Perturbed Markov Chains with Damping Component

Dmitrii Silvestrov¹ · Sergei Silvestrov² · Benard Abola² · Pitos Seleka Biganda²³ · Christopher Engström² · John Magero Mango⁴ · Godwin Kakuba⁴

Received: 19 March 2019 / Revised: 21 November 2019 / Accepted: 27 July 2020 / Published online: 13 August 2020 © The Author(s) 2020

Abstract
The paper is devoted to studies of regularly and singularly perturbed Markov chains with damping component. In such models, a matrix of transition probabilities is regularised by adding a special damping matrix multiplied by a small damping (perturbation) parameter $\varepsilon$. We perform a detailed perturbation analysis for such Markov chains, particularly, give effective upper bounds for the rate of approximation for stationary distributions of unperturbed Markov chains by stationary distributions of perturbed Markov chains with regularised matrices of transition probabilities, asymptotic expansions for approximating stationary distributions with respect to damping parameter, explicit coupling type upper bounds for the rate of convergence in ergodic theorems for $n$-step transition probabilities, as well as ergodic theorems in triangular array mode.

Keywords Markov chain · Damping component · Information network · Regular perturbation · Singular perturbation · Stationary distribution · Asymptotic expansion · Rate of convergence · Coupling · Ergodic theorem · Triangular array mode

Mathematics Subject Classification (2010) 60J10 · 60J22 · 65C40 · 90B15 · 68M11

1 Introduction
Perturbed Markov chains is one of the popular and important objects of studies in the theory of Markov processes and their applications to stochastic networks, queuing and reliability models, bio-stochastic systems, and many other stochastic models.

Dmitrii Silvestrov
silvestrov@math.su.se

¹ Department of Mathematics, Stockholm University, Stockholm, Sweden
² Division of Applied Mathematics, School of Education, Culture and Communication, Mälardalen University, Västerås, Sweden
³ University of Dar es Salaam, Dar es Salaam, Tanzania
⁴ Makerere University, Kampala, Uganda
We refer here to some books and recent papers devoted to perturbation problems for Markov type processes, (Avrachenkov et al. 2013, 2018; Bini et al. 2005; Englund 2001; Gambini et al. 2008; Gyllenberg and Silvestrov 2008; Hartfiel and Meyer 1998; Konstantinov et al. 2003; Koroliuk and Limnios 2005; Korolyuk and Korolyuk 1999; Ni 2011; Mitrophanov 2005; Ni et al. 2008; Petersson 2016; Silvestrov and Petersson 2014; Silvestrov et al. 2018; Silvestrov and Silvestrov 2016, 2017a, b, 2019a, b; Stewart 1998, 2001, 1994; Yin and Zhang 2005, 2013). In particular, we would like to mention works (Avrachenkov et al. 2013; Gyllenberg and Silvestrov 2008; Silvestrov and Silvestrov 2016, 2017a), where the extended bibliographies of works in the area and the corresponding methodological and historical remarks can be found.

We are especially interested in models of Markov chains commonly used for description of information networks. In such models an information network is represented by the Markov chain associated with the corresponding node links graph. Stationary distributions and other related characteristics of information Markov chains usually serve as basic tools for ranking of nodes in information networks. The ranking problem may be complicated by singularity of the corresponding information Markov chain, where its phase space is split into several weakly or completely non communicating groups of states. In such models, the matrix of transition probabilities $P_0$ of information Markov chain is usually regularised and approximated by matrix $P_\varepsilon = (1 - \varepsilon)P_0 + \varepsilon D$, where $D$ is a so-called damping stochastic matrix with identical rows and all positive elements, while $\varepsilon \in [0, 1]$ is a damping (perturbation) parameter.

The power method is often used to approximate the corresponding stationary distribution $\bar{\pi}_\varepsilon$ by rows of matrix $P_\varepsilon$. The damping parameter $\varepsilon \in (0, 1]$ should be chosen neither too small nor too large. In the first case, where $\varepsilon$ takes too small values, the damping effect will not work against absorbing and pseudo-absorbing effects, since the second eigenvalue for such matrices (determining the rate of convergence in the above mentioned ergodic approximation) take values approaching 1. In the second case, the ranking information (accumulated by matrix $P_0$ via the corresponding stationary distribution) may be partly lost, due to the deviation of matrix $P_\varepsilon$ from matrix $P_0$. This actualises the problem of construction asymptotic expansions for perturbed stationary distribution $\bar{\pi}_\varepsilon$ with respect to damping parameter $\varepsilon$, as well as studies of asymptotic behaviour of matrices $P_\varepsilon$ in triangular array mode, where $\varepsilon \to 0$ and $n \to \infty$, simultaneously.

In this paper, we perform the detailed perturbation analysis of Markov chains with damping component, using the procedure of artificial regeneration and renewal techniques for the approximating Markov chains and the coupling method in ergodic theorems for perturbed Markov chains with damping component. We get effective explicit series representations for the corresponding stationary distributions $\bar{\pi}_\varepsilon$, upper bounds for the deviation $\bar{\pi}_\varepsilon - \bar{\pi}_0$, and asymptotic expansions for $\bar{\pi}_\varepsilon$ with respect to the perturbation parameter $\varepsilon$, effective explicit upper bounds for the rate of convergence in ergodic theorems as well as ergodic theorems for Markov chains with damping component in triangular array mode.

Real world system consists of interacting units or components. These components constitute what is termed as information networks. With recent advancement in technology, filtering information has become a challenge in such systems. Moreover, their significance is visible as they find their applications in Internet search engines, biological, financial, transport, queuing networks and many others, (Andersson and Silvestrov 2008; Avrachenkov et al. 2013; Biganda et al. 2017; Brin and Page 1998; Engström 2016; Engström and Silvestrov 2014, 2016a, b, 2017; Gleich 2015; Langville and Meyer 2004, 2011; Sun and Han 2013).
PageRank is the link-based criteria that captures the importance of Web pages and provide rankings of the pages in the search engine Google (Avrachenkov et al. 2013; Biganda et al. 2017; Brin and Page 1998; Engström 2016; Engström and Silvestrov 2014, 2016a, b, 2017; Langville and Meyer 2004, 2011). The transition matrix (also called Google matrix $G$) of a Markov chain in PageRank problem is defined in Andersson and Silvestrov (2008), Haveliwala and Kamvar (2003) and Langville and Meyer (2004, 2011), as $G = cP + (1 - c)E$, where $P$ is an $n \times n$ row-stochastic matrix (also called hyperlink matrix), $E$ (the damping matrix) is the $n \times n$ rank-one stochastic matrix and $c \in (0, 1)$ is the damping parameter.

The fundamental concept of PageRank is to use the stationary distribution of the Markov chain on the network to rank Web pages. However, other algorithms similar to PageRank exist in literature, for instance, EigenTrust algorithm (Kamvar et al. 2003) and DeptRank algorithm (Battiston et al. 2012). In addition, variants of PageRank in relation to some specific networks has been studied, e.g. in Biganda et al. (2018a, b); and also updating PageRank due to changes in some network is in literature, for instance, in Abola et al. (2018a, b).

The parameter $c$ is very important in the PageRank definition, because it regulates the level of the uniform noise introduced to the system (Avrachenkov et al. 2013; Langville and Meyer 2011). If $c = 1$ there are several absorbing states for the random walk defined by $P$. However, if $0 < c < 1$, the Markov chain induced by matrix $G$ is ergodic (Avrachenkov et al. 2013). The authors of Langville and Meyer (2011) argued that the parameter $c$ controls the asymptotic rate of convergence of power method algorithm. Similar arguments were given in Andersson and Silvestrov (2008), where it is pointed out that the choice of $c$ is a delicate matter. It may result into convergence and stability problems, if not carefully chosen.

The damping factor $c$ may be denoted and interpreted differently depending on a model being studied. For instance, a model of Markov chain with restart is considered in Avrachenkov et al. (2018), where parameter $p$ is the probability to restart the move and $1 - p$ is the probability to follow the link according to the corresponding transition probability of the above Markov chain. Hence, one can argue that the parameter $p$ has the same interpretation as the damping factor in Google’s PageRank problem.

Our representation of perturbed Markov chains is traditional for perturbed stochastic processes. In fact, PageRank is the stationary distribution of the singularly perturbed Markov chain with perturbation parameter $\varepsilon = 1 - c$. Also, notation $D$ is usually used for the damping matrix instead of notation $E$ mentioned above. Hence, we wish to point out here that, representation of information network model by a Markov chain with matrix of transition probabilities $P_\varepsilon = (1 - \varepsilon)P_0 + \varepsilon D$ should not create any confusion to readers. We perform asymptotic analysis of such Markov chains, in particular, under the assumption that $\varepsilon \to 0$.

The paper includes 5 sections. In Section 2, we derive a special series representation for stationary distributions of Markov chains with damping component. In Section 3, we describe continuity properties of stationary distributions with respect to perturbation (damping) parameter, give explicit upper bounds for rates of convergence in approximations of the stationary distributions for Markov chain with damping component and present asymptotic expansions for stationary distribution of Markov chains with damping component. In Section 4, we get explicit coupling type upper bounds for the rate of convergence in ergodic theorems for Markov chain with damping component. In Section 5, we present ergodic theorems for Markov chains with damping component in a triangular array mode, where time tends to infinity and perturbation parameter tends to zero, simultaneously.
2 Stationary Distributions for Markov Chains with Damping Component

In this section, we introduce the model of Markov chains with damping component. We describe a procedure of embedding such Markov chains in the model of discrete time regenerative processes with special damping regenerative times and derive a special series representation for stationary distributions. This representation plays the key role in the following perturbation analysis. We also give formulas for stationary distributions of the corresponding limiting (unperturbed) Markov chains.

2.1 Markov Chains with Damping Component

Let (a) $X = \{1, 2, \ldots, m\}$ be a finite set, (b) $\bar{p} = \{p_1, \ldots, p_m\}$ and $\bar{d} = \{d_1, \ldots, d_m\}$, be two discrete probability distributions, (c) $P_0 = \|p_{0,ij}\|$ be a $m \times m$ stochastic matrix, (d) $D = \|d_{ij}\|$ be a $m \times m$ damping stochastic matrix with elements $d_{ij} > 0$, $i, j = 1, \ldots, m$, and (e) $P_\varepsilon = \|p_{\varepsilon,ij}\| = (1 - \varepsilon)P_0 + \varepsilon D$ is a stochastic matrix with elements $p_{\varepsilon,ij} = (1 - \varepsilon)p_{0,ij} + \varepsilon d_{ij}$, $i, j = 1, \ldots, m$, where $\varepsilon \in [0, 1]$.

We refer to a Markov chain $X_{\varepsilon,n}, n = 0, 1, \ldots$, with the phase space $X$, the initial distribution $\bar{p}$, and the matrix of transition probabilities $P_\varepsilon$ as a Markov chain with damping component.

Denote by $P_m$ the class of all initial distributions $\bar{p} = \{p_1, \ldots, p_m\}$.

Let $p_{\varepsilon,ij}(n) = P\{X_{\varepsilon,n} = j/X_{\varepsilon,0} = i\}, i, j \in X, n = 0, 1, \ldots$ be $n$-steps transition probabilities for the Markov chain $X_{\varepsilon,n}$. Obviously, $p_{\varepsilon,ij}(0) = I(i = j), i, j \in X$ and $p_{\varepsilon,ij}(1) = p_{\varepsilon,ij}, i, j \in X$. Let, also, $p_{\varepsilon,\bar{p},j}(n) = p_{\bar{p}}\{X_{\varepsilon,n} = j\} = \sum_{i \in X} p_i p_{\varepsilon,ij}(n), \bar{p} \in P_m, j \in X, n = 0, 1, \ldots$. Obviously, $p_{\varepsilon,\bar{p},j}(0) = p_j, j \in X$.

Here and henceforth, symbols $P_\bar{p}$ and $E_\bar{p}$ are used for probabilities and expectations related to a Markov chain with an initial distribution $\bar{p}$. In the case, where the initial distribution is concentrated in a state $i$ the above symbols take the forms $P_i$ and $E_i$.

The phase space $X$ of the Markov chain $X_{\varepsilon,n}$ is one aperiodic class of communicative states, for every $\varepsilon \in (0, 1]$.

The following theorem give a series representation for stationary distribution of Markov chain $X_{\varepsilon,n}$. This representation plays the key role in the asymptotic perturbation analysis for Markov chains with damping component.

**Theorem 1** The following ergodic relation takes place for any initial distribution $\bar{p} \in P_m, j \in X$, and $\varepsilon \in (0, 1]$,

$$p_{\varepsilon,\bar{p},j}(n) \rightarrow \pi_{\varepsilon,j} = \varepsilon \sum_{l=0}^{\infty} p_{0,\bar{d},j}(l)(1 - \varepsilon)^l \text{ as } n \rightarrow \infty. \quad (1)$$

**Proof** Let us describe a procedure of embedding the Markov chains with damping component in the model of discrete time regenerative processes with special damping regenerative times.

Let (a) $U$ be a random variable taking values in space $X$ and such that $P\{U = j\} = p_j, j \in X$; (b) $U_{i,n}, i, j \in X, n = 1, 2, \ldots$ be a family of independent random variables taking values in space $X$ and such that $P\{U_{i,n} = j\} = p_{0,ij}, i, j \in X, n = 1, 2, \ldots$; (c) $V_n, n = 0, 1, \ldots$ be a sequence of independent random variables taking values in space $X$ and such that $P\{V_n = j\} = d_j, j \in X, n = 1, 2, \ldots$; (d) $W$ is a binary random variable taking values 0 and 1 with probabilities, respectively $q_0$ and $q_1$; (e) $W_{\varepsilon,n}, n = 1, 2, \ldots$ be,
for every $\varepsilon \in [0,1]$, a sequence of independent binary random variables taking values 0 and 1 with probabilities, respectively, $1 - \varepsilon$ and $\varepsilon$, for $n=1,2,\ldots$; (d) the random variables $U, W$, the family of random variables $U_{i,n}, i \in \mathbb{X}, n = 1, 2, \ldots$, and the random sequences $V_{\varepsilon,n}, n = 1, 2, \ldots$ and $W_{\varepsilon,n}, n = 1, 2, \ldots$ are mutually independent, for every $\varepsilon \in [0,1]$.

Let us now define, for every $\varepsilon \in [0,1]$, the random sequence $X_{\varepsilon,n}, n = 0, 1, \ldots$ by the following recurrent relation,

$$X_{\varepsilon,n} = U_{X_{\varepsilon,n-1},1}I(W_{\varepsilon,n} = 0) + V_{n}I(W_{\varepsilon,n} = 1), n = 1, 2, \ldots, X_{\varepsilon,0} = U. \tag{2}$$

It is readily seen that the random sequence $X_{\varepsilon,n}, n = 0, 1, \ldots$ is, for every $\varepsilon \in [0,1]$, a homogeneous Markov chain with phase space $\mathbb{X}$, the initial distribution $\bar{p}$ and the matrix of transition probabilities $P_{\varepsilon}$.

Let us now consider the extended random sequence,

$$Y_{\varepsilon,n} = (X_{\varepsilon,n}, W_{\varepsilon,n}), n = 1, 2, \ldots, X_{\varepsilon,0} = U, W_{\varepsilon,0} = W. \tag{3}$$

This random sequence also is, for every $\varepsilon \in [0,1]$, a homogeneous Markov chain, with phase space $\mathbb{Y} = \mathbb{X} \times \{0,1\}$, the initial distribution $\bar{p}q = \langle p_{i}q_{r}, (i, r) \in \mathbb{Y}\rangle$ and transition probabilities,

$$p_{\varepsilon,ir,jk} = P\{X_{\varepsilon,1} = j, W_{\varepsilon,1} = k/X_{\varepsilon,0} = i, W_{\varepsilon,0} = r\}
= \begin{cases} 
(1 - \varepsilon)p_{0,ij} & \text{for } (i, r) \in \mathbb{Y}, j \in \mathbb{X}, k = 0, \\
\varepsilon d_{j} & \text{for } (i, r) \in \mathbb{Y}, j \in \mathbb{X}, k = 1. 
\end{cases} \tag{4}$$

Now, let us assume that $\varepsilon \in (0,1]$.

Let us define times of sequential hitting state 1 by the second component $W_{\varepsilon,n}$,

$$T_{\varepsilon,n} = \min(n > T_{\varepsilon,n-1}, W_{\varepsilon,n} = 1), n = 1, 2, \ldots, T_{\varepsilon,0} = 0. \tag{5}$$

The random sequence $Y_{\varepsilon,n}, n = 0, 1, \ldots$ is a discrete time regenerative process with “damping” regeneration times $T_{\varepsilon,n}, n = 0, 1, \ldots$.

This follows from independence of transition probabilities $p_{\varepsilon,ir,jk}$, given by relation (4), on $(i, r) \in \mathbb{Y}$ if $k = 1$.

Let us also denote by $P_{Q_{m}}$ the class of all initial distributions $\bar{p}q = \langle p_{i}q_{r}, (i, r) \in \mathbb{Y}\rangle$.

If the initial distribution $\bar{p}q = d\bar{q}_{0} = \langle d_{i}q_{o,ir}, (i, r) \in \mathbb{Y}\rangle$, where $\bar{q}_{0} = \langle q_{o,0} = 0, q_{o,1} = 1\rangle$, then $Y_{\varepsilon,n}$ is a standard regenerative process. If $\bar{p}q \neq d\bar{q}_{0}$, then $Y_{\varepsilon,n}$ is a regenerative process with the transition period $[0,T_{\varepsilon,1})$.

The inter-regeneration times $S_{\varepsilon,n} = T_{\varepsilon,n} - T_{\varepsilon,n-1}, n = 1, 2, \ldots$ are i.i.d. geometrically distributed random variables, with parameter $\varepsilon$, i.e., $P\{S_{\varepsilon,1} = n\} = \varepsilon(1 - \varepsilon)^{n-1}, n \geq 1$.

Let us introduce sets $Z_{j} = \{(j,0), (j,1)\}, j \in \mathbb{X}$. Obviously, $p_{\varepsilon,\bar{p}j,n} = P_{\bar{p}q}(Y_{\varepsilon,n} \in Z_{j}), n \geq 0$. That is why, probabilities $p_{\varepsilon,\bar{p}j,n}(n), n \geq 0$ are, for every $j \in \mathbb{X}$, the unique bounded solution for the following discrete time renewal type equation,

$$p_{\varepsilon,\bar{p}j,n}(n) = p_{0,\bar{p}j,n}(1 - \varepsilon)^{n} + \sum_{l=1}^{n} p_{\varepsilon,\bar{d}j,n-l}\varepsilon(1 - \varepsilon)^{l-1}, n \geq 0. \tag{6}$$

Equation 6 is the standard renewal equation for probabilities $p_{\varepsilon,\bar{d}j,n}(n)$ in the case, where the initial distribution $\bar{p}q = d\bar{q}_{0}$.

In the standard regeneration case, the geometrical distribution of the regeneration time $T_{\varepsilon,1} = S_{\varepsilon,1}$ is aperiodic and has expectation $\varepsilon^{-1}$. 

 Springer
This makes it possible to apply the discrete time renewal theorem (see, for example, Feller 1968) to the renewal (6). This yields the following ergodic relation, for $j \in \mathbb{X}$,

$$p_{\varepsilon, \delta, j}(n) \to \pi_{\varepsilon, j} = \varepsilon \sum_0^{\infty} p_{0, \delta, j}(l) (1 - \varepsilon)^l \text{ as } n \to \infty. \quad (7)$$

Let us define $p_{\varepsilon, \delta, j}(n - l) = 0$, for $l > n$. Relation (7) implies that $p_{\varepsilon, \delta, j}(n - l) \to \pi_{\varepsilon, j}$ as $n \to \infty$, for $l \geq 0$ and $j \in \mathbb{X}$. Using relation (7), and the Lebesgue dominated convergence theorem, we get, for $\overline{p} \in \mathcal{P}_m$, $j \in \mathbb{X}$,

$$\lim_{n \to \infty} p_{\varepsilon, \delta, j}(n) = \lim_{n \to \infty} p_{0, \delta, j}(n)(1 - \varepsilon)^n$$

$$+ \lim_{n \to \infty} \sum_1^{\infty} p_{\varepsilon, \delta, j}(n - l) \varepsilon (1 - \varepsilon)^{l-1} = \pi_{\varepsilon, j}. \quad (8)$$

The proof is complete. \(\Box\)

The phase space $\mathbb{X}$ is one aperiodic class of communicative states for the Markov chain $X_{\varepsilon, n}$, for every $\varepsilon \in (0, 1]$. In this case, its stationary distribution $\overline{\pi}_{\varepsilon} = \langle \pi_{\varepsilon, j}, j \in \mathbb{X} \rangle$ is the unique positive solution for the system of linear equations,

$$\sum_{i \in \mathbb{X}} \pi_{\varepsilon, i} p_{\varepsilon, ij} = \pi_{\varepsilon, j}, j \in \mathbb{X}, \sum_{j \in \mathbb{X}} \pi_{\varepsilon, j} = 1. \quad (9)$$

Also, the stationary probabilities $\pi_{\varepsilon, j}$ can be represented in the form $\pi_{\varepsilon, j} = e_{\varepsilon, j}^{-1}$, via the expected return times $e_{\varepsilon, j}$, with the use of regeneration property of the Markov chain $X_{\varepsilon, n}$ at moments of return in state $j$. The series representation (1) for the stationary distribution $\pi_{\varepsilon, j}$ is based on the use of alternative “shorter” damping regeneration times. This representation is, by our opinion, a more effective tool for performing asymptotic perturbation analysis for Markov chains with damping component.

It should be noted that the series representation (1) can also be obtained by alternative ways, for example, by expanding in series the matrix representation for stationary distribution of the Markov chain $X_{\varepsilon, n}$ given in Langville and Meyer (2004), or just by direct checking that limits $\pi_{\varepsilon, j}$ given by relation (1) satisfy the system of linear (9). We, however, do prefer to get it using embedding of the Markov chain $X_{\varepsilon, n}$ into the model of regenerative process with special damping regeneration moments and the corresponding renewal (6). This approach yields not only the series representation (1) for the stationary distribution but also the corresponding ergodic relation. Also the renewal (6) is essentially used in the proofs of Theorems 6, 7 and 8 and let us get more general convergence relations and, also, explicit upper bounds for the rate of convergence in these relations, pointed out in Remarks 8 and 9.

### 2.2 The Stationary Distribution of the Markov Chain $X_{0,n}$

Let us describe ergodic properties of Markov chains $X_{0,n}$.

Matrix $\| p_{\varepsilon, ij}(n) \| = P_{\varepsilon}^n$, for $n \geq 0$, where $P_{\varepsilon}^0 = \| I(i = j) \|$. Therefore, for $i, j \in \mathbb{X}, n \geq 0$,

$$p_{\varepsilon, ij}(n) \to p_{0, ij}(n) \text{ as } \varepsilon \to 0. \quad (10)$$

This relation let one consider the Markov chain $X_{\varepsilon, n}$, for $\varepsilon \in (0, 1]$, as a perturbed version of the Markov chain $X_{0,n}$ and to interpret the damping parameter $\varepsilon$ as a perturbation parameter.

In the present paper, we assume that the following condition holds for some $h \geq 1$:
A_{h,1}: The phase space $\mathbb{X} = \bigcup_{g=1}^{h} \mathbb{X}^{(g)}$, where: (a) $\mathbb{X}^{(g)}$, $g = 1, \ldots, h$ are non-intersecting subsets of $\mathbb{X}$, (b) $\mathbb{X}^{(g)}$, $g = 1, \ldots, h$ are non-empty, closed aperiodic classes of communicative states for the Markov chain $\mathcal{X}_{0,n}$.

The Markov chains $\mathcal{X}_{\varepsilon,n}$ has the same finite phase space $\mathbb{X} = \{1, \ldots, m\}$, for every $\varepsilon \in [0, 1]$. As it was pointed out in Section 2.1, this phase space is one aperiodic class of communicative states for the Markov chain $\mathcal{X}_{\epsilon,n}$, for every $\varepsilon \in (0, 1]$. However, the situation can be different for the limiting Markov chain $\mathcal{X}_{0,n}$, i.e., for the case where $\varepsilon = 0$. Condition $A_{h,1}$ describe possible alternatives, related either to the model with regular perturbation type ($h = 1$) or to the model with singular perturbation type ($h > 1$). In the first case ($h = 1$), the phase space $\mathbb{X}$ is one aperiodic class of communicative states for the Markov chain $\mathcal{X}_{0,n}$. In the second case ($h > 1$), the phase space $\mathbb{X}$ split in $h$ non-intersecting closed aperiodic classes of communicative states for the Markov chain $\mathcal{X}_{0,n}$.

If the initial distribution of the Markov chain $\mathcal{X}_{0,n}$ is concentrated at the set $\mathbb{X}^{(g)}$, for some $g = 1, \ldots, h$, then $\mathcal{X}_{0,n} = \mathcal{X}_{0,n}^{(g)}$, $n = 0, 1, \ldots$ can be considered as a Markov chain with the reduced phase space $\mathbb{X}^{(g)}$ and the matrix of transition probabilities $P_{0,g} = \| p_{0,rk}^{(g)} \|, r \in \mathbb{X} = \sum_{r \in \mathbb{X}} P_{0,rg}$ Let also, $P_{0,g}^{n} = \| p_{0,rg}^{n} \|, n = 0, 1, \ldots$ be matrices of $n$-steps transition probabilities for the Markov chain $\mathcal{X}_{0,n}^{(g)}$, for $g = 1, \ldots, h$.

According to condition $A_{h,1}$, the Markov chain $\mathcal{X}_{0,n}^{(g)}$ is, for every $g = 1, \ldots, h$, ergodic, i.e., the following ergodic relation holds for any $r, k \in \mathbb{X}^{(g)}$, $g = 1, \ldots, h$,

$$ p_{0,rg}^{(g)}(n) \rightarrow \pi_{0,k}^{(g)} \text{ as } n \rightarrow \infty, \quad (11) $$

where $\pi_{0,k}^{(g)} = \langle \pi_{0,k}^{(g)}, k \in \mathbb{X}^{(g)} \rangle$ is, for $g = 1, \ldots, h$, the stationary distribution of the Markov chain $\mathcal{X}^{(g)}$.

The stationary distribution $\pi_{0,k}^{(g)}$ is, for every $g = 1, \ldots, h$, the unique positive solution for the system of linear equations,

$$ \pi_{0,k}^{(g)} = \sum_{r \in \mathbb{X}^{(g)}} \pi_{0,r}^{(g)} p_{0,rg}^{(g)}, k \in \mathbb{X}^{(g)}, \quad \sum_{k \in \mathbb{X}^{(g)}} \pi_{0,k}^{(g)} = 1. \quad (12) $$

Let us introduce probabilities $f_{\tilde{p}}^{(g)} = \sum_{i \in \mathbb{X}^{(g)}} p_{i}$, for $\tilde{p} \in \mathcal{P}_{m}$, $g = 1, \ldots, h$. Note that some of these probabilities can be equal to 0, however, $\sum_{g=1}^{h} f_{\tilde{p}}^{(g)} = 1$.

The following simple lemma, which proof can be found, for example, in Avrachenkov and Litvak (2004) or Avrachenkov et al. (2008), presents the variant of ergodic theorem for the Markov chain $\mathcal{X}_{0,n}$.

**Lemma 1** Let condition $A_{h,1}$ holds, for some $h \geq 1$. Then, the following ergodic relation takes place, for any $\tilde{p} \in \mathcal{P}_{m}$, $k \in \mathbb{X}^{(g)}$, $g = 1, \ldots, h$,

$$ p^{(g)}_{0,\tilde{p},k}(n) \rightarrow \pi_{0,\tilde{p},k}^{(g)} = f_{\tilde{p}}^{(g)} \pi_{0,k}^{(g)} \text{ as } n \rightarrow \infty. \quad (13) $$

**Remark 1** Ergodic relation (13) shows that in the singular case, where condition $A_{h,1}$ holds for some $h > 1$, the stationary probabilities $\pi_{0,\tilde{p},k}^{(g)}$ defined by the asymptotic relation (13) may depend on the initial distribution. The stationary distributions $\tilde{\pi}_{0,\tilde{p}} = \langle \pi_{0,\tilde{p},k}, k \in \mathbb{X} \rangle$ and $\tilde{\pi}_{0,d} = \langle \pi_{0,d,k}, k \in \mathbb{X} \rangle$ coincide, if probabilities $f_{\tilde{p}}^{(g)} = f_{d}^{(g)}$, $g = 1, \ldots, h$. These relations obviously hold for any initial distribution $\tilde{p} \in \mathcal{P}_{m}$, in the regular case, where $h = 1$. 

© Springer
3 A Perturbation Analysis for Stationary Distributions of Markov Chains with Damping Component

In this section, we present results concerned with continuity of stationary distributions $\tilde{\pi}_\varepsilon$ with respect to damping (perturbation) parameter $\varepsilon \to 0$. We also give explicit upper bounds for the rate of convergence in the above continuity convergence relations for stationary probabilities and the corresponding asymptotic expansions with respect to perturbation parameter $\varepsilon$.

3.1 Continuity Property for Stationary Probabilities

In what follows, relation $\varepsilon \to 0$ is a reduced version of relation $0 < \varepsilon \to 0$.

**Theorem 2** Let condition $A_{h,1}$ holds. Then, the following asymptotic relation holds, for $k \in \mathbb{X}$,

$$\pi_{\varepsilon,k} \to \pi_{0,\tilde{d},k} \text{ as } \varepsilon \to 0.$$ (14)

**Proof** Let $S_\varepsilon$ be a random variable geometrically distributed with parameter $\varepsilon$, i.e., $P\{S_\varepsilon = n\} = \varepsilon(1 - \varepsilon)^{n-1}$, $n = 1, 2, \ldots$. Obviously, $S_\varepsilon - 1 \xrightarrow{p} \infty$ as $\varepsilon \to 0$. In the case, where condition $A_{h,1}$ holds, the above asymptotic relation and ergodic relation (13) imply that, for $k \in \mathbb{X}$,

$$p_{0,\tilde{d},k}(S_\varepsilon - 1) \xrightarrow{p} \pi_{0,\tilde{d},k} \text{ as } \varepsilon \to 0.$$ (15)

It is readily seen that the following representation takes place for the stationary probabilities $\pi_{\varepsilon,k}, k \in \mathbb{X}$,

$$\pi_{\varepsilon,k} = \varepsilon \sum_{l=0}^{\infty} p_{0,\tilde{d},k}(l)(1 - \varepsilon)^l = \mathbb{E}p_{0,\tilde{d},k}(S_\varepsilon - 1).$$ (16)

Since the sequence $p_{0,\tilde{d},k}(n), n = 0, 1, \ldots$ is a bounded, relations (15), and (16), and the corresponding variant of the Lebesgue dominated convergence theorem imply that the following asymptotic relation holds, for $k \in \mathbb{X}$,

$$\pi_{\varepsilon,k} = \mathbb{E}p_{0,\tilde{d},k}(S_\varepsilon - 1) \to \pi_{0,\tilde{d},k} \text{ as } \varepsilon \to 0.$$ (17)

The proof is complete. $\square$

**Remark 2** Theorem 2 implies that, in the case, where condition $A_{h,1}$ holds, the continuity property for stationary distributions $\tilde{\pi}_\varepsilon$, as $\varepsilon \to 0$, takes place under the additional assumption that $f_{\tilde{p}}^{(g)} = f_{\tilde{d}}^{(g)}, g = 1, \ldots, h$. In particular it is so, for the regular case, where $h = 1$.

3.2 Rate of Convergence for Stationary Distributions

Let assume that condition $A_{h,1}$ holds. In this case, the reduced Markov chain $X_{0,n}^{(g)}$ with the phase space $\mathbb{X}^{(g)}$ and the matrix of transition probabilities $P_{0,g}$ is, for every $g = 1, \ldots, h$, exponentially ergodic and the following inequalities take place, for $g = 1, \ldots, h$ and $n \geq 1$,

$$\max_{r,k \in \mathbb{X}^{(g)}} |p_{0,rk}^{(g)}(n) - \pi_{0,k}^{(g)}| \leq C_g \kappa_g^n.$$ (18)
with some constants $C_g = C_g(P_{0,g}) \in [0, \infty)$, $\lambda_g = \lambda_g(P_{0,g}) \in [0, 1)$, $g = 1, \ldots, h$ and distributions $\bar{\pi}^{(g)}_0 = (\pi^{(g)}_{0,k}, k \in X^{(g)})$, $g = 1, \ldots, h$, with all positive components.

According to the Perron-Frobenius theorem, the role of $\lambda_g$ can play, for every $g = 1, \ldots, h$, the absolute value of the second (by absolute value), eigenvalue of matrix $P_{0,g}$, and $C_g$ is the constant, which can be computed using the algorithm described, for example, in the book (Feller 1968).

Condition $A_{h,1}$ is, in fact, equivalent to the following condition:

$A_{h,2}$: The phase space $X = \bigcup_{g=1}^h X^{(g)}$, where: (a) $X^{(g)}$, $g = 1, \ldots, h$ are non-intersecting subsets of $X$, (b) $X^{(g)}$, $g = 1, \ldots, h$ are non-empty, closed classes of states for the Markov chain $X_{0,n}$, such that inequalities (18) hold.

Indeed, inequalities (18) imply that $P_{0,k}(n) \to \pi^{(g)}_{0,k}$ as $n \to \infty$, for $r, k \in X^{(g)}$, $g = 1, \ldots, h$, and, thus, $\bar{\pi}^{(g)}_0 = (\pi^{(g)}_{0,k}, k \in X^{(g)})$ is the stationary distribution for the Markov chain $X_{0,n}$, for $g = 1, \ldots, h$. Thus, probabilities $P_{0,k}(n) > 0$, $r, k \in X^{(g)}$, $g = 1, \ldots, h$ for all $n$ large enough. This implies that $X^{(g)}$, $g = 1, \ldots, h$ are closed, aperiodic classes of communicative states.

**Theorem 3** Let condition $A_{h,2}$ holds. Then the following relation holds, for $k \in X^{(g)}$, $g = 1, \ldots, h$ and $\varepsilon \in (0, 1]$,

$$\|\pi_{\varepsilon,k} - \pi_{0,\bar{d},k}\| \leq \varepsilon(\|d_k - \pi_{0,\bar{d},k}\| + \frac{f_d^{(g)}C_g\lambda_g}{1 - \lambda_g}). \quad (19)$$

**Proof** Inequalities (18) imply that the following relation holds, for $k \in X^{(g)}$, $g = 1, \ldots, h$ and $n \geq 1$,

$$|p_{0,\bar{d},k}(n) - \pi_{0,\bar{d},k}| = \left| \sum_{r \in X^{(g)}} d_r P_{0,rk}(n) - f_d^{(g)}\pi_{0,k} \right|$$

$$\leq \sum_{r \in X^{(g)}} d_r |P_{0,rk}(n) - \pi_{0,k}| \leq f_d^{(g)}C_g\lambda^n. \quad (20)$$

Using relations (1) and (20), we get the following relation, for $k \in X^{(g)}$, $g = 1, \ldots, h$,

$$\|\pi_{\varepsilon,k} - \pi_{0,\bar{d},k}\| = \left| \sum_{l=0}^{\infty} p_{0,\bar{d},k}(l)\varepsilon(1-\varepsilon)^l - \pi_{0,\bar{d},k} \right|$$

$$= \left| \sum_{l=0}^{\infty} p_{0,\bar{d},k}(l)\varepsilon(1-\varepsilon)^l \right| - \left| \sum_{l=0}^{\infty} \pi_{0,\bar{d},k}\varepsilon(1-\varepsilon)^l \right|$$

$$\leq \varepsilon|d_k - \pi_{0,\bar{d},k}| + \varepsilon f_d^{(g)}C_g \sum_{l=1}^{\infty} \lambda^l (1 - \varepsilon)^l$$

$$= \varepsilon|d_k - \pi_{0,\bar{d},k}| + \varepsilon \frac{f_d^{(g)}C_g\lambda(1 - \varepsilon)}{1 - \lambda_g}$$

$$\leq \varepsilon|d_k - \pi_{0,\bar{d},k}| + \frac{f_d^{(g)}C_g\lambda}{1 - \lambda_g}. \quad (21)$$

The proof is complete.
Remark 3 The quantities $|d_k - \pi_{0,\tilde{d},k}|$ appearing in inequality (19) are, in some sense, determined by a prior information about the stationary probabilities $\pi_{0,\tilde{d},k}$. They take smaller values if one can choose the damping distribution $\tilde{d}$ with smaller deviation from the stationary distribution $\tilde{\pi}_{0,\tilde{d}} = (\pi_{0,\tilde{d},k}, k \in X)$. Inequalities $|d_k - \pi_{0,\tilde{d},k}| \leq d_k \vee (1 - d_k) \leq 1$ let one replace the term $|d_k - \pi_{0,\tilde{d},k}|$ in inequality (19) by quantities independent on the corresponding stationary probabilities $\pi_{0,\tilde{d},k}$.

Remark 4 Theorem 2 remains valid even if we weaken condition $A_{h,2}$ by omitting in relation (18) the assumption of positivity for the distributions $\tilde{\pi}_0^{(g)} = (\pi_{0,i}^{(g)}, i \in X^{(g)})$, $g = 1, \ldots, h$. In this case, condition $A_{h,2}$ implies that the phase spaces $X^{(g)} = \bar{X}_1^{(g)} \cup X_0^{(g)}$, where $\bar{X}_1^{(g)} = \{i \in X^{(g)} : \pi_{0,i}^{(g)} > 0\}$ is the non-empty closed, aperiodic class of communicative states, while $X_0^{(g)} = \{i \in X^{(g)} : \pi_{0,i}^{(g)} = 0\}$ is the class (possibly empty) of transient states, for the Markov chain $X_{0,n}^{(g)}$. Note that $\tilde{\pi}_0^{(g)}$ still is the stationary distribution for the Markov chain $X_{0,n}^{(g)}$.

Remark 5 We would like also to refer to paper (Mitrophanov 2005), where one can find alternative upper bounds for the rate of convergence of stationary distributions for perturbed Markov chains and further related references.

3.3 Asymptotic Expansions for Stationary Distributions

Let condition $A_{h,1}$ holds, and assume that class $X^{(g)}$ includes $m_g \geq 1$ states, for $g = 1, \ldots, h$.

Let us denote by $\rho_{g,1}, \ldots, \rho_{g,m_g}$ the eigenvalues of the stochastic matrix $P_{0,g}$, for $g = 1, \ldots, h$.

Condition $A_{h,1}$ is, in fact, equivalent to the following condition:

$A_{h,3}$: The phase space $X = \bigcup_{g=1}^{h} X^{(g)}$, where: (a) $X^{(g)}$, $g = 1, \ldots, h$ are non-intersecting subsets of $X$, (b) $X^{(g)}$, $g = 1, \ldots, h$ are non-empty, closed classes of states for the Markov chain $X_{0,n}$, (c) inequalities $\rho_{g,1} = 1 > |\rho_{g,2}| \geq \cdots \geq |\rho_{g,m_g}|$, $g = 1, \ldots, h$, hold.

Condition $A_{h,1}$ implies that condition $A_{h,3}$ holds. This follows from the Perron-Frobenius theorem.

Condition $A_{h,3}$ imply that the following eigenvalues decomposition representations take place, for $r, k \in X^{(g)}$, $g = 1, \ldots, h$ and $n \geq 1$,

$$p_{0,rk}^{(g)}(n) = \pi_{0,k}^{(g)} + \rho_{g,2}^{n} \pi_{0,rk}^{(g)}[2] + \cdots + \rho_{g,m_g}^{n} \pi_{0,rk}^{(g)}[m_g].$$

(22)

where: (a) $\tilde{\pi}_0^{(g)} = (\pi_{0,k}^{(g)}, k \in X^{(g)})$ is a distribution with all positive components, for $g = 1, \ldots, h$, (b) $\pi_{0,rk}^{(g)}[l']$, $r, k \in X^{(g)}$, $l = 2, \ldots, m_g$, $g = 1, \ldots, h$ are some real- or complex-valued coefficients ($\pi_{0,rr'}[l''] = \pi_{0,rr'}[l'']$, if $\rho_{g,l'} = \rho_{g,l''}$, for some $2 \leq l' \leq l'' \leq m_g$).

Obviously, relation (22) implies that probabilities $p_{0,rk}^{(g)}(n) \to \pi_{0,k}^{(g)}$ as $n \to \infty$, for $r, k \in X^{(g)}$, $g = 1, \ldots, h$. Thus, $\tilde{\pi}_0^{(g)}$ is the stationary distribution for the Markov chain $X_{0,n}^{(g)}$, for $g = 1, \ldots, h$. Moreover, the above ergodic relation implies that probabilities $P_{0,rk}^{(g)}(n) > \ldots \ldots \ldots$
Theorem 4 Let condition \( \text{A}_{h,3} \) holds. Then, the following asymptotic expansions take place for every \( k \in \mathbb{X}, g = 1, \ldots, h, n \geq 1, \) and \( \varepsilon \to 0, \)

\[
\pi_{\varepsilon,k} = \pi_{0,d,k} + \pi_{0,d,k}^{(g)}[1] \varepsilon + \cdots + \pi_{0,d,k}^{(g)}[n] \varepsilon^n + O(\varepsilon^{n+1}).
\]  

Proof Relations (1) and (23) imply that the following relation holds, for \( k \in \mathbb{X}, g = 1, \ldots, h, n \geq 1, \)

\[
\pi_{\varepsilon,k} = \varepsilon \sum_{n=0}^{\infty} p_{0,d,k}(n)(1 - \varepsilon)^n
\]

\[
= \varepsilon d_k + \varepsilon \sum_{n=1}^{\infty} (\pi_{0,d,k} + \sum_{l=2}^{m_g} \rho_{g,l}^{n} \pi_{0,d,k}^{(g)}[l]) (1 - \varepsilon)^n
\]

\[
= \pi_{0,d,k} + \varepsilon (d_k - \pi_{0,d,k}) + \sum_{l=2}^{m_g} \pi_{0,d,k}^{(g)}[l] \varepsilon \sum_{n=1}^{\infty} \rho_{g,l}^{n} (1 - \varepsilon)^n
\]

\[
= \pi_{0,d,k} + \varepsilon (d_k - \pi_{0,d,k}) + \sum_{l=2}^{m_g} \pi_{0,d,k}^{(g)}[l] \varepsilon \frac{\rho_{g,l} \varepsilon}{(1 - \rho_{g,l})(1 - \varepsilon)}
\]

\[
= \pi_{0,d,k} + \varepsilon (d_k - \pi_{0,d,k}) + \sum_{l=2}^{m_g} \pi_{0,d,k}^{(g)}[l] \rho_{g,l} \varepsilon (1 - \varepsilon) (1 - \rho_{g,l} + \rho_{g,l} \varepsilon)^{-1}.
\]  

Function \( b \varepsilon (1 - \varepsilon)(a + b \varepsilon)^{-1}, \varepsilon \in [0, 1] \) admits, for any complex numbers \( a \neq 0 \) and \( b, \) the following asymptotic expansion, for every \( n \geq 1 \) and \( \varepsilon \to 0, \)

\[
b \varepsilon (1 - \varepsilon)(a + b \varepsilon)^{-1} = a^{-1} b \varepsilon - a^{-1} b (1 + a^{-1} b) \varepsilon^2 + a^{-2} b^2 (1 + a^{-1} b) \varepsilon^3 + \cdots + (-1)^{n-1} a^{-(n-1)} b^{n-1} (1 + a^{-1} b) \varepsilon^n + O(\varepsilon^{n+1}).
\]  

\( \odot \) Springer
Relations (26) and (27), used for $a = 1 - \rho_{g,l}, b = \rho_{g,l}$, let us write down the following asymptotic expansions for stationary probabilities $\pi_{\varepsilon,k}, k \in \mathbb{X}(g), g = 1, \ldots, h$, for every $n \geq 1$ and $\varepsilon \to 0$,

$$
\pi_{\varepsilon,k} = \pi_{0,d,k} + \varepsilon(d_k - \pi_{0,d,k}) + \sum_{l=2}^{m_\varepsilon} \pi^{(g)}_{0,d,k}[l] \rho_{g,l} \varepsilon (1 - \varepsilon)(1 - \rho_{g,l} + \rho_{g,l} \varepsilon)^{-1} \\
= \pi_{0,d,k} + \tilde{\pi}^{(g)}_{0,d,k}[1] \varepsilon + \cdots + \tilde{\pi}^{(g)}_{0,d,k}[n] \varepsilon^n + O(\varepsilon^{n+1}).
$$

(28)

The proof is complete.

It worth noting that some of eigenvalues $\rho_{g,l}$ and coefficients $\pi^{(g)}_{0,e_k}[l]$ can be complex numbers. Despite of this, coefficients $\tilde{\pi}^{(g)}_{0,d,k}[n], n \geq 1$, in the expansions given in relation (25), are real numbers.

Indeed, $\pi_{\varepsilon,k}, \varepsilon \in (0, 1]$ and $\pi_{0,d,k}$ are real numbers. Relation (25) implies that $\varepsilon^{-1}(\pi_{\varepsilon,k} - \pi_{0,d,k}) \to \tilde{\pi}^{(g)}_{0,d,k}[1] \varepsilon \to 0$. Thus, $\tilde{\pi}^{(g)}_{0,d,k}[1]$ is a real number. In the similar way, the above proposition can be proved for all coefficients in expansions (25). This implies that the remainders of these expansions $O(\varepsilon^{n+1})$ also are real-valued functions of $\varepsilon$.

Moreover, since $\tilde{\pi}_\varepsilon = \langle \pi_{\varepsilon,k}, k \in \mathbb{X} \rangle, \varepsilon \in (0, 1]$ and $\tilde{\pi}_{0,d} = \langle \pi_{0,d,k}, k \in \mathbb{X} \rangle$ are probability distributions, the equalities, $\sum_{g=1}^{h} \sum_{k \in \mathbb{X}(g)} \tilde{\pi}^{(g)}_{0,d,k}[n] = 0, n \geq 1$, connect coefficients in the asymptotic expansions (25).

4 Coupling and Ergodic Theorems for Perturbed Markov Chains with Damping Component

In this section, we present coupling algorithms and get the effective upper bounds for the rate of convergence in ergodic theorems for regularly and singularly perturbed Markov chains with damping component. We refer to books Lindvall (2002) and Thorisson (2000), where one can find historical and methodological remarks and extended bibliographies related to the coupling method.

4.1 Maximal Coupling for Discrete Distributions

Let $\tilde{\rho}' = \langle p'_i, i \in \mathbb{X} \rangle$ and $\tilde{\rho}'' = \langle p''_i, i \in \mathbb{X} \rangle$ be two discrete probability distributions. Let us denote by $\mathcal{L}[\tilde{\rho}', \tilde{\rho}'']$ the class of two-dimensional probability distribution $\tilde{P} = \langle P_{ij}, (i, j) \in \mathbb{X} \times \mathbb{X} \rangle$ which satisfy the following conditions (a) $P'_i = \sum_{j \in \mathbb{X}} P_{ij} = p'_i, i \in \mathbb{X}$; (b) $P''_j = \sum_{i \in \mathbb{X}} P_{ij} = p''_j, j \in \mathbb{X}$.

Let us also denote,

$$
Q_{\tilde{P}} = \sum_{i \in \mathbb{X}} P_{ii} \quad \text{and} \quad Q(\tilde{\rho}', \tilde{\rho}'') = \sup_{\tilde{P} \in \mathcal{L}[\tilde{\rho}', \tilde{\rho}'']} Q_{\tilde{P}}.
$$

(29)

The following lemma presents the well known “coupling” result, which variants can be found in Griffeth (1975), Kalashnikov (1989), Lindvall (2002), Pitman (1979), Thorisson (2000), Silvestrov (1983, 1984a, b, 1994) and Silvestrov and Pezinska (1985). The corresponding historical remarks can be found in Lindvall (2002).
Lemma 2 There exists the two-dimensional distribution \( \tilde{P}^* = (P_{ij}^*, i, j \in \mathbb{X}) \in \mathcal{L} [\tilde{p}', \tilde{p}''] \) such that:

\[
Q_{\tilde{p}'} = Q^* = \sum_{i \in \mathbb{X}} \min (p_i', p_i'') = Q(\tilde{p}', \tilde{p}''). \tag{30}
\]

The distribution \( \tilde{P}^* \) is given by the following relations:

(i) If \( Q^* \in (0, 1) \), then

\[
P_{ij}^* = \min (p_i', p_j'') I(i = j)
\]

\[
+ \frac{1}{1 - Q^*} (p_i' - \min (p_i', p_i'')) (p_j'' - \min (p_j', p_j'')), \quad i, j \in \mathbb{X}. \tag{31}
\]

(ii) \( Q^* = 1 \) if and only if \( p_k' = p_k'' \), \( k \in \mathbb{X} \) and,

\[
P_{ij}^* = \min (p_i', p_j'') I(i = j), \quad i, j \in \mathbb{X}. \tag{32}
\]

(iii) \( Q^* = 0 \) if and only if \( p_k' p_k'' = 0 \), \( k \in \mathbb{X} \) and,

\[
P_{ij}^* = p_i' p_j'', \quad i, j \in \mathbb{X}. \tag{33}
\]

Proof It can be found in the above mentioned works. In order, to improve self-readability of the present paper, we just give a short sketch of the proof. Obviously, probability \( P_{ii} \leq p_i' \wedge p_i'' \), \( i \in \mathbb{X} \), for any two-dimensional distribution \( \tilde{P} = (P_{ij}, (i, j) \in \mathbb{X} \times \mathbb{X}) \in \mathcal{L} [\tilde{p}', \tilde{p}'']. \) This relation implies that \( Q_{\tilde{p}} \leq Q^* = \sum_{i \in \mathbb{X}} p_i' \wedge p_i''. \) This is easily to check that every relation (31), (32), or Eq. 33 defines a two-dimensional distribution \( \tilde{P}^* \) from the class \( \mathcal{L} [\tilde{p}', \tilde{p}'']. \) Moreover, the corresponding quantity \( Q_{\tilde{p}'} = Q^* \). This is obvious for two cases presented in propositions (ii) and (iii). In the first case presented in proposition (i), this follows from relation, \( (p_i' - \min (p_i', p_i'')) (p_i'' - \min (p_i', p_i'')) = 0 \), \( i \in \mathbb{X}. \)

4.2 Coupling for Regularly Perturbed Markov Chains with Damping Component

Let \( \varepsilon \in (0, 1] \). Let us consider the random sequence \( X^{(N)}_{\varepsilon,n} = X_{\varepsilon,Nn}, n = 0, 1, \ldots \), for some natural \( N \geq 1 \). It is a homogeneous Markov chain, with an initial distribution \( \tilde{p} \), the phase space \( \mathbb{X} \), and the matrix of transition probabilities \( P^N_{\varepsilon} = ||p_{\varepsilon,ij}(N)|| \).

Let us define the quantities, for \( i, j \in \mathbb{X} \),

\[
Q_{\varepsilon,ij}^{(N)} = \sum_{r \in \mathbb{X}} \min (p_{\varepsilon,ir}(N), p_{\varepsilon,ij}(N)). \tag{34}
\]

Let us now use the multi-step coupling algorithm for construction a coupling Markov chain \( Z^{(N)}_{\varepsilon,n} = (X^{(N)}_{\varepsilon,n}, X^{(N)}_{\varepsilon,n}) \), \( n = 0, 1, \ldots \), with:

(i) the phase space \( \mathbb{Z} = \mathbb{X} \times \mathbb{X} \);

(ii) the initial distribution \( \tilde{P}_\varepsilon = (P_{\varepsilon,ij}, (i, j) \in \mathbb{Z}) \) constructing according to relation (31), (32), or Eq. 33 for distributions \( \tilde{p}' = \tilde{p} = (p_i, i \in \mathbb{X}) \) and \( \tilde{p}'' = \tilde{\pi}_\varepsilon = (\pi_{\varepsilon,i}, i \in \mathbb{X}); \)

(iii) transition probabilities \( P^{(N)}_{\varepsilon,ij,rk} \) defined by the following relations, for \( (i, j), (r, k) \in \mathbb{Z}; \)
(a) If $Q_{e,ij}^{(N)} \in (0, 1)$, then,
\[
P_{e,ij,rk}^{(N)} = P(X^{(N)}_{e,1} = k, X^{(N)}_{e,1} = r/X^{(N)}_{e,0} = i, X^{(N)}_{e,0} = j)
= \min(p_{e,ir}(N), p_{e,jk}(N))I(r = k)
+ \frac{1}{1 - Q_{e,ij}^{(N)}}(p_{e,ir}(N) - \min(p_{e,ir}(N), p_{e,jr}(N)))
\times (p_{e,jk}(N) - \min(p_{e,ik}(N), p_{e,jk}(N))),
\]
(35)

(b) If $Q_{e,ij}^{(N)} = 1$, then $p_{e,ir}(N) = p_{e,jr}(N), r \in \mathbb{X}$ and,
\[
P_{e,ij,rk}^{(N)} = \min(p_{e,ir}(N), p_{e,jk}(N))I(r = k), r, k \in \mathbb{X}.
\]
(36)

(c) If $Q_{e,ij}^{(N)} = 0$, then $p_{e,ir}(N)p_{e,jr}(N) = 0, r \in \mathbb{X}$ and,
\[
P_{e,ij,rk}^{(N)} = p_{e,ir}(N)p_{e,jk}(N), r, k \in \mathbb{X}.
\]
(37)

The above construction of coupling Markov chain and the following lemma originate from works (Griffeath 1975) and (Pitman 1979). It plays an important role in what follows.

**Lemma 3** Let $Z_{e,n}^{(N)} = (X^{(N)}_{e,n}, X^{(N)}_{e,n}), n = 0, 1, \ldots$, be a homogeneous Markov chain with the phase space $\mathbb{Z} = \mathbb{X} \times \mathbb{X}$, the initial distribution $\bar{P}_e$ and transition probabilities given by relations (35)–(37). Then:

(i) The first component, $X^{(N)}_{e,n}, n = 0, 1, \ldots$, is a homogeneous Markov chain with the phase space $\mathbb{X}$, the initial distribution $\bar{P}_e$ and the matrix of transition probabilities $P_e^{(N)}$.

(ii) The second component $X^{(N)}_{e,n}, n = 0, 1, \ldots$ is a homogeneous Markov chain with the phase space $\mathbb{X}$, the initial distribution $\bar{\pi}_e$ and the matrix of transition probabilities $P_{e}^{N}$.

(iii) The set $\mathbb{Z}_0 = \{(i, i), i \in \mathbb{X}\}$ is an absorbing set for the Markov chain $Z_{e,n}^{(N)}$, i.e., probabilities $P_{e,ii,rk}^{(N)} = 0$, for $i, r, k \in \mathbb{X}, r \neq k$.

**Proof** Variants of the proof can be found in the above mentioned works. In order, to improve self-readability of the present paper, we just give a short sketch of the proof. Let us consider the case, where transition probabilities $P_{e,ij,rk}^{(N)}$ are given by relation (35). According to Lemma 2, the following relation takes place, for $i, j, r \in \mathbb{X},$
\[
\sum_{k \in \mathbb{X}} P_{e,ij,rk}^{(N)} = \sum_{k \in \mathbb{X}} \min(p_{e,ir}(N), p_{e,jk}(N))I(r = k)
+ \frac{1}{1 - Q_{e,ij}^{(N)}}(p_{e,ir}(N) - \min(p_{e,ir}(N), p_{e,jr}(N)))
\times \sum_{k \in \mathbb{X}} (p_{e,jk}(N) - \min(p_{e,ik}(N), p_{e,jk}(N)))
= \min(p_{e,ir}(N), p_{e,jr}(N)) + p_{e,ir}(N)
- \min(p_{e,ir}(N), p_{e,jr}(N)) = p_{e,ir}(N).
\]
(38)

The proof of relation (38) is trivial in the cases, where transition probabilities $P_{e,ij,rk}^{(N)}$ are given by relation (36) or Eq. 37.
Using the assumption that $Z_{\varepsilon,n}^{(N)}$ is a Markov chain and relation (38), we get the following relation, for any chain of states $i_0, \ldots, i_n \in \mathbb{X}, n \geq 0$,

$$
\Pr\{X^{(N)}_{\varepsilon,l} = i_l, l = 0, \ldots, n\} = \sum_{j_0,\ldots,j_n \in \mathbb{X}} \Pr\{X^{(N)}_{\varepsilon,l} = j_l, l = 0, \ldots, n\} 
$$

$$
= \sum_{j_0 \in \mathbb{X}} \sum_{j_1 \in \mathbb{X}} \cdots \sum_{j_{n-1} \in \mathbb{X}} \sum_{j_n \in \mathbb{X}} \Pr\{X^{(N)}_{\varepsilon,l} = j_l, X^{(N)}_{\varepsilon,l} = j_{l-1}, \ldots, X^{(N)}_{\varepsilon,l} = j_n\}.
$$

Relation (39) proves proposition (i). The proof of proposition (ii) is analogous.

According to relation (32), the quantities $Q_{\varepsilon,i}^{(N)} = 1, i \in \mathbb{X}$. This proves proposition (iii).

Let $A = \|a_{ij}\|$ be a $m \times m$ a matrix with real-valued elements. Let us introduce functional,

$$
Q(A) = \min_{1 \leq i, j \leq m} \sum_{k=1}^{m} a_{ik} \land a_{jk}.
$$

The following simple lemma presents some basic properties of functional $Q(A)$

**Lemma 4** Functional $Q(A)$ possesses the following properties: (a) $Q(aA) = aQ(A)$, for any $a \geq 0$; (b) $Q(A) \geq a_1 Q(A_1) + \cdots + a_n Q(A_n)$, for any $m \times m$ matrices $A_1, \ldots, A_n$ with real-valued elements, numbers $a_1, \ldots, a_n \geq 0, a_1 \cdots + a_n = 1$, and matrix $A = a_1 A_1 + \cdots + a_n A_n$, for $n \geq 2$; (c) $Q(A) \in [0, 1]$, for any stochastic matrix $A$; (d) $Q(A) = 1$, for any $m \times m$ stochastic damping type matrix $A = \|a_{ij}\|$, with elements $a_{ij} = a_j \geq 0, i, j = 1, \ldots, m$.

The following useful proposition takes place.

**Lemma 5** The following inequality takes place, for $N \geq 1$ and $\varepsilon \in (0, 1]$,

$$
1 - Q(P_{\varepsilon}^{N}) \leq (1 - Q(P_{\varepsilon}^{0}))(1 - \varepsilon)^{N}.
$$

**Proof** Relation, $AB = B$, holds for any $m \times m$ stochastic matrix $A = \|a_{ij}\|$ and $m \times m$ stochastic damping type matrix $B = \|b_{ij}\|$, with elements $b_{ij} = b_j \geq 0, i, j = 1, \ldots, m$.

Also, matrix $C = BA$, which has elements, $c_{ij} = c_j = \sum_{k=1}^{m} b_k a_{kj} \geq 0, i, j = 1, \ldots, m$, is a stochastic damping type matrix, i.e., it has all rows the same.
Using these remarks, we get the following relation, for \( N \geq 1 \),
\[
\begin{align*}
P_N^\varepsilon &= ((1 - \varepsilon)P_0 + \varepsilon \Delta) \varepsilon \varepsilon 
\end{align*}
\]
\[
\begin{align*}
&= P_{N-1}^\varepsilon (1 - \varepsilon)P_0 + \varepsilon \Delta 
\end{align*}
\]
\[
\begin{align*}
&= P_{N-1}^\varepsilon (1 - \varepsilon)^2P_0^2 + \varepsilon \Delta (1 - \varepsilon)DP_0^2 + \varepsilon \Delta 
\end{align*}
\]
\[
\begin{align*}
&= \cdots = (1 - \varepsilon)^N P_0^N + \varepsilon (1 - \varepsilon)^{N-1} DP_0^{N-1} + \cdots + \varepsilon \Delta. 
\end{align*}
\]
(42)

Using relation (42) and Lemma 4, we get the following relation,
\[
\begin{align*}
Q(P_N^\varepsilon) \geq (1 - \varepsilon)^N Q(P_0^N) + \varepsilon (1 - \varepsilon)^{N-1} Q(DP_0^{N-1}) + \cdots + \varepsilon Q(\Delta) 
\end{align*}
\]
\[
\begin{align*}
&= (1 - \varepsilon)^N Q(P_0^N) + \varepsilon (1 - \varepsilon)^{N-1} + \cdots + \varepsilon 
\end{align*}
\]
\[
\begin{align*}
&= (1 - \varepsilon)^N Q(P_0^N) + 1 - (1 - \varepsilon)^N. 
\end{align*}
\]
(43)

This relation is equivalent to inequality (41).

Let us introduce, for \( N \geq 1 \), the coefficient of ergodicity originating from the measure of ergodicity introduced by Markov (1906) and related functionals used in Dobrushin (1956), Doeblin (1940), Loève (1978) and many other works,
\[
\Delta N (P_0) = \left(1 - Q(P_0^N)\right)^{1/N}. \quad \text{(44)}
\]

The given below Theorems 5 and 6 present effective coupling type upper bounds for the rate of convergence in the individual ergodic theorem for Markov chain with damping component. These theorems are based on corresponding general coupling results for Markov chains given in Griffeath (1975), Lindvall (2002), Pitman (1979) and specify and detail the corresponding coupling upper bounds for the rate of convergence in ergodic theorems for Markov chains with damping component.

Note that condition \( A_{h,1} \) is not required in Theorem 5 formulated below.

Also, we count \( \Delta N (P_0)^0 = 1 \), if \( \Delta N (P_0) = 0 \).

**Theorem 5** The following relation takes place, for every \( N \geq 1 \), \( \tilde{p} \in \mathcal{P}_m \), \( j \in \mathcal{X} \), \( n \geq 0 \), and \( \varepsilon \in (0, 1] \),
\[
|p_{\varepsilon, \tilde{p}, j}(n) - \pi_{\varepsilon, j}| \leq (1 - Q(\tilde{p}, \tilde{\pi}_\varepsilon)) \Delta N (P_0)^{n/|\mathcal{X}|} (1 - \varepsilon)^{|\mathcal{X}|n/\mathcal{X}|N}. \quad \text{(45)}
\]

**Proof** Obviously, for \( \tilde{p} \in \mathcal{P}_m \), \( j \in \mathcal{X} \), \( n \geq 0 \),
\[
\mathbb{P}\{X_{\varepsilon, n}^{(N)} = j\} = p_{\varepsilon, \tilde{p}, j}(Nn). \quad \text{(46)}
\]

Since, the initial distribution of Markov chain \( X_{\varepsilon, n}^{(N)} \) coincides with its stationary distribution, this Markov chain is a stationary random sequence and, thus, for \( j \in \mathcal{X} \), \( n \geq 0 \),
\[
\mathbb{P}\{X_{\varepsilon, n}^{(N)} = j\} = \pi_{\varepsilon, j}. \quad \text{(47)}
\]

Let us now define a hitting (coupling) time,
\[
T_{\varepsilon}^{(N)} = \min(n \geq 0 : X_{\varepsilon, n}^{(N)} = X_{\varepsilon, n}^{(N)}) = \min(n \geq 0 : Z_{\varepsilon, n}^{(N)} \in \mathcal{Z}_0). \quad \text{(48)}
\]

Since \( \mathcal{Z}_0 \) is an absorbing set for the Markov chain \( Z_{\varepsilon, n}^{(N)} \), the following relation holds,
\[
\mathbb{P}\{Z_{\varepsilon, n}^{(N)} \in \mathcal{Z}_0, n \geq T_{\varepsilon}^{(N)}\} = 1. \quad \text{(49)}
\]
Using the above remarks, we get the following relation, for \( j \in \mathbb{X}, n \geq 0 \),

\[
|p_{e,\bar{\pi},j}(Nn) - \pi_{e,j}| = |\mathbb{P}[X_{e,n}^{(N)} = j] - \mathbb{P}[X_{e,n}^{(N)} = j]| \\
= |\mathbb{P}[X_{e,n}^{(N)} = j, X_{e,n}^{(N)} \neq j] - \mathbb{P}[X_{e,n}^{(N)} = j]| \\
\leq \mathbb{P}[X_{e,n}^{(N)} = j, X_{e,n}^{(N)} \neq j] + \mathbb{P}[X_{e,n}^{(N)} = j]
\]

\[
\leq \mathbb{P}[T_{e}^{(N)} > n].
\] (50)

Analogously, we get, for \( \bar{\pi} \in \mathcal{P}_m, j \in \mathbb{X}, \)

\[
|p_{e,\bar{\pi},j}(0) - \pi_{e,j}| \leq \mathbb{P}[T_{e}^{(N)} > 0] = 1 - Q(\bar{\pi}, \bar{\pi}_e). \] (51)

Also, by continuing inequality (41), we get, for \( \bar{\pi} \in \mathcal{P}_m, j \in \mathbb{X}, n \geq 0, \)

\[
|p_{e,\bar{\pi},j}(Nn) - \pi_{e,j}| \leq \mathbb{P}[T_{e}^{(N)} > n] \\
= \sum_{i,j \in \mathbb{X}} \mathbb{P}[X_{e,n}^{(N)} \neq X_{e,n}^{(N)} / X_{e,n-1}^{(N)} = i, X_{e,n-1}^{(N)} = j] \\
\times \mathbb{P}[T_{e}^{(N)} > n - 1, X_{e,n-1}^{(N)} = i, X_{e,n-1}^{(N)} = j] \\
\leq \mathbb{P}[T_{e}^{(N)} > n - 1](1 - Q(\mathbb{P}_e^{N})) \\
\leq \cdots \leq \mathbb{P}[T_{e} > 0](1 - Q(\mathbb{P}_e^{N}))^n \\
\leq (1 - Q(\bar{\pi}, \bar{\pi}_e)(1 - Q(\mathbb{P}_0^{N}))^n(1 - \varepsilon)^{Nn} \\
= (1 - Q(\bar{\pi}, \bar{\pi}_e))\Delta_N(\mathbb{P}_0)^{Nn}(1 - \varepsilon)^{Nn}. \] (52)

Also, for \( \bar{\pi} \in \mathcal{P}_m, j \in \mathbb{X}, n \geq 0 \) and \( l = 0, \ldots, N - 1, \)

\[
|p_{e,\bar{\pi},j}(Nn + l) - \pi_{e,j}| = |\sum_{k \in \mathbb{X}} p_{e,\bar{\pi},k}(Nn + l) - \sum_{k \in \mathbb{X}} \pi_{e,k} p_{e,k}(l)| \\
\leq \sum_{k \in \mathbb{X}} |p_{e,\bar{\pi},k}(Nn) - \pi_{e,k}| p_{e,k}(l) \\
\leq \max_{k \in \mathbb{X}} |p_{e,\bar{\pi},k}(Nn) - \pi_{e,k}| \\
\leq (1 - Q(\bar{\pi}, \bar{\pi}_e))\Delta_N(\mathbb{P}_0)^{Nn}(1 - \varepsilon)^{Nn}. \] (53)

Inequalities (52) and (53) imply inequalities given in relation (45). The proof is complete.

Remark 6 The upper bounds given in relation (45) become better if quantities \( 1 - Q(\bar{\pi}, \bar{\pi}_e), \Delta_N(\mathbb{P}_0) \) and \( 1 - \varepsilon \) take smaller values. The factor \( 1 - Q(\bar{\pi}, \bar{\pi}_e) \), is determined by a prior information about the stationary probabilities. It takes smaller values if one can choose initial distribution \( \bar{\pi} \) with smaller deviation from the stationary distribution \( \bar{\pi}_e \). Relation (45) gives an effective upper bounds for the rate of convergence in the corresponding individual ergodic theorem for the Markov chain \( X_{e,n} \) even in the case, where factor \( \Delta_N(\mathbb{P}_0) = 1. \)

Remark 7 It’s worth noting that the weaker upper bound \( (1 - \varepsilon)^n \) on the right hand side of inequality (45) have been given for Markov chains with a general phase and damping component, in the recent paper (Avrachenkov et al. 2018).
In the case, where condition $A_{1,1}$ holds (i.e., the phase space $X$ is one aperiodic class of communicative states for the Markov chain $X_{0,n}$), $1 - Q(P_0^{N}) \to 0$ as $N \to \infty$, and, thus, the following condition holds for $N$ large enough:

$B_{N,1}: \quad \Delta_N(P_0) < 1.$

Also, condition $B_{N,1}$ is, for every $N \geq 1$, sufficient for holding the weaken variant of condition $A_{1,1}$ mentioned in Remark 4.

Indeed, probabilities $p_{\varepsilon,\tilde{p},j}(n) \to p_{0,\tilde{p},j}(n)$ as $\varepsilon \to 0$, for any $j \in X$, $n \geq 0$. Since stationary probabilities $p_{\varepsilon,\tilde{p}} \in [0,1]$, $j \in X$, any sequence $0 < \varepsilon_n \to 0$ as $n \to \infty$ contains a subsequence $0 < \varepsilon_{n_l} \to 0$ as $l \to \infty$ such that $p_{\varepsilon_{n_l},j} \to p_{0,j}$ as $l \to \infty$, for $j \in X$. By passing $\varepsilon \to 0$ in the inequality (45), we get the following relation holding for $\tilde{p} \in P_m$, $j \in X$, $n \geq 0$,

$$|p_{0,\tilde{p},j}(n) - \pi_{0,j}| \leq (1 - Q(\tilde{p}, \tilde{\pi}_0))\Delta_N(P_0)^{[n/N]N}, \quad (54)$$

where one should count $\Delta_N(P_0)^0 = 1$, if $\Delta_N(P_0) = 0$.

Relation (54) obviously implies that $p_{0,\tilde{p},j}(n) \to \pi_{0,j}$ as $n \to \infty$, for $j \in X$. Thus, limits $\pi_{0,j}, j \in X$ are the same for any subsequences $\varepsilon_n$ and $\varepsilon_{n_l}$ and, thus, stationary probabilities $p_{\varepsilon,\tilde{p}} \to \pi_{0,j}$ as $\varepsilon \to 0$, for $j \in X$. By passing $\varepsilon \to 0$ in the equations given in relation (9), we get that limits $\pi_{0,j}, j \in X$ satisfy the system of linear (9) and, thus, $\tilde{\pi}_0 = \langle \pi_{0,j}, j \in X \rangle$ is the stationary distribution for the Markov chain $X_{0,n}$. Some components of this stationary distribution can be equal to 0. In this case, set $X_1 = \{j \in X: \pi_{0,j} > 0\}$ is a closed, aperiodic class of communicative states, while set $X_0 = \{j \in X: \pi_{0,j} = 0\}$ is the class of transient states, for the Markov chain $X_{0,n}$.

If the stationary distribution $\tilde{\pi}_0 = \langle \pi_{0,j}, j \in X \rangle$ is positive, then $X_0 = \emptyset$. In this case, condition $B_{N,1}$ is sufficient for holding of condition $A_{1,1}$.

Relation (54) implies that the Markov chain $X_{0,n}$ is ergodic with an exponential rate of convergence in the corresponding ergodic theorem, if condition $B_{N,1}$ holds, for some $N \geq 1$.

As it was shown in Pitman (1979), the following formula takes place,

$$1 - Q(P_0^{N}) = \max_{i,j \in X}(1 - \sum_{k \in X} p_{0,ik}(N) \land p_{0,jk}(N))$$

$$= \frac{1}{2} \max_{i,j \in X} \left( \sum_{k \in X} (p_{0,ik}(N) - p_{0,ik}(N) \land p_{0,jk}(N)) \right.\left. + \sum_{k \in X} (p_{0,jk}(N) - p_{0,ik}(N) \land p_{0,jk}(N)) \right)$$

$$= \frac{1}{2} \max_{i,j \in X} \left( \sum_{k: p_{0,ik}(N) \geq p_{0,jk}(N)} (p_{0,ik}(N) - p_{0,jk}(N)) \right.\left. + \sum_{k: p_{0,jk}(N) \geq p_{0,ik}(N)} (p_{0,jk}(N) - p_{0,ik}(N)) \right)$$

$$= \frac{1}{2} \max_{i,j \in X} \left| \sum_{k \in X} (p_{0,ik}(N) - p_{0,jk}(N)) \right|.$$  \quad (55)
Thus, the following formula takes place,
\[
\Delta_N(P_0) = (1 - Q(P_0^N))^{1/N} = (\frac{1}{2} \max_{i,j \in \mathbb{X}} \sum_{k \in \mathbb{X}} |p_{0,ik}(N) - p_{0,jk}(N)|)^{1/N}.
\] (56)

In the case, where condition \( A_{1,1} \) holds, the eigenvalues decomposition representation (22) implies that the following relation holds, under the assumption that the second eigenvalue \( \rho_{1,2} \) and the factor which stays in front of \( |\rho_{1,2}| \) in relation (57) are not equal to 0,
\[
\Delta_N(P_0) \sim (\frac{1}{2} \max_{i,j \in \mathbb{X}} \sum_{k \in \mathbb{X}} |\sum_{2 \leq l \leq m; |\rho_{1,l}| = |\rho_{1,2}|} (\pi_{0,ik}[l] - \pi_{0,jk}[l])|^{1/N} |\rho_{1,2}| \to |\rho_{1,2}| \text{ as } N \to \infty. \] (57)

Relations (54) and (57) show that, in the case where condition \( A_{1,1} \) holds, the coupling upper bounds for rate of convergence in individual ergodic relations given in Theorem 5 usually are asymptotically equivalent with analogous upper bounds, which can be obtained with the use of eigenvalue decomposition representation for transition probabilities. At the same time, computing of coefficients of ergodicity \( \Delta_N(P_0) \) does not require solving of the polynomial equation, det(\( \rho \mathbf{I} - P_0 \)) = 0, that is required for finding eigenvalues.

### 4.3 Coupling for Singularly Perturbed Markov Chains with Damping Component

Let us assume that the following condition holds for some \( N \geq 1 \) and \( h > 1 \):

\[ \mathbf{B}_{N,h} : \text{ The phase space } \mathbb{X} = \bigcup_{g=1}^{h} \mathbb{X}^{(g)}, \text{ where: (a) } \mathbb{X}^{(g)}, \ g = 1, \ldots, h \text{ are non-intersecting subsets of } \mathbb{X}, \text{ (b) } \mathbb{X}^{(g)}, \ g = 1, \ldots, h \text{ are non-empty, closed classes of states for the Markov chain } X_{0,n}, \text{ (c) } \Delta_N(P_0,g) < 1, \text{ for } g = 1, \ldots, h. \]

Let us introduce the discrete distributions \( \tilde{f}_p = (\tilde{f}_p^{(g)}, \ g = 1, \ldots, h) \), where \( \tilde{f}_p^{(g)} = \sum_{k \in \mathbb{X}^{(g)}} p_k, \ g = 1, \ldots, h \) and, also, define, for \( g = 1, \ldots, h \), the conditional distributions \( \tilde{p}^{(g)}(k) = (p_k, k \in \mathbb{X}^{(g)}) \) using relation, \( p_k = p_k^{(g)} \tilde{f}_p^{(g)}, k \in \mathbb{X}^{(g)} \). Obviously \( p_k^{(g)} = p_k/\tilde{f}_p^{(g)}, k \in \mathbb{X}^{(g)} \), for \( 1 \leq g \leq h \) such that \( \tilde{f}_p^{(g)} > 0 \). As usual, one can also define the above conditional distributions in some standard way, for example, as \( p_k^{(g)} = \frac{1}{m_g}, k \in \mathbb{X}^{(g)} \), if \( \tilde{f}_p^{(g)} = 0 \).

We also use the above notations for the case, where distribution \( \tilde{p} \) coincides with distributions \( \tilde{d} = (d_k, k \in \mathbb{X}) \) or \( \tilde{\pi}_\varepsilon = (\pi_{\varepsilon,k}, k \in \mathbb{X}) \).

Let \( \varepsilon \in (0, 1] \), and \( \pi_{\varepsilon,k} = (\pi_{\varepsilon,k}, k \in \mathbb{X}) \) be the stationary distribution for the Markov chain \( X_{\varepsilon,n} \). Theorem 1 and condition \( \mathbf{B}_{N,h} \) imply that the following representation takes place, for \( k \in \mathbb{X}^{(g)}, \ g = 1, \ldots, h, \)
\[
\pi_{\varepsilon,k} = \varepsilon \sum_{l=0}^{\infty} p_{0,d,k}(l)(1 - \varepsilon)^l = \tilde{f}_d^{(g)} \pi_{\varepsilon,k}^{(g)}, \] (58)
where, for \( k \in \mathbb{X}^{(g)}, \ g = 1, \ldots, h, \)
\[
\pi_{\varepsilon,k}^{(g)} = \varepsilon \sum_{l=0}^{\infty} p_{0,d^{(g)},k}(l)(1 - \varepsilon)^l. \] (59)
It is readily seen that \( \bar{\pi}(g) = \langle \pi_{e,k}, k \in \mathbb{X} \rangle \) is, for every \( g = 1, \ldots, h \), the stationary distribution for the Markov chain \( X_{\varepsilon,n}^{(g)} \), with the phase space \( \mathbb{X}^{(g)} \) and the matrix of transition probabilities \( P_{\varepsilon,g} = (1 - \varepsilon)P_{0,g} + \varepsilon D_g \), where \( P_{0,g} = \|p_{0,rk}\|_{r,k \in \mathbb{X}^{(g)}} \) is, according to condition \( B_{N,h} \), a stochastic matrix, while \( D_g = \|d(g)\|_{r,k \in \mathbb{X}^{(g)}} \) is the stochastic damping matrix with elements \( d(g)_{r,k} = \frac{d(k)}{f(g)\bar{d}}, r, k \in \mathbb{X}^{(g)} \).

Note that relations (58) and (59) imply that, for \( g = 1, \ldots, h \),

\[
\pi(\varepsilon,k) \rightarrow \pi(0,k) \quad \text{as} \quad \varepsilon \rightarrow 0. \quad (61)
\]

Below, we count \( \Delta_N(P_{0,g})^0 = 1 \), if \( \Delta_N(P_{0,g}) = 0 \).

**Theorem 6** Let condition \( B_{N,h} \) holds. Then, the following relation takes place, for \( \bar{p} \in \mathcal{P}_m, k \in \mathbb{X}^{(g)}, g = 1, \ldots, h, n \geq 0, \) and \( \varepsilon \in (0, 1] \),

\[
|p_{\varepsilon,\bar{p},k}(n) - \pi_{\varepsilon,k}| \leq \left| (f_d^{(g)}(1 - Q(\bar{\pi}(\varepsilon), \bar{\pi}(g))) \\
+ f_d^{(g)}(1 - Q(\bar{\pi}(g), \bar{\pi}(g)))) \Delta_N(P_{0,g})^{[n/N]N} \\
+ |f_d^{(g)} - f_d^{(g)}\pi_{0,k}^{(g)}| (1 - \varepsilon)^n \right. \quad (62)
\]

**Proof** By using the renewal type relation (6), condition \( B_{N,h} \), and taking into account stationarity of the Markov chain \( X_{\varepsilon,n}^{(g)} \), with the initial distribution \( \pi_{\varepsilon} \), we get the following relation, for \( k \in \mathbb{X}^{(g)}, g = 1, \ldots, h, \) and \( n \geq 0, \)

\[
|p_{\varepsilon,\bar{p},k}(n) - \pi_{\varepsilon,k}| = |p_{\varepsilon,\bar{p},k}(n) - p_{\varepsilon,\pi}(k)| \\
= |p_{0,\bar{p},k}(n) - p_{0,\pi}(k)| (1 - \varepsilon)^n. \quad (63)
\]

Condition \( B_{N,h} \) implies that the following relations hold, for \( \bar{p} \in \mathcal{P}_m, k \in \mathbb{X}^{(g)}, g = 1, \ldots, h, \) \( n \geq 0, \)

\[
p_{0,\bar{p},k}(n) = f_d^{(g)}P_{\rho^{(g)}}(X_{0,n}^{(g)} = k). \quad (64)
\]

and

\[
\pi_{0,\bar{p},k} = f_d^{(g)}\pi_{0,k}^{(g)}. \quad (65)
\]
Relations (64), and (65) let us continue relation (63) and transform it to the following form, for \( k \in \mathbb{X}^{(g)}, g = 1, \ldots, h, n \geq 0, \)
\[
|p_{0,\bar{p},k}(n) - p_{0,\bar{\pi}_\varepsilon,k}(n)|(1 - \varepsilon)^n
\]
\[
= |f_{\bar{p}}^{(g)}(n)\mathbf{P}_{\bar{p}(e)}\{X_{0,n}^{(g)} = k\} - f_{\bar{\pi}_\varepsilon}^{(g)}\mathbf{P}_{\bar{\pi}_\varepsilon}^{(g)}\{X_{0,n}^{(g)} = k\}|(1 - \varepsilon)^n
\]
\[
\leq (f_{\bar{p}}^{(g)}|\mathbf{P}_{\bar{p}(e)}\{X_{0,n}^{(g)} = k\} - \pi_{0,k}^{(g)})
\]
\[
+ f_{\bar{\pi}_\varepsilon}^{(g)}|\mathbf{P}_{\bar{\pi}_\varepsilon}^{(g)}\{X_{0,n}^{(g)} = k\} - \pi_{0,k}^{(g)}|
\]
\[
+ |f_{\bar{p}}^{(g)} - f_{\bar{\pi}_\varepsilon}^{(g)}|\pi_{0,k}^{(g)}(1 - \varepsilon)^n. \tag{66}
\]

Relation (60) let us replace probabilities \( f_{\bar{\pi}_\varepsilon}^{(g)} \) by probabilities \( f_{\bar{d}}^{(g)} \), in relation (66). Also, condition \( \mathbf{B}_{X,h} \) let us apply inequality (54) to the Markov chains \( X_{0,n}^{(g)} \) and transforms relation (66) to the form of relation (62).

As in relation (54), by passing \( \varepsilon \to 0 \) in the first inequality given in relation (62) and taking into account Theorem 2 (which implies that \( Q(\pi_{\varepsilon}^{(g)}, \pi_0^{(g)}) \to 1 \) as \( \varepsilon \to 0 \)), we get the following relation holding for any \( \bar{p} \in \mathcal{P}_m, k \in \mathbb{X}^{(g)}, g = 1, \ldots, h, n \geq 0, \)
\[
|p_{0,\bar{p},k}(n) - \pi_{0,\bar{d},k}| \leq f_{\bar{p}}^{(g)}(1 - Q(\tilde{\pi}_\varepsilon^{(g)}, \tilde{\pi}_0^{(g)}))\Delta_N(\mathbf{P}_{0,g})^{[n/N]N}
\]
\[
+ |f_{\bar{p}}^{(g)} - f_{\bar{d}}^{(g)}|\pi_{0,k}^{(g)} \tag{67}
\]

This relation let one extend the conclusion made in Remark 2.

If \( f_{\bar{p}}^{(g)} = f_{\bar{d}}^{(g)} \), then the second term on the right hand side of relation (67) disappears, and, thus, probabilities \( p_{0,\bar{p},k}(n) \to \pi_{0,\bar{d},k} \) as \( n \to \infty \), with the exponential rate, for \( k \in \mathbb{X}^{(g)} \). In particular, \( f_{\bar{p}}^{(1)} = f_{\bar{d}}^{(1)} = 1 \), if \( h = 1 \).

If \( f_{\bar{p}}^{(g)} \neq f_{\bar{d}}^{(g)} \), then, Theorem 2 and relations (58), (64), and (65), imply that, for \( k \in \mathbb{X}^{(g)} \), probabilities \( p_{0,\bar{p},k}(n) \to \pi_{0,\bar{d},k} = \pi_{0,\bar{d},k} + (f_{\bar{p}}^{(g)} - f_{\bar{d}}^{(g)})\pi_{0,k}^{(g)} \) as \( n \to \infty \), with the exponential rate. More precisely, the above relations imply that the following relation holds, for \( k \in \mathbb{X}^{(g)}, g = 1, \ldots, h, n \geq 0, \)
\[
|p_{0,\bar{p},k}(n) - \pi_{0,\bar{d},k}| = |f_{\bar{p}}^{(g)}\mathbf{P}_{\bar{p}(e)}\{X_{0,n}^{(g)} = k\} - f_{\bar{p}}^{(g)}\pi_{0,k}^{(g)}|
\]
\[
\leq f_{\bar{p}}^{(g)}(1 - Q(\tilde{\pi}_\varepsilon^{(g)}, \tilde{\pi}_0^{(g)}))\Delta_N(\mathbf{P}_{0,g})^{[n/N]N}. \tag{68}
\]

### 5 Ergodic Theorems for Markov Chains with Damping Component in the Triangular Array Mode

In this section, we present ergodic theorems for Markov chains with damping component in the triangular array mode.

#### 5.1 Asymptotic Behaviour of Transition Probabilities in the Triangular Array Mode

In this mode, the asymptotic behaviour of transition probabilities \( p_{\varepsilon,\bar{p},k}(n) \) is studied when time \( n \to \infty \) and the damping parameter \( \varepsilon \to 0 \), simultaneously. One can assume that time \( n = n_\varepsilon \) depends on \( \varepsilon \) in such way that \( n_\varepsilon \to \infty \) as \( \varepsilon \to 0 \).
Asymptotic behaviour of transition probabilities \( p_{e, \bar{p}, k}(n_e) \) as \( \varepsilon \to 0 \) should be compared with repeated limits for \( p_{e, \bar{p}, k}(n) \), as \( n \to \infty \) and, then, \( \varepsilon \to 0 \) or, vice versa, as \( \varepsilon \to 0 \) and, then, \( n \to \infty \).

Let us, first, consider the case, where condition \( A_{1,1} \) holds.

In this case, according to Theorem 1, transition probabilities \( p_{e, \bar{p}, k}(n) \to \pi_{e,k} \) as \( n \to \infty \), for every \( \varepsilon \in (0, 1) \). Also, according to Theorem 2, stationary probabilities \( \pi_{e,k} \to \pi_{0,\bar{d},k} \) as \( \varepsilon \to 0 \), where \( \pi_{0,\bar{d},k} \) are stationary probabilities for the Markov chain \( X_{0,n} \), with the initial distribution \( \bar{d} \).

As follows from relation (10), in this case, transition probabilities \( p_{e, \bar{p}, k}(n) \to p_{0, \bar{p}, k}(n) \) as \( \varepsilon \to 0 \), for any \( n \geq 0 \). Also, according to Lemma 1 and Remark 1, transition probabilities \( p_{0, \bar{p}, k}(n) \to \pi_{0,\bar{d},k} \) as \( n \to 0 \).

The coincidence of the repeated limits for transition probabilities \( p_{e, \bar{p}, k}(n) \) let one expect that, in this case, transition probabilities \( p_{e, \bar{p}, k}(n_e) \) should converge to stationary probabilities \( \pi_{0,\bar{d},k} \) as \( \varepsilon \to 0 \), for an arbitrary \( n_e \to \infty \) as \( \varepsilon \to 0 \).

The situation is different, in the case, where condition \( A_{h,1} \) holds, for some \( h > 1 \).

Again, Theorem 1 implies that transition probabilities \( p_{e, \bar{p}, k}(n) \to \pi_{e,k} \) as \( n \to \infty \), for every \( \varepsilon \in (0, 1) \). Also, according to Theorem 2, stationary probabilities \( \pi_{e,k} \to \pi_{0,\bar{d},k} \) as \( \varepsilon \to 0 \), where \( \pi_{0,\bar{d},k} \) are stationary probabilities of the Markov chain \( X_{0,n} \), with the initial distribution \( \bar{d} \).

According to relation (10), transition probabilities \( p_{e, \bar{p}, k}(n) \to p_{0, \bar{p}, k}(n) \) as \( \varepsilon \to 0 \), for any \( n \geq 0 \). However, Lemma 1 implies that, in this case, transition probabilities \( p_{0, \bar{p}, k}(n) \to \pi_{0,\bar{d},k} \) as \( n \to \infty \).

Thus, the repeated limits for transition probabilities \( p_{e, \bar{p}, k}(n) \) coincides, under additional assumption that the initial distribution \( \bar{p} = \bar{d} \). But, they may not coincide, if \( \bar{p} \neq \bar{d} \).

The transition probabilities \( p_{e, \bar{p}, k}(n) \) satisfy the renewal type relation (6). This relation shows that the initial state \( X_{\varepsilon,0} \) influences the behaviour of the Markov chain \( X_{\varepsilon, n} \), only up to the first damping regeneration time \( T_{\varepsilon,1} \).

The random variables \( T_{\varepsilon,1} \) has the geometric distribution with parameter \( \varepsilon \). Its rate of growth is \( \varepsilon^{-1} \) as \( \varepsilon \to 0 \). Moreover, random variables \( \varepsilon T_{\varepsilon,1} \xrightarrow{d} T_1 \) as \( \varepsilon \to 0 \), where \( T_1 \) is a random variable, which has the exponential distribution with parameter 1.

This hints one that it would be natural to study the asymptotic behaviour of transition probabilities \( p_{e, \bar{p}, k}(n) \) for \( n \to \infty \) as \( \varepsilon \to 0 \) such that \( \varepsilon n \to t \in [0, \infty] \) as \( \varepsilon \to 0 \).

Moreover, it can be expected that in two extremal cases, where \( t = \infty \) or \( t = 0 \), the transition probabilities \( p_{e, \bar{p}, k}(n_e) \) converge as \( \varepsilon \to 0 \) to the corresponding repeated limits, respectively, \( \pi_{0,\bar{d},k} \) or \( \pi_{0,\bar{p},k} \).

The question arises about the asymptotic behaviour of transition probabilities \( p_{e, \bar{p}, k}(n_e) \) in the intermediate case, where the above limit \( t \in (0, \infty) \).

In conclusion of this informal discussion, we would like to mention works Englund (2001), Englund and Silvestrov (1997), Gyllenberg and Silvestrov (2000, 2008), Petersson (2016), Silvestrov (1978, 1979, 2018) and Silvestrov and Petersson (2014), which contain results related to ergodic theorems in triangular array mode and to so-called quasi-stationary ergodic theorems for perturbed regenerative processes, Markov chains and semi-Markov processes.

### 5.2 Ergodic Theorems for Regularly Perturbed Markov Chains with Damping Component in the Triangular Array Mode

The following theorem takes place.
Theorem 7 Let condition $B_{N,1}$ holds for some $N \geq 1$. Then, for $\bar{p} \in \mathcal{P}_m$, $k \in \mathbb{X}$ and any $n_\varepsilon \to \infty$ as $\varepsilon \to 0$,

$$p_{\varepsilon,\bar{p},k}(n_\varepsilon) \to \pi_{0,k} = \pi_{0,\bar{d},k} \text{ as } \varepsilon \to 0. \quad (69)$$

Proof Using the renewal type relation (6) and inequality (54), we get that the following relation holds, for $k \in \mathbb{X}$ and any $n_\varepsilon \to \infty$ as $\varepsilon \to 0$,

$$|p_{\varepsilon,\bar{p},k}(n_\varepsilon) - \pi_{0,k}| = |p_{\varepsilon,\bar{p},k}(n_\varepsilon) - \pi_{0,k}(1 - \varepsilon)^{n_\varepsilon} - \pi_{0,k}(1 - (1 - \varepsilon)^{n_\varepsilon})|$$

$$= |(p_{0,\bar{p},k}(n_\varepsilon) - \pi_{0,k})(1 - \varepsilon)^{n_\varepsilon} + \sum_{l=1}^{n_\varepsilon} (p_{0,\bar{d},k}(n_\varepsilon - l) - \pi_{0,k})\varepsilon(1 - \varepsilon)^{l-1}|$$

$$\leq (1 - Q(\bar{p}, \bar{\pi}_0))\Delta_N(\mathbf{P}_0)^{[n_\varepsilon/N]}N$$

$$+ (1 - Q(\bar{d}, \bar{\pi}_0))\varepsilon \sum_{l=1}^{n_\varepsilon} \Delta_N(\mathbf{P}_0)^{[n_\varepsilon-l]/N}N$$

$$\leq (1 - Q(\bar{p}, \bar{\pi}_0))\Delta_N(\mathbf{P}_0)^{[n_\varepsilon/N]}N$$

$$+ (1 - Q(\bar{d}, \bar{\pi}_0))\varepsilon N(1 - \Delta_N(\mathbf{P}_0)^N)^{-1} \to 0 \text{ as } \varepsilon \to 0. \quad (70)$$

The proof is complete.

Remark 8 Relation (70) gives, in fact, explicit upper bounds for the rate of convergence in the ergodic relation given in Theorem 7.

5.3 Ergodic Theorems for Singularly Perturbed Markov Chains with Damping Component in the Triangular Array Mode

As follows from remarks made in Section 5.1, one should expect that ergodic relations take in this case much more complex forms.

Theorem 8 Let condition $B_{N,h}$ holds for some $N \geq 1$ and $h > 1$. Then the following ergodic relations take place, for $\bar{p} \in \mathcal{P}_m$, $k \in \mathbb{X}(g)$, $g = 1, \ldots, h$:

(i) If $n_\varepsilon \to \infty$ and $\varepsilon n_\varepsilon \to \infty$ as $\varepsilon \to 0$, then,

$$p_{\varepsilon,\bar{p},k}(n_\varepsilon) \to \pi_{0,\bar{p},k}(\infty) = \pi_{0,\bar{d},k} \text{ as } \varepsilon \to 0. \quad (71)$$

(ii) If $n_\varepsilon \to \infty$ and $\varepsilon n_\varepsilon \to t \in (0, \infty)$ as $\varepsilon \to 0$, then,

$$p_{\varepsilon,\bar{p},k}(n_\varepsilon) \to \pi_{0,\bar{p},k}(t) = \pi_{0,\bar{p},k}e^{-t} + \pi_{0,\bar{d},k}(1 - e^{-t}) \text{ as } \varepsilon \to 0. \quad (72)$$

(iii) If $n_\varepsilon \to \infty$ and $\varepsilon n_\varepsilon \to 0$ as $\varepsilon \to 0$, then,

$$p_{\varepsilon,\bar{p},k}(n_\varepsilon) \to \pi_{0,\bar{p},k}(0) = \pi_{0,\bar{p},k} \text{ as } \varepsilon \to 0. \quad (73)$$

Proof The renewal type relation (6) written for $n = n_\varepsilon$ takes the following form, for $\bar{p} \in \mathcal{P}_m$, $k \in \mathbb{X}$,

$$p_{\varepsilon,\bar{p},k}(n_\varepsilon) = p_{0,\bar{p},k}(n_\varepsilon)(1 - \varepsilon)^{n_\varepsilon} + \sum_{l=1}^{n_\varepsilon} p_{0,\bar{d},k}(n_\varepsilon - l)\varepsilon(1 - \varepsilon)^{l-1}. \quad (74)$$
By applying inequality (68) to the transition probabilities appearing in the above renewal type relation, we get the following inequality, \( \bar{p} \in P_m, k \in X(g), g = 1, \ldots, h, \)

\[
|p_{\epsilon, \bar{p}, k}(n_{\epsilon}) - \pi_{0, \bar{p}, k}(1 - \epsilon)^{n_{\epsilon}} - \pi_{0, \bar{d}, k}(1 - (1 - \epsilon)^{n_{\epsilon}})| \\
= |(p_{0, \bar{p}, k}(n_{\epsilon}) - \pi_{0, \bar{p}, k})(1 - \epsilon)^{n_{\epsilon}} \\
+ \sum_{l=1}^{n_{\epsilon}} (p_{0, \bar{d}, k}(n_{\epsilon} - l) - \pi_{0, \bar{d}, k})\epsilon(1 - \epsilon)^{l-1}|
\]

\[
\leq f^{(g)}_{\bar{p}} (1 - Q(\bar{p}^{(g)}, \bar{\pi}_0^{(g)})) \Delta_N(P_{0, g})^{n_{\epsilon}/N} \\
+ f^{(g)}_{\bar{d}} (1 - Q(\bar{d}^{(g)}, \bar{\pi}_0^{(g)}))\epsilon \sum_{l=1}^{n_{\epsilon}} \Delta_N(P_{0, g})^{(n_{\epsilon} - l)/N} \\
\leq f^{(g)}_{\bar{p}} (1 - Q(\bar{p}^{(g)}, \bar{\pi}_0^{(g)})) \Delta_N(P_{0, g})^{n_{\epsilon}/N} \\
+ f^{(g)}_{\bar{d}} (1 - Q(\bar{d}^{(g)}, \bar{\pi}_0^{(g)}))\epsilon N(1 - \Delta_N(P_{0, g})^N)^{-1}.
\] (75)

Let us introduce function \( R_{\epsilon}(t) = |(1 - \epsilon)^{n_{\epsilon}} - e^{-t}|, t \in [0, \infty]. \)

If \( n_{\epsilon} \to \infty \) and \( \epsilon n_{\epsilon} \to t \in [0, \infty) \) as \( \epsilon \to 0, \) then,

\[ R_{\epsilon}(t) \to 0 \] as \( \epsilon \to 0. \) (76)

The following inequality takes place,

\[
|\pi_{0, \bar{p}, k} e^{-t} + \pi_{0, \bar{d}, k}(1 - e^{-t}) - \pi_{0, \bar{p}, k}(1 - \epsilon)^{n_{\epsilon}} - \pi_{0, \bar{d}, k}(1 - (1 - \epsilon)^{n_{\epsilon}})| \\
\leq |\pi_{0, \bar{p}, k} - \pi_{0, \bar{d}, k}| R_{\epsilon}(t) = |f^{(g)}_{\bar{p}} - f^{(g)}_{\bar{d}}| |\pi_{0, k}| R_{\epsilon}(t).
\] (77)

Relations (75)–(77) obviously imply that the following relation holds, for \( k \in X(g), k = 1, \ldots, h, \) if \( n_{\epsilon} \to \infty \) and \( \epsilon n_{\epsilon} \to t \in [0, \infty) \) as \( \epsilon \to 0, \)

\[
|p_{\epsilon, \bar{p}, k}(n_{\epsilon}) - \pi_{0, \bar{p}, k}(t)| \\
\leq f^{(g)}_{\bar{p}} (1 - Q(\bar{p}^{(g)}, \bar{\pi}_0^{(g)})) \Delta_N(P_{0, g})^{n_{\epsilon}/N} \\
+ f^{(g)}_{\bar{d}} (1 - Q(\bar{d}^{(g)}, \bar{\pi}_0^{(g)}))\epsilon N(1 - \Delta_N(P_{0, g})^N)^{-1} \\
+ |f^{(g)}_{\bar{p}} - f^{(g)}_{\bar{d}}| |\pi_{0, k}| R_{\epsilon}(t) \to 0 \] as \( \epsilon \to 0. \) (78)

This relation proves the theorem. \( \square \)

Remark 9 Inequality (78) gives, in fact, explicit upper bounds for the rate of convergence in ergodic relation given in Theorem 8. Of course, it is possible to get some simple explicit upper bounds for \( R_{\epsilon}(t) \) in terms of quantities \( \epsilon n_{\epsilon} \) and \( t. \)

It should be noted that analogous asymptotic relations for Cezaro limits of transition probabilities have been obtained in Filar et al. (2002). The asymptotic relations given in Theorem 8 relate to usual limits for transition probabilities (which are stronger than Cezaro limits). Moreover, the proof of Theorem 8 yields the explicit upper bounds for the rate of convergence in this theorem, which are indicated in Remark 9.
5.4 Concluding Comments

One of the main reasons for approximation of the Markov chain $X_{0,n}$ (with the matrix of transition probabilities $P_0$), by perturbed (regularised) Markov chains with damping component $X_{\varepsilon,n}$ (with the matrix of transition probabilities $P_{\varepsilon} = (1 - \varepsilon)P_0 + \varepsilon D$), is to use it for approximation of the stationary distribution $\pi_0 = \{\pi_{0,j}, j \in \mathbb{X}\}$ of the Markov chain $X_{0,n}$ by the stationary distribution $\pi_{\varepsilon} = \{\pi_{\varepsilon,j}, j \in \mathbb{X}\}$ of the Markov chain $X_{\varepsilon,n}$.

Since the corresponding phase space $\mathbb{X} = \{1, \ldots, m\}$ can be large, the power method can be used for approximative computing of stationary distribution $\pi_{\varepsilon}$. In this case, its components $\pi_{\varepsilon,j}$ are approximated by probabilities $p_{\varepsilon,\hat{p},j}(n) = \sum_{i \in \mathbb{X}} p_i P_{\varepsilon,ij}(n)$, where $p_{\varepsilon,ij}(n)$ are elements of the matrix $P_{\varepsilon}^n$ and $\hat{p} = (p_j, j \in \mathbb{X})$ is some initial distribution.

The results given in Theorems 1–8 show that the situation significantly differs for two models, where: (a) the phase space $\mathbb{X}$ is one class of communicative states for the Markov chain $X_{0,n}$ (condition $A_{1,1}$ holds), i.e., this Markov chain is ergodic, and, (b) the phase space $\mathbb{X}$ splits in several closed classes of communicative states for the Markov chain $X_{0,n}$ (condition $A_{h,1}$ holds, for some $h > 1$), and, thus, this Markov chain is not ergodic.

Rates of convergence for stationary probabilities $\pi_{\varepsilon,j}$ to $\pi_{0,j}$ as $\varepsilon \to 0$ and probabilities $p_{\varepsilon,\hat{p},j}(n)$ to stationary probabilities $\pi_{0,j}$, $\pi_{\varepsilon,j}$ as $n \to \infty$ play the key role in the above method. We give explicit upper bounds for rates of convergence of stationary probabilities $\pi_{\varepsilon,j}$ to $\pi_{0,j}$ and asymptotic expansions for $\pi_{\varepsilon,j}$, with respect to damping parameter $\varepsilon$, in Theorems 3 and 4. We also give explicit upper bounds for rates of convergence of probabilities $p_{\varepsilon,\hat{p},j}(n)$ to stationary probabilities $\pi_{\varepsilon,j}$, in Theorems 5 and 6.

These results let one construct the two-stage effective algorithms for approximating the stationary distribution $\pi_0$ in the case, where the phase space $\mathbb{X}$ is one class of communicative states for the Markov chain $X_{0,n}$ (condition $A_{1,1}$ holds), and, thus, this Markov chain is ergodic.

At the first stage, one approximates stationary probabilities $\pi_{\varepsilon,j}$ by probabilities $p_{\varepsilon,\hat{p},j}(n)$. The rate of this approximation has order $O(D_N(P_0)^n(1 - \varepsilon)^n)$, where $D_N(P_0)$ is some initial distribution. The effectiveness of this approximation declines for small values of damping parameter $\varepsilon$ and values of ergodicity coefficient $D_N(P_0)$ closed to 1 that is typical for models with the phase space $\mathbb{X}$ nearly decomposed in several closed classes of states.

At the second stage, one approximate stationary probabilities $\pi_{0,j}$ by stationary probabilities $\pi_{\varepsilon,j}$. The rate of approximation has order $O(\varepsilon)$, which, also, can be improved by using the approximations based on the corresponding asymptotic expansions.

Here, some dual effect takes place. At the second stage it would be better to use small values of regularisation parameter $\varepsilon$, while at the first stage using small values of $\varepsilon$ are not desirable.

Theorem 7 let one also to use the one-step variant of the algorithm described above and approximate the stationary probabilities $\pi_{0,j}$ by probabilities $p_{\varepsilon,\hat{p},j}(n_x)$, with an arbitrary positive integers $n_x \to \infty$ as $\varepsilon \to 0$. Moreover, the explicit upper bounds for $|p_{\varepsilon,\hat{p},j}(n_x) - \pi_{0,j}|$ pointed out in Remark 8 let one balance the choice of $\varepsilon$ and $n_x$.

The case, where the phase space $\mathbb{X}$ splits in several closed classes of communicative states for the Markov chain $X_{0,n}$, (condition $A_{h,1}$ holds, for some $h > 1$) and, thus, this Markov chain is not ergodic, is more complex.

As a matter of fact, in this case, (c) stationary probabilities $\pi_{0,\hat{p},k}$ for the Markov chain $X_{0,n}$ depend on the initial distribution $\hat{p}$, (d) the stationary probabilities $\pi_{\varepsilon,k}$ for the Markov chain $X_{\varepsilon,n}$ converge to the stationary probabilities $\pi_{0,\hat{d},k}$ as $\varepsilon \to 0$. 

© Springer
If the initial distribution $\bar{p} = \bar{d}$ then the two-stage algorithm as well as its one-stage variant described above can be applied.

However, if $\bar{p} \neq \bar{d}$, one should be more careful, since in this case it may be that the stationary probability $\pi_{0,\bar{d},k} \neq \pi_{0,\bar{p},k}$ and, thus, the two-stage algorithm described above does not work.

In this case, Theorem 8 answers the question about applicability probabilities $p_{\epsilon,\bar{p},j}(n_{\epsilon})$ as approximations for stationary probabilities for the Markov chain $X_{0,n}$. In fact, these probabilities converge to some mixture of stationary probabilities $\pi_{0,\bar{p},k}$ and $\pi_{0,\bar{d},k}$, namely,

$$\pi_{0,\bar{p},k}(t) = \pi_{0,\bar{p},k}e^{-t} + \pi_{0,\bar{d},k}(1 - e^{-t}),$$

as $n_{\epsilon} \to \infty$ in such way that $\epsilon n_{\epsilon} \to t \in [0, \infty]$ as $\epsilon \to 0$. Moreover, the explicit upper bounds for $|p_{\epsilon,\bar{p},k}(n_{\epsilon}) - \pi_{0,\bar{p},k}(t)|$ pointed out in Remark 9 let one balance the choice of $\epsilon$ and $n_{\epsilon}$ and, in some sense, predict the value of limit $\pi_{0,\bar{p},k}(t)$ depending on the value of quantity $\epsilon n_{\epsilon}$.

In conclusion, we would like to note that Theorems 1–8 present results of the perturbation analysis for Markov chains with damping component for the basic cases, where the phase space of the unperturbed Markov chain $X_{0,n}$, either is one aperiodic class of communicative states, or split in several closed aperiodic classes of communicative states. Despite some technical complications, analogous results can also be obtained for the cases, where the phase space of the unperturbed Markov chain $X_{0,n}$, either is one periodic class of communicative states, or split in several closed periodic classes of communicative states plus possibly a class of transient states. We are going to present such results as well as results of experimental studies supporting and illustrating the results of the present paper in future publications.

Acknowledgements This research was supported by the Swedish International Development Cooperation Agency (Sida), International Science Programme (ISP) in Mathematical Sciences (IPMS) and Sida Bilateral Research Programmes for research and education capacity development in Mathematics in Uganda and Tanzania. The authors are also grateful to the research environment Mathematics and Applied Mathematics (MAM), Division of Applied Mathematics, Mälardalen University for providing an excellent and inspiring environment for research education and research.

Funding Information Open access funding provided by Stockholm University.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article’s Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article’s Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit http://creativecommons.org/licenses/by/4.0/.

References

Abola B, Biganda PS, Engström C, Mango JM, Kakuba G, Silvestrov S (2018a) Pagerank in evolving tree graphs. In: Silvestrov S, Rančić M, Malyarenko A (eds) Stochastic processes and applications, Chapter 16. Springer proceedings in mathematics & statistics, vol 271. Springer, Cham, pp 375–390
Abola B, Biganda PS, Engström C, Mango JM, Kakuba G, Silvestrov S (2018b) Updating of PageRank in evolving tree graphs. In: Skiadas CH (ed) Proceedings of the 5th stochastic modeling techniques and data analysis international conference with demographics workshop. Chania, Crete, Greece, 2018. ISAST: international society for the advancement of science and technology, pp 15–26
Andersson FK, Silvestrov SD (2008) The mathematics of internet search engines. Acta Appl Math 104(2):211–242
Avrachenkov K, Litvak N (2004) Decomposition of the google PageRank and optimal linking strategy. Research report, INRIA, 15
Avrachenkov K, Litvak N, Pham KS (2008) A singular perturbation approach for choosing the PageRank damping factor. Internet Math 5(1–2):47–69
Avrachenkov K, Filar J, Howlett PG (2013) Analytic perturbation theory and its applications. SIAM, Philadelphia, p xii+372
Avrachenkov K, Piotnovskiy A, Zhang Y (2018) Hitting times in Markov chains with restart and their application to network centrality. Methodol Comput Appl Probab 20(4):1173–1188
Battiston S, Puliga M, Kaushik R, Tasca P, Caldarelli G (2012) Debrank: too central to fail? Financial networks, the fed and systemic risk. Scientific Reports 2:541
Biganda PS, Abola B, Engström C, Silvestrov S (2017) Pagerank, connecting a line of nodes with multiple complete graphs. In: Skiadas CH (ed) Proceedings of the 17th applied stochastic models and data analysis international conference with the 6th demographics workshop. London, UK, 2017. ISAST: international society for the advancement of science and technology, pp 113–126
Biganda PS, Abola B, Engström C, Mango JM, Kakuba G, Silvestrov S (2018a) Exploring the relationship between ordinary PageRank, lazy PageRank and random walk with backstep PageRank for different graph structures. In: Skiadas CH (ed) Proceedings of the 5th stochastic modeling techniques and data analysis international conference with demographics workshop. Chania, Crete, Greece, 2018. ISAST: international society for the advancement of science and technology, pp 71–85
Biganda PS, Abola B, Engström C, Mango JM, Kakuba G, Silvestrov S (2018b) Traditional and lazy Pageranks for a line of nodes connected with complete graphs. In: Silvestrov S, Rančić M, Malyarenko A (eds) Stochastic processes and applications, Chapter 17. Springer proceedings in mathematics & statistics, vol 271. Springer, Cham, pp 391–412
Bini DA, Latouche G, Meini B (2005) Numerical methods for structured Markov chains. Numerical mathematics and scientific computation, Oxford Science Publications. Oxford University Press, New York, p xii+327
Brin S, Page L (1998) The anatomy of a large-scale hypertextual web search engine. Comp Networks, ISDN Syst 30(1-7):107–117
Dobrushin R (1956) Central limit theorem for non-stationary Markov chains. I Teor Veroyatnost i Primenen 1(1):72–89. (English translation in Theory Probab Appl 1(1):65–80)
Doeblin W (1940) Éléments d’une théorie générale des chaînes simples constantes de markoff. (French) Ann École Norm 57(3):61–111
Englund E (2001) Nonlinearly perturbed renewal equations with applications. Doctoral dissertation, Umeå University
Englund E, Silvestrov D (1997) Mixed large deviation and ergodic theorems for regenerative processes with discrete time. In: Jagers P, Kullendorff G, Portenko N, Silvestrov D (eds) Proceedings of the second Scandinavian–Ukrainian conference in mathematical statistics, Vol. I, Umeå, 1997. Theory Stoch. Process., 3(19), no. 1–2, 164–176
Engström C (2016) Pagerank in evolving networks and applications of graphs in natural language processing and biology. Doctoral dissertation 217, Mälardalen University, Västerås
Engström C, Silvestrov S (2014) Generalisation of the damping factor in PageRank for weighted networks. In: Silvestrov D, Martin-Löf A (eds) Modern problems in insurance mathematics, chapter 19. EAA series. Springer, Cham, pp 313–334
Engström C, Silvestrov S (2016a) Pagerank, a look at small changes in a line of nodes and the complete graph. In: Silvestrov S, Rančić M (eds) Engineering mathematics II. Algebraic, stochastic and analysis structures for networks, data classification and optimization, Chapter 11. Springer Proceedings in Mathematics & Statistics, vol 179. Springer, Cham, pp 223–248
Engström C, Silvestrov S (2016b) Pagerank, connecting a line of nodes with a complete graph. In: Silvestrov S, Rančić M (eds) Engineering mathematics II. Algebraic, stochastic and analysis structures for networks, data classification and optimization, Chapter 12. Springer Proceedings in Mathematics & Statistics, vol 179. Springer, Cham, pp 249–274
Engström C, Silvestrov S (2017) Pagerank for networks, graphs, and Markov chains. Teor Ĭmovîrn Mat Stat 96:61–83. (Also, in Theor Probab Math Statist 96:59–82)
Feller W (1968) An introduction to probability theory and its applications, vol I, 3rd edn. Wiley, New York, p xviii+509
Filar J, Krieger HA, Syed Z (2002) Cesaro limits of analytically perturbed stochastic matrices. Linear Algeb Appl 353(1–3):227–243
Methodology and Computing in Applied Probability

Gambini A, Krzyanowski P, Pokarowski P (2008) Aggregation algorithms for perturbed Markov chains with applications to networks modeling. SIAM J Sci Comput 31(1):45–73

Gleich DF (2015) Pagerank beyond the Web. SIAM Rev 57(3):321–363

Griffeath D (1975) A maximal coupling for Markov chains. Z Wahrsch Verw Gebiete 31:95–106

Gyllenberg M, Silvestrov D (2000) Nonlinearly perturbed regenerative processes and pseudo-stationary phenomena for stochastic systems. Stoch Process Appl 86:1–27

Gyllenberg M, Silvestrov D (2008) Quasi-stationary phenomena in nonlinearly perturbed stochastic systems. de gruyter expositions in mathematics, 44. Walter de Gruyter, Berlin, p ix+x+579

Hartfiel DJ, Meyer CD (1998) On the structure of stochastic matrices with a subdominant eigenvalue near 1. Linear Algebra Appl 272(1-3):193–203

Haveliwala T, Kamvar S (2003) The second eigenvalue of the google matrix. Technical report 2003-20. Stanford University

Kalashnikov VV (1989) Coupling method, development and applications. Supplement to the Russian edition of the book by Nummelin, E. (1984). General Irreducible Markov Chains and Non-negative Operators. Cambridge University Press (Russian edition, MIR, Moscow, 1989, 176–190)

Kamvar SD, Schlosser MT, Garcia-Molina H (2003) The eigentrust algorithm for reputation management in p2p networks. In: Proceedings of the 12th international conference on world wide web. ACM, pp 640–651

Konstantinov M, Gu DW, Mehrmann V, Petkov P (2003) Perturbation theory for matrix equations. Studies in computational mathematics, 9, North-Holland, Amsterdam, pp xii+429

Koroliuk VS, Limnios N (2005) Stochastic systems in merging phase space. World Scientific, Singapore, p xv+331

Korolyuk VS, Korolyuk VV (1999) Stochastic models of systems. Mathematics and its applications, 469. Kluwer, Dordrecht, p xii+185

Langville AN, Meyer CD (2004) Deeper inside PageRank. Internet Math 1(3):335–338

Langville AN, Meyer CD (2011) Google’s pagerank and beyond: the science of search engine rankings. Princeton University Press, Princeton, p x+224

Lindvall T (2002) Lectures on the coupling method. Wiley series in probability and mathematical statistics: probability and mathematical statistics. Wiley, New York, p xiv+257. (A revised reprint of the 1992 original)

Loève M (1978) Probability theory II. Fourth edition. Graduate texts in mathematics, springer, xvi+416pp. (Extended version of the first (1955) edition)

Markov AA (1906) Generalisation of the law of large numbers on dependent trials. Izv Kazan Fiz-Mat Obsch 15(4):135–156

Mitrophanov AY (2005) Sensitivity and convergence of uniformly ergodic Markov chains. J Appl Prob 42:1003–1014

Ni Y (2011) Nonlinearly perturbed renewal equations: asymptotic results and applications. Doctoral dissertation 106, Mälardalen University, Västerås

Ni Y, Silvestrov D, Malyarenko A (2008) Exponential asymptotics for nonlinearly perturbed renewal equation with non-polynomial perturbations. J Numer Appl Math 1(96):173–197

Petersson M (2016) Perturbed discrete time stochastic models. Doctoral dissertation. Stockholm University

Peterson M (1979) On coupling of Markov chains. Z Wahrsch Verw Gebiete 35:315–322

Silvestrov D (1978) The renewal theorem in a series scheme. 1. Teor Veroyatn Mat Stat 18:144–161. (English translation in Theory Probab Math Statist 18:155–172)

Silvestrov D (1979) The renewal theorem in a series scheme 2. Teor Veroyatn Mat Stat 20:97–116. (English translation in Theory Probab Math Statist 20:113–130)

Silvestrov D (1983) Method of a single probability space in ergodic theorems for regenerative processes 1. Math Operat Statist, Ser Optim 14:285–299

Silvestrov D (1984a) Method of a single probability space in ergodic theorems for regenerative processes 2. Math Operat Statist, Ser Optim 15:601–612

Silvestrov D (1984b) Method of a single probability space in ergodic theorems for regenerative processes 3. Math Operat Statist, Ser Optim 15:613–622

Silvestrov D (1994) Coupling for Markov renewal processes and the rate of convergence in ergodic theorems for processes with semi-Markov switchings. Acta Appl Math 34:109–124

Silvestrov D (2018) Individual ergodic theorems for perturbed alternating regenerative processes. In: Silvestrov S, Rančić M, Malyarenko A (eds) Stochastic processes and applications, Chapter 3. Springer Proceedings in Mathematics & Statistics, vol 271. Springer, Cham, pp 23–89

Silvestrov D, Petersson M (2014) Exponential expansions for perturbed discrete time renewal equations. In: Karagrigoriou A, Lisiński A, Kleyner A, Frenkel I (eds) Applied reliability engineering and risk analysis, Chapter 23. Probabilistic models and statistical inference. Wiley, New York, pp 349–362
Methodology and Computing in Applied Probability

Silvestrov D, Petersson M, Hössjer O (2018) Nonlinearly perturbed birth-death-type models. In: Silvestrov S, Rančić M, Malyarenko A (eds) Stochastic processes and applications, Chapter 3. Springer proceedings in mathematics & statistics, vol 271. Springer, Cham, pp 189–244

Silvestrov D, Pezinska G (1985) On maximally coinciding random variables. Teor Veroyatn Mat Stat 32:102–105. (English translation in Theory Probab Math Statist 32:113–115)

Silvestrov D, Silvestrov S (2016) Asymptotic expansions for stationary distributions of perturbed semi-Markov processes. In: Silvestrov S, Rančić M (eds) Engineering mathematics II. Algebraic, stochastic and analysis structures for networks, data classification and optimization, Chapter 10. Springer Proceedings in Mathematics & Statistics, vol 179. Springer, Cham, pp 151–222

Silvestrov D, Silvestrov S (2017a) Nonlinearly perturbed semi-Markov processes. Springer briefs in probability and mathematical statistics. Springer, Cham, p xiv+143

Silvestrov D, Silvestrov S (2017b) Asymptotic expansions for power-exponential moments of hitting times for nonlinearly perturbed semi-Markov processes. Teor Īmovır Mat Stat 97:171–187. (Also, in Theory Probab Math Statist 97:183–200)

Silvestrov D, Silvestrov S (2019a) Asymptotic expansions for stationary distributions of nonlinearly perturbed semi-Markov processes. 1. Methodol Comput Appl Probab 21(3):945–964. (First Online 2017)

Silvestrov D, Silvestrov S (2019b) Asymptotic expansions for stationary distributions of nonlinearly perturbed semi-Markov processes. 2. Methodol Comput Appl Probab 21(3):965–984. (First Online 2017)

Stewart GW (1998) Matrix algorithms. Vol. I. Basic decompositions. SIAM, Philadelphia, p xx+458

Stewart GW (2001) Matrix algorithms. Vol. II. Eigensystems. SIAM, Philadelphia, p xx+469

Stewart WJ (1994) Introduction to the numerical solution of markov chains. Princeton University Press, Princeton, p xx+539

Sun Y, Han J (2013) Mining heterogeneous information networks: a structural analysis approach. ACM SIGKDD Explor Newslet 14(2):20–28

Thorisson H (2000) Coupling, stationarity and regeneration. Probability and its applications. Springer, New York, p ivx+517

Yin GG, Zhang Q (2005) Discrete-time markov chains. Two-time-scale methods and applications. Stochastic modelling and applied probability, 55. Springer, New York, p xix+348

Yin GG, Zhang Q (2013) Continuous-time markov chains and applications. a two-time-scale approach. Second edition, stochastic modelling and applied probability, 37. Springer, New York, p xxii+427. (An extended variant of the first (1998) edition)

Publisher’s Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.