{m'-1}$.

In the sequel, the state (m, η_m) will denote an arbitrary but fixed state in Δ_m.

Given the partition $\mathcal{D}_m$ we successively construct the following Markov chains.

a) A Markov chain $Z_m(t)$ with state space Δ_m and transition matrix $\underline{\underline{U}}_{\Delta_m}$ with

$$u_{\Delta_m}(m,j \,|\, m,i) = \begin{cases} p_m(m,j \,|\, m,i) + \sum\limits_{(k,i') \notin \Delta_m} p_m(k,i' \,|\, m,i), & \text{if } (m,j) = (m,\eta_m), \\ p_m(m,j \,|\, m,i), & \text{otherwise.} \end{cases} \tag{6}$$

Suppose that the chains $Z_m(t)$ is irreducible and positive recurrent. Then its steady state probabilities will be denoted by $\upsilon_{\Delta_m}(m,i)$, i.e.,

$$\upsilon_{\Delta_m}(m,i) = \lim_{t \to \infty} \mathbf{Pr}[Z_m(t) = (m,i)].$$

b) A Markov chain $X_{m+1}(t)$ with state space $\mathcal{X}_{m+1} = \Delta_{m+1} \cup D_{m+2} \cup \ldots \cup D_M$ and transition matrix $\underline{\underline{P}}_{m+1}$, with elements $((k,j),(k',j'))$ defined by Eq. (7) below if $(k,j) = (k',j') = (m+1,0)$ and by Eq. (8) otherwise.

$$p_{m+1}(m+1,0 \,|\, m+1,0) = \sum_{(m,i'),\,(m,i) \in \Delta_m} p_m(m,i' \,|\, m,i)\upsilon_{\Delta_m}(m,i), \tag{7}$$

$$p_{m+1}(k',j'\,|\,k,j) = \begin{cases} \displaystyle\sum_{(m,i)\in\Delta_m} p_m(k',j'\,|\,m,i)\upsilon_{\Delta_m}(m,i), & \text{if } (k,j)=(m+1,0), \\ \displaystyle\sum_{(m,i)\in\Delta_m} p_m(m,i\,|\,k,j), & \text{if } (k',j')=(m+1,0), \\ p_m(k',j'\,|\,k,j), & \text{otherwise.} \end{cases} \tag{8}$$

Note that in order to compute $p_{m+1}(\cdot\,|\,\cdot)$ we first need to compute $\upsilon_{\Delta_m}(\cdot)$. The vector of the steady state probabilities of the chain X_{m+1} will be denoted by:

$$\underline{\pi}_{m+1} = \big(\pi_{m+1}(m,0);\pi_{m+1}(m,1),\dots,\pi_{m+1}(m,\ell_m),\dots,\pi_{m+1}(M,1),\dots,\pi_{m+1}(M,\ell_M)\big).$$

We will use the notation:

$$\underline{\underline{\mathrm{U}}}_{\Delta_m} = [\underline{u}'_{\Delta_m}(m,1),\dots,\underline{u}'_{\Delta_m}(m,\ell_m)],$$

and

$$\underline{\underline{\mathrm{P}}}_m = [\underline{p}'_m(m,0);\underline{p}'_m(m,1),\dots,\underline{p}'_m(m,\ell_m),\dots,\underline{p}'_m(M,1),\dots,\underline{p}'_m(M,\ell_M)].$$

Remark 2. Every Markov chain is successively lumpable with respect to a partition $\mathcal{D}=\{D_0,D_1\}$ when $D_0=\{(0,\varepsilon_0(\mathcal{D}))\}$ is any single state and D_1 contains the remaining states.

We can now state the following proposition regarding successively lumpable Markov chains.

Proposition 2 *If Markov chain $X_0(t)$ is successively lumpable with respect to partition $\mathcal{D}_0$, then $X_m(t)$ is successively lumpable with respect to partition $\mathcal{D}_m$, for all $m=1,\dots,M$.*

Proof. To complete an induction proof we need to show that if $X_m(t)$ is successively lumpable with respect to partition $\mathcal{D}_m$, then $X_{m+1}(t)$ is successively lumpable with respect to partition $\mathcal{D}_{m+1}$.

For $m=0$, Definition 2 holds by assumption on $\underline{\underline{\mathrm{P}}}_0(=\underline{\underline{\mathrm{P}}})$. We assume the induction holds for $k=0,\dots,m$ and we show it holds for $m+1$. We have defined $\underline{\underline{\mathrm{P}}}_{m+1}$ in Eqs. (7)-(8).

To prove that $X_{m+1}(t)$ is successively lumpable with respect to $\mathcal{D}_{m+1}$ we first show that Δ_{m+1} has an entrance state $(m+1,\epsilon_{m+1}(\mathcal{D}_{m+1}))$.

By induction we know that $\Delta_m\cup D_{m+1}$ has an entrance state in $X_m(t)$: either $(m,\epsilon_m(\mathcal{D}_m))$ or $(m+1,i_1)$, a state in D_{m+1}. Furthermore, we know by Eq. (8) that for $i\neq 0$, with $k>m+1$:

$$p_{m+1}(m+1,i\,|\,k,j) = p_m(m+1,i\,|\,k,j), \tag{9}$$

and for $i=0$:

$$p_{m+1}(m+1,0\,|\,k,j) = \sum_{(m,i')\in\Delta_m} p_m(m,i'\,|\,k,j). \tag{10}$$

Now, if $(m,\epsilon_m(\mathcal{D}_m))$ is the entrance state of $\Delta_m \cup D_{m+1}$ in $X_m(t)$ we get by Eq. (9) that $p_{m+1}(m+1,i\,|\,k,j) = 0$ for all $i > 0, k > m+1$ and thus that $(m+1,0)$ is the entrance state of Δ_{m+1} in $X_{m+1}(t)$.

If $(m+1,i_1)$ is the entrance state of $\Delta_m \cup D_{m+1}$ in $X_m(t)$, we know by Eq. (9) that $p_{m+1}(m+1,i\,|\,k,j) = 0$ for all i except i_1 and by Eq.(10) that $p_{m+1}(m+1,0\,|\,k,j) = 0$. Thus $(m+1,i_1)$ is the entrance state of Δ_{m+1} in $X_{m+1}(t)$.

With a similar argument we can prove that $\Delta_{m+1} \cup \ldots \cup D_i$ has an entrance state in $X_{m+1}(t)$ for all i. Thus $X_{m+1}(t)$ is successively lumpable with respect to $\mathcal{D}_{m+1}$ when $X_m(t)$ is successively lumpable with respect to $\mathcal{D}_m$. $\square$

Remark 3. Because of Proposition 2 we know that Δ_m has an entrance state in $X_m(t)$ for all $m \leqslant M$. In the construction of $Z_m(t)$, (m,η_m) was chosen arbitrarily. From now on we choose $(m,\eta_m) = (m,\epsilon_m(\mathcal{D}_m))$. Then $Z_m(t)$ is irreducible and positive recurrent as can be easily seen from a graphical representation.

We can now state the following.

Theorem 2 *Under the assumption of Proposition 2 the following are true:*

i)

$$\upsilon_{\Delta_m}(m,i) = \frac{\pi_m(m,i)}{\sum_{(m,i')\in\Delta_m} \pi_m(m,i')}. \tag{11}$$

ii)

$$\pi_{m+1}(k,j) = \begin{cases} \displaystyle\sum_{(m,i')\in\Delta_m} \pi_m(m,i'), & \text{if } (k,j) = (m+1,0), \\ \pi_m(k,j), & \text{otherwise.} \end{cases} \tag{12}$$

Proof. The proof is easy to complete by induction using a similar derivation as in Proposition 1 and Theorem 1, combined with the induction result of Proposition 2. $\square$

The previous results imply that the following theorem holds.

Theorem 3 *If $X_0(t)$ is successively lumpable with $|\mathcal{X}_0| < \infty$ the following is true:*

$$\pi_0(m,j) = \upsilon_{\Delta_m}(m,j) \prod_{k=m+1}^{M} \upsilon_{\Delta_k}(k,0),\ \forall (m,j) \in \mathcal{X}_0.$$

Proof. The proof follows by induction on decreasing values of $n = M, M-1, \ldots, 0$ for fixed M; note that $|\mathcal{X}_0| < \infty$ implies that M is finite.

For $n = M$, we need to show that

$$\pi_0(M, j) = \upsilon_{\Delta_M}(M, j), \forall (M, j) \in D_M.$$

Indeed, by Theorem 2, we have $\upsilon_{\Delta_M}(M, j) = \pi_M(M, j)/1$, where the denominator is 1 because Δ_M contains all states of $\mathcal{X}_M$. Since $j \neq 0$, (i.e. (M, j) has never been lumped by our lumping procedure) by using Theorem 2 repeatedly we obtain $\pi_M(M, j) = \pi_{M-1}(M, j) = \ldots = \pi_0(M, j)$, and the proof is complete for $n = M$.

We next show that the claim is true for $n = M-1$, assuming it is true for $n = M$. So we will show that:

$$\pi_0(M-1, j) = \upsilon_{\Delta_{M-1}}(M-1, j) \prod_{k=M}^{M} \upsilon_{\Delta_k}(k, 0).$$

The right hand side of the above is

$$\begin{aligned}
\upsilon_{\Delta_{M-1}}(M-1, j)\upsilon_{\Delta_M}(M, 0) &= \frac{\pi_{M-1}(M-1, j)}{\sum_{(M-1,j')\in\Delta_{M-1}} \pi_{M-1}(M-1, j')} \upsilon_{\Delta_M}(M, 0) \\
&= \frac{\pi_{M-1}(M-1, j)}{\sum_{(M-1,j')\in\Delta_{M-1}} \pi_{M-1}(M-1, j')} \frac{\pi_M(M, 0)}{\sum_{(M,j')\in\Delta_M} \pi_M(M, j')} \\
&= \pi_{M-1}(M-1, j),
\end{aligned}$$

where the first two equalities follow from Theorem 2, Eq. (11). The last equality uses Eq. (12) and the fact that $\sum_{(M,\ell)\in\Delta_M} \pi_M(M, \ell) = 1$, as before. The proof for $n = M-1$ is complete when we observe that $\pi_{M-1}(M-1, j) = \pi_{M-2}(M-1, j) = \ldots = \pi_0(M-1, j)$ since $j \neq 0$, as in the case when $n = M$.

The induction step from n to $n-1$ is easy to complete using similar algebra with albeit more cluttered equations. □

2.3. THE ALGORITHM AND AN EXAMPLE

Using the construction and results of Theorem 3 of the previous section, we can now state an algorithm for computing the stationary probability vector $\underline{\pi}$, of a successively lumpable Markov chain with respect to partition $\mathcal{D}$ as below.

Algorithm SL

1 Construct $\underline{\underline{U}}_{D_0}$, cf., Eq. (1).
2 Calculate $\underline{\upsilon}_{D_0}$.
3 Lump D_0 to $(1,0)$ and let $\Delta_1 = \{(1,0)\} \cup D_1$.
Set $m = 1$.
While $m \leqslant M$
4.1 Construct $\underline{\underline{U}}_{\Delta_m}$ cf., Eq. (6).
4.2 Calculate $\underline{u'}_{\Delta_m}$.
4.3 Lump Δ_m to $(m+1,0)$ and let $\Delta_{m+1} = (m+1,0) \cup D_m$.
$m = m + 1$
End
5 Calculate $\underline{\pi}$, cf., Theorem 3.

We next clarify the previous results with a small example.

Example 1. For clarity we will number the state space directly according to the notation introduced in Section 2, so we take:
$\mathcal{X} = \{(0,1),(0,2),(1,1),(1,2),(2,1),(2,2),(3,1),(3,2),(3,3)\}$,
with a partition $\mathcal{D} = \{D_0, D_1, D_2, D_3\}$ where $D_0 = \{(0,1),(0,2)\}, D_1 = \{(1,1),(1,2)\}$, $D_2 = \{(2,1),(2,2)\}$ $D_3 = \{(3,1),(3,2),(3,3)\}$ and transition matrix $\underline{\underline{P}}$:

$$\underline{\underline{P}} = \left[\begin{array}{c|ccccccccc} & (0,1) & (0,2) & (1,1) & (1,2) & (2,1) & (2,2) & (3,1) & (3,2) & (3,3) \\ \hline (0,1) & 0 & 1/3 & 5/9 & 0 & 0 & 0 & 0 & 1/9 & 0 \\ (0,2) & 0 & 0 & 1/3 & 2/3 & 0 & 0 & 0 & 0 & 0 \\ (1,1) & 0 & 0 & 0 & 1/6 & 2/3 & 0 & 1/6 & 0 & 0 \\ (1,2) & 0 & 0 & 0 & 0 & 1/6 & 3/4 & 0 & 0 & 1/12 \\ (2,1) & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ (2,2) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ (3,1) & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 \\ (3,2) & 1/2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1/2 \\ (3,3) & 1/2 & 0 & 0 & 0 & 0 & 0 & 1/2 & 0 & 0 \end{array}\right]$$

The transition diagram of the corresponding Markov chain $X(t)$ is given in Figure 1. It is easy to see that $X(t)$ is successively lumpable with respect to the partition $\mathcal{D}$. The first steps of the algorithm are:

1 $\underline{\underline{U}}_{\Delta_0} = \begin{bmatrix} 2/3 & 1/3 \\ 1 & 0 \end{bmatrix}$.

2 $\underline{\upsilon}_{\Delta_0} = [3/4, 1/4]$.

Next, we continue with $D_1 \neq \emptyset$, and note that for every positive $p(k', j' \mid k, j)$ with $(k', j') \in D_1$ we have $(k, j) \in D_0 \cup D_1$.

3 Lump $\{(0,1),(0,2)\}$ to $(1,0)$ and let $\Delta_1 = \{(1,0),(1,1),(1,2)\}$.

4.1 $\underline{\underline{U}}_{\Delta_1} = \begin{bmatrix} 1/3 & 1/2 & 1/6 \\ 5/6 & 0 & 1/6 \\ 1 & 0 & 0 \end{bmatrix}$.

4.2 $\underline{v}_{\Delta_1} = [4/7, 2/7, 1/7]$.

Figure 2b illustrates the transition diagram of this $Z_1(t)$ chain.

4.3 Lump $\{(1,0),(1,1),(1,2)\}$ to $(2,0)$ and let $\Delta_2 = \{(2,0),(2,1),(2,2)\}$. Note that since we know $\underline{v}_{\Delta_1}$ we can construct transition probabilities of a set Δ_2 without knowledge of $\underline{\pi}$ with the use of the previous states $\{(1,0),(1,1),(1,2)\}$.

4.1 $\underline{\underline{\mathrm{U}}}_{\Delta_2} = \begin{bmatrix} 19/28 & 3/14 & 3/28 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{bmatrix}$.

4.2 $\underline{v}_{\Delta_2} = [28/43, 6/43, 9/43]$.

Next we look at subset D_3 and repeat the previous.

4.3 Lump $\{(2,0),(2,1),(2,2)\}$ to $(3,0)$ and let $\Delta_3 = \{(3,0),(3,1),(3,2),(3,3)\}$.

4.1 $\underline{\underline{\mathrm{U}}}_{\Delta_3} = \begin{bmatrix} 93/129 & 4/129 & 4/129 & 4/129 \\ 1/2 & 0 & 1/2 & 0 \\ 1/2 & 0 & 0 & 1/2 \\ 1/2 & 1/2 & 0 & 0 \end{bmatrix}$.

4.2 $\underline{v}_{\Delta_3} = [43/67, 142/1407, 104/1407, 248/1407]$.

We advance to step 5 and calculate $\underline{\pi}$:

$$
\begin{aligned}
\pi(0,1) &= v_{\Delta_3}(3,0)v_{\Delta_2}(2,0)v_{\Delta_1}(1,0)v_{\Delta_0}(0,1) &&= \frac{43}{67}\frac{28}{43}\frac{4}{7}\frac{3}{4} &&= \frac{12}{67} \\
\pi(0,2) &= v_{\Delta_3}(3,0)v_{\Delta_2}(2,0)v_{\Delta_1}(1,0)v_{\Delta_0}(0,2) &&= \frac{43}{67}\frac{28}{43}\frac{4}{7}\frac{1}{4} &&= \frac{4}{67} \\
\pi(1,1) &= v_{\Delta_3}(3,0)v_{\Delta_2}(2,0)v_{\Delta_1}(1,1) &&= \frac{43}{67}\frac{28}{43}\frac{2}{7} &&= \frac{8}{67} \\
\pi(1,2) &= v_{\Delta_3}(3,0)v_{\Delta_2}(2,0)v_{\Delta_1}(1,2) &&= \frac{43}{67}\frac{28}{43}\frac{1}{7} &&= \frac{4}{67} \\
\pi(2,1) &= v_{\Delta_3}(3,0)v_{\Delta_2}(2,1) &&= \frac{43}{67}\frac{6}{43} &&= \frac{6}{67} \\
\pi(2,2) &= v_{\Delta_3}(3,0)v_{\Delta_2}(2,2) &&= \frac{43}{67}\frac{9}{43} &&= \frac{9}{67} \\
\pi(3,1) &= v_{\Delta_3}(3,1) &&= \frac{142}{1407} \\
\pi(3,2) &= v_{\Delta_3}(3,2) &&= \frac{104}{1407} \\
\pi(3,3) &= v_{\Delta_3}(3,3) &&= \frac{248}{1407}.
\end{aligned}
$$

Remark 4.

i) To illustrate the fact that a Markov chain can be successively lumped with respect to different partitions, Figure 4 shows its transition diagram of the chain of example 1, where highlighted areas represent the sets of a different partition $\mathcal{D}'$ and it is easy to see that the chain is also successively lumpable with respect to partition $\mathcal{D}'$.

ii) Figure 5 illustrates the transition diagram of a Markov chain. An arrow from a state (m, j) to a state (m', j') is present only if the corresponding transition probability $p(m', j' \,|\, m, j)$ is positive; where we ignore "loop" transitions with $p(m, j \,|\, m, j) > 0$ that do not play a role in determining successive lumpability. The Markov chain corresponding to this transition diagram is successively lumpable with respect to partition $\mathcal{D}$ that consists of the four highlighted sets of states. This picture is interesting since it is easy to see that "adding" any additional (non-loop) arrows (i.e., transition(s), with positive probability) will result in a transition diagram of a chain for which the successive lumpability property does not hold.

3. MULTIPLE SUCCESSIVELY LUMPABLE MARKOV CHAINS

The main result of Section 2 is that when a Markov chain is successively lumpable its steady state probability vector can be calculated using successive lumping. In Section 3.1 we will show that it is possible to have multiple lumpable structures in one Markov chain. In this case it is possible to calculate the steady state vector using multiple times successive lumping. We also establish a product form expression, for finite state spaces. We present the algorithm and an example in Section 3.2.

3.1. DEFINITIONS AND PROOFS FOR MULTIPLE SUCCESSIVE LUMPING

Let $X(t)$ be a Markov chain on a finite (or countable) state space $\mathcal{X}$ with transition matrix $\underline{\underline{P}}$. We will assume that the state space $\mathcal{X}$ is composed of $N \leqslant \infty$ mutually exclusive and exhaustive sets, i.e.

$$\mathcal{X} = \bigcup_{n=1}^{N} \mathcal{X}^n,$$

where each subset $\mathcal{X}^n$ can be partitioned into a (possibly infinite) sequence of

$$\mathcal{D}^n = \{D_0^n, \dots, D_{M_n}^n\}.$$

Alternatively the partition

$$\mathcal{D} = \{D_0^1, \dots, D_{M_1}^1, \dots, D_0^N, \dots, D_{M_N}^N\}$$

is a sequence of $N \leqslant \infty$ subpartitions of $\mathcal{X}$. For notational convenience, the elements of each set D_m^n will be relabelled to a triple-notation as $\{(n, m, 1), (n, m, 2), \dots, (n, m, \ell_{(n,m)})\}$, for given constants $\ell_{(n,m)} \leqslant \infty$. After this state relabeling, the transition matrix of $X(t)$ will be denoted by $\underline{\underline{\mathrm{P}}} = [p(n', m', j' \,|\, n, m, j)]$, where the $((n, m, j)\,,\,(n', m', j'))$ element is given by

$$p(n', m', j' \,|\, n, m, j) = \mathbf{Pr}[X(t+1) = (n', m', j') \,|\, X(t) = (n, m, j)].$$

In the sequel, to avoid trivial cases we assume that $N \geqslant 2$, i.e., the partition $\mathcal{D}$ has at least two subsets.

The definition of an **entrance state** in this triple index notation is as follows.

Definition 3 *A subset D^n_m of $\mathcal{D}$ has an* ***entrance state*** $(n, m, \varepsilon_{n,m}(\mathcal{D})) \in D^n_m$ *iff for all* $n' \neq n$, $m' \neq m$ *and* $j \neq \varepsilon_{n,m}(\mathcal{D})$,

$$p(n, m, j \,|\, n', m', j') = 0.$$

Given a Markov chain $X(t)$ with partition $\mathcal{D}$ for which every subset $\mathcal{X}^n$ has an entrance state $(n, m, \epsilon_{n,m}(\mathcal{D}))$, we next study the Markov chains $X^n(t)$ on state space $\mathcal{X}^n$. We state the following definition about their transition matrix $\underline{\underline{\mathrm{P}}}^{\,n}$.

Definition 4 *We call $X^n(t)$ the* ***$\mathbf{n^{th}}$- component Markov chain*** *of $X(t)$ when the $(n, m', j' \,|\, n, m, j)$-element of $\underline{\underline{\mathrm{P}}}^{\,n}$ is defined as follows:*

a) If $(n, m', j') = (n, m', \varepsilon_{n,m'}(\mathcal{D}))$, then

$$p^n(n, m', j' \,|\, n, m, j) = p(n, m', j' \,|\, n, m, j) + \sum_{(n', m'', j'') \notin \mathcal{X}^n} p(n', m'', j'' \,|\, n, m, j).$$

b) Otherwise,

$$p^n(n, m', j' \,|\, n, m, j) = p(n, m', j' \,|\, n, m, j).$$

We can now state the following.

Definition 5 *A Markov chain $X(t)$ is called* ***multiple successively lumpable*** *with respect to partition $\mathcal{D} = \{\mathcal{D}^1, \ldots, \mathcal{D}^N\}$ if and only if the following conditions hold.*

a) $D^n_0 \cup D^n_1 \cup \ldots \cup D^n_{M_n}$ has an entrance state in $X(t)$ for all $n = 1, \ldots, N$.

b) $D^n_0 \cup D^n_1 \cup \ldots \cup D^n_{m_n}$ has an entrance state in $X^n(t)$ for all $n = 1, \ldots, N$ and for all $m_n = 0, \ldots, M_n$.

Note that condition (a) makes the assertion that any state (n', m', j') in $D^{n'}_{m'}$ can not be entered from a state (n, m, j) in D^n_m, except when $(n', m', j') = (n', m', \varepsilon_{n',m'}(\mathcal{D}))$. Condition (b) asserts that a state (n, m', j') in $D^n_{m'}$ can not be entered from a state (n, m, j) in D^n_m when $m' < m$, except when (n, m', j') is the entrance state of $D^n_0 \cup \ldots \cup D^n_{m'}$ in $X^n(t)$.

We can now state the following lemma which shows that a multiple successively lumpable Markov chain is indeed several times successively lumpable.

Lemma 1 *When $X(t)$ is multiple successively lumpable with respect to $\mathcal{D}$, the n^{th}- component-Markov chain $X^n(t)$, is successively lumpable with respect to $\mathcal{D}^n$ for all $n \leqslant N$.*

Proof. Using their construction (cf. Definition 4 (a) and (b)) the transition probabilities $p^n(n, m', j' \,|\, n, m, j)$ can be shown to satisfy the conditions of Definition 5. □

We next introduce some notation, extending that of Section 2.

i) On the n^{th}- component-Markov chain $X^n(t)$ of $X(t)$ we define $\Delta_0^n = D_0^n$, $\Delta_1^n = \{(n,1,0)\} \cup D_1^n$, $\Delta_m^n = \{(n,m,0)\} \cup D_m^n$, where states $(n,m,0)$ are lumped states representing $\bigcup_{k=0}^{m-1} D_k^n$.

   For notational consistency we will use the notation: $X_0^n(t) = X^n(t)$, $\mathcal{X}_0^n = \mathcal{X}^n$, $\mathcal{D}_0^n = \mathcal{D}^n$, and $\underline{\underline{\mathrm{P}}}{}_0^n = \underline{\underline{\mathrm{P}}}$. Further, we consider $X_m^n(t)$ on $\mathcal{X}_m^n$ with transitions $\underline{\underline{\mathrm{P}}}{}_m^n$.

ii) Analogously to the chains $Z_m(t)$ defined in Section 2, we define Markov chains $Z_m^n(t)$ with state space Δ_m^n and transition matrix $\underline{\underline{\mathrm{U}}}{}_{\Delta_n^m}^n$.

iii) Also, let $\pi(n,m,j) = \lim_{t\to\infty} \mathbf{Pr}[X(t) = (n,m,j)]$, $\pi^n(n,m,j) = \lim_{t\to\infty} \mathbf{Pr}[X^n(t) = (n,m,j)]$ and $\upsilon_{\Delta_n^m}^n(j) = \lim_{t\to\infty} \mathbf{Pr}[Z_m^n(t) = (j)]$.

iv) Similarly, we define $\underline{\pi}$, $\underline{\pi}^n$, $\underline{\upsilon}{}_{\Delta_n^m}^n$ to be the corresponding probability vectors; with dimensions $\prod_{n=1}^N \prod_{m=0}^{M_n} \ell_m^n$, $\prod_{m=0}^{M_n} \ell_m^n$, $\ell_m^n + \delta(m)$ respectively, where in the last expression the term "$\delta(m)$" is equal to one if $m > 0$, and equal to zero when $m = 0$ (note that no artificial state in Δ_m^n is used when $m = 0$).

   The elements of $\underline{\underline{\mathrm{U}}}{}_{\Delta_n^m}^n$ and $\underline{\underline{\mathrm{P}}}{}_{m+1}^n$ are akin to the elements of $\underline{\underline{\mathrm{U}}}{}_{\Delta_m}$ and $\underline{\underline{\mathrm{P}}}{}_{m+1}$, cf., Eqs. (6)-(8).

v) Finally, we define a chain $Y(t)$ with state space $\mathcal{E} = \{1, \ldots, N\}$ and transition matrix $\underline{\underline{\mathrm{Q}}}$ with its (n, n') element being equal to:

$$q(n' \,|\, n) = \sum_{(n',m',j') \in \mathcal{X}^{n'}} \sum_{(n,m,j) \in \mathcal{X}^n} \pi^n(n,m,j) p(n', m', j' \,|\, n, m, j). \tag{13}$$

   Note that the chain $Y(t)$ can be viewed as a chain between the different "lumped" successively lumpable chains. We will use the notation $\sigma(n)$ for the steady state probabilities of the above chain, i.e., $\sigma(n) = \lim_{t\to\infty} \mathbf{Pr}[\,Y(t) = n\,]$.

We will next show the following:

Lemma 2 *Assuming that $X(t)$ is a multiple successively lumpable Markov chain with its n^{th}- component Markov chain $X^n(t)$ defined as above, the following is true:*

$$\pi^n(n,m,j) = \frac{\pi(n,m,j)}{\sum_{(n,m',j') \in \mathcal{X}^n} \pi(n,m',j')}.$$

Proof. It is clear that $\underline{\pi}^n \, \underline{1}' = 1$. Now from Definition 5 we see that $\mathcal{X}^n$ has an entrance state $(n, m, \varepsilon_{n,m}(\mathcal{D}))$ and therefore we can use a similar derivation as is used in Proposition 1 to complete the proof. $\square$

Proposition 3 *For a multiple successively lumpable Markov chain $X(t)$ and with $Y(t)$ defined as above, the following is true:*

$$\sigma(n) = \sum_{(n,m,j)\in\mathcal{X}^n} \pi(n, m, j).$$

Proof. It is clear that $\underline{\sigma}\,\underline{1}' = 1$. It suffices to prove that the above choice of $\underline{\sigma}$ is the solution of the steady state equations, of the $Y(t)$ process, below:

$$\sigma(n') = \sum_{n=1}^{N} \sigma(n) q(n'|n) \quad \text{for } n' = 1, 2, \ldots, N.$$

Indeed:

$$\begin{aligned}
\sum_{n=1}^{N} \sigma(n) q(n'|n) &= \sum_{n=1}^{N} \sigma(n) \sum_{(n',m',j')\in\mathcal{X}^{n'}} \sum_{(n,m,j)\in\mathcal{X}^n} \pi^n(n, m, j) p(n', m', j' \,|\, n, m, j) \\
&= \sum_{n=1}^{N} \sum_{(n,m,j)\in\mathcal{X}^n} \pi(n, m, j) \sum_{(n',m',j')\in\mathcal{X}^{n'}} \sum_{(n,m,j)\in\mathcal{X}^n} \frac{\pi(n, m, j) p(n', m', j' \,|\, n, m, j)}{\sum_{(n,m',j')\in\mathcal{X}^n} \pi(n, m', j')} \\
&= \sum_{n=1}^{N} \sum_{(n',m',j')\in\mathcal{X}^{n'}} \sum_{(n,m,j)\in\mathcal{X}^n} \pi(n, m, j) p(n', m', j' \,|\, n, m, j) \qquad (14) \\
&= \sum_{(n',m',j')\in\mathcal{X}^{n'}} \sum_{n=1}^{N} \sum_{(n,m,j)\in\mathcal{X}^n} \pi(n, m, j) p(n', m', j' \,|\, n, m, j) \qquad (15) \\
&= \sum_{(n',m',j')\in\mathcal{X}^{n'}} \sum_{(n,m,j)\in\mathcal{X}} \pi(n, m, j) p(n', m', j' \,|\, n, m, j) \\
&= \sum_{(n',m',j')\in\mathcal{X}^{n'}} \pi(n', m', j') \\
&= \sigma(n').
\end{aligned}$$

The second equality above follows from Lemma 2. It is clear that the summations in Eqs. (14) and (15) can be interchanged freely. $\square$

The main result of this section is the next theorem, for a multiple successively lumpable Markov chain $X(t)$ with respect to a partition $\mathcal{D}$ and with $|\mathcal{X}| < \infty$.

Theorem 4 *If $X(t)$ is multiple successively lumpable with respect to partition $\mathcal{D}$ and $|\mathcal{X}| < \infty$ then:*

$$\pi(n,m,j) = \sigma(n)\upsilon^n_{\Delta^m_n}(n,m,j) \prod_{k=m+1}^{M} \upsilon^n_{\Delta^k_n}(n,k,0) \quad \textit{for all } (n,m,j) \in \mathcal{X}.$$

Proof. Since by Lemma 1, $X^n(t)$ is a successively lumpable Markov chain with respect to partition $\mathcal{D}^n$ we know by Theorem 3 that for all n

$$\pi^n(n,m,j) = \upsilon^n_{\Delta^m_n}(n,m,j) \prod_{k=m+1}^{M} \upsilon^n_{\Delta^k_n}(n,k,0).$$

The proof is easy to complete using Lemma 2 and Proposition 3. □

Remark 5. When $M_n = 1$ for all n, then Theorem 4 becomes the main result in Feinberg and Chui [2].

Remark 6. For a multiple successively lumpable Markov chain we can solve $\prod_{n=1}^{N} M_n$ Markov chains of sizes $\ell^n_{m_n} + \delta(m_n)$ each, instead of one big system of size $\prod_{n=1}^{N}\prod_{m=0}^{M_n} \ell^n_m$. For example, if $N = 10^4$, $M_n = 10^2$ for all n and $l^n_{m_n} = 10^4$ for all n, m, we need to solve 10^6 systems of size 10^4 instead of 1 of size 10^{10}.

3.2. THE ALGORITHM AND AN EXAMPLE

Similarly to algorithm SL for a successively lumpable Markov chain presented in Section 2.3, we state an algorithm for a Markov chain that is multiple successively lumpable with respect to a partition $\mathcal{D} = \{\mathcal{D}^1, \ldots, \mathcal{D}^N\}$. Again, this algorithm does not require a proof, it is a direct result of Theorem 4.

Algorithm MSL

For $n = 1, \ldots N$
1.1 Construct $X^n(t)$ with Def. 4.
1.2 Call Algorithm SL and solve $X^n(t)$.
End
2 Construct $\underline{\underline{Q}}$, cf., Eq. (13).
3 Calculate $\underline{\sigma}$ with Proposition 3.
4 Calculate $\underline{\pi}$, cf., Theorem 4.

To clarify the algorithm, Figure 6 shows a multiple successively lumpable Markov chain, with $N = 2, M_1 = 2, M_2 = 2, \ell^1_{1_1} = 2, \ell^2_{1_2} = 3, \ell^1_{2_1} = 2, \ell^2_{2_2} = 3$.

It is easy to see that in this example both $(1,0,1)$ and $(2,0,1)$ are the *entrance states* of D^1_0 and D^2_0 respectively. The procedure will solve $\mathcal{D}^1$ and $\mathcal{D}^2$ separately as successive lumpable Markov chains and conclude with solving a chain $Y(t)$ between two states representing $\mathcal{D}^1$

and $\mathcal{D}^2$. To do so, the arrows from $(1,1,2)$ and $(1,1,3)$ to $(2,0,1)$ will be "redirected" to $(1,0,1)$ and the arrow from $(2,1,3)$ to $(1,0,1)$ redirected to $(2,0,1)$ as described above.

4. EXTENSION TO SEMI-MARKOV AND CONTINUOUS TIME PROCESSES

It is possible to extend the successive lumpable and the multiple successive lumpable theory to semi-Markov or a continuous time Markov chains either directly or using the following construction and notation, cf., Ross [15]. A process $\{\mathcal{Z}(t),\ t \geqslant 0\}$ is a semi-Markov chain, on a state space $\mathcal{X}$, if it can be constructed as follows.

i) Transitions from state to state are generated from an "embedded" discrete time chain $\{X_n\ ,\ n = 0, 1, \ldots\}$ with state space $\mathcal{X}$ and transition matrix $\underline{\underline{P}} = [p_{ij}]$.

ii) The sojourn times (durations of a visit) for any state i are i.i.d. non-negative random variables T_i^m, $m \geqslant 1$, having an arbitrary distribution F_i, i.e. they are distributed as a random variable T_i with $F_i(t) = \mathbf{Pr}\,(T_i \leqslant t \,|\, \mathcal{Z}(0) = i)$.

A notable special case is that of a continuous time Markov process, in which case the F_i's are all exponential distributions with parameters λ_i. To avoid trivial cases with instantaneous states, it is assumed that the expected durations $\mu_i = \mathbf{E}T_i$ are finite positive constants.

Let $\mathcal{O}_i^n$ denote the time spent in states other than state i between the n^{th} and $n{+}1^{st}$ visits to state i. The above assumptions imply that $\mathcal{O}_i^n$, $n = 1, 2, \ldots$ are i.i.d. random variables, i.e., they are distributed as a random variable $\mathcal{O}_i$, having a distribution $G_i(t)$. The main results concerning the long run steady state probabilities ϖ_i, and π_i of $\mathcal{Z}(t)$ and X_n, respectively, are summarized in the next Theorem, cf., Ross [15].

Theorem 5 *If $\underline{\underline{P}}$ is irreducible and if the distribution of $T_i + \mathcal{O}_i$ is nonlattice then the limit $\lim_{t\to\infty} \mathbf{Pr}(\mathcal{Z}(t) = i | \mathcal{Z}(0) = j)$ exists, it is independent of the initial state $\mathcal{Z}(0) = j$ and it is equal to:*

$$\varpi_i = \frac{\mathbf{E}T_i}{\mathbf{E}T_i + \mathbf{E}\mathcal{O}_i} \tag{16}$$

$$= \frac{\pi_i \mathbf{E}T_i}{\sum_{j\in\mathcal{X}} \pi_j \mathbf{E}T_j}. \tag{17}$$

Equation (17) provides a method for computing the steady states probabilities ϖ_i, from π_i and the expected sojourn times $\mathbf{E}T_i$. It also makes possible a similar successive construction for continuous time or Semi-Markov chains using the expected sojourn times for all states.

For completeness, we next describe the successively lumpable process for a continuous time Markov chain $X(t)$ on a state space $\mathcal{X}$ with a partition $\mathcal{D}$, of size M.

We revert back to the double index state notation of Section 2. The transition rates of $X(t)$ will be denoted by $\mu(k', j' \,|\, k, j)$ where $\mu(k, i \,|\, k, i) = -\sum_{(k',i')\neq(k,i)} \mu(k', i' \,|\, k, i)$.

It is easy to see that $X(t)$ is successively lumpable with respect to $\mathcal{D}$ if the conditions of Definition 2 are valid with $p(\cdot\,|\,\cdot)$ replaced by $\mu(\cdot\,|\,\cdot)$.

We provide below the rates of the corresponding processes $Z_m(t)$ and $X_m(t)$.

For the chain $Z_0(t)$ on $D_0(=\Delta_0)$, corresponding to the process described in Eq. (1), its $((0,i)\,,(0,j))-$element can be shown to be as follows:

$$\lambda_{\Delta_0}(0,j\,|\,0,i) = \begin{cases} \mu(0,j\,|\,0,i), & \text{if } i \neq \varepsilon_0(\mathcal{D}), j \neq \varepsilon_0(\mathcal{D}), \\ \sum\limits_{(k,j')\notin D_0} \mu(k,j'\,|\,0,i) + \mu(0,\varepsilon_0(\mathcal{D})\,|\,0,i), & \text{if } i \neq \varepsilon_0(\mathcal{D}), j = \varepsilon_0(\mathcal{D}), \\ \mu(0,j\,|\,0,\varepsilon_0(\mathcal{D})), & \text{if } i = \varepsilon_0(\mathcal{D}), j \neq \varepsilon_0(\mathcal{D}), \\ -\sum\limits_{(0,j')\in D_0} \mu(0,j'\,|\,0,\varepsilon_0(\mathcal{D})), & \text{if } i = j = \varepsilon_0(\mathcal{D}). \end{cases}$$

The transition rate matrix of the above rates is denoted as $\underline{\underline{\Lambda}}_{\Delta_0}$ and the steady state equations are:

$$\underline{\upsilon}_{\Delta_0}\underline{\underline{\Lambda}}_{\Delta_0} = 0,$$

and

$$\sum_{(0,i)\in D_0} \upsilon_{\Delta_0}(0,i) = 1,$$

where $\underline{\upsilon}_{\Delta_0}$ as before denotes the steady state probability vector.

As in Section 2 we can construct successively a sequence of processes $X_m(t)$ on $\mathcal{X}_m = \Delta_m \cup D_{m+1} \cup \ldots \cup D_M$ and $Z_m(t)$ on $\Delta_m = (m,0) \cup D_m$ (with steady state vector $\underline{\upsilon}_{\Delta_m}$) as follows for $m = 1, 2 \ldots M$.

i) For the Markov process $X_m(t)$, the transition rates $\mu_m(k',j'\,|\,k,j)$ are defined as follows.

a) If $(k,j) = (m,0)$ and $(k',j') \neq (m,0)$:

$$\mu_m(k',j'\,|\,m,0) = \sum_{(m-1,i)\in\Delta_{m-1}} \mu_{m-1}(k',j'\,|\,m-1,i)\upsilon_{\Delta_{m-1}}(m-1,i).$$

b) If $(k,j) \neq (m,0)$:

$$\mu_m(k',j'\,|\,k,j) = \begin{cases} \mu_{m-1}(m-1,0\,|\,k,j), & \text{if } (k',j') = (m,0), \\ \mu_{m-1}(k',j'\,|\,k,j), & \text{if } (k',j') \neq (m,0). \end{cases}$$

c) And if $(k,j) = (k',j') = (m,0)$:

$$\mu_m(m,0\,|\,m,0) = -\sum_{(m,j)\in D_m} \mu_m(m,j).$$

ii) For $Z_m(t)$:

$$\lambda_{\Delta_m}(m,j\,|\,m,i) = \begin{cases} \mu_m(m,j\,|\,m,i), & \text{if } i \neq 0, j \neq 0, \\ \displaystyle\sum_{(k,j') \notin \Delta_m} \mu_m(k,j'\,|\,0,i) + \mu_m(m,0\,|\,m,i), & \text{if } i \neq 0, j = 0, \\ \mu_m(m,j\,|\,m,0), & \text{if } i = 0, j \neq 0, \\ -\displaystyle\sum_{(m,j') \in D_m} \mu_m(m,j'\,|\,m,0), & \text{if } i = j = 0. \end{cases}$$

5. APPLICATIONS TO RELIABILITY AND QUEUEING SYSTEMS

Consider the classical reliability problem where a system of known structure is composed of N components and it operates continuously. The time to failure of component $i = 1,\ldots,N$ is exponentially distributed with rate $\mu(i)$ and it is independent of the state of the other components. This type of systems has been studied in Derman et al. [1], Katehakis and Derman [5], Frostig [3] and Hooghiemstra and Koole [4] as well as Katehakis and Melolidakis [7], Righter [14], Koole and Spieksma [11] and Ungureanu, Melamed, Katehakis and Bradford [17].

In this section we assume that when the system fails it is restored (or replaced) to a state "as good as new" and the time it takes for this restoration is exponentially distributed with rate λ.

These assumptions imply that at any point in time the state of the system can be identified by a boolean M-vector $\underline{x} = (x_1,\ldots,x_M)$, with $x_i = 1$ if the $i-$th component is working, else $x_i = 0$. Hence $\mathcal{X} = \{0,1\}^M$ is the set of all possible states. Under these conditions the time evolution of the state of the system can be described by a continuous time Markov chain. The structure of the system is specified by a binary function ϕ defined on $\mathcal{X}$. Let $G = \{\underline{x} \, : \, \phi(\underline{x}) = 1\}$ denote the set of all operational (good) states of the system and let $B = \{\underline{x} \, : \, \phi(\underline{x}) = 0\}$ denote all failed states of the system. For such a system it is important to compute measures of performance such as the availability of the system defined as $\alpha_\phi = \sum_{\underline{x} \in G} \pi(\underline{x})$. Regardless of the choice of the structure ϕ it is easy to see that the corresponding chain is successively lumpable.

For example, for the parallel system we have $B = \{\,(0,\ldots,0)\,\}$. Figure 7 illustrates the transition diagram of the corresponding Markov process for the parallel system when $M = 3$. It is clear that this process is successively lumpable with respect to partition $\mathcal{D}$, of size $M+1$, with $\forall \underline{x} \in \mathcal{X}$:

$$\underline{x} \in D_m, \text{ if } \sum_i x_i = m.$$

It is important to note that the successively lumpable property of the process results in the following computational gains: instead of solving a system of size 2^M we only need to solve M systems the largest of which is of size $\binom{M}{\lfloor M/2 \rfloor} + 1$.

Another application of the successive lumping procedure is the $M/PH/1$ queueing model, for both a finite and countable state space. In this model customers arrive according to a Poisson process with parameter λ and are served in multiple (finite) phases. The duration of each phase is exponentially distributed with parameter $\mu(i)$ for phase i and each customer has to complete all L phases of service before he can leave the system. More information about the $M/PH/1$ queueing model can be found in Neuts [13].
Figure 8 shows the transition diagram for an $M/PH/1$ model when there are 3 or less customers in queue. In state (i, j) there are i customers in the system and the first customer in line has j phases to go before he is completely served and able to leave the system. This figure clearly shows that the $M/PH/1$ model is successively lumpable with (i, L) being the entrance state for the set $D_0 \cup \ldots \cup D_i$ for all i.

6. FURTHER RESEARCH

We are investigating the possibility that more complex structures can be handled analogously and the class of successively lumpable Markov chains can be expanded. Further, we are working on supply chain applications of the models described herein.

Acknowledgement: The authors would like to thank Dr. Flora Spieksma, University of Leiden, for many useful comments. Research partially supported by Rutgers Business School Research Committee.

References

[1] Derman, C., Lieberman, G.J. & Ross, S.M. (1980). On the optimal assignment of servers and a repairman . *Journal of Applied Probability*, 17, 577–581.

[2] Feinberg, B.N. & Chui, S.S. (1987). A method to calculate steady state distributions of large Markov chains. *Operations Research*, 35, 282–290.

[3] Frostig, E. (1999). Jointly optimal allocation of a repairman and optimal control of service rate for machine repairman problem. *European Journal of Operational Research*, 116, 274–280.

[4] Hooghiemstra, G. & Koole, G. (2000). On the convergence of the power series algorithm. *Performance Evaluation*, 42, 21–39.

[5] Katehakis, M.N. & Derman, C. (1984). Optimal repair allocation in a series system. *Mathematics of Operations Research*, 9, 615–623.

[6] Katehakis, M.N. & Derman, C. (1989). On the maintenance of systems composed of highly reliable components. *Management Science*, 551–560.

[7] Katehakis, M.N. & Melolidakis, C. (1995). On the optimal maintenance of systems and control of arrivals in queues. *Stochastic Analysis and Applications*, 13, 137–164.

[8] Kemeny, J.G. & Snell, J.L. (1960). *Finite Markov Chains.* D. van Nostrand Company, inc., Princeton, N.J.

[9] Kim, D.S. & Smith, R.L. (1989). An exact aggregation algorithm for a special class of markov chains. *Technical Report.*

[10] Kim, D.S. & Smith, R.L. (1990), An exact aggregation-disaggregation algortihm for mandatory set decomposable markovchains. *Numerical Solution of Markov Chains*, 89–104.

[11] Koole, G.M. & Spieksma, F. M. (2001). On deviation matrices for birth - death processes. *Probability in the Engineering and Informational Sciences*, 15, 239–258.

[12] Miranker, W.L. & Pan, V.Ya. (1980). Methods of aggregation. *Linear Algebra and its Application*, 29, 231–257.

[13] Neuts, M.F. (1981). *Matrix-geometric solutions in stochastic models.* The Johns Hopkins University Press.

[14] Righter, R. (1996). Optimal policies for scheduling repairs and allocating heterogeneous servers. *Journal of applied probability*, 436–547.

[15] Ross, S.M. (1996). *Stochastic Processes.* John Wiley and Sons, NY.

[16] Schweitzer, P.J., Puterman, M.L. & Kindle, W.L. (1984). Iterative aggregation-disaggregation procedures for discounted semi-markov reward processes. *Operations Research*, 33, 589–605.

[17] Ungureanu, V., Melamed, B., Katehakis, M.N. & Bradford, P.G. (2006). Deferred assignment scheduling in cluster-based servers. *Cluster Computing*, 9, 57–65.

[18] Yap, V.B. (2009). Similar states in continuous time markov chains. *Journal of Applied Probability*, 46, 497–506.

[19] Zhang, Q., Mi, N., Riska, A. & Smirni, E. (2007), Performance-guided load (un) balancing under autocorrelated flows. *IEEE Transactions on Parallel and Distributed Systems*, 652–665.

7. FIGURES

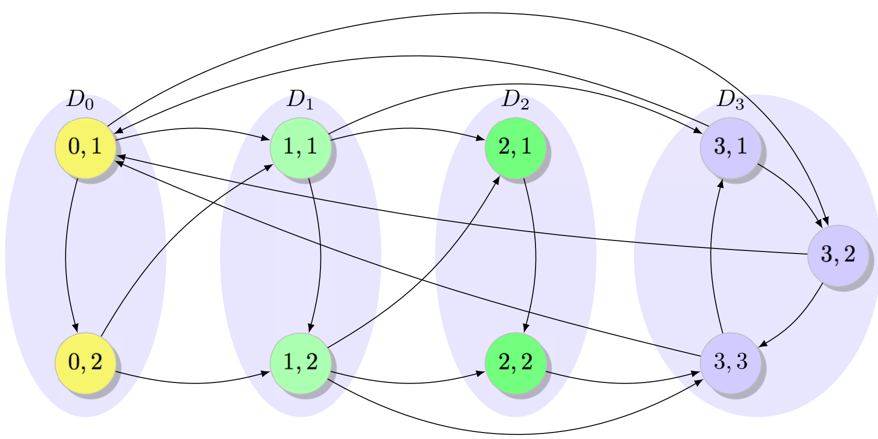
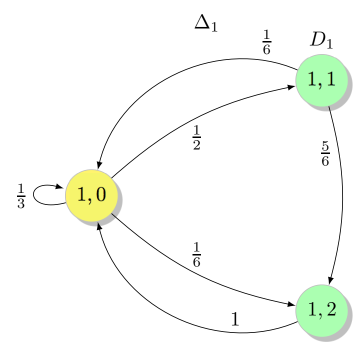


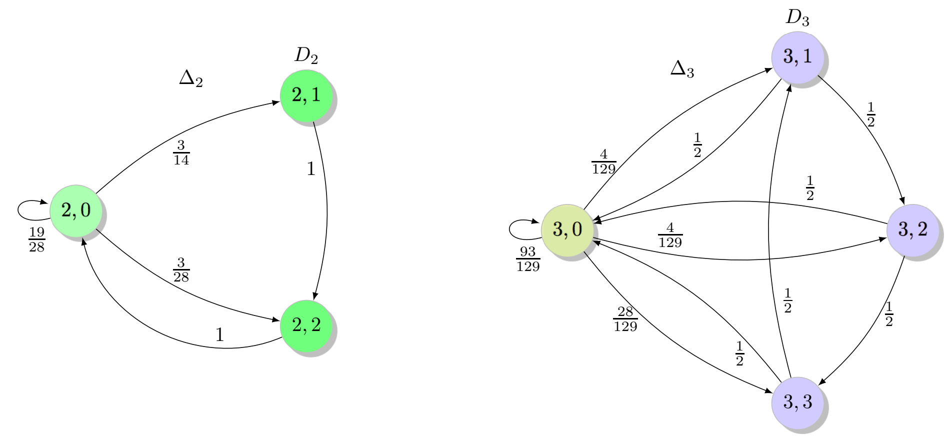
Figure 1: Transition diagram of a successively lumpable Markov chain $X(t)$, arrows represent possible transitions under $\underline{\underline{P}}$.



(a) Transition diagram of $\underline{\underline{U}}_{\Delta_0}$.

(b) Transition diagram of $\underline{\underline{U}}_{\Delta_1}$.

Figure 2: First iteration.



(a) Graphical representation of $\underline{\underline{U}}_{\Delta_2}$.

(b) Transition diagram of $\underline{\underline{U}}_{\Delta_3}$.

Figure 3: Second iteration.

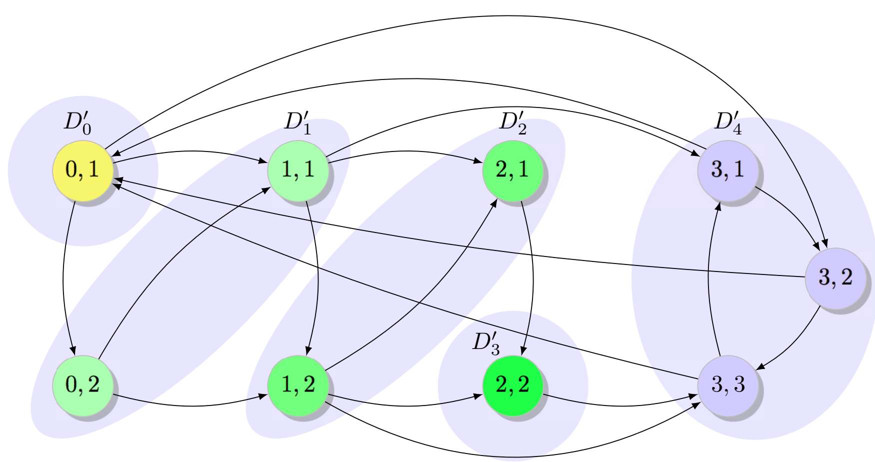


Figure 4: Transition diagram of $X(t)$, of Figure 1, with state space $\mathcal{X}$ partitioned by $\mathcal{D}' = \{D'_0, D'_1, D'_2, D'_3, D'_4\}$.

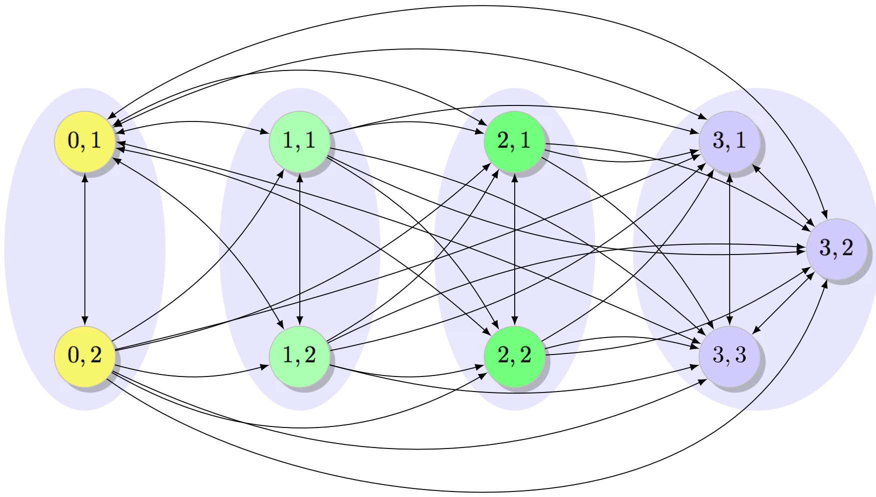


Figure 5: Transition diagram of a process with maximum number of positive probability transitions that is successively lumpable, cf., Remark 4.

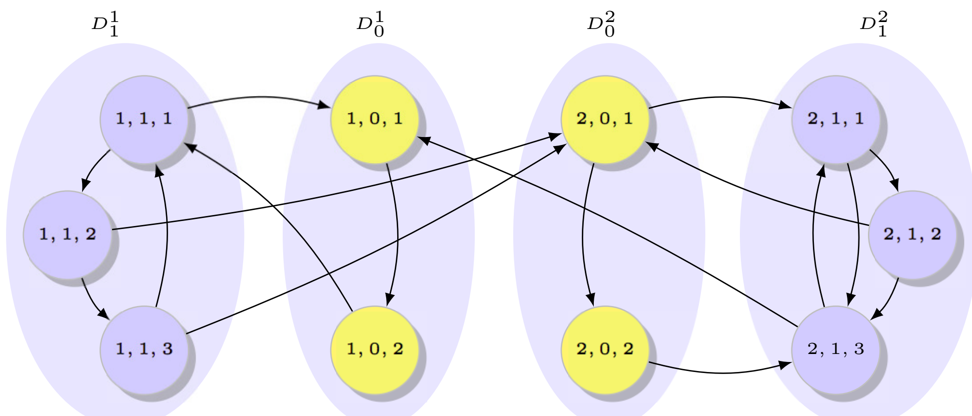


Figure 6: Transition diagram of a multiple successively lumpable Markov chain.

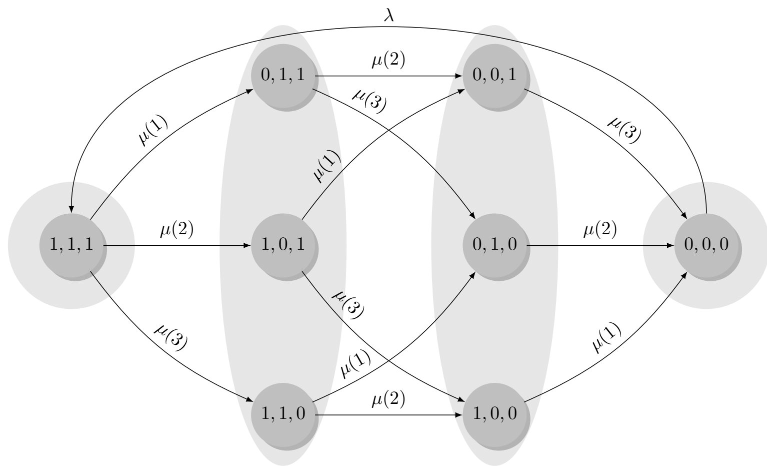


Figure 7: Transition diagram for parallel system with $M = 3$.

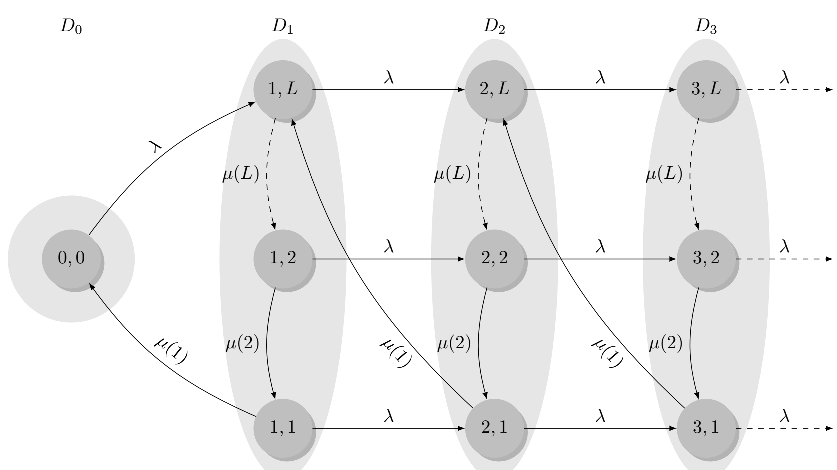


Figure 8: M/PH/1 model.