Chapter 11
Nonlinearly Perturbed Birth-Death-Type Models



Dmitrii Silvestrov, Mikael Petersson and Ola Hössjer

Abstract Asymptotic expansions are presented for stationary and conditional quasi-stationary distributions of nonlinearly perturbed birth-death-type semi-Markov models, as well as algorithms for computing the coefficients of these expansions. Three types of applications are discussed in detail. The first is a model of population growth, where either an isolated population is perturbed by immigration, or a sink population with immigration is perturbed by internal births. The second application is epidemic spread of disease, in which a closed population is perturbed by infected individuals from outside. The third model captures the time dynamics of the genetic composition of a population with genetic drift and selection, that is perturbed by various mutation scenarios.

Keywords Semi-Markov birth-death process · Quasi-stationary distribution Nonlinear perturbation · Population dynamics model · Population genetics model Epidemic model

11.1 Introduction

Models of perturbed Markov chains and semi-Markov processes attracted attention of researchers in the mid of the 20th century, in particular the most difficult cases of perturbed processes with absorption and so-called singularly perturbed processes. An interest in these models has been stimulated by applications to control, queuing systems, information networks, and various types of biological systems. As a rule,

D. Silvestrov (✉) · O. Hössjer
Department of Mathematics, Stockholm University, 106 91 Stockholm, Sweden
e-mail: silvestrov@math.su.se

O. Hössjer
e-mail: ola@math.su.se

M. Petersson
Statistics Sweden, Stockholm, Sweden
e-mail: mikpe274@gmail.com

© Springer Nature Switzerland AG 2018

S. Silvestrov et al. (eds.), *Stochastic Processes and Applications*,
Springer Proceedings in Mathematics & Statistics 271,
https://doi.org/10.1007/978-3-030-02825-1_11

Markov-type processes with singular perturbations appear as natural tools for mathematical analysis of multi-component systems with weakly interacting components.

In this paper, we present new algorithms for construction of asymptotic expansions for stationary and conditional quasi-stationary distributions of nonlinearly perturbed semi-Markov birth-death processes with a finite phase space. We consider models that include a positive perturbation parameter that tends to zero as the unperturbed null model is approached. It is assumed that the phase space is one class of communicative states, for the embedded Markov chains of pre-limiting perturbed semi-Markov birth-death processes, whereas the limiting unperturbed model either consists of one closed class of communicative states, or of one class of communicative transient internal states that has one or both end points as absorbing states.

These new algorithms are applied to several perturbed birth-death models of biological nature. The first application is population size dynamics in a constant environment with a finite carrying capacity. It is assumed that one individual at a time is born, immigrates or dies, see, for instance, Lande, Engen and Saether [26]. In order to study the impact of immigration or births, it is possible to either view the immigration rate as a perturbation parameter of an isolated population, or the birth rate as a perturbation parameter of a sink population in which no individuals are born. The first analysis depends heavily on the ratio between the birth and death rates for the null model, whereas the second analysis involves the corresponding ratio of the immigration and death rates.

The second application is epidemic spread of a disease, reviewed, for instance, in Hethcote [14] and Nåsell [34]. Here one individual at a time gets infected or recovers, and recovered individuals become susceptible for new infections. We perturb an isolated population with no immigration, by including the possibility of occasional infected immigrants to arrive, and obtain a special case of the population dynamics model with occasional immigration.

The third application is population genetic models, treated extensively in Crow and Kimura [7] and Ewens [9]. We focus in particular on models with overlapping generations, introduced by Moran [27]. These Moran type models describe the time dynamics of the genetic composition of a population, represented as the frequency distribution of two variants of a certain gene. It is assumed that one copy of the gene is replaced for one individual at a time, and the model includes genetic drift, mutation, and various types of selection. The mutation rates between the two variants are perturbed, and the analysis depends heavily on the mutation rates and selection scheme of the unperturbed model.

The general setting of perturbed semi-Markov birth-death processes used in the paper can be motivated as follows: First, it makes it possible to consider models where inter-event times have more general non-geometric/non-exponential distributions. Second, the semi-Markov setting is a necessary element of the proposed method of sequential phase space reduction, which yields effective recurrent algorithms for computing asymptotic expansions. Third, the proposed method has a universal character. We are quite sure that it can be applied to more general models, for example, to meta-population models with several sub-populations possessing birth-death-type dynamics.

In this paper, we present asymptotic expansions of the second order and give explicit formulas for the coefficients of these expansions. The coefficients of such asymptotic expansions have a clear meaning. The first coefficients describe the asymptotic behaviour of stationary and quasi-stationary probabilities and their continuity properties with respect to small perturbations of transition characteristics of the corresponding semi-Markov birth-death processes. The second coefficients determine sensitivity of stationary and quasi-stationary probabilities with respect to small perturbations of transition characteristics.

However, it is worth to note that the proposed method can also be used for constructions of asymptotic expansions of higher orders, which also can be useful and improve accuracy of the numerical computations based on the corresponding asymptotic expansions, especially, for the models, where actual values of the perturbation parameter are not small enough to neglect the high order terms in the corresponding asymptotic expansions.

We refer here to the book by Gyllenberg and Silvestrov [13], where one can find results on asymptotic expansions for stationary and quasi-stationary distributions for perturbed semi-Markov processes, that created the background for our studies. Other recent books containing results on asymptotic expansions for perturbed Markov chains and semi-Markov processes are Korolyuk, V.S. and Korolyuk, V.V. [23], Stewart [42, 43], Konstantinov, Gu, Mehrmann and Petkov [22], Bini, Latouche and Meini [4], Koroliuk and Limnios [24], Yin and Zhang [48, 49], Avrachenkov, Filar and Howlett [3], and Silvestrov, D. and Silvestrov, S. [41]. Readers can find comprehensive bibliographies of this research area in the above books, the papers by Silvestrov, D. and Silvestrov, S. [39], Petersson [36], the doctoral dissertation of Petersson [37], and book Silvestrov, D. and Silvestrov, S. [41].

The paper includes 8 sections. In Sect. 11.2, we give examples of perturbed population dynamics, epidemic and population genetic models, which can be described in the framework of birth-death-type Markov chains and semi-Markov processes. In Sect. 11.3, we introduce a more general model of perturbed semi-Markov birth-death processes, define stationary and conditional quasi-stationary distributions for such processes and formulate basic perturbation conditions. In Sect. 11.4, we illustrate this framework for the biological models of Sect. 11.2. In Sect. 11.5, we present time-space screening procedures of phase space reduction for perturbed semi-Markov processes and recurrent algorithms for computing expectations of hitting times and stationary and conditional quasi-stationary distributions for semi-Markov birth-death processes. In Sect. 11.6, we describe algorithms for construction of the second order asymptotic expansions for stationary and conditional quasi-stationary distributions of perturbed semi-Markov birth-death processes. In Sect. 11.7, we apply the above asymptotic results to the perturbed birth-death models of biological nature defined in Sect. 11.2, and present results of related numerical studies. In Sect. 11.8, we give concluding remarks and comments.

11.2 Examples of Perturbed Birth-Death Processes

In this section, we consider a number of examples of perturbed birth-death processes that represent the time dynamics of a biological system, such as size variations of a population with a finite carrying capacity, the spread of an epidemic, or changes of the genetic composition of a population.

We let $\eta^{(\varepsilon)}(t) \in \mathbb{X} = \{0, \ldots, N\}$ denote the value of the process at time $t \geq 0$, with N a fixed (and typically large) positive integer that corresponds to the size or maximal size of the population. The perturbation parameter $\varepsilon \in (0, \varepsilon_0]$ is typically small. It either represents an immigration rate for an almost isolated population, or the mutation rate of a population in which several genetic variants segregate.

We assume that $\eta^{(\varepsilon)}(t)$ is a piecewise constant and right-continuous semi-Markov process, with discontinuities at time points

$$\zeta_n^{(\varepsilon)} = \kappa_1^{(\varepsilon)} + \cdots + \kappa_n^{(\varepsilon)}, \quad n = 0, 1, \ldots. \tag{11.1}$$

At inner points ($0 < \eta^{(\varepsilon)}(t) < N$) the process changes by one unit up or down. This either corresponds to birth or death of one individual, recovery or infection of one individual, or a change of the population's genetic decomposition. At boundary points ($\eta^{(\varepsilon)}(t) \in \{0, N\}$), any jump out of the state space is projected back to $\mathbb{X}$, so that, for instance, a "jump" from 0 ends at 0 or 1.

The time $\kappa_n^{(\varepsilon)}$ between the n:th and $(n+1)$:th jumps of $\eta^{(\varepsilon)}(t)$ will be referred to as the n:th transition time. Its distribution function

$$F_i^{(\varepsilon)}(t) = \mathsf{P}\{\kappa_n^{(\varepsilon)} \leq t / \eta^{(\varepsilon)}(\zeta_{n-1}^{(\varepsilon)}) = i\} \tag{11.2}$$

only depends on the state $i \in \mathbb{X}$ from which a jump occurs.

In this section, we consider two examples of transition time distributions (11.2). The first one is geometric,

$$F_i^{(\varepsilon)} \sim \mathrm{Ge}\,[\lambda_i(\varepsilon)] \Longrightarrow F_i^{(\varepsilon)}(t) = 1 - [1 - \lambda_i(\varepsilon)]^{[t]}, \tag{11.3}$$

where $[t]$ is the integer part of t and $0 < \lambda_i(\varepsilon) \leq 1$ represents the probability that a jump occurs in one time step. The second example corresponds to a continuous time Markov process, with an exponential transition time distribution

$$F_i^{(\varepsilon)} \sim \mathrm{Exp}\,[\lambda_i(\varepsilon)] \Longrightarrow F_i^{(\varepsilon)}(t) = 1 - e^{-\lambda_i(\varepsilon)t}, \tag{11.4}$$

with $0 < \lambda_i(\varepsilon) < \infty$ the rate at which a jump occurs. It is convenient to decompose

$$\lambda_i(\varepsilon) = \lambda_{i,-}(\varepsilon) + \lambda_{i,+}(\varepsilon) \tag{11.5}$$

as a sum of two terms, where $\lambda_{i,-}(\varepsilon)$ represents the probability of death in one time step in (11.3), or the rate at which a death occurs in (11.4) ($i \to i - 1$ when $i > 0$,

$0 \to 0$ when $i = 0$). Similarly, $\lambda_{i,+}(\varepsilon)$ is the probability or rate of a birth event ($i \to i+1$ when $i < N$, $N \to N$ when $i = N$). For both models (11.3) or (11.4), $\eta_n^{(\varepsilon)} = \eta^{(\varepsilon)}(\zeta_n^{(\varepsilon)}), \quad n = 0, 1, 2, \ldots$ is an embedded discrete time Markov chain, with transition probabilities

$$p_{i,+}(\varepsilon) = 1 - p_{i,-}(\varepsilon) = \frac{\lambda_{i,+}(\varepsilon)}{\lambda_i(\varepsilon)} \tag{11.6}$$

of jumping upwards or downwards.

It is assumed that $\mathbb{X}$ is one single class of communicative states for each $\varepsilon > 0$. The behaviour of the limiting $\varepsilon = 0$ model will satisfy one of the following three conditions:

$$\begin{aligned}
&\mathbf{H}_1 : \text{The } \varepsilon = 0 \text{ model has one class } \mathbb{X} \text{ of communicative states,} \\
&\mathbf{H}_2 : \text{The } \varepsilon = 0 \text{ model has one absorbing state 0 and one class} \\
&\qquad {}_0\mathbb{X} = \mathbb{X} \setminus \{0\} \text{ of communicative transient states,} \\
&\mathbf{H}_3 : \text{The } \varepsilon = 0 \text{ model has two absorbing states 0 and } N, \text{ and one} \\
&\qquad \text{class } {}_{0,N}\mathbb{X} = \mathbb{X} \setminus \{0, N\} \text{ of communicative transient states.}
\end{aligned} \tag{11.7}$$

These three perturbation scenarios can be rephrased in terms of the birth and death rates (11.5) as follows:

$$\begin{aligned}
&\mathbf{H}_1 : \lambda_{0,+}(0) > 0, \lambda_{N,-}(0) > 0, \\
&\mathbf{H}_2 : \lambda_{0,+}(0) = 0, \lambda_{N,-}(0) > 0, \\
&\mathbf{H}_3 : \lambda_{0,+}(0) = 0, \lambda_{N,-}(0) = 0.
\end{aligned} \tag{11.8}$$

This will be utilised in Sects. 11.2.1–11.2.3 in order to characterise the various perturbed models that we propose.

11.2.1 Perturbed Population Dynamics Models

Let N denote the maximal size of a population, and let $\eta^{(\varepsilon)}(t)$ be its size at time t. In order to model the dynamics of the population, we introduce births, deaths, and immigration from outside, according to a parametric model with

$$\lambda_{i,+}(\varepsilon) = \lambda i \left[1 - \alpha_1 \left(\frac{i}{N}\right)^{\theta_1}\right] + \nu \left[1 - \left(\frac{i}{N}\right)^{\theta_2}\right] \tag{11.9}$$

and

$$\lambda_{i,-}(\varepsilon) = \mu i \left[1 + \alpha_2 \left(\frac{i}{N}\right)^{\theta_3}\right]. \tag{11.10}$$

For a small population ($i \ll N$), we interpret the three parameters $\lambda > 0$, $\mu > 0$ and $\nu > 0$ as a birth rate per individual, a death rate per individual, and an immigration rate, whereas α_k, θ_k are density regulation parameters that model decreased birth/immigration and increased death for a population close to its maximal size. They satisfy $\theta_k > 0$, $\alpha_1 \leq 1$, and $\alpha_1, \alpha_2 \geq 0$, where the last inequality is strict for at least one of α_1 and α_2. A more general model would allow birth, death, and immigration rates to vary non-parametrically with i.

The expected growth rate of the population, when $0 < i < N$, is

$$\begin{aligned}
&\mathsf{E}\left[\eta^{(\varepsilon)}(t+\Delta t) - \eta^{(\varepsilon)}(t) | \eta^{(\varepsilon)}(t) = i\right] = \Delta t \left[\lambda_{i,+}(\varepsilon) - \lambda_{i,-}(\varepsilon)\right] \\
&\quad = \Delta t \left\{\lambda i \left[1 - \alpha_1 \left(\tfrac{i}{N}\right)^{\theta_1}\right] + \nu \left[1 - \left(\tfrac{i}{N}\right)^{\theta_2}\right] - \mu i \left[1 + \alpha_2 \left(\tfrac{i}{N}\right)^{\theta_3}\right]\right\},
\end{aligned}$$

where $\Delta t = 1$ in discrete time (11.3), and $\Delta t > 0$ is infinitesimal in continuous time (11.4). When $\theta_1 = \theta_2 = \theta_3 = \theta$, this expression simplifies to

$$\begin{aligned}
&\mathsf{E}\left[\eta^{(\varepsilon)}(t+\Delta t) - \eta^{(\varepsilon)}(t) | \eta^{(\varepsilon)}(t) = i\right] \\
&\quad = \Delta t \left\{\lambda i \left[1 - \alpha_1 \left(\tfrac{i}{N}\right)^{\theta}\right] - \mu i \left[1 + \alpha_2 \left(\tfrac{i}{N}\right)^{\theta}\right]\right\} \\
&\quad + \nu \left[1 - \left(\tfrac{i}{N}\right)^{\theta}\right] \\
&\quad = \Delta t \cdot \begin{cases}
(\lambda - \mu) i \left[1 - \frac{\alpha_1 \lambda + \alpha_2 \mu}{\lambda - \mu} \left(\frac{i}{N}\right)^{\theta}\right] & \\
+ \nu \left[1 - \left(\frac{i}{N}\right)^{\theta}\right], & \text{if } \lambda \neq \mu \\
-\mu i (\alpha_1 + \alpha_2) \left(\frac{i}{N}\right)^{\theta} + \nu \left[1 - \left(\frac{i}{N}\right)^{\theta}\right], & \text{if } \lambda = \mu.
\end{cases}
\end{aligned} \tag{11.11}$$

We shall consider two perturbation scenarios. The first one has

$$\mathbf{H}_2 : \quad \nu = \nu(\varepsilon) = \varepsilon, \tag{11.12}$$

whereas all other parameters are kept fixed, not depending on ε. It is also possible to consider more general nonlinear functions $\nu(\varepsilon)$, but this will hardly add more insight to how immigration affects population dynamics.

The unperturbed $\varepsilon = 0$ model corresponds to an isolated population that only increases through birth events. For small ε, we can think of a population that resides on an island and faces subsequent extinction and recolonisation events. After the population temporarily dies out, the island occasionally receives new immigrants at rate or probability ε. We shall find in Sect. 11.4.1 that for small migration rates ε, the properties of the model are highly dependent on whether the basic reproduction number

$$R_0 = \frac{\lambda}{\mu} \tag{11.13}$$

exceeds 1 or not.

A second perturbation scenario has a birth rate

$$\mathbf{H}_1: \quad \lambda = \lambda(\varepsilon) = \varepsilon \tag{11.14}$$

that equals ε, whereas all other parameters are kept fixed, not depending on ε. Again, more general nonlinear functions $\lambda(\varepsilon)$ can be studied, but for simplicity assume that (11.14) holds. The unperturbed $\varepsilon = 0$ model corresponds to a sink population that only increases through immigration, and its properties depend heavily on ν/μ.

11.2.2 Perturbed Epidemic Models

In order to model an epidemic in a population of size N, we let $\eta^{(\varepsilon)}(t)$ refer to the number of infected individuals at time t, whereas the remaining $N - \eta^{(\varepsilon)}(t)$ are susceptible. We assume that

$$\lambda_{i,+}(\varepsilon) = \lambda i \left(1 - \frac{i}{N}\right) + \nu(N - i), \tag{11.15}$$

and

$$\lambda_{i,-}(\varepsilon) = \mu i, \tag{11.16}$$

where the first parameter $\lambda(N-1)/N \approx \langle$ is the total contact rate between each individual and the other members of the population. The first term on the right hand side of (11.15) may be written as the product of the force of infection $\lambda i/N$ caused by i infected individuals, and the number of susceptibles $N - i$. The second parameter of the model, ν, is the contact rate between each individual and the group of infected ones outside of the population. The third parameter μ is the recovery rate per individual. It may also include a combined death and birth of an infected and susceptible individual. The model in (11.15)–(11.16) is an SIS-epidemic, since infected individuals become susceptible after recovery. It is essentially a special case of (11.9)–(11.10), with $\theta_1 = \theta_2 = \theta_3 = 1, \alpha_1 = 1$ and $\alpha_2 = 0$, although immigration is parameterised differently in (11.9) and (11.15).

Assume that the external contact rate

$$\mathbf{H}_2: \quad \nu = \nu(\varepsilon) = \varepsilon \tag{11.17}$$

equals the perturbation parameter, whereas all other parameters are kept fixed, not depending on ε. The unperturbed $\varepsilon = 0$ model refers to an isolated population without external contagion. The epidemic will then, sooner or later, die out and reach the only absorbing state 0.

Weiss and Dishon [46] first formulated the SIS-model as a continuous time birth-death Markov process (11.4) without immigration ($\varepsilon = 0$). It has since then been extended in a number of directions, see, for instance, Cavender [5], Kryscio and

Lefévre [25], Jacquez and O'Neill [18], Jacquez and Simon [19], Nåsell [29, 30] and Allen and Burgin [2]. The quasi-stationary distribution of $\eta^{(\varepsilon)}(t)$ is studied in several of these papers. In this work, we generalise previously studied models of epidemic spread by treating discrete and continuous time in a unified manner through semi-Markov processes.

The expected growth rate of the null model $\varepsilon = 0$ satisfies

$$\mathsf{E}\left[\eta^{(0)}(t+\Delta t) - \eta^{(0)}(t) | \eta^{(0)}(t) = i\right] = \Delta t \cdot ri\left(1 - \frac{i}{K(0)}\right), \tag{11.18}$$

if $0 < i < N$, when the basic reproduction ratio $R_0 = \lambda/\mu$ exceeds 1. This implies that the expected number of infected individuals follows Verhulst's logistic growth model (Verhulst [45]), with intrinsic growth rate $r = \mu(R_0 - 1)$, and a carrying capacity $K(0) = N(1 - R_0^{-1})$ of the environment.

11.2.3 Perturbed Models of Population Genetics

Let N be a positive even integer, and consider a one-sex population with $N/2$ individuals, each one of which carries two copies of a certain gene. This gene exists in two variants (or alleles); A_1 and A_2. Let $\eta^{(\varepsilon)}(t)$ be the number of gene copies with allele A_1 at time t. Consequently, the remaining $N - \eta^{(\varepsilon)}(t)$ gene copies have the other allele A_2 at time t. At each moment $\zeta_n^{(\varepsilon)}$ of jump in (11.1), a new gene copy replaces an existing one, so that

$$\eta^{(\varepsilon)}(\zeta_n^{(\varepsilon)}) = \begin{cases} \eta^{(\varepsilon)}(\zeta_n^{(\varepsilon)}-) + 1, & \text{if } A_1 \text{replaces } A_2, \\ \eta^{(\varepsilon)}(\zeta_n^{(\varepsilon)}-), & \text{if } A_k \text{replaces } A_k, \\ \eta^{(\varepsilon)}(\zeta_n^{(\varepsilon)}-) - 1, & \text{if } A_2 \text{replaces } A_1. \end{cases} \tag{11.19}$$

In discrete time (11.3), we define $\lambda_{ij}(\varepsilon)$ as the probability that the number of A_1 alleles changes from i to j when a gene copy is replaced, at each time step. In continuous time (11.4), we let $\lambda_{ij}(\varepsilon)$ be the rate at which the number of A_1 alleles changes from i to j when a gene copy replacement occurs. Let x^{**} refer to the probability that the new gene copy has variant A_1 when the fraction of A_1-alleles before replacement is $x = i/N$. We further assume that the removed gene copy is chosen randomly among all N gene copies, with equal probabilities $1/N$, so that

$$\lambda_{ij}(\varepsilon) = \begin{cases} x^{**}(1-x), & j = i+1, \\ (1-x^{**})x, & j = i-1, \\ 1 - x^{**}(1-x) - (1-x^{**})x, & j = i. \end{cases} \tag{11.20}$$

Notice that in order to make $\eta^{(\varepsilon)}(t)$ a semi-Markov process of birth-death type that satisfies (11.6), we do not regard instances when the new gene copy replaces a gene copy with the same allele as a moment of jump, if the current number i of A_1 alleles

satisfies $0 < i < N$. That is, the second line on the right hand side of (11.19) is only possible in a homogeneous population where all gene copies have the same allele A_1 or A_2, and therefore $\lambda_{ii}(\varepsilon)$ is not included in the probability or rate $\lambda_i(\varepsilon)$ to leave state i in (11.5), when $0 < i < N$.

The choice of x^{**} will determine the properties of the model. The new gene copy is formed in two steps. In the first step, a pair of genes is drawn randomly with replacement, so that its genotype is A_1A_1, A_1A_2 and A_2A_2 with probabilities x^2, $2x(1-x)$ and $(1-x)^2$ respectively. Since the gene pair is drawn with replacement, this corresponds to a probability $2/N$ that the two genes originate from the same individual (self fertilisation). A gene pair survives with probabilities proportional to $1+s_1$, 1 and $1+s_2$ for these three genotypes, where $1+s_1 \geq 0$ and $1+s_2 \geq 0$ determine the fitnesses of genotypes A_1A_1 and A_2A_2 relative to that of genotype A_1A_2. This is repeated until a surviving gene pair appears, from which a gene copy is picked randomly. Consequently, the probability is

$$x^* = \frac{1 \cdot (1+s_1)x^2 + \frac{1}{2} \cdot 2x(1-x)}{(1+s_1)x^2 + 2x(1-x) + (1+s_2)(1-x)^2} \tag{11.21}$$

that the chosen allele is A_1. In the second step, before the newly formed gene copy is put into the population, an A_1 allele mutates with probability $u_1 = \mathsf{P}(A_1 \to A_2)$, and an A_2 allele with probability $u_2 = \mathsf{P}(A_2 \to A_1)$. This implies that

$$x^{**} = (1-u_1)x^* + u_2(1-x^*). \tag{11.22}$$

By inserting (11.22) into (11.20), and (11.20) into (11.6) we get a semi-Markov process of Moran type that describes the time dynamics of two alleles in a one-sex population in the presence of selection and mutation. A special case of it was originally introduced by Moran [27], and some of its properties can be found, for instance, in Karlin and McGregor [20] and Durrett [8]. The model incorporates a number of different selection scenarios. A selectively neutral model corresponds to all three genotypes having the same fitness ($s_1 = s_2 = 0$), for directional selection, one of the two alleles is more fit than the other ($s_1 < 0 < s_2$ or $s_1 > 0 > s_2$), an underdominant model has a heterozygous genotype A_1A_2 with smaller fitness than the two homozygous genotypes A_1A_1 and A_2A_2 ($s_1, s_2 > 0$), whereas overdominance or balancing selection means that the heterozygous genotype is the one with highest fitness ($s_1, s_2 < 0$).

In continuous time (11.4), the expected value of the Moran model satisfies a differential equation

$$\begin{aligned}
&\mathsf{E}\left[\eta^{(\varepsilon)}(t+\Delta t)-\eta^{(\varepsilon)}(t)|\eta^{(\varepsilon)}(t)=Nx\right]=\Delta t\left[\lambda_{i,+}(\varepsilon)-\lambda_{i,-}(\varepsilon)\right]\\
&\quad=\Delta t\left[x^{**}(1-x)-x(1-x^{**})\right]\\
&\quad=\Delta t\,(x^{**}-x)\\
&\quad=\Delta t\left[(1-u_1-u_2)x^{*}+u_2-x\right]\\
&\quad=\Delta t\left[(1-u_1-u_2)\tfrac{x+s_1x^2}{1+s_1x^2+s_2(1-x)^2}+u_2-x\right]\\
&\quad=:\Delta t\left[N^{-1}m(x)+o(N^{-1})\right],
\end{aligned}\tag{11.23}$$

whenever $0<x<1$, with $\Delta t>0$ infinitesimal. The discrete time Moran model (11.3) also satisfies (11.23), interpreted as a difference equation, with $\Delta t=1$. In the last step of (11.23), we assumed that all mutation and selection parameters are inversely proportional to population size;

$$\begin{aligned}
u_1&=U_1/N,\\
u_2&=U_2/N,\\
s_1&=S_1/N,\\
s_2&=S_2/N,
\end{aligned}\tag{11.24}$$

and introduced an infinitesimal drift function

$$m(x)=U_2(1-x)-U_1x+[(S_1+S_2)x-S_2]\,x(1-x).$$

The corresponding infinitesimal variance function $v(x)=2x(1-x)$ follows similarly from (11.24), according to

$$\begin{aligned}
&\mathsf{V}\left[\eta^{(\varepsilon)}(t+\Delta t)|\eta^{(\varepsilon)}(t)=Nx\right]=\Delta t\left[\lambda_{i,+}(\varepsilon)+\lambda_{i,-}(\varepsilon)+O(N^{-1})\right]\\
&\quad=\Delta t\left[x^{**}(1-x)+x(1-x^{**})+O(N^{-1})\right]\\
&\quad=\Delta t\left[2x(1-x)+O(u_1+u_2+|s_1|+|s_2|)+O(N^{-1})\right]\\
&\quad=:\Delta t\left[v(x)+O(N^{-1})\right].
\end{aligned}\tag{11.25}$$

Assume that N is fixed, whereas the perturbation parameter ε varies. We let the two selection parameters s_1 and s_2, and hence also the rescaled selection parameters S_1 and S_2, be independent of ε, whereas the rescaled mutation parameters satisfy

$$\begin{aligned}
U_1&=U_1(\varepsilon)=C_1+D_1\varepsilon,\\
U_2&=U_2(\varepsilon)=C_2+D_2\varepsilon,
\end{aligned}\tag{11.26}$$

for some non-negative constants C_1, D_1, C_2, D_2, where at least one of D_1 and D_2 is strictly positive. It follows from (11.8) and (11.20) that the values of $0\le C_1, C_2<1$ will determine the properties of the unperturbed $\varepsilon=0$ model, according to the three distinct scenarios

$$\begin{aligned}
&\mathbf{H}_1: C_1>0, C_2>0,\\
&\mathbf{H}_2: C_1>0, C_2=0,\\
&\mathbf{H}_3: C_1=0, C_2=0.
\end{aligned}\tag{11.27}$$

The null model $\varepsilon = 0$ incorporates two-way mutations $A_1 \to A_2$ and $A_2 \to A_1$ for Perturbation scenario $\mathbf{H}_1$, with no absorbing state, it has one-way mutations $A_1 \to A_2$ for Perturbation scenario $\mathbf{H}_2$, with $i = 0$ as absorbing state, and no mutations for Perturbation scenario $\mathbf{H}_3$, with $i = 0$ and $i = N$ as the two absorbing states.

11.3 Nonlinearly Perturbed Semi-Markov Birth-Death Processes

In this section, we will generalise the framework of Sect. 11.2 and introduce a model of perturbed semi-Markov birth-death processes, define stationary and conditional quasi-stationary distributions for such processes and formulate basic perturbation conditions.

11.3.1 Perturbed Semi-Markov Birth-Death Processes

Let $(\eta_n^{(\varepsilon)}, \kappa_n^{(\varepsilon)})$, $n = 0, 1, \ldots$ be, for every value of a perturbation parameter $\varepsilon \in (0, \varepsilon_0]$, where $0 < \varepsilon_0 \le 1$, a Markov renewal process, i.e., a homogeneous Markov chain with the phase space $\mathbb{X} \times [0, \infty)$, where $\mathbb{X} = \{0, 1, \ldots, N\}$, an initial distribution $\bar{p}^{(\varepsilon)} = \langle p_i^{(\varepsilon)} = \mathsf{P}\{\eta_0^{(\varepsilon)} = i, \kappa_0^{(\varepsilon)} = 0\} = \mathsf{P}\{\eta_0^{(\varepsilon)} = i\}, i \in \mathbb{X}\rangle$ and transition probabilities, defined for $(i, s), (j, t) \in \mathbb{X} \times [0, \infty)$,

$$Q_{ij}^{(\varepsilon)}(t) = \mathsf{P}\{\eta_1^{(\varepsilon)} = j, \kappa_1^{(\varepsilon)} \le t / \eta_0^{(\varepsilon)} = i, \kappa_0^{(\varepsilon)} = s\}$$

$$= \begin{cases} F_{0,\pm}^{(\varepsilon)}(t) p_{0,\pm}(\varepsilon) & \text{if } j = 0 + \frac{1 \pm 1}{2}, \quad \text{for } i = 0, \\ F_{i,\pm}^{(\varepsilon)}(t) p_{i,\pm}(\varepsilon) & \text{if } j = i \pm 1, \qquad \text{for } 0 < i < N, \\ F_{N,\pm}^{(\varepsilon)}(t) p_{N,\pm}(\varepsilon) & \text{if } j = N - \frac{1 \mp 1}{2}, \text{ for } i = N, \\ 0 & \text{otherwise,} \end{cases} \tag{11.28}$$

where: (a) $F_{i,\pm}^{(\varepsilon)}(t)$, $i \in \mathbb{X}$ are distribution functions concentrated on $[0, \infty)$, for every $\varepsilon \in (0, \varepsilon_0]$; (b) $p_{i,\pm}(\varepsilon) \ge 0$, $p_{i,-}(\varepsilon) + p_{i,+}(\varepsilon) = 1$, $i \in \mathbb{X}$, for every $\varepsilon \in (0, \varepsilon_0]$.

In this case, the random sequence $\eta_n^{(\varepsilon)}$ is also a homogeneous (embedded) Markov chain with the phase space $\mathbb{X}$ and the transition probabilities, defined for $i, j \in \mathbb{X}$,

$$p_{ij}(\varepsilon) = \mathsf{P}\{\eta_1^{(\varepsilon)} = j / \eta_0^{(\varepsilon)} = i\} = Q_{ij}^{(\varepsilon)}(\infty)$$

$$= \begin{cases} p_{0,\pm}(\varepsilon) & \text{if } j = 0 + \frac{1 \pm 1}{2}, \quad \text{for } i = 0, \\ p_{i,\pm}(\varepsilon) & \text{if } j = i \pm 1, \qquad \text{for } 0 < i < N, \\ p_{N,\pm}(\varepsilon) & \text{if } j = N - \frac{1 \mp 1}{2}, \text{ for } i = N, \\ 0 & \text{otherwise.} \end{cases} \tag{11.29}$$

We assume that the following condition holds:

A: $p_{i,\pm}(\varepsilon) > 0,\ i \in \mathbb{X}$, for every $\varepsilon \in (0, \varepsilon_0]$.

Condition **A** obviously implies that the phase space $\mathbb{X}$ is a communicative class of states for the embedded Markov chain $\eta_n^{(\varepsilon)}$, for every $\varepsilon \in (0, \varepsilon_0]$.

We exclude instant transitions and assume that the following condition holds:

B: $F_{i,\pm}^{(\varepsilon)}(0) = 0,\ i \in \mathbb{X}$, for every $\varepsilon \in (0, \varepsilon_0]$.

Let us now introduce a semi-Markov process,

$$\eta^{(\varepsilon)}(t) = \eta^{(\varepsilon)}_{\nu^{(\varepsilon)}(t)},\ t \geq 0, \tag{11.30}$$

where $\nu^{(\varepsilon)}(t) = \max(n \geq 0 : \zeta_n^{(\varepsilon)} \leq t)$ is the number of jumps in the time interval $[0, t]$, for $t \geq 0$, and $\zeta_n^{(\varepsilon)}$ are sequential moments of jumps for the semi-Markov process $\eta^{(\varepsilon)}(t)$. This process has the phase space $\mathbb{X}$, the initial distribution $\bar{p} = \langle p_i = \mathsf{P}\{\eta^{(\varepsilon)}(0) = i\}, i \in \mathbb{X}\rangle$ and transition probabilities $Q_{ij}^{(\varepsilon)}(t), t \geq 0, i, j \in \mathbb{X}$.

Due to the specific assumptions imposed on the transition probabilities $p_{ij}^{(\varepsilon)}, i, j \in \mathbb{X}$ in relation (11.29), one can refer to $\eta^{(\varepsilon)}(t)$ as a semi-Markov birth-death process.

If $F_{i,\pm}^{(\varepsilon)}(t) = \mathrm{I}(t \geq 1), t \geq 0, i, j \in \mathbb{X}$, then $\eta^{(\varepsilon)}(t) = \eta^{(\varepsilon)}_{[t]}, t \geq 0$ is a discrete time homogeneous Markov birth-death chain embedded in continuous time.

If $F_{ij}^{(\varepsilon)}(t) = (1 - e^{-\lambda_i(\varepsilon)t}), t \geq 0, i, j \in \mathbb{X}$ (here, $0 < \lambda_i(\varepsilon) < \infty, i \in \mathbb{X}$), then $\eta^{(\varepsilon)}(t), t \geq 0$ is a continuous time homogeneous Markov birth-death process.

Let us define expectations of transition times, for $i, j \in \mathbb{X}$,

$$e_{ij}(\varepsilon) = \mathsf{E}_i\{\kappa_1^{(\varepsilon)} I(\eta_1^{(\varepsilon)} = j)\} = \int_0^\infty t\, Q_{ij}^{(\varepsilon)}(dt) \tag{11.31}$$

$$= \begin{cases} e_{0,\pm}(\varepsilon) & \text{if } j = 0 + \frac{1\pm1}{2}, \quad \text{for } i = 0, \\ e_{i,\pm}(\varepsilon) & \text{if } j = i \pm 1, \quad\ \text{for } 0 < i < N, \\ e_{N,\pm}(\varepsilon) & \text{if } j = N - \frac{1\mp1}{2}, \text{ for } i = N, \\ 0 & \text{otherwise,} \end{cases} \tag{11.32}$$

and

$$e_i(\varepsilon) = \mathsf{E}_i\kappa_1^{(\varepsilon)} = e_{i,-}(\varepsilon) + e_{i,+}(\varepsilon). \tag{11.33}$$

Here and henceforth, the notations P_i and E_i are used for conditional probabilities and expectations under the condition $\eta^{(\varepsilon)}(0) = i$.

We also assume that the following condition holds:

C: $e_{i,\pm}(\varepsilon) < \infty,\ i, j \in \mathbb{X}$, for $\varepsilon \in (0, \varepsilon_0]$.

It is useful to note that conditions **B** and **C** imply that all expectations $e_i(\varepsilon) \in (0, \infty), i \in \mathbb{X}$.

In the case of discrete time Markov birth-death chain, $e_i(\varepsilon) = 1, i \in \mathbb{X}$, whereas in the case of continuous time Markov birth-death process, $e_i(\varepsilon) = \lambda_i^{-1}(\varepsilon), i \in \mathbb{X}$.

Conditions **A**–**C** imply that the semi-Markov birth-death process $\eta^{(\varepsilon)}(t)$ is, for every $\varepsilon \in (0, \varepsilon_0]$, ergodic in the sense that the following asymptotic relation holds,

$$\mu_i^{(\varepsilon)}(t) = \frac{1}{t}\int_0^t I(\eta^{(\varepsilon)}(s) = i)ds \xrightarrow{a.s.} \pi_i(\varepsilon) \text{ as } t \to \infty,\ i \in \mathbb{X}. \tag{11.34}$$

The ergodic relation (11.34) holds for any initial distribution $\bar{p}^{(\varepsilon)}$ and the stationary probabilities $\pi_i(\varepsilon)$, $i \in \mathbb{X}$ do not depend on the initial distribution. Moreover, $\pi_i(\varepsilon) > 0$, $i \in \mathbb{X}$ and these probabilities are the unique solution of the following system of linear equations,

$$\pi_i(\varepsilon)e_i^{-1}(\varepsilon) = \sum_{j\in\mathbb{X}} \pi_j(\varepsilon)e_j^{-1}(\varepsilon)p_{ji}(\varepsilon),\ i \in \mathbb{X},\ \sum_{i\in\mathbb{X}} \pi_i(\varepsilon) = 1. \tag{11.35}$$

11.3.2 Perturbation Conditions for Semi-Markov Birth-Death Processes

Let us assume that that the following perturbation conditions hold:

D: $p_{i,\pm}(\varepsilon) = \sum_{l=0}^{1+l_{i,\pm}} a_{i,\pm}[l]\varepsilon^l + o_{i,\pm}(\varepsilon^{1+l_{i,\pm}})$, $\varepsilon \in (0, \varepsilon_0]$, for $i \in \mathbb{X}$, where: **(a)** $|a_{i,\pm}[l]| < \infty$, for $0 \le l \le 1 + l_{i,\pm}$, $i \in \mathbb{X}$; **(b)** $l_{i,\pm} = 0$, $a_{i,\pm}[0] > 0$, for $0 < i < N$; **(c)** $l_{0,\pm} = 0$, $a_{0,\pm}[0] > 0$ or $l_{0,+} = 1$, $a_{0,+}[0] = 0$, $a_{0,+}[1] > 0$, $l_{0,-} = 0$, $a_{0,-}[0] > 0$; **(d)** $l_{N,\pm} = 0$, $a_{N,\pm}[0] \ge 0$ or $l_{N,+} = 0$, $a_{N,+}[0] > 0$, $l_{N,-} = 1$, $a_{N,-}[0] = 0$, $a_{N,-}[1] > 0$; **(e)** $o_{i,\pm}(\varepsilon^{1+l_{i,\pm}})/\varepsilon^{1+l_{i,\pm}} \to 0$ as $\varepsilon \to 0$, for $i \in \mathbb{X}$.

and

E: $e_{i,\pm}(\varepsilon) = \sum_{l=0}^{1+l_{i,\pm}} b_{i,\pm}[l]\varepsilon^l + \dot{o}_{i,\pm}(\varepsilon^{1+l_{i,\pm}})$, $\varepsilon \in (0, \varepsilon_0]$, for $i \in \mathbb{X}$, where: **(a)** $|b_{i,\pm}[l]| < \infty$, for $0 \le l \le L + l_{i,\pm}$, $i \in \mathbb{X}$; **(b)** $l_{i,\pm} = 0$, $b_{i,\pm}[0] > 0$, for $0 < i < N$; **(c)** $l_{0,\pm} = 0$, $b_{0,\pm}[0] > 0$ or $l_{0,+} = 1$, $b_{0,+}[0] = 0$, $b_{0,+}[1] > 0$, $l_{0,-} = 0$, $b_{0,-}[0] > 0$; **(d)** $l_{N,\pm} = 0$, $b_{N,\pm}[0] > 0$ or $l_{N,+} = 0$, $b_{N,+}[0] > 0$, $l_{N,-} = 1$, $b_{N,-}[0] = 0$, $b_{N,-}[1] > 0$; **(e)** $\dot{o}_{i,\pm}(\varepsilon^{1+l_{i,\pm}})/\varepsilon^{1+l_{i,\pm}} \to 0$ as $\varepsilon \to 0$, for $i \in \mathbb{X}$.

It is useful to explain the role played by the parameters $l_{i,\pm}$ in conditions **D** and **E**. These parameters equalise the so-called length of asymptotic expansions penetrating these conditions.

The length of an asymptotic expansion is defined as the number of coefficients for powers of ε in this expansion, beginning from the first non-zero coefficient and up to the coefficient for the largest power of ε in this expansion.

The asymptotic expansions penetrating conditions **D** and **E** can be rewritten in the following form, $p_{i,\pm}(\varepsilon) = \sum_{l=l_{i,\pm}}^{1+l_{i,\pm}} a_{i,\pm}[l]\varepsilon^l + o_{i,\pm}(\varepsilon^{1+l_{i,\pm}})$, $\varepsilon \in (0, \varepsilon_0]$ and $e_{i,\pm}(\varepsilon) = \sum_{l=l_{i,\pm}}^{1+l_{i,\pm}} b_{i,\pm}[l]\varepsilon^l + \dot{o}_{i,\pm}(\varepsilon^{1+l_{i,\pm}})$, $\varepsilon \in (0, \varepsilon_0]$, for $i \in \mathbb{X}$. According to conditions **D**

and **E**, these asymptotic expansions have non-zero first coefficients. Therefore, all asymptotic expansions penetrating conditions **D** and **E** have the length 2. As we shall see, this makes it possible to represent the stationary and conditional quasi-stationary probabilities in the form of asymptotic expansions of the length 2.

Note that conditions **D** and **E** imply that there exists $\varepsilon_0' \in (0, \varepsilon_0]$ such that the probabilities $p_{i,\pm}(\varepsilon) > 0, i \in \mathbb{X}$ and the expectations $e_{i,\pm}(\varepsilon) > 0, i \in \mathbb{X}$ for $\varepsilon \in (0, \varepsilon_0']$. Therefore, let us just assume that $\varepsilon_0' = \varepsilon_0$.

The model assumption, $p_{i,-}(\varepsilon) + p_{i,+}(\varepsilon) = 1, \varepsilon \in (0, \varepsilon_0]$, also implies that the following condition should hold:

F: $a_{i,-}[0] + a_{i,+}[0] = 1, a_{i,-}[1] + a_{i,+}[1] = 0$, for $i \in \mathbb{X}$.

We also assume that the following natural consistency condition for asymptotic expansions penetrating perturbation conditions **D** and **E** holds:

G: $b_{i,\pm}[0] > 0$ if and only if $a_{i,\pm}[0] > 0$, for $i = 0, N$.

There are three basic variants of the model that correspond to (11.7) and (11.8). For the more general setup of semi-Markov chains in this section, we formulate this a bit differently and assume that one of the following conditions holds:

$\mathbf{H_1}$: $a_{0,+}[0] > 0, a_{N,-}[0] > 0$.

$\mathbf{H_2}$: $a_{0,+}[0] = 0, a_{N,-}[0] > 0$.

$\mathbf{H_3}$: $a_{0,+}[0] = 0, a_{N,-}[0] = 0$.

The case $a_{0,+}[0] > 0, a_{N,-}[0] = 0$ is analogous to the case where condition $\mathbf{H_2}$ holds and we omit its consideration.

Condition **D** implies that there exist $\lim_{\varepsilon \to 0} p_{i,\pm}(\varepsilon) = p_{i,\pm}(0), i \in \mathbb{X}$ and, thus, there also exist $\lim_{\varepsilon \to 0} p_{ij}(\varepsilon) = p_{ij}(0), i, j \in \mathbb{X}$. Condition **E** implies that there exist $\lim_{\varepsilon \to 0} e_{i,\pm}(\varepsilon) = e_{i,\pm}(0), i \in \mathbb{X}$ and, thus, there also exist $\lim_{\varepsilon \to 0} e_{ij}(\varepsilon) = e_{ij}(0), i, j \in \mathbb{X}$.

The limiting birth-death type Markov chain $\eta_n^{(0)}$ with the matrix of transition probabilities $\|p_{ij}(0)\|$ has: **(a)** one class of communicative states $\mathbb{X}$, if condition $\mathbf{H_1}$ holds, **(b)** one communicative class of transient states ${}_{\langle 1,N \rangle}\mathbb{X} = \mathbb{X} \setminus \{0\}$ and an absorbing state 0, if condition $\mathbf{H_2}$ holds, and **(c)** one communicative class of transient states ${}_{\langle 1,N-1 \rangle}\mathbb{X} = \mathbb{X} \setminus \{0, N\}$ and two absorbing states 0 and N, if condition $\mathbf{H_3}$ holds.

In this paper, we get, under conditions **A**–**G** and $\mathbf{H}_i$ (for $i = 1, 2, 3$), asymptotic expansions for stationary probabilities, as $\varepsilon \to 0$,

$$\pi_i(\varepsilon) = \sum_{l=\dot{l}_i}^{1+\dot{l}_i} c_i[l]\varepsilon^l + o_i(\varepsilon^{1+\dot{l}_i}), \ i \in \mathbb{X}, \tag{11.36}$$

where: (a) $\dot{l}_i = 0, i \in \mathbb{X}$ and the limiting stationary probabilities $\pi_i(0) > 0, i \in \mathbb{X}$, if condition $\mathbf{H_1}$ holds, (b) $\dot{l}_i = I(i \neq 0), i \in \mathbb{X}$ and $\pi_0(0) = 1, \pi_i(0) = 0, i \in$

$_{\langle 1,N\rangle}\mathbb{X}$, if condition $\mathbf{H_2}$ holds, and (c) $\dot{l}_i = I(i \neq 0, N)$, $i \in \mathbb{X}$ and $\pi_0(0), \pi_N(0) > 0$, $\pi_0(0) + \pi_N(0) = 1$, $\pi_i(0) = 0$, $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$, if condition $\mathbf{H_3}$ holds.

This implies that there is sense to consider so-called conditional quasi-stationary probabilities, which are defined as,

$$\tilde{\pi}_i(\varepsilon) = \frac{\pi_i(\varepsilon)}{1 - \pi_0(\varepsilon)} = \frac{\pi_i(\varepsilon)}{\sum_{j \in {}_0\mathbb{X}} \pi_j(\varepsilon)}, \ i \in {}_{\langle 1,N\rangle}\mathbb{X}, \tag{11.37}$$

in the case where condition $\mathbf{H_2}$ holds, or as,

$$\hat{\pi}_i(\varepsilon) = \frac{\pi_i(\varepsilon)}{1 - \pi_0(\varepsilon) - \pi_N(\varepsilon)} = \frac{\pi_i(\varepsilon)}{\sum_{j \in {}_{0,N}\mathbb{X}} \pi_j(\varepsilon)}, \ \ i \in {}_{\langle 1,N-1\rangle}\mathbb{X}, \tag{11.38}$$

in the case where condition $\mathbf{H_3}$ holds.

We also get, under conditions **A**–**G** and $\mathbf{H_2}$, asymptotic expansions for conditional quasi-stationary probabilities,

$$\tilde{\pi}_i(\varepsilon) = \sum_{l=0}^{1} \tilde{c}_i[l]\varepsilon^l + \tilde{o}_i(\varepsilon), \ i \in {}_{\langle 1,N\rangle}\mathbb{X}, \tag{11.39}$$

and, under conditions **A**–**G** and $\mathbf{H_3}$, asymptotic expansions for conditional quasi-stationary probabilities,

$$\hat{\pi}_i(\varepsilon) = \sum_{l=0}^{1} \hat{c}_i[l]\varepsilon^l + \hat{o}_i(\varepsilon), \ i \in {}_{\langle 1,N-1\rangle}\mathbb{X}. \tag{11.40}$$

The coefficients in the above asymptotic expansions are given by explicit formulas via coefficients in asymptotic expansions given in initial perturbation conditions **D** and **E**.

As it was mentioned in the introduction, the first coefficients $\pi_i(\varepsilon) = c_i[0]$, $\tilde{\pi}_i(0) = \tilde{c}_i[0]$ and $\hat{\pi}_i(0) = \hat{c}_i[0]$ describe the asymptotic behaviour of stationary and quasi-stationary probabilities and their continuity properties with respect to small perturbations of transition characteristics of the corresponding semi-Markov processes. The second coefficients $c_i[1]$, $\tilde{c}_i[1]$ and $\hat{c}_i[1]$ determine sensitivity of stationary and quasi-stationary probabilities with respect to small perturbations of transition characteristics.

We also would like to comment the use of the term "conditional quasi-stationary probability" for quantities defined in relations (11.37) and (11.38). As a matter of fact, the term "quasi-stationary probability (distribution)" is traditionally used for limits,

$$q_j(\varepsilon) = \lim_{t \to \infty} \mathsf{P}_i\{\eta^{(\varepsilon)}(t) = j / \eta^{(\varepsilon)}(s) \notin A, 0 \leq s \leq t\}, \tag{11.41}$$

where A is some special subset of $\mathbb{X}$.

A detailed presentation of results concerning quasi-stationary distributions and comprehensive bibliographies of works in this area can be found in the books by Gyllenberg and Silvestrov [13], Nåsell [34] and Collet, Martínez and San Martín [6]. We would also like to mention the paper by Allen and Tarnita [1], where one can find a discussion concerning the above two forms of quasi-stationary distributions for some bio-stochastic systems.

11.4 Examples of Stationary Distributions

In this section, we will revisit the examples of Sect. 11.2 and illustrate how to compute, approximate and expand the various stationary and conditional quasi-stationary distributions that were introduced in Sect. 11.3. Since all the models of Sect. 11.2 have a geometric or exponential transition time distribution (11.3)–(11.4), and since the transition probabilities satisfy (11.28), it follows that the stationary distribution (11.34)–(11.35) has a very explicit expression,

$$\pi_i(\varepsilon) \propto \begin{cases} 1, & i = 0, \\ \frac{\lambda_{0,+}(\varepsilon)\cdot\ldots\cdot\lambda_{i-1,+}(\varepsilon)}{\lambda_{1,-}(\varepsilon)\cdot\ldots\cdot\lambda_{i,-}(\varepsilon)}, & i = 1, \ldots, N, \end{cases} \tag{11.42}$$

for $0 < \varepsilon \leq \varepsilon_0$, with a proportionality constant chosen so that $\sum_{i=0}^{N} \pi_i(\varepsilon) = 1$.

Our goal is to find a series representation of (11.42). Since the models of Sect. 11.2 are formulated in terms of the death and birth rates in (11.5), we will assume that these rates admit expansions

$$\lambda_{i,\pm}(\varepsilon) = \sum_{l=0}^{L_{i,\pm}} g_{i,\pm}[l]\varepsilon^l + o_{i,\pm}(\varepsilon^{L+l_{i,\pm}}) \tag{11.43}$$

for $\varepsilon \in (0, \varepsilon_0]$, and then check the regularity conditions of Sect. 11.3 that are needed to hold. From Eqs. (11.3)–(11.4), (11.6), and (11.31), we deduce that

$$e_{i,\pm}(\varepsilon) = \frac{1}{\lambda_i(\varepsilon)} \cdot \frac{\lambda_{i,\pm}(\varepsilon)}{\lambda_i(\varepsilon)}. \tag{11.44}$$

Inserting (11.43) into (11.44), we find that

$$g_{i,-}[0] + g_{i,+}[0] > 0 \tag{11.45}$$

must hold for all $i \in \mathbb{X}$ in order for the series expansion of $e_{i,\pm}(\varepsilon)$ to satisfy condition $\mathbf{E}_L$. It therefore follows from (11.6) that $p_{i,\pm}(\varepsilon)$ will satisfy perturbation condition $\mathbf{D}_L$, with $L + l_{i,+} = L + l_{i,-} = \min(L_{i,-}, L_{i,+})$, and

$$a_{i,\pm}[0] = \frac{g_{i,\pm}[0]}{g_{i,-}[0] + g_{i,+}[0]}. \tag{11.46}$$

Because of (11.45) and (11.46), we can rephrase the three perturbation scenarios $\mathbf{H}_1$ – $\mathbf{H}_3$ of Sect. 11.3.2 as

$$\begin{aligned} \mathbf{H}_1 &: g_{0,+}[0] > 0,\, g_{N,-}[0] > 0, \\ \mathbf{H}_2 &: g_{0,+}[0] = 0,\, g_{N,-}[0] > 0, \\ \mathbf{H}_3 &: g_{0,+}[0] = 0,\, g_{N,-}[0] = 0, \end{aligned} \tag{11.47}$$

in agreement with (11.8). Under $\mathbf{H}_2$, the exact expression for the conditional quasi-stationary distribution (11.37) is readily obtained from (11.42). It equals

$$\tilde{\pi}_i(\varepsilon) \propto \frac{\lambda_{1,+}(\varepsilon) \cdot \ldots \cdot \lambda_{i-1,+}(\varepsilon)}{\lambda_{1,-}(\varepsilon) \cdot \ldots \cdot \lambda_{i,-}(\varepsilon)} \tag{11.48}$$

for $i \in {}_0\mathbb{X}$ and $0 < \varepsilon \leq \varepsilon_0$, with the numerator equal to 1 when $i = 1$, and a proportionality constant chosen so that $\sum_{i=1}^{N} \tilde{\pi}_i(\varepsilon) = 1$. As $\varepsilon \to 0$, this expression converges to

$$\tilde{\pi}_i(0) \propto \frac{\lambda_{1,+}(0) \cdot \ldots \cdot \lambda_{i-1,+}(0)}{\lambda_{1,-}(0) \cdot \ldots \cdot \lambda_{i,-}(0)}. \tag{11.49}$$

If scenario $\mathbf{H}_3$ holds, we find analogously that the conditional quasi-stationary distribution (11.38) is given by

$$\hat{\pi}_i(\varepsilon) \propto \frac{\lambda_{1,+}(\varepsilon) \cdot \ldots \cdot \lambda_{i-1,+}(\varepsilon)}{\lambda_{1,-}(\varepsilon) \cdot \ldots \cdot \lambda_{i,-}(\varepsilon)} \tag{11.50}$$

for $i \in_{0,N} \mathbb{X}$, with a limit

$$\hat{\pi}_i(0) \propto \frac{\lambda_{1,+}(0) \cdot \ldots \cdot \lambda_{i-1,+}(0)}{\lambda_{1,-}(0) \cdot \ldots \cdot \lambda_{i,-}(0)}. \tag{11.51}$$

11.4.1 Stationary Distributions for Perturbed Population Dynamics Models

For the population dynamics model (11.9) of Sect. 11.2.1, we considered two perturbation scenarios. Recall that the first one in (11.12) has a varying immigration parameter $\nu(\varepsilon) = \varepsilon$, whereas all other parameters are kept fixed. Since $\lambda_{0,-}(\varepsilon) = 0$ and $\lambda_{0,+}(\varepsilon) = \varepsilon$, it follows that $g_{0,-}[0] = g_{0,+}[0] = 0$, and therefore formula (11.45) is violated for $i = 0$. But the properties of $\eta^{(\varepsilon)}$ remain the same if we put $\lambda_{0,-}(\varepsilon) = 1$ instead. With this modification, formula (11.47) implies that condition $\mathbf{H}_2$ of

Sect. 11.3.2 holds, and hence the $\varepsilon \to 0$ limit of the stationary distribution in (11.34) and (11.42) is concentrated at state 0 ($\pi_0(0) = 1$).

Let $\tau_0^{(\varepsilon)}$ be the time it takes for the population to get temporarily extinct again, after an immigrant has entered an empty island. It then follows from a slight modification of Eq. (11.90) in Sect. 11.5.2 and the relation $\lambda_{0,+}(\varepsilon) = \varepsilon$, that a first order expansion of the probability that the island is empty at stationarity, is

$$\pi_0(\varepsilon) = \frac{1/\lambda_{0,+}(\varepsilon)}{1/\lambda_{0,+}(\varepsilon) + \mathsf{E}_1(\tau_0^{(\varepsilon)})} = \frac{1/\varepsilon}{1/\varepsilon + E_{10}(\varepsilon)} = 1 - E_{10}(\varepsilon)\varepsilon + o(\varepsilon). \quad (11.52)$$

This expansion is accurate when the perturbation parameter is small ($\varepsilon \ll 1/E_{10}(\varepsilon)$), otherwise higher order terms in (11.52) are needed. The value of $E_{10}(\varepsilon)$ will be highly dependent on the value of the basic reproduction number R_0 in (11.13). When $R_0 > 1$, the expected time to extinction will be very large, and $\pi_0(\varepsilon)$ will be close to 0 for all but very small ε. On the other hand, (11.52) is accurate for a larger range of ε when $R_0 < 1$, since $E_{10}(\varepsilon)$ is then small.

In order to find useful approximations of the conditional quasi-stationary distribution $\tilde{\pi}_i(\varepsilon)$ in (11.48), we will distinguish between whether R_0 is larger than or smaller than 1. When $R_0 > 1$, or equivalently $\lambda > \mu$, we can rewrite (11.11) as

$$\mathsf{E}\left[\eta^{(\varepsilon)}(t + \Delta t) - \eta^{(\varepsilon)}(t) | \eta^{(\varepsilon)}(t) = i\right] = \Delta t \cdot Nm\left(\frac{i}{N}\right), \quad (11.53)$$

where

$$m(x) = rx + \frac{\varepsilon}{N} - \left[rx(0)^{-\theta} \cdot x + \frac{\varepsilon}{N}\right] x^{\theta} \quad (11.54)$$

is a rescaled mean function of the drift, $r = \mu(R_0 - 1)$ is the intrinsic growth rate, or growth rate per capita, of a small population without immigration ($\varepsilon = 0$), and

$$x(0) = \left(\frac{R_0 - 1}{\alpha_1 R_0 + \alpha_2}\right)^{1/\theta}.$$

We assume that α_1 and α_2 are large enough so that $x(0) < 1$. A sufficient condition for this is $\alpha_1 + \alpha_2 = 1$. The carrying capacity $K(\varepsilon) = Nx(\varepsilon)$ of the environment is the value of i such that the right hand side of (11.53) equals zero. We can write $x = x(\varepsilon)$ as the unique solution of $m(x) = 0$, or equivalently

$$x^{\theta} = \frac{rx + \varepsilon N^{-1}}{rx(0)^{-\theta}x + \varepsilon N^{-1}},$$

with $x(\varepsilon) \searrow x(0)$ as $\varepsilon \to 0$. The conditional quasi-stationary distribution (11.48) will be centred around $K(\varepsilon)$. In order to find a good approximation of this distribution, we look at the second moment

$$\begin{aligned}\mathsf{E}\left\{\left[\eta^{(\varepsilon)}(t+\Delta t)-\eta^{(\varepsilon)}(t)\right]^2|\eta^{(\varepsilon)}(t)=i\right\}&=\Delta t\left[\lambda_{i,+}(\varepsilon)+\lambda_{i,-}(\varepsilon)\right]\\&=\Delta t\cdot Nv\left(\tfrac{i}{N}\right),\end{aligned}$$

of the drift of $\eta^{(\varepsilon)}$, with

$$v(x)=\lambda x(1-\alpha_1 x^\theta)+\frac{\varepsilon}{N}(1-x^\theta)+\mu x(1+\alpha_2 x^\theta). \tag{11.55}$$

When N is large, we may approximate the conditional quasi-stationary distribution

$$\begin{aligned}\tilde{\pi}_i(\varepsilon)&\approx\textstyle\int_{i-}^{i+}f^{(\varepsilon)}(k)dk\\&=\textstyle\int_{i-}^{i+}f^{(0)}(k)dk+\int_{i-}^{i+}\left.\frac{df^{(\varepsilon)}(k)}{d\varepsilon}\right|_{\varepsilon=0}dk\cdot\varepsilon+o(\varepsilon),\end{aligned} \tag{11.56}$$

by integrating a density function $f^{(\varepsilon)}$ on $[0,N]$ between $i_-=\max(0,i-1/2)$ and $i_+=\min(N,i+1/2)$. This density function can be found through a diffusion argument as the stationary density

$$\begin{aligned}f^{(\varepsilon)}(k)&\propto\tfrac{1}{Nv(\frac{k}{N})}\exp\left(2\textstyle\int_{K(\varepsilon)}^{k}\frac{Nm(\frac{y}{N})}{Nv(\frac{y}{N})}dy\right)\\&\propto\tfrac{1}{v(\frac{k}{N})}\exp\left(2\textstyle\int_{K(\varepsilon)}^{k}\frac{m(\frac{y}{N})}{v(\frac{y}{N})}dy\right)\end{aligned} \tag{11.57}$$

of Kolmogorov's forward equation, with a proportionality constant chosen so that $\int_0^N f^{(\varepsilon)}(k)dk=1$ (see, for instance, Chap. 9 of Crow and Kimura [7]). A substitution of variables $x=y/N$ in (11.57), and a Taylor expansion of $m(x)$ around $x(\varepsilon)$ reveals that the diffusion density has approximately a normal distribution

$$f^{(\varepsilon)}\sim N\left(K(\varepsilon),N\frac{v\left[x(\varepsilon)\right]}{2|m'\left[x(\varepsilon)\right]|}\right). \tag{11.58}$$

Expansion (11.56) is valid for small migration rates ε, and its linear term quantifies how sensitive the conditional quasi-stationary distribution is to a small amount of immigration.

It follows from (11.53) that the expected population size

$$\mathsf{E}\left[\eta^{(0)}(t+\Delta t)-\eta^{(0)}(t)|\eta^{(0)}(t)=i\right]=\Delta t\cdot ri\left[1-\left(\frac{i}{K(0)}\right)^\theta\right] \tag{11.59}$$

of an isolated population varies according to a theta logistic model (Gilpin and Ayala [12]), which is a special case of the generalised growth curve model in Tsoularis and Wallace [44]. The theta logistic model has a carrying capacity $K(0)$ of the environment to accommodate new births. When $\theta=1$, we obtain the logistic growth model of Verhulst [45]. Pearl [35] used such a curve to approximate population growth in the United States, and Feller [10] introduced a stochastic version of the

logistic model in terms of a Markov birth-death process (11.4) in continuous time. Feller's approach has been extended for instance by Kendall [21], Whittle [47], and Nåsell [31, 33]. In particular, Nåsell studied the quasi-stationary distribution (11.41) of $\eta^{(\varepsilon)}$, with $A = \{1, \ldots, N\}$. In this paper the previously studied population growth models are generalised in two directions; we consider semi-Markov processes and allow for theta logistic expected growth.

When $0 < R_0 < 1$, or equivalently $0 < \lambda < \mu$, we rewrite (11.11) as

$$\begin{aligned} &\mathsf{E}\left[\eta^{(\varepsilon)}(t + \Delta t) - \eta^{(\varepsilon)}(t) | \eta^{(\varepsilon)}(t) = i\right] \\ &\quad = \Delta t \cdot \left[\nu - ri - (\nu + ri\tilde{x}^{-\theta})\left(\tfrac{i}{N}\right)^{\theta}\right], \end{aligned} \tag{11.60}$$

where $r = (1 - R_0)\mu$ quantifies per capita decrease for a small population without immigration, and $\tilde{x} = [(1 - R_0)/(\alpha_1 R_0 + \alpha_2)]^{1/\theta}$ is the fraction of the maximal population size at which the per capita decrease of an isolated $\varepsilon = 0$ population has doubled to $2r$. For large N, we can neglect all $O(N^{-\theta})$ terms, and it follows from (11.49) that

$$\tilde{\pi}_i(\varepsilon) \approx \frac{1}{\log(1 - R_0)} \cdot \frac{R_0^i}{i} + \tilde{c}_i[1]\varepsilon + o(\varepsilon),$$

for $i = 1, \ldots, N$.

Recall that the second perturbation scenario (11.14) has a varying birth rate $\langle(\varepsilon) = \varepsilon$, whereas all other parameters are kept fixed, not depending on ε. In view of (11.47), it satisfies condition $\mathbf{H}_1$ of Sect. 11.3.2. Suppose N is large. If $\nu = o(N)$, it follows from (11.42) that the stationary distribution for small values of ε is well approximated by

$$\pi_i(\varepsilon) \approx \frac{(\nu/\mu)^i}{i!} e^{-\nu/\mu} + c_i[1]\varepsilon + o(\varepsilon)$$

for $i = 0, \ldots, N$, a Poisson distribution with mean ν/μ, corrupted by a sensitivity term $c_i[1]\varepsilon$ due to births. If $\nu = VN$, the carrying capacity of the environment is $K(\varepsilon) = Nx(\varepsilon)$, where $x = x(\varepsilon)$ is the value of i/N in (11.60) such that the right hand side vanishes, i.e. the unique solution of the equation

$$rx + Vx^{\theta} + r\tilde{x}^{-\theta}x^{\theta+1} = V,$$

with $r = r(\varepsilon) = \mu - \varepsilon$. The stationary distribution (11.42) is well approximated by a discretised normal distribution (11.56)–(11.58), but with a mean drift function $m(x)$ obtained from (11.60), and a variance function $v(x)$ derived similarly.

11.4.2 Stationary Distributions for Perturbed Epidemic Models

For the epidemic models of Sect. 11.2.2, we considered one perturbation scenario (11.18), with a varying external contact rate $\nu(\varepsilon) = \varepsilon$. When the basic reproduction model $R_0 = \mu/\mu$ exceeds one, the expected growth rate follows a logistic model (11.18) when $\varepsilon = 0$, which is a special case of the theta logistic mean growth curve model (11.59), with $\theta = 1$. When $R_0 < 1$, we similarly write the expected population decline as in (11.60), with $\theta = 1$. Since the SIS model is a particular case of the population dynamic models of Sect. 11.2.1 (Nåsell, [34]), the stationary and conditional quasi-stationary distributions are approximated in the same way as in Sect. 11.4.1.

11.4.3 Stationary Distributions for Perturbed Models of Population Genetics

For the population genetics model of Sect. 11.2.3, we recall there were three different perturbation scenarios (11.27). For all of them, the rescaled mutation rates $U_1(\varepsilon) = N\mathsf{P}(A_1 \to A_2)$ and $U_2(\varepsilon) = N\mathsf{P}(A_2 \to A_1)$ between the two alleles A_1 and A_2 are linear functions of ε.

The stationary distribution is either found by first inserting (11.20) into (11.5), and then (11.5) into (11.42), or, for large N, it is often more convenient to use a diffusion approximation,

$$\pi_i(\varepsilon) \approx \int_{x_{i,-}}^{x_{i,+}} f^{(\varepsilon)}(x)dx. \tag{11.61}$$

It is obtained by integrating the density function

$$\begin{aligned} f^{(\varepsilon)}(x) &\propto \tfrac{1}{\nu(x)} \exp\left(2\textstyle\int_{1/2}^{x} \frac{m(y)}{\nu(y)}dy\right) \\ &\propto (1-x)^{-1+U_1} x^{-1+U_2} \exp\left[\tfrac{1}{2}(S_1+S_2)x^2 - S_2 x\right] \end{aligned} \tag{11.62}$$

between $x_{i,-} = \max[0, (i-1/2)/N]$ and $x_{i,+} = \min[1, (i+1/2)/N]$. This density is defined in terms of the infinitesimal drift and variance functions $m(x)$ and $\nu(x)$ in (11.23)–(11.25), with a constant of proportionality chosen to ensure that $\int f^{(\varepsilon)}(x)dx = 1$. See, for instance, Chap. 9 of Crow and Kimura [7] and Chap. 7 of Durrett [8] for details.

For $\mathbf{H}_1$, we use this diffusion argument to find an approximate first order series expansion

$$\pi_i(\varepsilon) \approx \int_{x_{i,-}}^{x_{i,+}} f^{(0)}(x)dx + \int_{x_{i,-}}^{x_{i,+}} \left.\frac{df^{(\varepsilon)}(x)}{d\varepsilon}\right|_{\varepsilon=0} dx \cdot \varepsilon + o(\varepsilon)$$

of the stationary distribution by inserting (11.26) into (11.61)–(11.62). The null density $f^{(0)}(x)$ is defined by (11.62), with C_1 and C_2 instead of U_1 and U_2. For a neutral model ($S_1 = S_2 = 0$), the stationary null distribution is approximately beta with parameters C_1 and C_2, and expected value $C_2/(C_1 + C_2)$. A model with $S_1 > 0 > S_2$ corresponds to directional selection, with higher fitness for A_1 compared to A_2. It can be seen from (11.62) that the stationary null distribution is further skewed to the right than for a neutral model. A model with balancing selection or overdominance has negative S_1 and S_2, so that the heterozygous genotype A_1A_2 has a selective advantage. The stationary null distribution will then have a peak around $S_2/(S_1 + S_2)$. On the other hand, for an underdominant model where S_1 and S_2 are both positive, the heterozygous genotype will have a selective disadvantage. Then $S_2/(S_1 + S_2)$ functions as a repelling point of the stationary null distribution.

For scenario $\mathbf{H}_2$, the null model has one absorbing state 0. In analogy with (11.52), we find that the series expansion of the stationary probability of no A_1 alleles in the population is

$$\pi_0(\varepsilon) = 1 - \mathsf{E}_1(\tau_0^{(\varepsilon)}) \cdot \frac{D_2\varepsilon}{N} + o(\varepsilon)$$

when $D_2 > 0$, for small values of the perturbation parameter. Here $D_2\varepsilon/N$ is the probability that a mutation $A_2 \to A_1$ occurs in a homogeneous A_2 population, and $\tau_0^{(\varepsilon)}$ is the time it takes for the A_1 allele to disappear again.

Because of the singularity at $i = 0$ for small ε, we avoid the diffusion argument and find the conditional quasi-stationary distribution (11.37) directly by first inserting (11.20) into (11.5), and then (11.5) into (11.48)–(11.49). After some computations, this leads to

$$\begin{aligned}\tilde{\pi}_i(\varepsilon) \approx{}& \tilde{c}_1[0] i^{-1} \left(1 - \tfrac{i-1}{N}\right)^{C_1 - 1} \exp\left[\tfrac{1}{2}(S_1 + S_2)\tfrac{i-1}{N}\tfrac{i}{N} - S_2\tfrac{i-1}{N}\right] \\ &+ \tilde{c}_1[1]\varepsilon + o(\varepsilon)\end{aligned} \tag{11.63}$$

for $i = 1, \ldots, N$, where $\tilde{c}_1[0]$ is chosen so that $\sum_{i=1}^{N} \tilde{\pi}_i(0) = 1$, and $\tilde{c}_1[1]$ will additionally involve D_1 and D_2. If $D_2 = 0$, we have that $\pi_0(\varepsilon) = 1$ for all $0 < \varepsilon \leq \varepsilon_0$, so that the conditional quasi-stationary distribution (11.37) is not well defined. However, the time to reach absorption is very large for small $U_1 > 0$. It is shown in Hössjer, Tyvand and Miloh [17] that $\eta^{(\varepsilon)}$ may be quasi-fixed for a long time at the other boundary point $i = N$, before eventual absorption at $i = 0$ occurs.

For scenario $\mathbf{H}_3$, the null model is mutation free, and the asymptotic distribution

$$P_j(0; i) = \lim_{t \to \infty} \mathsf{P}_i(\eta^{(0)}(t) = j)$$

is supported on the two absorbing states ($j \in \{0, N\}$), and it is dependent on the state i at which the process starts. For a neutral model ($s_1 = s_2 = 0$), we have that

$$P_N(0; i) = 1 - P_0(0; i) = \frac{i}{N}. \tag{11.64}$$

A particular case of directional selection is multiplicative fitness, with $1 + s_1 = (1 + s_2)^{-1}$. It is mathematically simpler since selection operates directly on alleles, not on genotypes, with selective advantages 1 and $1 + s_2$ for A_1 and A_2. It follows for instance from Sect. 6.1 of Durrett [8] that

$$P_N(0; i) = 1 - P_0(0; i) = \frac{1 - (1 + s_2)^i}{1 - (1 + s_2)^N} \tag{11.65}$$

for multiplicative fitness. Notice that $P_0(0; i)$ and $P_N(0; i)$ will differ from $\pi_j(0) = \lim_{\varepsilon \to 0} \pi_0(\varepsilon)$ at the two boundaries. Indeed, by ergodicity (11.34) for each $\varepsilon > 0$, the latter two probabilities are not functions of $i = \eta^{(0)}(0)$. From (11.61)–(11.62), we find that

$$\pi_N(0) = 1 - \pi_0(0) \approx \frac{D_2}{\exp\left[-\frac{1}{2}(S_1 - S_2)\right] D_1 + D_2}. \tag{11.66}$$

Similarly as in (11.63), we find after some computations that the conditional quasi-stationary distribution (11.38) admits an approximate expansion

$$\begin{aligned} \hat{\pi}_i(\varepsilon) \approx{}& \hat{c}_1[0] i^{-1} \left(1 - \tfrac{i-1}{N}\right)^{-1} \exp\left[\tfrac{1}{2}(S_1 + S_2)\tfrac{i-1}{N}\tfrac{i}{N} - S_2 \tfrac{i-1}{N}\right] \\ &+ \hat{c}_1[1]\varepsilon + o(\varepsilon) \end{aligned} \tag{11.67}$$

for $i = 1, \ldots, N - 1$, where $\hat{c}_1[0]$ is chosen so that $\sum_{i=1}^{N-1} \hat{\pi}_i(0) = 1$, and $\hat{c}_1[1]$ will additionally involve D_1 and D_2. Notice that the limiting fixation probabilities in (11.66) are functions of the mutation probability ratio $D_1/(D_1 + D_2)$, but the limiting conditional quasi-stationary distribution $\hat{\pi}_i(0)$ in (11.67) does not involve any of D_1 or D_2.

11.5 Reduced Semi-Markov Birth-Death Processes

In this section, we present a time-space screening procedure of phase space reduction for perturbed semi-Markov birth-death processes and recurrent algorithms for computing expectations of hitting times and stationary and conditional quasi-stationary distributions for such processes.

11.5.1 Phase Space Reduction for Semi-Markov Birth-Death Processes

Let us assume that $N \geq 1$. Let $0 \leq k \leq i \leq r \leq N$ and define the reduced phase space ${}_{\langle k,r \rangle}\mathbb{X} = \{k, \ldots, r\}$. Note that, by the definition, ${}_{\langle 0,N \rangle}\mathbb{X} = \mathbb{X}$. Let us also assume

that the initial distribution $\bar{p}^{(\varepsilon)}$ is concentrated on the phase space ${}_{\langle k,r\rangle}\mathbb{X}$, i.e. $p_i^{(\varepsilon)} = 0, i \notin {}_{\langle k,r\rangle}\mathbb{X}$.

Let us define the sequential moments of hitting the reduced space ${}_{\langle k,r\rangle}\mathbb{X}$, by the embedded Markov chain $\eta_n^{(\varepsilon)}$,

$$ {}_{\langle k,r\rangle}\xi_n^{(\varepsilon)} = \min(k > {}_{\langle k,r\rangle}\xi_{n-1}^{(\varepsilon)},\ \eta_k^{(\varepsilon)} \in {}_{\langle k,r\rangle}\mathbb{X}),\ n \geq 1,\ {}_{\langle k,r\rangle}\xi_0^{(\varepsilon)} = 0. \tag{11.68} $$

Now, let us define the random sequence,

$$ ({}_{\langle k,r\rangle}\eta_n^{(\varepsilon)},\ {}_{\langle k,r\rangle}\kappa_n^{(\varepsilon)}) = \begin{cases} (\eta_0^{(\varepsilon)}, 0) & \text{for } n = 0, \\ (\eta^{(\varepsilon)}_{{}_{\langle k,r\rangle}\xi_n^{(\varepsilon)}}, \sum_{l={}_{\langle k,r\rangle}\xi_{n-1}^{(\varepsilon)}+1}^{{}_{\langle k,r\rangle}\xi_n^{(\varepsilon)}} \kappa_l^{(\varepsilon)}) & \text{for } n \geq 1. \end{cases} \tag{11.69} $$

This sequence is a Markov renewal process with a phase space ${}_{\langle k,r\rangle}\mathbb{X} \times [0,\infty)$, the initial distribution $\bar{p}^{(\varepsilon)}$, and transition probabilities defined for $(i,s),(j,t) \in \mathbb{X} \times [0,\infty)$,

$$ {}_{\langle k,r\rangle}Q_{ij}^{(\varepsilon)}(t) = \mathsf{P}\{ {}_{\langle k,r\rangle}\eta_1^{(\varepsilon)} = j,\ {}_{\langle k,r\rangle}\kappa_1^{(\varepsilon)} \leq t / {}_{\langle k,r\rangle}\eta_0^{(\varepsilon)} = i,\ {}_{\langle k,r\rangle}\kappa_0^{(\varepsilon)} = s\}. \tag{11.70} $$

We define a reduced semi-Markov process by

$$ {}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t) = {}_{\langle k,r\rangle}\eta^{(\varepsilon)}_{{}_{\langle k,r\rangle}\nu^{(\varepsilon)}(t)},\ t \geq 0, \tag{11.71} $$

where ${}_{\langle k,r\rangle}\nu^{(\varepsilon)}(t) = \max(n \geq 0 :\ {}_{\langle k,r\rangle}\zeta_n^{(\varepsilon)} \leq t)$ is the number of jumps in the time interval $[0,t]$, for $t \geq 0$, and ${}_{\langle k,r\rangle}\zeta_n^{(\varepsilon)} = {}_{\langle k,r\rangle}\kappa_1^{(\varepsilon)} + \cdots + {}_{\langle k,r\rangle}\kappa_n^{(\varepsilon)}$, $n = 0,1,\ldots$ are sequential moments of jumps, for the semi-Markov process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$.

In particular, the initial semi-Markov process $\eta^{(\varepsilon)}(t) = {}_{\langle 0,N\rangle}\eta^{(\varepsilon)}(t)$.

It is readily seen that ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$ is also a semi-Markov birth-death process, i.e. the time-space screening procedure of phase space reduction described above preserves the birth-death structure of the semi-Markov birth-death process $\eta^{(\varepsilon)}(t)$.

11.5.2 Expectations of Hitting Times for Reduced Semi-Markov Birth-Death Processes

Let us now introduce hitting times for semi-Markov birth-death process $\eta^{(\varepsilon)}(t)$. We define hitting times, which are random variables given by the following relation, for $j \in \mathbb{X}$,

$$ \tau_j^{(\varepsilon)} = \sum_{n=1}^{\nu_j^{(\varepsilon)}} \kappa_n^{(\varepsilon)}, \tag{11.72} $$

where $\nu_j^{(\varepsilon)} = \min(n \geq 1 : \eta_n^{(\varepsilon)} = j)$.

Let us denote,

$$E_{ij}(\varepsilon) = \mathsf{E}_i \tau_j^{(\varepsilon)},\ i, j \in \mathbb{X}. \tag{11.73}$$

As is known, conditions **A**–**C** imply that, for every $\varepsilon \in (0, \varepsilon_0]$, expectations of hitting times are finite, i.e,

$$0 < E_{ij}(\varepsilon) < \infty,\ i, j \in \mathbb{X}. \tag{11.74}$$

We also denote by ${}_{\langle k,r\rangle}\tau_j^{(\varepsilon)}$ the hitting time to the state $j \in {}_{\langle k,r\rangle}\mathbb{X}$ for the reduced semi-Markov birth-death process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$.

The following theorem, which proof can be found, for example, in Silvestrov and Manca [38], plays the key role in what follows.

Theorem 11.1 *Let conditions* **A**–**C** *hold for the semi-Markov birth-death process* $\eta^{(\varepsilon)}(t)$. *Then, for any state* $j \in {}_{\langle k,r\rangle}\mathbb{X}$, *the first hitting times* $\tau_j^{(\varepsilon)}$ *and* ${}_{\langle k,r\rangle}\tau_j^{(\varepsilon)}$ *to the state* j, *respectively, for semi-Markov processes* $\eta^{(\varepsilon)}(t)$ *and* ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$, *coincide, and, thus, the expectations of hitting times* $E_{ij}(\varepsilon) = \mathsf{E}_i \tau_j^{(\varepsilon)} = \mathsf{E}_i\, {}_{\langle k,r\rangle}\tau_j^{(\varepsilon)}$, *for any* $i, j \in {}_{\langle k,r\rangle}\mathbb{X}$ *and* $\varepsilon \in (0, \varepsilon_0]$.

11.5.3 Sequential Reduction of Phase Space for Semi-Markov Birth-Death Processes

Let us consider the case, where the left end state 0 is excluded from the phase space $\mathbb{X}$. In this case, the reduced phase space ${}_{\langle 1,N\rangle}\mathbb{X} = \{1, \ldots, N\}$.

We assume that the initial distribution of the semi-Markov process $\eta^{(\varepsilon)}(t)$ is concentrated on the reduced phase space ${}_{\langle 1,N\rangle}\mathbb{X}$.

The transition probabilities of the reduced semi-Markov process ${}_{\langle 1,N\rangle}\eta^{(\varepsilon)}(t)$ have, for every $\varepsilon \in (0, \varepsilon_0]$, the following form, for $t \geq 0$,

$${}_{\langle 1,N\rangle}Q_{ij}^{(\varepsilon)}(t) = \begin{cases} F_{1,+}^{(\varepsilon)}(t)\, p_{1,+}(\varepsilon) & \text{if } j = 2, i = 1, \\ {}_{\langle 1,N\rangle}F_{1,-}^{(\varepsilon)}(t)\, p_{1,-}(\varepsilon) & \text{if } j = 1, i = 1, \\ F_{i,\pm}^{(\varepsilon)}(t)\, p_{i,\pm}(\varepsilon) & \text{if } j = i \pm 1, 1 < i < N, \\ F_{N,\pm}^{(\varepsilon)}(t)\, p_{N,\pm}(\varepsilon) & \text{if } j = N - \frac{1\mp 1}{2}, i = N, \\ 0 & \text{otherwise,} \end{cases} \tag{11.75}$$

where

$${}_{\langle 1,N\rangle}F_{1,-}^{(\varepsilon)}(t) = \sum_{n=0}^{\infty} F_{1,-}^{(\varepsilon)}(t) * F_{0,-}^{(\varepsilon)*n}(t) * F_{0,+}^{(\varepsilon)}(t) \cdot p_{0,-}(\varepsilon)^n p_{0,+}(\varepsilon). \tag{11.76}$$

This relation implies, for every $\varepsilon \in (0, \varepsilon_0]$, the following relation for transition probabilities of the reduced embedded Markov chain ${}_{\langle 1,N\rangle}\eta_n^{(\varepsilon)}$,

$$
{}_{\langle 1,N\rangle}p_{ij}^{(\varepsilon)} = \begin{cases} {}_{\langle 1,N\rangle}p_{1,\pm}(\varepsilon) = p_{1,\pm}(\varepsilon) & \text{if } j = 1 + \frac{1\pm 1}{2}, i = 1, \\ {}_{\langle 1,N\rangle}p_{i,+}(\varepsilon) = p_{i,\pm}(\varepsilon) & \text{if } j = i \pm 1, 1 < i < N, \\ {}_{\langle 1,N\rangle}p_{N,+}(\varepsilon) = p_{N,\pm}(\varepsilon) & \text{if } j = N - \frac{1\mp 1}{2}, i = N, \\ 0 & \text{otherwise,} \end{cases} \tag{11.77}
$$

and the following relation for transition expectations of the reduced embedded semi-Markov process ${}_{\langle 1,N\rangle}\eta^{(\varepsilon)}(t)$,

$$
{}_{\langle 1,N\rangle}e_{ij}^{(\varepsilon)} = \begin{cases} {}_{\langle 1,N\rangle}e_{1,+}(\varepsilon) = e_{1,+}(\varepsilon) & \text{if } j = 2, i = 1, \\ {}_{\langle 1,N\rangle}e_{1,-}(\varepsilon) = e_{1,-}(\varepsilon) & \\ \quad + e_0(\varepsilon) \cdot \frac{p_{1,-}(\varepsilon)}{p_{0,+}(\varepsilon)} & \text{if } j = 1, i = 1, \\ {}_{\langle 1,N\rangle}e_{i,\pm}(\varepsilon) = e_{i,\pm}(\varepsilon) & \text{if } j = i \pm 1, 1 < i < N, \\ {}_{\langle 1,N\rangle}e_{N,\pm}(\varepsilon) = e_{N,\pm}(\varepsilon) & \text{if } j = N - \frac{1\mp 1}{2}, i = N, \\ 0 & \text{otherwise.} \end{cases} \tag{11.78}
$$

Note that, by Theorem 11.1, the following relation takes place, for $i, j \in {}_{\langle 1,N\rangle}\mathbb{X}$ and every $\varepsilon \in (0, \varepsilon_0]$,

$$
\mathsf{E}_i \tau_j^{(\varepsilon)} = \mathsf{E}_i \, {}_{\langle 1,N\rangle}\tau_j^{(\varepsilon)}. \tag{11.79}
$$

Analogously, the right end state N can be excluded from the phase space $\mathbb{X}$. In this case, the reduced phase space ${}_{\langle 0,N-1\rangle}\mathbb{X} = \{0, \ldots, N-1\}$.

As was mentioned above, the reduced semi-Markov processes ${}_{\langle 1,N\rangle}\eta^{(\varepsilon)}(t)$ and ${}_{\langle 0,N-1\rangle}\eta^{(\varepsilon)}(t)$ also have a birth-death type.

Let $0 \le k \le i \le r \le N$. The states $0, \ldots, k-1$ and $N, \ldots, r+1$ can be sequentially excluded from the phase space $\mathbb{X}$ of the semi-Markov process $\eta^{(\varepsilon)}(t)$.

Let us describe the corresponding recurrent procedure.

The reduced semi-Markov process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$ can be obtained by excluding the state $k-1$ from the phase space ${}_{\langle k-1,j\rangle}\mathbb{X}$ of the reduced semi-Markov process ${}_{\langle k-1,r\rangle}\eta^{(\varepsilon)}(t)$ or by excluding state $r+1$ from the phase space ${}_{\langle k,r+1\rangle}\mathbb{X}$ of the reduced semi-Markov process ${}_{\langle k,r+1\rangle}\eta^{(\varepsilon)}(t)$.

The sequential exclusion of the states $0, \ldots, k-1$ and $N, \ldots, r+1$ can be realized in an arbitrary order of choice of one of these sequences and then by excluding the corresponding next state from the chosen sequence.

The simplest variants for the sequences of excluded states are $0, \ldots, k-1, N, \ldots, r+1$ and $N, \ldots, r+1, 0, \ldots, k-1$.

The resulting reduced semi-Markov process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$ will be the same and it will have a birth-death type.

Here, we also accept the reduced semi-Markov process ${}_{\langle i,i\rangle}\eta^{(\varepsilon)}(t)$ with one-state phase space ${}_{\langle i,i\rangle}\mathbb{X} = \{i\}$ as a semi-Markov birth-death process.

This process has transition probability for the embedded Markov chain,

$$ {}_{\langle i,i\rangle}p_{ii}^{(\varepsilon)} = {}_{\langle i,i\rangle}p_{i,+}(\varepsilon) + {}_{\langle i,i\rangle}p_{i,-}(\varepsilon) = 1, \tag{11.80} $$

and the semi-Markov transition probabilities,

$$ \begin{aligned} {}_{\langle i,i\rangle}Q_{ii}^{(\varepsilon)}(t) &= {}_{\langle i,i\rangle}F_{i,+}^{(\varepsilon)}(t)\,{}_{\langle i,i\rangle}p_{i,+}(\varepsilon) + {}_{\langle i,i\rangle}F_{i,-}^{(\varepsilon)}(t)\,{}_{\langle i,i\rangle}p_{i,-}(\varepsilon). \\ &= \mathsf{P}_i\{\tau_i^{(\varepsilon)} \le t\}. \end{aligned} \tag{11.81} $$

The following relations, which are, in fact, variants of relations (11.77) and (11.78), express the transition probabilities ${}_{\langle k,r\rangle}p_{ij}^{(\varepsilon)}$ and the expectations of transition times ${}_{\langle k,r\rangle}e_{ij}^{(\varepsilon)}$ for the reduced semi-Markov process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$, via the transition probabilities ${}_{\langle k-1,r\rangle}p_{ij}^{(\varepsilon)}$ and the expectations of transition times ${}_{\langle k-1,r\rangle}e_{ij}^{(\varepsilon)}$ for the reduced semi-Markov process ${}_{\langle k-1,r\rangle}\eta^{(\varepsilon)}(t)$, for $1 \le k \le r \le N$ and, for every $\varepsilon \in (0, \varepsilon_0]$,

$$ {}_{\langle k,r\rangle}p_{ij}^{(\varepsilon)} = \begin{cases} {}_{\langle k,r\rangle}p_{k,\pm}(\varepsilon) = {}_{\langle k-1,r\rangle}p_{k,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = k + \frac{1\pm1}{2}, i = k, \\ {}_{\langle k,r\rangle}p_{i,\pm}(\varepsilon) = {}_{\langle k-1,r\rangle}p_{i,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = i \pm 1, k < i < r, \\ {}_{\langle k,r\rangle}p_{r,\pm}(\varepsilon) = {}_{\langle k-1,r\rangle}p_{r,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = r - \frac{1\mp1}{2}, i = r, \\ 0 \qquad\qquad \text{otherwise}, \end{cases} \tag{11.82} $$

and

$$ {}_{\langle k,r\rangle}e_{ij}^{(\varepsilon)} = \begin{cases} {}_{\langle k,r\rangle}e_{k,+}(\varepsilon) = {}_{\langle k-1,r\rangle}e_{k,+}(\varepsilon) \\ \qquad\qquad \text{if } j = k+1, i = k, \\ {}_{\langle k,r\rangle}e_{k,-}(\varepsilon) = {}_{\langle k-1,r\rangle}e_{k,-}(\varepsilon) \\ \qquad\qquad + {}_{\langle k-1,r\rangle}e_{k-1}(\varepsilon) \cdot \frac{{}_{\langle k-1,r\rangle}p_{k,-}(\varepsilon)}{{}_{\langle k-1,r\rangle}p_{k-1,+}(\varepsilon)} \\ \qquad\qquad \text{if } j = k, i = k, \\ {}_{\langle k,r\rangle}e_{i,\pm}(\varepsilon) = {}_{\langle k-1,r\rangle}e_{i,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = i \pm 1, k < i < r, \\ {}_{\langle k,r\rangle}e_{r,\pm}(\varepsilon) = {}_{\langle k-1,r\rangle}e_{r,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = r - \frac{1\mp1}{2}, i = r, \\ 0 \qquad\qquad \text{otherwise}, \end{cases} \tag{11.83} $$

where

$$ {}_{\langle k,r\rangle}e_i(\varepsilon) = {}_{\langle k,r\rangle}e_{i,+}(\varepsilon) + {}_{\langle k,r\rangle}e_{i,-}(\varepsilon). \tag{11.84} $$

The transition probabilities ${}_{\langle k,r\rangle}p_{ij}^{(\varepsilon)}$ and the expectations of transition times ${}_{\langle k,r\rangle}e_{ij}^{(\varepsilon)}$ for the reduced semi-Markov process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$ can also be expressed via the transition probabilities ${}_{\langle k,r+1\rangle}p_{ij}^{(\varepsilon)}$ and the expectations of transition times ${}_{\langle k,r+1\rangle}e_{ij}^{(\varepsilon)}$ for the reduced semi-Markov process ${}_{\langle k,r+1\rangle}\eta^{(\varepsilon)}(t)$, for $0 \le k \le r \le N-1$ in an analogous way.

11.5.4 Explicit Formulas for Expectations of Hitting Times for Semi-Markov Birth-Death Processes

By iterating the recurrent formulas (11.82)–(11.83) and their right hand analogues, we get the following explicit formulas for the transition probabilities ${}_{\langle k,r\rangle}p_{ij}^{(\varepsilon)}$ and the expectations of transition times ${}_{\langle k,r\rangle}e_{ij}^{(\varepsilon)}$ for the reduced semi-Markov process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$ expressed in terms of the transition characteristics for the initial semi-Markov process $\eta^{(\varepsilon)}(t)$, for $0 \le k \le r \le N$ and, for every $\varepsilon \in (0, \varepsilon_0]$,

$$
{}_{\langle k,r\rangle}p_{ij}^{(\varepsilon)} = \begin{cases} {}_{\langle k,r\rangle}p_{k,\pm}(\varepsilon) = p_{k,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = k + \frac{1\pm 1}{2}, i = k, \\ {}_{\langle k,r\rangle}p_{i,\pm}(\varepsilon) = p_{i,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = i \pm 1, k < i < r, \\ {}_{\langle k,r\rangle}p_{r,+}(\varepsilon) = p_{r,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = r - \frac{1\mp 1}{2}, i = r, \\ 0 \qquad\qquad \text{otherwise.} \end{cases} \tag{11.85}
$$

and

$$
{}_{\langle k,r\rangle}e_{ij}^{(\varepsilon)} = \begin{cases} {}_{\langle k,r\rangle}e_{k,+}(\varepsilon) = e_{k,+}(\varepsilon) \\ \qquad\qquad \text{if } j = k+1, i = k, \\ {}_{\langle k,r\rangle}e_{k,-}(\varepsilon) = e_{k,-}(\varepsilon) + e_{k-1}(\varepsilon) \cdot \frac{p_{k,-}(\varepsilon)}{p_{k-1,+}(\varepsilon)} \\ \qquad\qquad + \cdots + e_0(\varepsilon) \cdot \frac{p_{1,-}(\varepsilon)\cdots p_{k,-}(\varepsilon)}{p_{0,+}(\varepsilon)\cdots p_{k-1,+}(\varepsilon)} \\ \qquad\qquad \text{if } j = k, i = k, \\ {}_{\langle k,r\rangle}e_{i,+}(\varepsilon) = e_{i,\pm}(\varepsilon) \\ \qquad\qquad \text{if } j = i \pm 1, k < i < r, \\ {}_{\langle k,r\rangle}e_{r,+}(\varepsilon) = e_{r,+}(\varepsilon) + e_{r+1}(\varepsilon) \cdot \frac{p_{r,+}(\varepsilon)}{p_{r+1,-}(\varepsilon)} \\ \qquad\qquad + \cdots + e_N(\varepsilon) \cdot \frac{p_{N-1,+}(\varepsilon)\cdots p_{r,+}(\varepsilon)}{p_{N,-}(\varepsilon)\cdots p_{r+1,-}(\varepsilon)} \\ \qquad\qquad \text{if } j = r, i = r, \\ {}_{\langle k,r\rangle}e_{r,-}(\varepsilon) = e_{r,-}(\varepsilon) \\ \qquad\qquad \text{if } j = r-1, i = r, \\ 0 \qquad\qquad \text{otherwise.} \end{cases} \tag{11.86}
$$

Recall that ${}_{\langle k,r\rangle}\tau_j^{(\varepsilon)}$ is the hitting time for the state $j \in {}_{\langle k,r\rangle}\mathbb{X}$ for the reduced semi-Markov process ${}_{\langle k,r\rangle}\eta^{(\varepsilon)}(t)$.

By Theorem 11.1, the following relation takes place, for $i, j \in {}_{\langle k,r\rangle}\mathbb{X}$ and, for every $\varepsilon \in (0, \varepsilon_0]$,

$$
\mathsf{E}_i \tau_j^{(\varepsilon)} = \mathsf{E}_i \, {}_{\langle k,r\rangle}\tau_j^{(\varepsilon)}. \tag{11.87}
$$

Let us now choose $k = r = i \in \mathbb{X}$. In this case, the reduced phase space ${}_{\langle i,i\rangle}\mathbb{X} = \{i\}$ is a one-state set. In this case, the process ${}_{\langle i,i\rangle}\eta^{(\varepsilon)}(t)$ returns to the state i after every jump. This implies that, in this case, for every $\varepsilon \in (0, \varepsilon_0]$,

$$E_{ii}(\varepsilon) = \mathsf{E}_i \tau_i^{(\varepsilon)} = \mathsf{E}_{i\,\langle i,i\rangle} \tau_i^{(\varepsilon)} = {}_{\langle i,i\rangle} e_i(\varepsilon). \tag{11.88}$$

Thus, the following formulas take place, for every $i \in \mathbb{X}$ and, for every $\varepsilon \in (0, \varepsilon_0]$,

$$\begin{aligned} E_{ii}(\varepsilon) &= e_i(\varepsilon) \\ &+ e_{i-1}(\varepsilon)\frac{p_{i,-}(\varepsilon)}{p_{i-1,+}(\varepsilon)} + e_{i-2}(\varepsilon)\frac{p_{i-1,-}(\varepsilon)p_{i,-}(\varepsilon)}{p_{i-2,+}(\varepsilon)p_{i-1,+}(\varepsilon)} \\ &+ \cdots + e_0(\varepsilon)\frac{p_{1,-}(\varepsilon)p_{2,-}(\varepsilon)\cdots p_{i,-}(\varepsilon)}{p_{0,+}(\varepsilon)p_{1,+}(\varepsilon)\cdots p_{i-1,+}(\varepsilon)} \\ &+ e_{i+1}(\varepsilon)\frac{p_{i,+}(\varepsilon)}{p_{i+1,-}(\varepsilon)} + e_{i+2}(\varepsilon)\frac{p_{i+1,+}(\varepsilon)p_{i,+}(\varepsilon)}{p_{i+2,-}(\varepsilon)p_{i+1,-}(\varepsilon)} \\ &+ \cdots + e_N(\varepsilon)\frac{p_{N-1,+}(\varepsilon)p_{N-2,+}(\varepsilon)\cdots p_{i,+}(\varepsilon)}{p_{N,-}(\varepsilon)p_{N-1,-}(\varepsilon)\cdots p_{i+1,-}(\varepsilon)}. \end{aligned} \tag{11.89}$$

In what follows, we use the following well known formula for the stationary probabilities $\pi_i(\varepsilon)$, $i \in \mathbb{X}$, which takes place, for every $\varepsilon \in (0, \varepsilon_0]$,

$$\pi_i(\varepsilon) = \frac{e_i(\varepsilon)}{E_{ii}(\varepsilon)},\ i \in \mathbb{X}. \tag{11.90}$$

It should be noted that such formulas for stationary distributions of Markov birth-death chains are well known and can be found, for example, in Feller [11]. In context of our studies, a special value has the presented above recurrent algorithm for getting such formulas, based on sequential reduction of the phase space for semi-Markov birth-death processes.

As far as explicit expressions for conditional quasi-stationary probabilities are concerned, they can be obtained by substituting stationary probabilities $\pi_i(\varepsilon)$, $i \in \mathbb{X}$ given by formula (11.90) into formulas (11.37) and (11.38).

11.6 First and Second Order Asymptotic Expansions

In this section, we give explicit the first and the second order asymptotic expansions for stationary and conditional quasi-stationary distributions for perturbed semi-Markov birth-death processes.

The results of the present section are based on the explicit formula (11.89) for expected return times and the expressions which connect these quantities with stationary and conditional quasi-stationary distributions given respectively in formulas (11.90) and (11.37)–(11.38). We obtain the first and second order asymptotic expansions from these formulas by using operational rules for Laurent asymptotic expansions presented in Lemmas 11.1 and 11.2 given below.

It will be convenient to use the following notation,

$$\Gamma_{i,j,\pm}(\varepsilon) = p_{i,\pm}(\varepsilon)p_{i+1,\pm}(\varepsilon)\cdots p_{j,\pm}(\varepsilon),\ 0 \le i \le j \le N. \tag{11.91}$$

Using (11.91), we can write formula (11.89) as

$$E_{ii}(\varepsilon) = e_i(\varepsilon) + \sum_{k=0}^{i-1} e_k(\varepsilon)\frac{\Gamma_{k+1,i,-}(\varepsilon)}{\Gamma_{k,i-1,+}(\varepsilon)} + \sum_{k=i+1}^{N} e_k(\varepsilon)\frac{\Gamma_{i,k-1,+}(\varepsilon)}{\Gamma_{i+1,k,-}(\varepsilon)},\ i \in \mathbb{X}. \tag{11.92}$$

In particular, we have

$$E_{00}(\varepsilon) = e_0(\varepsilon) + \sum_{k \in {}_0\mathbb{X}} e_k(\varepsilon)\frac{\Gamma_{0,k-1,+}(\varepsilon)}{\Gamma_{1,k,-}(\varepsilon)}, \tag{11.93}$$

and

$$E_{NN}(\varepsilon) = e_N(\varepsilon) + \sum_{k \in {}_N\mathbb{X}} e_k(\varepsilon)\frac{\Gamma_{k+1,N,-}(\varepsilon)}{\Gamma_{k,N-1,+}(\varepsilon)}. \tag{11.94}$$

We will compute the desired asymptotic expansions by applying operational rules for Laurent asymptotic expansions in relations (11.92)–(11.94). In order for the presentation to not be too repetitive, we will directly compute the second order asymptotic expansions which contain the first order asymptotic expansions as special cases. In particular, this gives us limits for stationary and conditional quasi-stationary distributions.

The formulas for computing the asymptotic expansions are different depending on whether condition $\mathbf{H}_1$, $\mathbf{H}_2$ or $\mathbf{H}_3$ holds. We consider these three cases in Sects. 11.6.2, 11.6.3 and 11.6.4, respectively.

Each of these sections will have the same structure: First, we present a ma which successively constructs asymptotic expansions for the quantities given in relations (11.92)–(11.94). Then, using these expansions, we construct the first and the second order asymptotic expansions for stationary (Sects. 11.6.2–11.6.4) and conditional quasi-stationary distributions (Sects. 11.6.3–11.6.4).

11.6.1 Laurent Asymptotic Expansions

In this subsection, we present some operational rules for Laurent asymptotic expansions given in Silvestrov, D. and Silvestrov, S. [39–41] which are used in the present paper for constructions of asymptotic expansions for stationary and conditional quasi-stationary distributions of perturbed semi-Markov birth-death processes.

A real-valued function $A(\varepsilon)$, defined on an interval $(0, \varepsilon_0]$ for some $0 < \varepsilon_0 \le 1$, is a Laurent asymptotic expansion if it can be represented in the following form, $A(\varepsilon) = a_{h_A}\varepsilon^{h_A} + \cdots + a_{k_A}\varepsilon^{k_A} + o_A(\varepsilon^{k_A})$, $\varepsilon \in (0, \varepsilon_0]$, where **(a)** $-\infty < h_A \le k_A <$

∞ are integers, **(b)** coefficients $a_{h_A}, \dots, a_{k_A}$ are real numbers, **(c)** the function $o_A(\varepsilon^{k_A})/\varepsilon^{k_A} \to 0$ as $\varepsilon \to 0$. Such an expansion is pivotal if it is known that $a_{h_A} \neq 0$.

The above paper presents operational rules for Laurent asymptotic expansions. Let us shortly formulate some of these rules, in particular, for summation, multiplication and division of Laurent asymptotic expansions.

Lemma 11.1 *Let* $A(\varepsilon) = a_{h_A}\varepsilon^{h_A} + \cdots + a_{k_A}\varepsilon^{k_A} + o_A(\varepsilon^{k_A})$ *and* $B(\varepsilon) = b_{h_B}\varepsilon^{h_B} + \cdots + b_{k_B}\varepsilon^{k_B} + o_B(\varepsilon^{k_B})$ *be two pivotal Laurent asymptotic expansions. Then*:

(i) $C(\varepsilon) = cA(\varepsilon) = c_{h_C}\varepsilon^{h_C} + \cdots + c_{k_C}\varepsilon^{k_C} + o_C(\varepsilon^{k_C})$, *where a constant* $c \neq 0$, *is a pivotal Laurent asymptotic expansion and* $h_C = h_A, k_C = k_A, c_{h_C+r} = ca_{h_C+r}, r = 0, \dots, k_C - h_C$,

(ii) $D(\varepsilon) = A(\varepsilon) + B(\varepsilon) = d_{h_D}\varepsilon^{h_D} + \cdots + d_{k_D}\varepsilon^{k_D} + o_D(\varepsilon^{k_D})$ *is a pivotal Laurent asymptotic expansion and* $h_D = h_A \wedge h_B, k_D = k_A \wedge k_B, d_{h_D+r} = a_{h_D+r} + b_{h_D+r}, r = 0, \dots, k_D - h_D$, *where* $a_{h_D+r} = 0, r < h_A - h_D, b_{h_D+r} = 0, r < h_B - h_D$,

(iii) $E(\varepsilon) = A(\varepsilon) \cdot B(\varepsilon) = e_{h_E}\varepsilon^{h_E} + \cdots + e_{k_E}\varepsilon^{k_E} + o_E(\varepsilon^{k_E})$ *is a pivotal Laurent asymptotic expansion and* $h_E = h_A + h_B, k_E = (k_A + h_B) \wedge (k_B + h_A), e_{h_E+r} = \sum_{l=0}^{r} a_{h_A+l} \cdot b_{h_B+r-l}, r = 0, \dots, k_E - h_E$,

(iv) $F(\varepsilon) = A(\varepsilon)/B(\varepsilon) = f_{h_F}\varepsilon^{h_F} + \cdots + f_{k_F}\varepsilon^{k_F} + o_F(\varepsilon^{k_F})$ *ia a pivotal Laurent asymptotic expansion and* $h_F = h_A - h_B, k_F = (k_A - h_B) \wedge (k_B - 2h_B + h_A), f_{h_F+r} = \frac{1}{b_{h_B}}(a_{h_A+r} + \sum_{l=1}^{r} b_{h_B+l} \cdot f_{h_F+r-l}), r = 0, \dots, k_F - h_F$.

The following lemma presents useful multiple generalizations of summation and multiplication rules given in Lemma 11.1.

Lemma 11.2 *Let* $A_i(\varepsilon) = a_{i,h_{A_i}}\varepsilon^{h_A} + \cdots + a_{i,k_{A_i}}\varepsilon^{k_{A_i}} + o_{A_i}(\varepsilon^{k_{A_i}}), i = 1, \dots, m$ *be pivotal Laurent asymptotic expansions. Then*:

(i) $D(\varepsilon) = \sum_{i=1}^{m} A_i(\varepsilon) = d_{h_D}\varepsilon^{h_D} + \cdots + d_{k_D}\varepsilon^{k_D} + o_D(\varepsilon^{k_D})$ *is a pivotal Laurent asymptotic expansion and* $h_D = \min_{1 \le l \le m} h_{A_l}$, $k_D = \min_{1 \le l \le m} k_{A_l}$, $d_{h_D+l} = a_{1,h_D+l} + \cdots + a_{m,h_D+l}$, $l = 0, \dots, k_D - h_D$, *where* $a_{i,h_D+l} = 0$ *for* $0 \le l < h_{A_i} - h_D, i = 1, \dots, m$,

(ii) $E(\varepsilon) = \prod_{i=1}^{m} A_i(\varepsilon) = e_{h_E}\varepsilon^{h_E} + \cdots + e_{k_E}\varepsilon^{k_E} + o_E(\varepsilon^{k_E})$ *is a pivotal Laurent asymptotic expansion and*
$h_E = \sum_{l=1}^{m} h_{A_l}$, $k_E = \min_{1 \le l \le m}(k_{A_l} + \sum_{1 \le r \le m, r \neq l} h_{A_r})$,
$e_{h_E+l} = \sum_{l_1+\cdots+l_m=l, 0 \le l_i \le k_{A_i} - h_{A_i}, i=1,\dots,m} \prod_{1 \le i \le m} a_{i,h_{A_i}+l_i}$,
$l = 0, \dots, k_E - h_E$.

11.6.2 First and Second Order Asymptotic Expansions for Stationary Distributions Under Condition $\mathbf{H}_1$

In the case where condition $\mathbf{H}_1$ holds, the semi-Markov process has no asymptotically absorbing states. In this case, all quantities in relations (11.92)–(11.94) are of order $O(1)$ and the construction of asymptotic expansions are rather straightforward.

In the following lemma we successively construct asymptotic expansions for the quantities given in relations (11.92)–(11.94).

Lemma 11.3 *Assume that conditions* **A**–**G** *and* $\mathbf{H}_1$ *hold. Then:*

(i) *For* $i \in \mathbb{X}$, *we have,*

$$e_i(\varepsilon) = b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $\dot{o}_i(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$b_i[0] = b_{i,-}[0] + b_{i,+}[0] > 0, \quad b_i[1] = b_{i,-}[1] + b_{i,+}[1].$$

(ii) *For* $0 \le i \le j \le N$, *we have,*

$$\Gamma_{i,j,\pm}(\varepsilon) = A_{i,j,\pm}[0] + A_{i,j,\pm}[1]\varepsilon + o_{i,j,\pm}(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $o_{i,j,\pm}(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$A_{i,j,\pm}[0] = a_{i,\pm}[0]a_{i+1,\pm}[0]\cdots a_{j,\pm}[0] > 0,$$

$$A_{i,j,\pm}[1] = \sum_{n_i+n_{i+1}+\cdots+n_j=1} a_{i,\pm}[n_i]a_{i+1,\pm}[n_{i+1}]\cdots a_{j,\pm}[n_j].$$

(iii) *For* $0 \le k \le i-1,\ i \in {}_{\langle 1,N\rangle}\mathbb{X}$, *we have*

$$\frac{\Gamma_{k+1,i,-}(\varepsilon)}{\Gamma_{k,i-1,+}(\varepsilon)} = A^*_{k,i}[0] + A^*_{k,i}[1]\varepsilon + o^*_{k,i}(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $o^*_{k,i}(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$A^*_{k,i}[0] = \frac{A_{k+1,i,-}[0]}{A_{k,i-1,+}[0]} > 0,$$

$$A^*_{k,i}[1] = \frac{A_{k+1,i,-}[1]A_{k,i-1,+}[0] - A_{k+1,i,-}[0]A_{k,i-1,+}[1]}{A_{k,i-1,+}[0]^2}.$$

(iv) *For* $i+1 \le k \le N,\ i \in {}_{\langle 0,N-1\rangle}\mathbb{X}$, *we have*

$$\frac{\Gamma_{i,k-1,+}(\varepsilon)}{\Gamma_{i+1,k,-}(\varepsilon)} = A^*_{k,i}[0] + A^*_{k,i}[1]\varepsilon + o^*_{k,i}(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $o^*_{k,i}(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$A^*_{k,i}[0] = \frac{A_{i,k-1,+}[0]}{A_{i+1,k,-}[0]} > 0,$$

$$A^*_{k,i}[1] = \frac{A_{i,k-1,+}[1]A_{i+1,k,-}[0] - A_{i,k-1,+}[0]A_{i+1,k,-}[1]}{A_{i+1,k,-}[0]^2}.$$

(**v**) *For $i \in \mathbb{X}$, we have*

$$E_{ii}(\varepsilon) = B_{ii}[0] + B_{ii}[1]\varepsilon + \dot{o}_{ii}(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $\dot{o}_{ii}(\varepsilon)/\varepsilon \to 0$ as $\varepsilon \to 0$ and

$$B_{ii}[0] = b_i[0] + \sum_{k \in {}_i\mathbb{X}} b_k[0]A^*_{k,i}[0] > 0,$$

$$B_{ii}[1] = b_i[1] + \sum_{k \in {}_i\mathbb{X}} (b_k[0]A^*_{k,i}[1] + b_k[1]A^*_{k,i}[0]).$$

Proof Since $e_i(\varepsilon) = e_{i,-}(\varepsilon) + e_{i,+}(\varepsilon), i \in \mathbb{X}$, part (**i**) follows immediately from condition **E**.

For the proof of part (**ii**) we notice that it follows from the definition (11.91) of $\Gamma_{i,j,\pm}(\varepsilon)$ and condition **D** that

$$\Gamma_{i,j,\pm}(\varepsilon) = \prod_{k=i}^{j} (a_{k,\pm}[0] + a_{k,\pm}[1]\varepsilon + o_{k,\pm}(\varepsilon)),\ 0 \le i \le j \le N.$$

By applying the multiple product rule for asymptotic expansions, we obtain the asymptotic relation given in part (**ii**) where the coefficients $A_{i,j,\pm}[0], 0 \le i \le j \le n$, are positive since condition $\mathbf{H}_1$ holds.

In order to prove parts (**iii**) and (**iv**) we use the result in part (**ii**). For $0 \le k \le i-1$, $i \in {}_0\mathbb{X}$, this gives us

$$\frac{\Gamma_{k+1,i,-}(\varepsilon)}{\Gamma_{k,i-1,+}(\varepsilon)} = \frac{A_{k+1,i,-}[0] + A_{k+1,i,-}[1]\varepsilon + o_{k+1,i,-}(\varepsilon)}{A_{k,i-1,+}[0] + A_{k,i-1,+}[1]\varepsilon + o_{k,i-1,+}(\varepsilon)}, \tag{11.95}$$

and, for $i+1 \le k \le N$, $i \in {}_N\mathbb{X}$, we get

$$\frac{\Gamma_{i,k-1,+}(\varepsilon)}{\Gamma_{i+1,k,-}(\varepsilon)} = \frac{A_{i,k-1,+}[0] + A_{i,k-1,+}[1]\varepsilon + o_{i,k-1,+}(\varepsilon)}{A_{i+1,k,-}[0] + A_{i+1,k,-}[1]\varepsilon + o_{i+1,k,-}(\varepsilon)}. \tag{11.96}$$

Using the division rule for asymptotic expansions in relations (11.95) and (11.96) we get the asymptotic expansions given in parts (**iii**) and (**iv**).

Finally, we can use relation (11.92) to prove part (**v**). This relation together with the results in parts (**i**), (**iii**) and (**iv**) yield

$$E_{ii}(\varepsilon) = b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon) + \sum_{k \in \mathbb{X}\setminus\{i\}} (b_k[0] + b_k[1]\varepsilon + \dot{o}_k(\varepsilon)) \times (A^*_{k,i}[0] + A^*_{k,i}[1]\varepsilon + o^*_{k,i}(\varepsilon)),\ i \in \mathbb{X}.$$

A combination of the product rule and the multiple summation rule for asymptotic expansions gives the asymptotic relation in part (**v**). □

The following theorem gives second order asymptotic expansions for stationary probabilities. In particular, this theorem shows that there exist limits for stationary probabilities,

$$\pi_i(0) = \lim_{\varepsilon \to 0} \pi_i(\varepsilon),\ i \in \mathbb{X},$$

where $\pi_i(0) > 0$, $i \in \mathbb{X}$.

Theorem 11.2 *Assume that conditions* **A**–**G** *and* $\mathbf{H}_1$ *hold. Then, we have the following asymptotic relation for the stationary probabilities* $\pi_i(\varepsilon)$, $i \in \mathbb{X}$,

$$\pi_i(\varepsilon) = c_i[0] + c_i[1]\varepsilon + o_i(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $o_i(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$c_i[0] = \frac{b_i[0]}{B_{ii}[0]} > 0, \quad c_i[1] = \frac{b_i[1]B_{ii}[0] - b_i[0]B_{ii}[1]}{B_{ii}[0]^2},$$

where $B_{ii}[0]$, $B_{ii}[1]$, $i \in \mathbb{X}$, *can be computed from the formulas given in Lemma 11.3.*

Proof It follows from condition **E** and part (**v**) of Lemma 11.3 that, for $i \in \mathbb{X}$,

$$\pi_i(\varepsilon) = \frac{e_i(\varepsilon)}{E_{ii}(\varepsilon)} = \frac{b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon)}{B_{ii}[0] + B_{ii}[1]\varepsilon + \dot{o}_{ii}(\varepsilon)}.$$

The result now follows from the division rule for asymptotic expansions **(iv)**, given in Lemma 11.1. □

11.6.3 First and Second Order Asymptotic Expansions for Stationary and Conditional Quasi-stationary Distributions Under Condition $\mathbf{H}_2$

In the case where condition $\mathbf{H}_2$ holds, the semi-Markov process has one asymptotically absorbing state, namely state 0. This means that $p_{0,+}(\varepsilon) \sim O(\varepsilon)$ and since this quantity is involved in relations (11.92)–(11.94), the pivotal properties of the

expansions are less obvious. Furthermore, since some terms now tends to infinity, we partly need to operate with Laurent asymptotic expansions.

In order to separate cases where $i = 0$ or $i \in {}_{\langle 1,N\rangle}\mathbb{X}$ we will use the indicator function $\gamma_i = I(i = 0)$, that is, $\gamma_0 = 1$ and $\gamma_i = 0$ for $i \in {}_{\langle 1,N\rangle}\mathbb{X}$.

The following lemma gives asymptotic expansions for quantities in relations (11.92)–(11.94).

Lemma 11.4 *Assume that conditions* **A**–**G** *and* $\mathbf{H}_2$ *hold. Then:*

(i) *For* $i \in \mathbb{X}$, *we have,* $e_i(\varepsilon) = b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon)$, $\varepsilon \in (0, \varepsilon_0]$, *where* $\dot{o}_i(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$b_i[0] = b_{i,-}[0] + b_{i,+}[0] > 0, \quad b_i[1] = b_{i,-}[1] + b_{i,+}[1].$$

(ii) *For* $0 \le i \le j \le N$, *we have,*

$$\Gamma_{i,j,+}(\varepsilon) = A_{i,j,+}[\gamma_i]\varepsilon^{\gamma_i} + A_{i,j,+}[\gamma_i + 1]\varepsilon^{\gamma_i+1} + o_{i,j,+}(\varepsilon^{\gamma_i+1}),\ \varepsilon \in (0, \varepsilon_0],$$

where $o_{i,j,+}(\varepsilon^{\gamma_i+1})/\varepsilon^{\gamma_i+1} \to 0$ *as* $\varepsilon \to 0$ *and*

$$A_{i,j,+}[\gamma_i] = a_{i,+}[\gamma_i]a_{i+1,+}[0]\cdots a_{j,+}[0] > 0,$$

$$A_{i,j,+}[\gamma_i + 1] = \sum_{n_i+n_{i+1}+\cdots+n_j=1} a_{i,+}[\gamma_i + n_i]a_{i+1,+}[n_{i+1}]\cdots a_{j,+}[n_j].$$

(iii) *For* $0 \le i \le j \le N$, *we have,*

$$\Gamma_{i,j,-}(\varepsilon) = A_{i,j,-}[0] + A_{i,j,-}[1]\varepsilon + o_{i,j,-}(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $o_{i,j,-}(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$A_{i,j,-}[0] = a_{i,-}[0]a_{i+1,-}[0]\cdots a_{j,-}[0] > 0,$$

$$A_{i,j,-}[1] = \sum_{n_i+n_{i+1}+\cdots+n_j=1} a_{i,-}[n_i]a_{i+1,-}[n_{i+1}]\cdots a_{j,-}[n_j].$$

(iv) *For* $0 \le k \le i-1$, $i \in {}_{\langle 1,N\rangle}\mathbb{X}$, *we have,*

$$\frac{\Gamma_{k+1,i,-}(\varepsilon)}{\Gamma_{k,i-1,+}(\varepsilon)} = A^*_{k,i}[-\gamma_k]\varepsilon^{-\gamma_k} + A^*_{k,i}[-\gamma_k + 1]\varepsilon^{-\gamma_k+1} + o^*_{k,i}(\varepsilon^{-\gamma_k+1}),\ \varepsilon \in (0, \varepsilon_0],$$

where $o^*_{k,i}(\varepsilon^{-\gamma_k+1})/\varepsilon^{-\gamma_k+1} \to 0$ *as* $\varepsilon \to 0$ *and*

$$A^*_{k,i}[-\gamma_k] = \frac{A_{k+1,i,-}[0]}{A_{k,i-1,+}[\gamma_k]} > 0,$$

$$A^*_{k,i}[-\gamma_k + 1] = \frac{A_{k+1,i,-}[1]A_{k,i-1,+}[\gamma_k] - A_{k+1,i,-}[0]A_{k,i-1,+}[\gamma_k + 1]}{A_{k,i-1,+}[\gamma_k]^2}.$$

(v) *For $i + 1 \le k \le N$, $i \in {}_{\langle 0,N-1\rangle}\mathbb{X}$, we have,*

$$\frac{\Gamma_{i,k-1,+}(\varepsilon)}{\Gamma_{i+1,k,-}(\varepsilon)} = A^*_{k,i}[\gamma_i]\varepsilon^{\gamma_i} + A^*_{k,i}[\gamma_i + 1]\varepsilon^{\gamma_i+1} + o^*_{k,i}(\varepsilon^{\gamma_i+1}),\ \varepsilon \in (0, \varepsilon_0],$$

where $o^*_{k,i}(\varepsilon^{\gamma_i+1})/\varepsilon^{\gamma_i+1} \to 0$ as $\varepsilon \to 0$ and

$$A^*_{k,i}[\gamma_i] = \frac{A_{i,k-1,+}[\gamma_i]}{A_{i+1,k,-}[0]} > 0,$$

$$A^*_{k,i}[\gamma_i + 1] = \frac{A_{i,k-1,+}[\gamma_i + 1]A_{i+1,k,-}[0] - A_{i,k-1,+}[\gamma_i]A_{i+1,k,-}[1]}{A_{i+1,k,-}[0]^2}.$$

(vi) *For $i \in \mathbb{X}$, we have,*

$$E_{ii}(\varepsilon) = B_{ii}[\gamma_i - 1]\varepsilon^{\gamma_i-1} + B_{ii}[\gamma_i]\varepsilon^{\gamma_i} + \dot{o}_{ii}(\varepsilon^{\gamma_i}),\ \varepsilon \in (0, \varepsilon_0],$$

where $\dot{o}_{ii}(\varepsilon^{\gamma_i})/\varepsilon^{\gamma_i} \to 0$ as $\varepsilon \to 0$ and

$$B_{00}[0] = b_0[0] > 0, \quad B_{00}[1] = b_0[1] + \sum_{k \in {}_0\mathbb{X}} b_k[0]A^*_{k,0}[1],$$

$$B_{ii}[-1] = b_0[0]A^*_{0,i}[-1] > 0,\ i \in {}_{\langle 1,N\rangle}\mathbb{X},$$

$$B_{ii}[0] = b_0[1]A^*_{0,i}[-1] + b_i[0] + \sum_{k \in {}_i\mathbb{X}} b_k[0]A^*_{k,i}[0],\ i \in {}_{\langle 1,N\rangle}\mathbb{X}.$$

Proof Let us first note that the quantities in parts (**i**) and (**iii**) do not depend on $p_{0,+}(\varepsilon)$, so the proofs for these parts are the same as the proofs for parts (**i**) and (**ii**) in Lemma 11.3, respectively.

We now prove part (**ii**). Notice that it follows from conditions **D** and $\mathbf{H}_2$ that $p_{i,+}(\varepsilon) = a_{i,+}[\gamma_i]\varepsilon^{\gamma_i} + a_{i,+}[\gamma_i + 1]\varepsilon^{\gamma_i+1} + o_{i,+}(\varepsilon^{\gamma_i+1})$, $i \in \mathbb{X}$. Using this and the definition (11.91) of $\Gamma_{i,j,+}(\varepsilon)$ gives

$$\begin{aligned}\Gamma_{i,j,+}(\varepsilon) &= (a_{i,+}[\gamma_i]\varepsilon^{\gamma_i} + a_{i,+}[\gamma_i+1]\varepsilon^{\gamma_i+1} + o_{i,+}(\varepsilon^{\gamma_i+1})) \\ &\quad \times (a_{i+1,+}[0] + a_{i+1,+}[1]\varepsilon + o_{i+1,+}(\varepsilon)) \\ &\quad \times \cdots \times \\ &\quad \times (a_{j,+}[0] + a_{j,+}[1]\varepsilon + o_{j,+}(\varepsilon)),\ 0 \le i \le j \le N.\end{aligned}$$

An application of the multiple product rule for asymptotic expansions shows that part (**ii**) holds.

Now, using the results in parts (**ii**) and (**iii**) we get, for $0 \le k \le i-1$, $i \in {}_{\langle 1,N\rangle}\mathbb{X}$,

$$\frac{\Gamma_{k+1,i,-}(\varepsilon)}{\Gamma_{k,i-1,+}(\varepsilon)} = \frac{A_{k+1,i,-}[0] + A_{k+1,i,-}[1]\varepsilon + o_{k+1,i,-}(\varepsilon)}{A_{k,i-1,+}[\gamma_k]\varepsilon^{\gamma_k} + A_{k,i-1,+}[\gamma_k+1]\varepsilon^{\gamma_k+1} + o_{k,i-1,+}(\varepsilon^{\gamma_k+1})},$$

and, for $i+1 \le k \le N$, $i \in {}_{N}\mathbb{X}$,

$$\frac{\Gamma_{i,k-1,+}(\varepsilon)}{\Gamma_{i+1,k,-}(\varepsilon)} = \frac{A_{i,k-1,+}[\gamma_i]\varepsilon^{\gamma_i} + A_{i,k-1,+}[\gamma_i+1]\varepsilon^{\gamma_i+1} + o_{i,k-1,+}(\varepsilon^{\gamma_i+1})}{A_{i+1,k,-}[0] + A_{i+1,k,-}[1]\varepsilon + o_{i+1,k,-}(\varepsilon)}.$$

Notice that it is possible that the quantity in the first of the above equations tends to infinity as $\varepsilon \to 0$. Applying the division rule for Laurent asymptotic expansions in the above two relations yields the asymptotic relations given in parts (**iv**) and (**v**).

In order to prove part (**vi**), we consider the cases $i = 0$ and $i \in {}_{\langle 1,N\rangle}\mathbb{X}$ separately. First note that it follows from relation (11.93) and the results in parts (**i**) and (**iv**) that

$$\begin{aligned}E_{00}(\varepsilon) &= b_0[0] + b_0[1]\varepsilon + \dot{o}_0(\varepsilon) \\ &\quad + \sum_{k \in {}_0\mathbb{X}} (b_k[0] + b_k[1]\varepsilon + \dot{o}_k(\varepsilon))(A^*_{k,0}[1]\varepsilon + A^*_{k,0}[2]\varepsilon^2 + o^*_{k,0}(\varepsilon^2)).\end{aligned}$$

Using the product rule and the multiple summation rule for asymptotic expansions we obtain the asymptotic relation in part (**vi**) for the case $i = 0$.

If $i \in {}_{\langle 1,N\rangle}\mathbb{X}$, relation (11.92) implies together with parts (**i**), (**iv**) and (**v**) that

$$\begin{aligned}E_{ii}(\varepsilon) &= b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon) \\ &\quad + (b_0[0] + b_0[1]\varepsilon + \dot{o}_0(\varepsilon))(A^*_{0,i}[-1]\varepsilon^{-1} + A^*_{0,i}[0] + o^*_{0,i}(1)) \\ &\quad + \sum_{k \in {}_{0,i}\mathbb{X}} (b_k[0] + b_k[1]\varepsilon + \dot{o}_k(\varepsilon))(A^*_{k,i}[0] + A^*_{k,i}[1]\varepsilon + o^*_{k,i}(\varepsilon)).\end{aligned}$$

Notice that the term corresponding to $k = 0$ is of order $O(\varepsilon^{-1})$ while all other terms in the sum are of order $O(1)$. We can again apply the product rule and multiple summation rule for Laurent asymptotic expansions and in this case, the asymptotic relation in part (**vi**) is obtained for $i \in {}_{\langle 1,N\rangle}\mathbb{X}$. □

The following theorem gives second order asymptotic expansions for stationary and conditional quasi-stationary probabilities. In particular, part (**i**) of this theorem

shows that there exist limits for stationary probabilities, $\pi_i(0) = \lim_{\varepsilon \to 0} \pi_i(\varepsilon), i \in \mathbb{X}$, where $\pi_0(0) = 1$ and $\pi_i(0) = 0$ for $i \in {}_{\langle 1,N \rangle}\mathbb{X}$.

Furthermore, part (**ii**) of the theorem shows, in particular, that there exist limits for conditional quasi-stationary probabilities $\tilde{\pi}_i(0) = \lim_{\varepsilon \to 0} \tilde{\pi}_i(\varepsilon)$, $i \in {}_{\langle 1,N \rangle}\mathbb{X}$, where $\tilde{\pi}_i(0) > 0, i \in {}_{\langle 1,N \rangle}\mathbb{X}$.

Theorem 11.3 *Assume that conditions* **A**–**G** *and* $\mathbf{H}_2$ *hold. Then:*

(**i**) *We have the following asymptotic relation for the stationary probabilities* $\pi_i(\varepsilon)$, $i \in \mathbb{X}$,

$$\pi_i(\varepsilon) = c_i[\tilde{l}_i]\varepsilon^{\tilde{l}_i} + c_i[\tilde{l}_i + 1]\varepsilon^{\tilde{l}_i+1} + o_i(\varepsilon^{\tilde{l}_i+1}),\ \varepsilon \in (0, \varepsilon_0],$$

where $\tilde{l}_i = I(i \neq 0)$, $o_i(\varepsilon^{\tilde{l}_i+1})/\varepsilon^{\tilde{l}_i+1} \to 0$ *as* $\varepsilon \to 0$, *and*

$$c_i[\tilde{l}_i] = \frac{b_i[0]}{B_{ii}[-\tilde{l}_i]} > 0, \quad c_i[\tilde{l}_i + 1] = \frac{b_i[1]B_{ii}[-\tilde{l}_i] - b_i[0]B_{ii}[-\tilde{l}_i + 1]}{B_{ii}[-\tilde{l}_i]^2},$$

where $B_{ii}[-1]$, $i \in {}_{\langle 1,N \rangle}\mathbb{X}$, $B_{ii}[0]$, $i \in \mathbb{X}$, *and* $B_{00}[1]$, *can be computed from the formulas given in Lemma 11.4.*

(**ii**) *We have the following asymptotic relation for the conditional quasi-stationary probabilities* $\tilde{\pi}_i(\varepsilon)$, $i \in {}_{\langle 1,N \rangle}\mathbb{X}$,

$$\tilde{\pi}_i(\varepsilon) = \tilde{c}_i[0] + \tilde{c}_i[1]\varepsilon + \tilde{o}_i(\varepsilon),\ \varepsilon \in (0, \varepsilon_0],$$

where $\tilde{o}_i(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$\tilde{c}_i[0] = \frac{c_i[1]}{d[1]} > 0, \quad \tilde{c}_i[1] = \frac{c_i[2]d[1] - c_i[1]d[2]}{d[1]^2},$$

where $d[l] = \sum_{j \in {}_0\mathbb{X}} c_i[l],\ l = 1, 2.$

Proof It follows from parts (**i**) and (**vi**) in Lemma 11.4 that, for $i \in \mathbb{X}$,

$$\pi_i(\varepsilon) = \frac{e_i(\varepsilon)}{E_{ii}(\varepsilon)} = \frac{b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon)}{B_{ii}[\gamma_i - 1]\varepsilon^{\gamma_i - 1} + B_{ii}[\gamma_i]\varepsilon^{\gamma_i} + \dot{o}_{ii}(\varepsilon^{\gamma_i})}. \tag{11.97}$$

We also have $\gamma_i = I(i = 0) = 1 - I(i \neq 0) = 1 - \tilde{l}_i$. By changing the indicator function and then using the division rule for Laurent asymptotic expansions in relation (11.97), we obtain the asymptotic expansion given in part (**i**).

In order to prove part (**ii**) we first use part (**i**) for $i \in {}_{\langle 1,N \rangle}\mathbb{X}$ to get

$$\tilde{\pi}_i(\varepsilon) = \frac{\pi_i(\varepsilon)}{\sum_{j \in {}_0\mathbb{X}} \pi_j(\varepsilon)} = \frac{c_i[1]\varepsilon + c_i[2]\varepsilon^2 + o_i(\varepsilon^2)}{\sum_{j \in {}_0\mathbb{X}}(c_j[1]\varepsilon + c_j[2]\varepsilon^2 + o_j(\varepsilon^2))},$$

and then we apply the multiple summation rule **(i)** for asymptotic expansions, given in Lemma 11.2, and the division rule for asymptotic expansions, given in Lemma 11.1. □

11.6.4 First and Second Order Asymptotic Expansions for Stationary and Conditional Quasi-stationary Distributions Under Condition $\mathbf{H_3}$

In the case where condition $\mathbf{H}_3$ holds, both state 0 and state N are asymptotically absorbing for the semi-Markov process. This means that $p_{0,+}(\varepsilon) \sim O(\varepsilon)$ and $p_{N,-}(\varepsilon) \sim O(\varepsilon)$ which makes the asymptotic analysis of relations (11.92)–(11.94) even more involved.

In order to separate cases where $i = 0$, $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$ or $i = N$, we will use the indicator functions $\gamma_i = I(i = 0)$ and $\beta_i = I(i = N)$.

The following lemma gives asymptotic expansions for quantities given in relations (11.92)–(11.94).

Lemma 11.5 *Assume that conditions* **A**–**G** *and* $\mathbf{H}_3$ *hold. Then:*

(i) *For* $i \in \mathbb{X}$, *we have* $e_i(\varepsilon) = b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon)$, $\varepsilon \in (0, \varepsilon_0]$, *where* $\dot{o}_i(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$b_i[0] = b_{i,-}[0] + b_{i,+}[0] > 0, \quad b_i[1] = b_{i,-}[1] + b_{i,+}[1].$$

(ii) *For* $0 \le i \le j \le N$, *we have*

$$\Gamma_{i,j,+}(\varepsilon) = A_{i,j,+}[\gamma_i]\varepsilon^{\gamma_i} + A_{i,j,+}[\gamma_i + 1]\varepsilon^{\gamma_i+1} + o_{i,j,+}(\varepsilon^{\gamma_i+1}), \ \varepsilon \in (0, \varepsilon_0],$$

where $o_{i,j,+}(\varepsilon^{\gamma_i+1})/\varepsilon^{\gamma_i+1} \to 0$ *as* $\varepsilon \to 0$ *and*

$$A_{i,j,+}[\gamma_i] = a_{i,+}[\gamma_i]a_{i+1,+}[0] \cdots a_{j,+}[0] > 0,$$

$$A_{i,j,+}[\gamma_i + 1] = \sum_{n_i+n_{i+1}+\cdots+n_j=1} a_{i,+}[\gamma_i + n_i]a_{i+1,+}[n_{i+1}] \cdots a_{j,+}[n_j].$$

(iii) *For* $0 \le i \le j \le N$, *we have*

$$\Gamma_{i,j,-}(\varepsilon) = A_{i,j,-}[\beta_j]\varepsilon^{\beta_j} + A_{i,j,-}[\beta_j + 1]\varepsilon^{\beta_j+1} + o_{i,j,-}(\varepsilon^{\beta_j+1}), \ \varepsilon \in (0, \varepsilon_0],$$

where $o_{i,j,-}(\varepsilon^{\beta_j+1})/\varepsilon^{\beta_j+1} \to 0$ *as* $\varepsilon \to 0$ *and*

$$A_{i,j,-}[\beta_j] = a_{i,-}[0] \cdots a_{j-1,-}[0]a_{j,-}[\beta_j] > 0,$$

$$A_{i,j,-}[\beta_j+1]=\sum_{n_i+\cdots+n_{j-1}+n_j=1} a_{i,-}[n_i]\cdots a_{j-1,-}[n_{j-1}]a_{j,-}[\beta_j+n_j].$$

(**iv**) *For* $0\le k\le i-1,\ i\in {}_{\langle 1,N\rangle}\mathbb{X}$, *we have*

$$\frac{\Gamma_{k+1,i,-}(\varepsilon)}{\Gamma_{k,i-1,+}(\varepsilon)}=A^*_{k,i}[\beta_i-\gamma_k]\varepsilon^{\beta_i-\gamma_k}+A^*_{k,i}[\beta_i-\gamma_k+1]\varepsilon^{\beta_i-\gamma_k+1}+o^*_{k,i}(\varepsilon^{\beta_i-\gamma_k+1}),\ \varepsilon\in(0,\varepsilon_0],$$

where $o^*_{k,i}(\varepsilon^{\beta_i-\gamma_k+1})/\varepsilon^{\beta_i-\gamma_k+1}\to 0$ *as* $\varepsilon\to 0$ *and*

$$A^*_{k,i}[\beta_i-\gamma_k]=\frac{A_{k+1,i,-}[\beta_i]}{A_{k,i-1,+}[\gamma_k]}>0,$$

$$A^*_{k,i}[\beta_i-\gamma_k+1]=\frac{A_{k+1,i,-}[\beta_i+1]A_{k,i-1,+}[\gamma_k]-A_{k+1,i,-}[\beta_i]A_{k,i-1,+}[\gamma_k+1]}{A_{k,i-1,+}[\gamma_k]^2}.$$

(**v**) *For* $i+1\le k\le N,\ i\in {}_{\langle 0,N-1\rangle}\mathbb{X}$, *we have*

$$\frac{\Gamma_{i,k-1,+}(\varepsilon)}{\Gamma_{i+1,k,-}(\varepsilon)}=A^*_{k,i}[\gamma_i-\beta_k]\varepsilon^{\gamma_i-\beta_k}+A^*_{k,i}[\gamma_i-\beta_k+1]\varepsilon^{\gamma_i-\beta_k+1}+o^*_{k,i}(\varepsilon^{\gamma_i-\beta_k+1}),\ \varepsilon\in(0,\varepsilon_0],$$

where $o^*_{k,i}(\varepsilon^{\gamma_i-\beta_k+1})/\varepsilon^{\gamma_i-\beta_k+1}\to 0$ *as* $\varepsilon\to 0$ *and*

$$A^*_{k,i}[\gamma_i-\beta_k]=\frac{A_{i,k-1,+}[\gamma_i]}{A_{i+1,k,-}[\beta_k]}>0,$$

$$A^*_{k,i}[\gamma_i-\beta_k+1]=\frac{A_{i,k-1,+}[\gamma_i+1]A_{i+1,k,-}[\beta_k]-A_{i,k-1,+}[\gamma_i]A_{i+1,k,-}[\beta_k+1]}{A_{i+1,k,-}[\beta_k]^2}.$$

(**vi**) *For* $i\in\mathbb{X}$, *we have*

$$E_{ii}(\varepsilon)=B_{ii}[\gamma_i+\beta_i-1]\varepsilon^{\gamma_i+\beta_i-1}+B_{ii}[\gamma_i+\beta_i]\varepsilon^{\gamma_i+\beta_i}+\dot o_{ii}(\varepsilon^{\gamma_i+\beta_i}),\ \varepsilon\in(0,\varepsilon_0],$$

where $\dot o_{ii}(\varepsilon^{\gamma_i+\beta_i})/\varepsilon^{\gamma_i+\beta_i}\to 0$ *as* $\varepsilon\to 0$ *and*

$$B_{ii}[0]=b_i[0]+b_{N-i}[0]A^*_{N-i,i}[0]>0,\ i=0,N,$$

$$B_{ii}[1] = b_{N-i}[1]A^*_{N-i,i}[0] + b_i[1] + \sum_{k \in {}_i\mathbb{X}} b_k[0]A^*_{k,i}[1],\ i = 0, N,$$

$$B_{ii}[-1] = b_0[0]A^*_{0,i}[-1] + b_N[0]A^*_{N,i}[-1] > 0,\ i \in {}_{\langle 1,N-1\rangle}\mathbb{X},$$

$$B_{ii}[0] = b_0[1]A^*_{0,i}[-1] + b_N[1]A^*_{N,i}[-1] + b_i[0]$$

$$+ \sum_{k \in {}_i\mathbb{X}} b_k[0]A^*_{k,i}[0],\ i \in {}_{\langle 1,N-1\rangle}\mathbb{X}.$$

Proof We first note that the quantities in parts (**i**) and (**ii**) do not depend on $p_{N,-}(\varepsilon)$, so the proofs for these parts are the same as the proofs for parts (**i**) and (**ii**) in Lemma 11.4, respectively.

In order to prove part (**iii**) we notice that it follows from conditions **D** and $\mathbf{H}_3$ that $p_{i,-}(\varepsilon) = a_{i,-}[\beta_i]\varepsilon^{\beta_i} + a_{i,-}[\beta_i + 1]\varepsilon^{\beta_i+1} + o_{i,-}(\varepsilon^{\beta_i+1})$, $i \in \mathbb{X}$. From this and the definition (11.91) of $\Gamma_{i,j,-}(\varepsilon)$ we get, for $0 \le i \le j \le N$,

$$\begin{aligned}\Gamma_{i,j,-}(\varepsilon) &= (a_{i,-}[0] + a_{i,-}[1]\varepsilon + o_{i,-}(\varepsilon)) \times \cdots \times \\ &\times (a_{j-1,-}[0] + a_{j-1,-}[1]\varepsilon + o_{j-1,-}(\varepsilon)) \\ &\times (a_{j,-}[\beta_j]\varepsilon^{\beta_j} + a_{j,-}[\beta_j + 1] + \varepsilon^{\beta_j+1} + o_{j,-}(\varepsilon^{\beta_j+1})).\end{aligned}$$

By applying the multiple product rule for asymptotic expansions we obtain the asymptotic relation given in part (**iii**).

From parts (**ii**) and (**iii**) it follows that, for $0 \le k \le i-1$, $i \in {}_{\langle 1,N\rangle}\mathbb{X}$,

$$\frac{\Gamma_{k+1,i,-}(\varepsilon)}{\Gamma_{k,i-1,+}(\varepsilon)} = \frac{A_{k+1,i,-}[\beta_i]\varepsilon^{\beta_i} + A_{k+1,i,-}[\beta_i + 1]\varepsilon^{\beta_i+1} + o_{k+1,i,-}(\varepsilon^{\beta_i+1})}{A_{k,i-1,+}[\gamma_k]\varepsilon^{\gamma_k} + A_{k,i-1,+}[\gamma_k + 1]\varepsilon^{\gamma_k+1} + o_{k,i-1,+}(\varepsilon^{\gamma_k+1})},$$

and, for $i + 1 \le k \le N$, $i \in {}_{\langle 0,N-1\rangle}\mathbb{X}$,

$$\frac{\Gamma_{i,k-1,+}(\varepsilon)}{\Gamma_{i+1,k,-}(\varepsilon)} = \frac{A_{i,k-1,+}[\gamma_i]\varepsilon^{\gamma_i} + A_{i,k-1,+}[\gamma_i + 1]\varepsilon^{\gamma_i+1} + o_{i,k-1,+}(\varepsilon^{\gamma_i+1})}{A_{i+1,k,-}[\beta_k]\varepsilon^{\beta_k} + A_{i+1,k,-}[\beta_k + 1]\varepsilon^{\beta_k+1} + o_{i+1,k,-}(\varepsilon^{\beta_k+1})}.$$

Notice that in the above two relations it is possible that the corresponding quantity tends to infinity as $\varepsilon \to 0$.

The asymptotic relations given in parts (**iv**) and (**v**) are obtained by using the division rule for Laurent asymptotic expansions in the above two relations.

We finally give the proof of part (**vi**). For the case $i = 0$, it follows from relation (11.93) and parts (**i**) and (**v**) that

$$E_{00}(\varepsilon) = b_0[0] + b_0[1]\varepsilon + \dot{o}_0(\varepsilon)$$
$$+ (b_N[0] + b_N[1]\varepsilon + \dot{o}_N(\varepsilon))(A^*_{N,0}[0] + A^*_{N,0}[1]\varepsilon + o^*_{N,0}(\varepsilon))$$
$$+ \sum_{k \in {}_{0,N}\mathbb{X}} (b_k[0] + b_k[1]\varepsilon + \dot{o}_k(\varepsilon))(A^*_{k,0}[1]\varepsilon + A^*_{k,0}[2]\varepsilon^2 + o^*_{k,0}(\varepsilon^2)).$$

The product rule and multiple summation rule for asymptotic expansions now proves part (**vi**) for the case $i = 0$.

If $i = N$, it follows from relation (11.94) and parts (**i**) and (**iv**) that

$$E_{NN}(\varepsilon) = b_N[0] + b_N[1]\varepsilon + \dot{o}_N(\varepsilon)$$
$$+ (b_0[0] + b_0[1]\varepsilon + \dot{o}_0(\varepsilon))(A^*_{0,N}[0] + A^*_{0,N}[1]\varepsilon + o^*_{0,N}(\varepsilon))$$
$$+ \sum_{k \in {}_{0,N}\mathbb{X}} (b_k[0] + b_k[1]\varepsilon + \dot{o}_k(\varepsilon))(A^*_{k,N}[1]\varepsilon + A^*_{k,N}[2]\varepsilon^2 + o^*_{k,N}(\varepsilon^2)).$$

Again, we can use the product rule and multiple summation rule in order to prove part (**vi**), in this case, for $i = N$.

For the case where $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$, we use relation (11.92) and parts (**i**), (**iv**) and (**v**) to get

$$E_{ii}(\varepsilon) = b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon)$$
$$+ \sum_{k \in \{0,N\}} (b_k[0] + b_k[1]\varepsilon + \dot{o}_k(\varepsilon))(A^*_{k,i}[-1]\varepsilon^{-1} + A^*_{k,i}[0] + o^*_{k,i}(1))$$
$$+ \sum_{k \in {}_{0,i,N}\mathbb{X}} (b_k[0] + b_k[1]\varepsilon + \dot{o}_k(\varepsilon))(A^*_{k,i}[0] + A^*_{k,i}[1]\varepsilon + o^*_{k,i}(\varepsilon)).$$

Here we can note that the terms corresponding to $k \in \{0, N\}$ are of order $O(\varepsilon^{-1})$ while all other terms are of order $O(1)$. By using the product rule and multiple summation rule for Laurent asymptotic expansions, we conclude that the asymptotic relation given in part (**vi**) also holds for $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$. □

The following theorem gives second order asymptotic expansions for stationary and conditional quasi-stationary probabilities. In particular, part (**i**) of this theorem shows that there exist limits for stationary probabilities, $\pi_i(0) = \lim_{\varepsilon \to 0} \pi_i(\varepsilon), i \in \mathbb{X}$, where $\pi_0(0) > 0, \pi_N(0) > 0$, and $\pi_i(0) = 0$ for $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$. Furthermore, part (**ii**) of the theorem shows, in particular, that there exist limits for conditional quasi-stationary probabilities, $\hat{\pi}_i(0) = \lim_{\varepsilon \to 0} \hat{\pi}_i(\varepsilon)$, $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$, where $\hat{\pi}_i(0) > 0$, $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$.

Theorem 11.4 *Assume that conditions* **A**–**G** *and* $\mathbf{H}_3$ *hold. Then:*

(**i**) *We have the following asymptotic relation for the stationary probabilities* $\pi_i(\varepsilon)$, $i \in \mathbb{X}$,

$$\pi_i(\varepsilon) = c_i[\hat{l}_i]\varepsilon^{\hat{l}_i} + c_i[\hat{l}_i + 1]\varepsilon^{\hat{l}_i+1} + o_i(\varepsilon^{\hat{l}_i+1}), \ \varepsilon \in (0, \varepsilon_0],$$

where $\hat{l}_i = I(i \neq 0, N)$, $o_i(\varepsilon^{\hat{l}_i+1})/\varepsilon^{\hat{l}_i+1} \to 0$ *as* $\varepsilon \to 0$, *and*

$$c_i[\hat{l}_i] = \frac{b_i[0]}{B_{ii}[-\hat{l}_i]} > 0, \quad c_i[\hat{l}_i + 1] = \frac{b_i[1]B_{ii}[-\hat{l}_i] - b_i[0]B_{ii}[-\hat{l}_i + 1]}{B_{ii}[-\hat{l}_i]^2},$$

where $B_{ii}[-1]$, $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$, $B_{ii}[0]$, $i \in \mathbb{X}$, *and* $B_{ii}[1]$, $i = 0, N$, *can be computed from the formulas given in Lemma 11.5.*

(ii) *We have the following asymptotic relation for the conditional quasi-stationary probabilities,* $\hat{\pi}_i(\varepsilon)$, $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$,

$$\hat{\pi}_i(\varepsilon) = \hat{c}_i[0] + \hat{c}_i[1]\varepsilon + \hat{o}_i(\varepsilon), \ \varepsilon \in (0, \varepsilon_0],$$

where $\hat{o}_i(\varepsilon)/\varepsilon \to 0$ *as* $\varepsilon \to 0$ *and*

$$\hat{c}_i[0] = \frac{c_i[1]}{d[1]} > 0, \quad \hat{c}_i[1] = \frac{c_i[2]d[1] - c_i[1]d[2]}{d[1]^2},$$

where $d[l] = \sum_{j \in {}_{0,N}\mathbb{X}} c_i[l]$, $l = 1, 2$.

Proof It follows from parts **(i)** and **(vi)** in Lemma 11.5 that, for $i \in \mathbb{X}$,

$$\pi_i(\varepsilon) = \frac{b_i[0] + b_i[1]\varepsilon + \dot{o}_i(\varepsilon)}{B_{ii}[\gamma_i + \beta_i - 1]\varepsilon^{\gamma_i+\beta_i-1} + B_{ii}[\gamma_i + \beta_i]\varepsilon^{\gamma_i+\beta_i} + \dot{o}_{ii}(\varepsilon^{\gamma_i+\beta_i})}. \tag{11.98}$$

We also have $\gamma_i + \beta_i = I(i = 0) + I(i = N) = 1 - I(i \neq 0, N) = 1 - \hat{l}_i$. Using this relation for indicator functions and the division rule for Laurent asymptotic expansions in relation (11.98) we obtain the asymptotic relation given in part **(i)**.

For the proof of part **(ii)**, we first use part **(i)** for $i \in {}_{\langle 1,N-1\rangle}\mathbb{X}$ to get

$$\hat{\pi}_i(\varepsilon) = \frac{\pi_i(\varepsilon)}{\sum_{j \in {}_{0,N}\mathbb{X}} \pi_j(\varepsilon)} = \frac{c_i[1]\varepsilon + c_i[2]\varepsilon^2 + o_i(\varepsilon^2)}{\sum_{j \in {}_{0,N}\mathbb{X}}(c_j[1]\varepsilon + c_j[2]\varepsilon^2 + o_j(\varepsilon^2))},$$

and then we apply the multiple summation rule **(i)** given in Lemma 11.2 and the division rule **(iv)** given in Lemma 11.1. □

11.6.5 Asymptotic Expansions of Higher Orders for Stationary and Conditional Quasi-stationary Distributions

In Sects. 11.6.2–11.6.4, we give asymptotic expansions of the first and second orders (with length, respectively, 1 and 2) for stationary and conditional quasi-stationary distributions of perturbed semi-Markov birth-death processes, under the assumption that conditions **A**–**G** and $\mathbf{H}_i$ $(i = 1, 2, 3)$ hold.

It is readily seen from the proofs of Lemmas 11.3–11.5 and Theorems 11.2–11.4 that the perturbation conditions **D** and **E** can be weaken in the case of first order asymptotics. The asymptotic expansions of the length 2 appearing in these conditions can be replaced by the analogous asymptotic expansions of the length 1. Namely, the upper indices $1 + l_{i,\pm}$ in the sums representing these asymptotic expansions and power indices in the corresponding remainders should be, just, replaced by indices $l_{i,\pm}$.

Moreover, the method of construction of asymptotic expansions for stationary distributions and conditional quasi-stationary distribution of perturbed semi-Markov birth-death processes based on the use of operational rules for Laurent asymptotic expansions presented in Lemmas 11.1 and 11.2 let one also construct the corresponding asymptotic expansions of higher orders, with the length larger than 2. In this case, the asymptotic expansions of the length 2 appearing in the perturbation conditions **D** and **E** should be replaced by the analogous asymptotic expansions of the corresponding length larger than 2. Namely, the upper indices $1 + l_{i,\pm}$ in the sums representing these asymptotic expansions and power indices for the corresponding remainders should be formally replaced by indices $L + l_{i,\pm}$, with parameter $L > 1$. In this case, the length of the corresponding asymptotic expansions will be $L + 1$.

The algorithm for construction of the corresponding asymptotic expansions, with length $L + 1 > 2$ is absolutely analogous to those used in Theorems 11.2–11.4. The difference is that at all steps the asymptotic expansions, with length $L + 1$, are constructed for the corresponding intermediate quantities, $\Gamma_{i,j,\pm}(\varepsilon)$, etc., using operational rules for Laurent asymptotic expansions given in Lemmas 11.1 and 11.2.

This program is realised in book by Silvestrov, D. and Silvestrov, S. [41].

11.7 Numerical Examples

In this section, the results of the present paper are illustrated by numerical examples for some of the perturbed models of birth-death-type discussed in Sect. 11.2. Let us first note that each model presented in Sect. 11.2 is defined in terms of intensities for a continuous time Markov chain and the perturbation scenarios considered give intensities which are linear functions of the perturbation parameter, that is,

$$\lambda_{i,\pm}(\varepsilon) = g_{i,\pm}[0] + g_{i,\pm}[1]\varepsilon, \ i \in \mathbb{X}, \tag{11.99}$$

where the coefficients $g_{i,\pm}[l]$ depend on the model under consideration. Consequently, the higher order ($l \geq 2$) terms in (11.43) all vanish.

In order to use the algorithm based on successive reduction of the phase space, we first need to calculate the coefficients in perturbation conditions **D** and **E**. This can be done from relations (11.5), (11.6) and (11.44) by applying the operational rules for Laurent asymptotic expansions given in Lemmas 11.1 and 11.2. By relation (11.5), we have $\lambda_i(\varepsilon) = \lambda_{i,-}(\varepsilon) + \lambda_{i,+}(\varepsilon)$, so it follows immediately from (11.99) that

$$\lambda_i(\varepsilon) = g_i[0] + g_i[1]\varepsilon,\ i \in \mathbb{X}, \tag{11.100}$$

where $g_i[l] = g_{i,-}[l] + g_{i,+}[l]$, $l = 0, 1$. From (11.6), (11.99), (11.100) and Lemma 11.1 we deduce the following asymptotic series expansions, for $i \in \mathbb{X}$,

$$\begin{aligned} p_{i,\pm}(\varepsilon) &= \frac{\lambda_{i,\pm}(\varepsilon)}{\lambda_i(\varepsilon)} = \frac{g_{i,\pm}[0] + g_{i,\pm}[1]\varepsilon}{g_i[0] + g_i[1]\varepsilon} \\ &= \sum_{l=l_{i,\pm}}^{1+l_{i,\pm}} a_{i,\pm}[l]\varepsilon^l + o_{i,\pm}(\varepsilon^{1+l_{i,\pm}}). \end{aligned} \tag{11.101}$$

The expansion (11.101) exists and its coefficients can be calculated from the division rule for asymptotic expansions. Then, using (11.44), (11.100), (11.101) and Lemma 11.1, the following asymptotic series expansions can be constructed, for $i \in \mathbb{X}$,

$$e_{i,\pm}(\varepsilon) = \frac{p_{i,\pm}(\varepsilon)}{\lambda_i(\varepsilon)} = \sum_{l=l_{i,\pm}}^{1+l_{i,\pm}} b_{i,\pm}[l]\varepsilon^l + \dot{o}_{i,\pm}(\varepsilon^{1+l_{i,\pm}}). \tag{11.102}$$

Once the coefficients in the expansions (11.101) and (11.102) have been calculated, we can use the algorithm described in Sect. 11.6, in order to construct asymptotic expansions for stationary and conditional quasi-stationary probabilities.

The remainder of this section is organised as follows. In Sect. 11.7.1 we illustrate our results with numerical calculations for the perturbed models of population genetics discussed in Sect. 11.2.3. We first consider an example where condition $\mathbf{H}_1$ holds and then an example where condition $\mathbf{H}_3$ is satisfied. Numerical examples for the perturbed model of epidemics presented in Sect. 11.2.2 are discussed in Sect. 11.7.2. This provides an example where condition $\mathbf{H}_2$ holds. All illustrations for the numerical examples are placed in a special subsection at the end of this section for convenience.

11.7.1 Numerical Examples for Perturbed Models of Population Genetics

Recall that the perturbation conditions for the model in Sect. 11.2.3 are formulated in terms of the mutation parameters as

$$U_1(\varepsilon) = C_1 + D_1\varepsilon, \quad U_2(\varepsilon) = C_2 + D_2\varepsilon. \tag{11.103}$$

Additionally, the model depends on the size $N/2$ of the population and the selection parameters S_1 and S_2 which are assumed to be independent of ε. Thus, there are in total seven parameters to choose.

In our first example, we choose the following values for the parameters: $N = 100$, $C_1 = C_2 = 5$, $D_1 = 0$, $D_2 = N$ and $S_1 = S_2 = 0$. Recall that the mutation probabilities are related to the mutation parameters by $u_1(\varepsilon) = U_1(\varepsilon)/N$ and $u_2(\varepsilon) = U_2(\varepsilon)/N$. It follows from (11.103) that $u_1(\varepsilon) = 0.05$ and $u_2(\varepsilon) = 0.05 + \varepsilon$. Thus, in the limiting model, a chosen allele mutates with probability 0.05 for both types A_1 and A_2. In this case, we have no absorbing states which means that condition $\mathbf{H}_1$ holds.

Since we have no selection, the stationary distribution for the limiting model will be symmetric around state 50. The perturbation parameter ε can be interpreted as an increase in the probability that a chosen allele of type A_2 mutates to an allele of type A_1. Increasing the perturbation parameter will shift the mass of the stationary distribution to the right.

With model parameters given above, we first used relations (11.20), (11.21), (11.22), (11.24), and (11.26) to calculate the coefficients in (11.99) for the intensities. Then, these coefficients were used to compute the coefficients in the perturbation conditions **D** and **E** as described above. After this, we used the algorithm outlined in Sect. 11.6 to calculate the asymptotic expansions for the stationary distribution given in Theorem 11.2, with parameters $L = 0, 1$, i.e., lengths $L + 1 = 1, 2$. Moreover, we also computed the analogous asymptotic expansions, with parameter $L = 2, 3$, i.e., lengths $L + 1 = 3, 4$, using the higher order variant of the corresponding algorithm described in Sect. 11.6.5. Approximations for the stationary distribution based on these expansions were obtained by approximating the corresponding remainders by zero.

Let us first compare our approximations with the exact stationary distribution for some particular values of the perturbation parameter. Figure 11.1a shows the stationary distribution for the limiting model ($\varepsilon = 0$) and, as already mentioned above, we see that it is symmetric around state 50. The stationary distribution for the model with $\varepsilon = 0.01$ and the approximation corresponding to $L = 1$ are shown in Fig. 11.1b. Here, the approximation seems the match the exact distribution very well. The approximation for $L = 2$ is not included here since it will not show any visible difference from the exact stationary distribution. In Fig. 11.1c, d, corresponding to the models where $\varepsilon = 0.02$ and $\varepsilon = 0.03$, respectively, we also include the approximations for $L = 2$. As expected, the approximations for the stationary distribution get worse as the perturbation parameter increases. However, it seems that even for higher values of the perturbation parameter, some parts of our approximations fit better to the exact stationary distribution. In this example, it seems that the approximations are in general better for states that belong to the right part of the distribution.

In order to illustrate that the quality of the approximations differs depending on which states we consider, let us compare the stationary probabilities for the states 40 and 80. The stationary probabilities of these two states are approximately of the same magnitude and we can compare them in plots with the same scale on both the horizontal and the vertical axes. Figure 11.2a shows the stationary probability for state 40 as a function of the perturbation parameter and its approximations for $L = 1, 2, 3$. The corresponding quantities for state 80 are shown in Fig. 11.2b where we have omitted the approximation for $L = 3$ since the approximation is very good

already for $L = 2$. When $L = 2$, the approximation for state 80 is clearly better compared to the approximation for state 40.

Another point illustrated by Figs. 11.1 and 11.2 is that for a fixed value of the perturbation parameter, the quality of an approximation based on a higher order asymptotic expansion is not necessarily better. For instance, in Fig. 11.2a we see that for $\varepsilon \in [0.04, 0.05]$ the approximations for $L = 1$ is better compared to both $L = 2$ and $L = 3$. However, asymptotically as $\varepsilon \to 0$, the higher order approximations are better. For example, we see in Fig. 11.2a that when $\varepsilon \in [0, 0.02]$ the approximations for $L = 3$ are the best.

Let us now consider a second example for the perturbed model of population genetics. We now choose the parameters as follows: $N = 100$, $C_1 = C_2 = 0$, $D_1 = D_2 = N$ and $S_1 = S_2 = 0$. In this case, both types of mutations have the same probabilities and are equal to the perturbation parameter, that is, $u_1(\varepsilon) = u_2(\varepsilon) = \varepsilon$. This means that both boundary states will be asymptotically absorbing, so condition $\mathbf{H}_3$ holds. In this case, we calculated the asymptotic expansions for the stationary and conditional quasi-stationary stationary distribution, given in Theorem 11.4.

Let us illustrate the numerical results for conditional quasi-stationary distributions. Figure 11.3a shows the conditional quasi-stationary distribution for $\varepsilon = 0.005$ and some of its approximations. Since it is quite hard to see the details near the boundary states for this plot, we also show the same curves restricted to the states 1–20 in Fig. 11.3b. As in the previous example, it can be seen that the qualities of the approximations differ between the states. In this case, we see that the approximations for states close to the boundary are not as good as for interior states. Similar type of behaviour also appears for different choices of the selection parameters S_1 and S_2. We omit the plots showing this since they do not contribute with more understanding of the model.

Let us instead study the limiting conditional quasi-stationary distributions (11.51) for some different values of the selection parameters S_1 and S_2. These types of distributions are interesting in their own right and are studied, for instance, by Allen and Tarnita [1], where they are called *rare-mutation dimorphic distributions*. In our example, if mutations are rare (i.e., ε is very small), the probabilities of such a distribution can be interpreted as the likelihoods for different allele frequencies to appear during periods of competition which are separated by long periods of fixation.

Figure 11.4a shows the limiting conditional quasi-stationary distribution in the case $S_1 = S_2 = 0$, that is, for a selectively neutral model. Now, let the selection parameters be given by $S_1 = 10$ and $S_2 = -10$. In this case, the gene pairs with genotypes A_1A_1, A_2A_2 and A_1A_2 have survival probabilities approximately equal to 0.37, 0.30 and 0.33, respectively. Thus, allele A_1 has a selective advantage and this is reflected in Fig. 11.4b where the limiting conditional quasi-stationary distribution is shown in this case. The mass of the distribution is now shifted to the right compared to a selectively neutral model. Next, we take the selection parameters as $S_1 = S_2 = 10$ which implies that gene pairs with genotypes A_1A_1, A_2A_2 and A_1A_2 have survival probabilities approximately equal to 0.345, 0.345 and 0.31, respectively. This means that we have a model with underdominance and we see in Fig. 11.4c that the limiting conditional quasi-stationary distribution then has more of its mass near the boundary

compared to a selectively neutral model. Finally, we set the selection parameters as $S_1 = S_2 = -10$. Then, gene pairs with genotypes A_1A_1, A_2A_2 and A_1A_2 have survival probabilities approximately equal to 0.32, 0.32 and 0.36, respectively. This gives us a model with overdominace or balancing selection and in this case we see in Fig. 11.4d that the limiting conditional quasi-stationary distribution has more mass concentrated to the interior states compared to a selectively neutral model.

11.7.2 Numerical Examples for Perturbed Epidemic Models

In our last numerical example, we consider the perturbed epidemic model described in Sect. 11.2.2. Recall from the variant of condition $\mathbf{H}_2$ given in this subsection that the contact rate ν for each individual and the group of infected individuals outside the population is considered as a perturbation parameter, that is, $\nu = \nu(\varepsilon) = \varepsilon$. In this case, state 0 is asymptotically absorbing which means that condition $\mathbf{H}_2$ holds.

It follows directly from (11.15) and (11.16) that the intensities of the Markov chain describing the number of infected individuals are linear functions of ε given by $\lambda_{i,+}(\varepsilon) = \lambda i(1 - i/N) + (N - i)\varepsilon$, $\lambda_{i,-}(\varepsilon) = \mu i$, $i \in \mathbb{X}$. In this model, we only have three parameters to choose: N, λ, and μ. As in the previous examples, let us take $N = 100$ which here corresponds to the size of the population. Furthermore, we let $\mu = 1$ so that the expected time for an infected individual to be infectious is equal to one time unit. Numerical illustrations will be given for the cases where $\lambda = 0.5$ and $\lambda = 1.5$. For the limiting model, we have in the former case that the basic reproduction ratio $R_0 = 0.5$ and in the latter case $R_0 = 1.5$. The properties of the model are quite different depending on which of these two cases we consider. For the two choices of model parameters given above, we calculated asymptotic expansions for stationary and conditional quasi-stationary distributions given in Theorem 11.3.

Let us first compare the limiting conditional quasi-stationary distributions in (11.49). Figure 11.5a shows this distribution for the case where $\lambda = 0.5$ and $\mu = 1$ and in Fig. 11.5b it is shown for the case where $\lambda = 1.5$ and $\mu = 1$. In the former case, the limiting conditional quasi-stationary distribution has most of its mass concentrated near zero and in the latter case the distribution has a shape which resembles a normal curve and most of its mass is distributed on the states between 0 and 60.

We can also study plots of the type given in Figs. 11.1, 11.2 and 11.3. Also in this example, intervals for the perturbation parameter, where the approximations are good, depend on which state is considered. In this case, states close to zero are more sensitive to perturbations. Let us here just show two of the plots for illustration. For the model with $\lambda = 1.5$ and $\mu = 1$, Fig. 11.6a shows the conditional quasi-stationary distribution for $\varepsilon = 0.02$ and the corresponding approximations for $L = 1$ and $L = 2$. For the same model parameters, the quasi-stationary probability for state 10 is shown in Fig. 11.6b as a function of the perturbation parameter together with some of its approximations.

Finally, let us compare the stationary probabilities for state 0. Note that, despite that the limiting conditional quasi-stationary distribution is very different depending

on whether $R_0 = 0.5$ or $R_0 = 1.5$ for the model with $\varepsilon = 0$, the limiting stationary distribution is concentrated at state 0 in both these cases. Figure 11.7a shows the stationary probability of state 0 as a function of the perturbation parameter and some of its approximations in the case where $\lambda = 0.5$ and $\mu = 1$. The corresponding quantities for the case where $\lambda = 1.5$ and $\mu = 1$ are shown in Fig. 11.7b.

Qualitatively the plots show approximately the same behavior, but note that the scales on the horizontal axes are very different. We see that the stationary probability of state 0 for the limiting model is much more sensitive to perturbations in the case where $R_0 = 1.5$. It follows from (11.52) that this is due to fact that the expected time $E_{10}(\varepsilon)$ for the infection to (temporarily) die out after one individual gets infected, is much larger for the model with $R_0 = 1.5$.

Illustrations for Numerical Examples

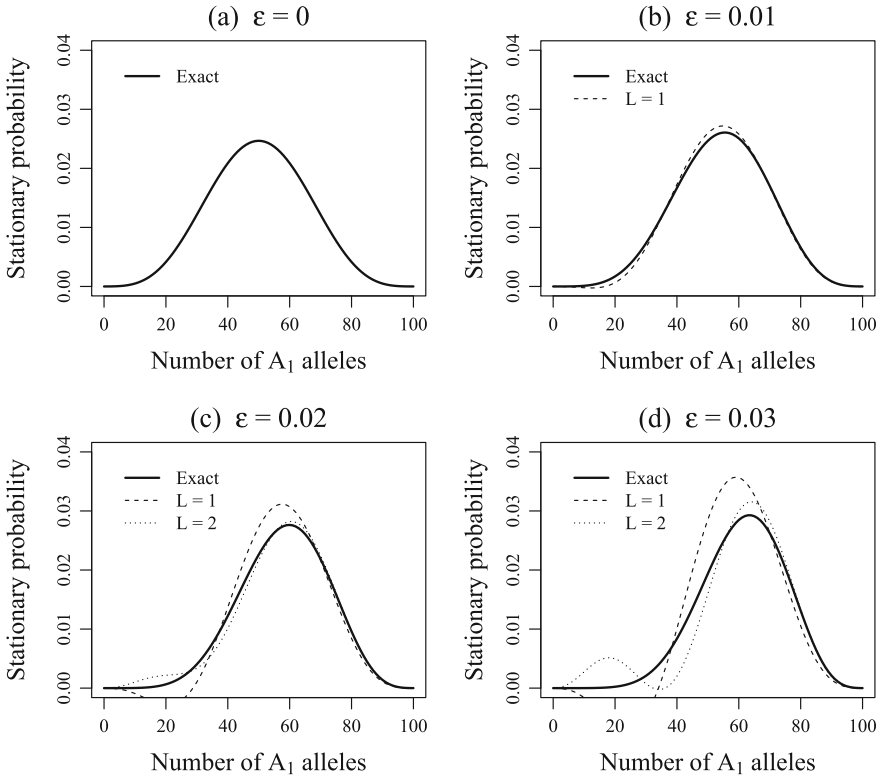


Fig. 11.1 Comparison of the stationary distribution $\pi_i(\varepsilon)$ and some of its approximations for the population genetic example of Sect. 11.2.3. The plots are functions of the number of A_1 alleles i, for different values of the perturbation parameter ε, with $N = 100$, $C_1 = C_2 = 5$, $D_1 = 0$, $D_2 = N$ and $S_1 = S_2 = 0$

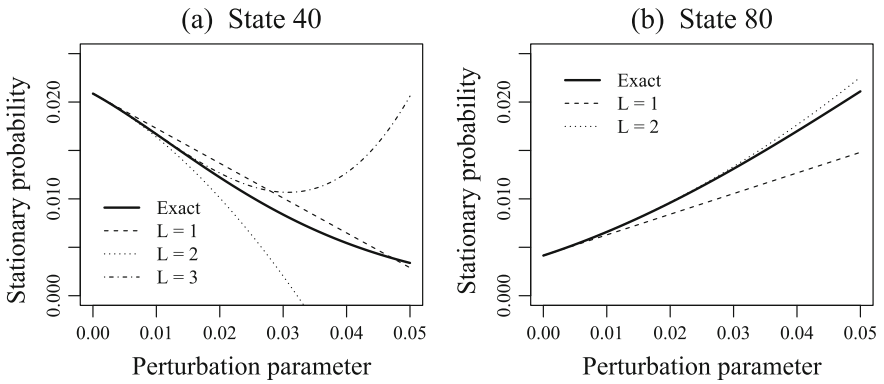


Fig. 11.2 Comparison of stationary probabilities $\pi_i(\varepsilon)$ for states $i = 40$ and $i = 80$ and some of its approximations considered as a function of the perturbation parameter ε. The model is based on the population genetic example of Sect. 11.2.3, with the same parameter values as in Fig. 11.1

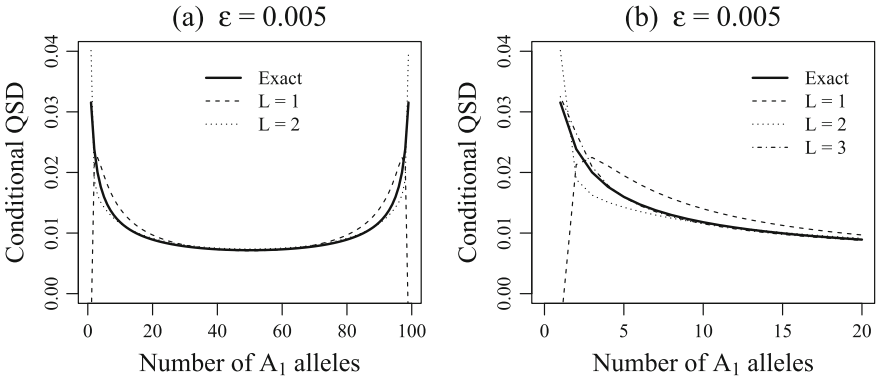


Fig. 11.3 The conditional quasi-stationary distribution $\hat{\pi}_i(\varepsilon)$ and some of its approximations for the population genetic example of Sect. 11.2.3. The plots are functions of the number of A_1 alleles i, with the perturbation parameter $\varepsilon = 0.005$ fixed. Plot **a** shows the distribution for all states while plot **b** is restricted to states 1–20. The parameter values of the model are $N = 100$, $C_1, C_2 = 0$, $D_1, D_2 = N$ and $S_1, S_2 = 0$

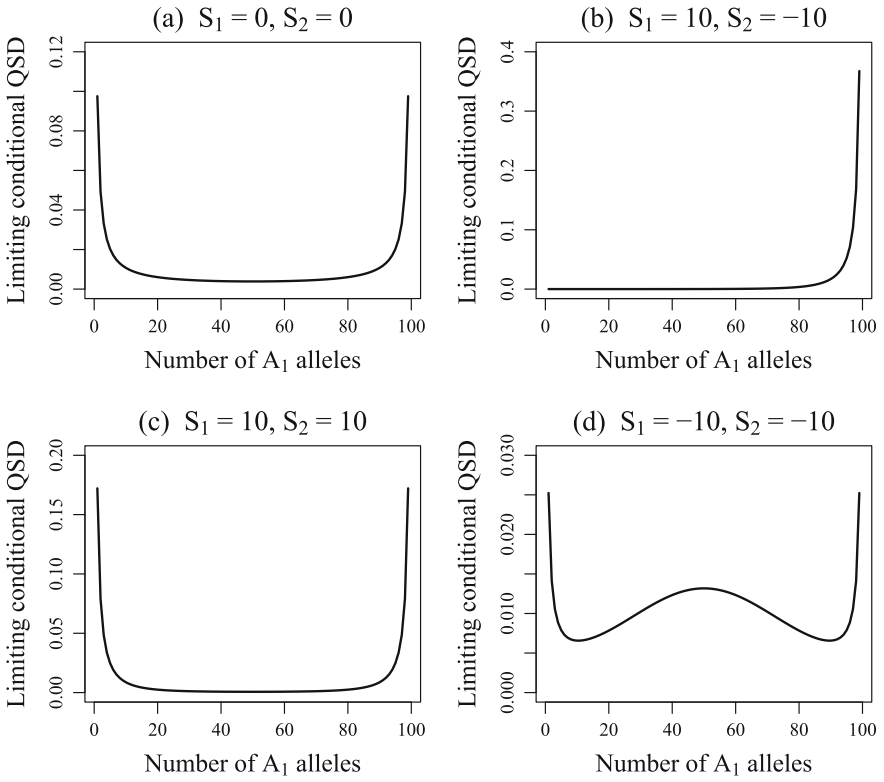


Fig. 11.4 Plots of the limiting conditional quasi-stationary distribution $\hat{\pi}_i(0)$ for the population genetic example of Sect. 11.2.3, as a function of the number of A_1-alleles i, for different values of the selection parameters. The model parameters N, C_1, C_2, D_1 and D_2 are the same as in Fig. 11.3. Note that the scales of the vertical axes differ between the plots

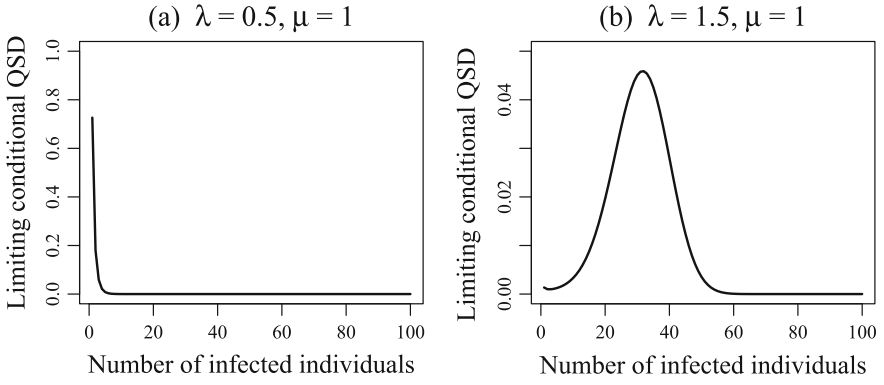


Fig. 11.5 Comparison of the limiting conditional quasi-stationary distribution $\tilde{\pi}_i(0)$ for the epidemic model of Sect. 11.2.2, as a function of the number of infected individuals i, for a population of size $N = 100$ with recovery rate $\mu = 1$. The force of infection parameter is $\lambda = 0.5$ in **a** and $\lambda = 1.5$ in **b**. Note that the scales of the vertical axes differ between the two plots

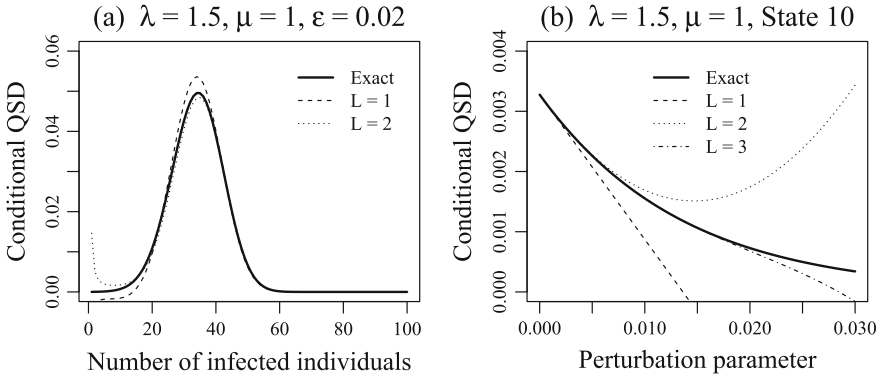


Fig. 11.6 Conditional quasi-stationary probabilities $\tilde{\pi}_i(\varepsilon)$ and some approximations for the epidemic model of Sect. 11.2.2, with $N = 100$, $\lambda = 1.5$ and $\mu = 1$. Note that the horizontal axes in the two plots represent different quantities; the number of infected individuals i in **a** and the perturbation parameter ε in **b**

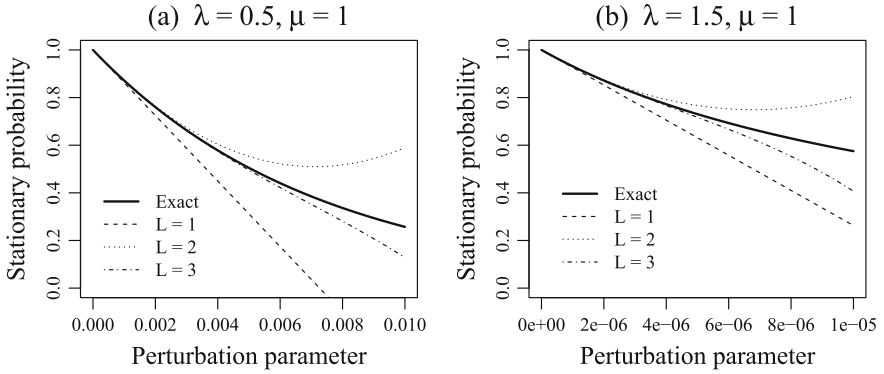


Fig. 11.7 Comparison of the stationary probability $\pi_i(\varepsilon)$ of state $i = 0$ as a function of the perturbation parameter ε for the epidemic model of Sect. 11.2.2 when $N = 100$, $\mu = 1$, and the contact rate parameter equals **a** $\lambda = 0.5$ and **b** $\lambda = 1.5$. Note that the scales of the horizontal axes differ between the two plots

11.8 Discussion

The present paper is devoted to studies of asymptotic expansions for stationary and conditional quasi-stationary distributions for perturbed semi-Markov birth-death processes. We employ the algorithms of sequential phase space reduction for perturbed semi-Markov processes combined with techniques of Laurent asymptotic expansions developed in the recent works by Silvestrov, D. and Silvestrov, S. [39–41], and apply them to semi-Markov birth-death processes. In this model, the proposed algorithms of phase space reduction preserve the birth-death structure for reduced semi-Markov processes. This made it possible to get, in the present paper, explicit

formulas for coefficients in the corresponding asymptotic expansions of the first and the second orders, for stationary and conditional quasi-stationary distributions of perturbed semi-Markov birth-death processes.

The above results are applied to three types of perturbed models from biology; population dynamics, epidemic models and models of population genetics. We supplement theoretical results by computations, illustrating numerical accuracy of the corresponding asymptotic expansions for stationary and quasi-stationary distributions of varying form. Even though exact expressions for the (quasi-)stationary distributions of these biological models are available, the asymptotic expansions may still be preferable when the state space is large and (quasi-)stationary distributions are computed for several values of the perturbation parameter, since only the first coefficients of the appropriate Laurent expansions are needed.

It should be mentioned that the semi-Markov setting is an adequate and necessary element of the proposed method. Even in the case where the initial birth-death-type process is a discrete or continuous time Markov chain, the time-space screening procedure of phase space reduction results in a semi-Markov birth-death process, since times between sequential hitting of the reduced space by the initial process have distributions which can differ from geometrical or exponential ones.

Also, the use of Laurent asymptotic expansions for expectations of sojourn times of perturbed semi-Markov processes is a necessary element of the proposed method. Indeed, even when expectations of sojourn times for all states of the initial semi-Markov birth-death process are asymptotically bounded and represented by Taylor asymptotic expansions, the exclusion of an asymptotically absorbing state from the initial phase space can generate states with asymptotically unbounded expectations of sojourn times represented by Laurent asymptotic expansions, for the reduced semi-Markov birth-death processes.

Several extensions of our work are possible. We have considered semi-Markov processes defined on a finite and linearly ordered state space $\mathbb{X}$, that is a subset of a one-dimensional lattice. We also confined ourselves to processes of birth-death type, where only jumps to neighboring states are possible.

For population dynamics models, it is noted by Lande, Engen and Saether [26] that one needs to go beyond birth-death processes though and incorporate larger jumps in order to account for a changing environment. State spaces that are subsets of higher-dimensional lattices are of interest in a number of applications, for instance SIR-models of epidemic spread where some recovered individuals get immune, Nåsell [32], population genetic models with two sexes, Moran [28], Hössjer and Tyvand [16], and population dynamics or population genetics models with several species or subpopulations, see Lande, Engen and Saether [26], Hössjer et al. [15] and references therein. It is an interesting topic of further research to apply the methodology of this paper to such models.

The method of sequential phase space reduction proposed in this paper can be applied to get asymptotic expansions for high order power and mixed power-exponential moments of hitting times and, in sequel, for more complex

quasi-stationary distributions (given by relation (11.41)) for nonlinearly perturbed semi-Markov birth-death processes and, thus, for models of population dynamics, epidemic spread and population genetics, which are the objects of interest in the present paper. We hope to present such results in the future.

References

1. Allen, B., Tarnita, C.E.: Measures of success in a class of evolutionary models with fixed population size and structure. J. Math. Biol. **68**, 109–143 (2014)
2. Allen, L.J.S., Burgin, A.M.: Comparison of deterministic and stochastic SIS and SIR models in discrete time. Math. Biosci. **163**, 1–33 (2000)
3. Avrachenkov, K.E., Filar, J.A., Howlett, P.G.: Analytic Perturbation Theory and Its Applications, xii+372 pp. SIAM, Philadelphia (2013)
4. Bini, D.A., Latouche, G., Meini, B.: Numerical Methods for Structured Markov Chains. Numerical Mathematics and Scientific Computation. Oxford Science Publications, xii+327 pp. Oxford University Press, New York (2005)
5. Cavender, J.A.: Quasi-stationary distributions of birth-and-death processes. Adv. Appl. Probab. **10**, 570–586 (1978)
6. Collet, P., Martínez, S., San Martín, J.: Quasi-Stationary Distributions. Markov Chains, Diffusions and Dynamical Systems. Probability and its Applications, xvi+280 pp. Springer, Heidelberg (2013)
7. Crow, J.F., Kimura, M.: An Introduction to Population Genetics Theory, xiv+591 pp. Harper and Row Publishers, New York (1970) (The Blackburn Press, Caldwell, NJ, pp. 608 (2009))
8. Durrett, R.: Probability Models for DNA Sequence Evolution, xii+431 pp. Springer, New York (2008) (2nd revised edition of Probability Models for DNA Sequence Evolution. Springer, New York, viii+240 pp. (2002))
9. Ewens, W.J.: Mathematical Population Genetics I. Theoretical Introduction, xx+417 pp. Springer, New York (2004) (2nd revised edition of Mathematical Population Genetics. Biomathematics, vol. 9, xii+325 pp. Springer, Berlin (1979))
10. Feller, W.: Die Grundlagen Volterraschen Theorie des Kampes ums Dasein in wahrscheinlichkeitsteoretischer Behandlung. Acta Biotheor. **5**, 11–40 (1939)
11. Feller, W.: An Introduction to Probability Theory and Its Applications, xviii+509 pp. Wiley, New York (1968) (3rd edition of An Introduction to Probability Theory and Its Applications, vol. I, xii+419 pp. Wiley, New York (1950))
12. Gilpin, M.E., Ayala, F.J.: Global models of growth and competition. Proc. Natl. Acad. Sci. USA **70**, 3590–3593 (1973)
13. Gyllenberg, M., Silvestrov, D.S.: Quasi-Stationary Phenomena in Nonlinearly Perturbed Stochastic Systems. De Gruyter Expositions in Mathematics, vol. 44, ix+579 pp. Walter de Gruyter, Berlin (2008)
14. Hethcote, H.W.: The mathematics of infectious diseases. SIAM Rev. **42**(4), 599–653 (2000)
15. Hössjer, O., Olsson, F., Laikre, L., Ryman, N.: A new general analytical approach for modeling patterns of genetic differentiation and effective size of subdivided populations over time. Math. Biosci. **258**, 113–133 (2014)
16. Hössjer, O., Tyvand, P.: A monoecious and diploid Moran model of random mating. J. Theor. Biol. **394**, 182–196 (2016)
17. Hössjer, O., Tyvand, P.A., Miloh, T.: Exact Markov chain and approximate diffusion solution for haploid genetic drift with one-way mutation. Math. Biosci. **272**, 100–112 (2016)
18. Jacquez, J.A., O'Neill, P.: Reproduction numbers and thresholds in stochastic epidemic models I. Homogenous populations. Math. Biosci. **107**, 161–186 (1991)
19. Jacquez, J.A., Simon, C.P.: The stochastic SI model with recruitment and deaths I. Comparisons with the closed SIS model. Math. Biosci. **117**, 77–125 (1993)

20. Karlin, S., McGregor, J.: On a genetics model of Moran. Proc. Camb. Philos. Soc. **58**, 299–311 (1962)
21. Kendall, D.G.: Stochastic processes and population growth. J. R. Stat. Soc. Ser. B **11**, 230–264 (1949)
22. Konstantinov, M., Gu, D.W., Mehrmann, V., Petkov, P.: Perturbation Theory for Matrix Equations. Studies in Computational Mathematics, vol. 9, xii+429 pp. North-Holland, Amsterdam (2003)
23. Korolyuk, V.S., Korolyuk, V.V.: Stochastic Models of Systems. Mathematics and its Applications, vol. 469, xii+185 pp. Kluwer, Dordrecht (1999)
24. Koroliuk, V.S., Limnios, N.: Stochastic Systems in Merging Phase Space, xv+331 pp. World Scientific, Singapore (2005)
25. Kryscio, R.J., Lefévre, C.: On the extinction of the S-I-S stochastic logistic epidemic. J. Appl. Probab. **27**, 685–694 (1989)
26. Lande, R., Engen, S., Saether, B.-E.: Stochastic Population Dynamics in Ecology and Conservation. Oxford Series and Ecology and Evolution, x+212 pp. Oxford University Press, Oxford (2003)
27. Moran, P.A.P.: Random processes in genetics. Proc. Camb. Philos. Soc. **54**, 60–71 (1958)
28. Moran, P.A.P.: A general theory of the distribution of gene frequencies I. Overlapping generations. Proc. Camb. Philos. Soc. **B149**, 102–112 (1958)
29. Nåsell, I.: The quasi-stationary distribution of the closed endemic SIS model. Adv. Appl. Probab. **28**, 895–932 (1996)
30. Nåsell, I.: On the quasi-stationary distribution of the stochastic logistic epidemic. Math. Biosci. **156**, 21–40 (1999)
31. Nåsell, I.: Extinction and quasi-stationarity of the Verhulst logistic model. J. Theor. Biol. **211**, 11–27 (2001)
32. Nåsell, I.: Stochastic models of some endemic infections. Math. Biosci. **179**, 1–19 (2002)
33. Nåsell, I.: Moment closure and the stochastic logistic model. Theor. Popul. Biol. **63**, 159–168 (2003)
34. Nåsell, I.: Extinction and Quasi-Stationarity in the Stochastic Logistic SIS Model. Lecture Notes in Mathematics, Mathematical Biosciences Subseries, xii+199 pp. Springer, Heidelberg (2011)
35. Pearl, R.: The growth of populations. Q. Rev. Biol. **2**, 532–548 (1927)
36. Petersson, M.: Asymptotics for quasi-stationary distributions of perturbed discrete time semi-Markov processes. In: Silvestrov, S., Rančić, M. (eds.) Engineering Mathematics II. Algebraic, Stochastic and Analysis Structures for Networks, Data Classification and Optimization. Springer Proceedings in Mathematics and Statistics, vol. 179, pp. 131–149. Springer, Heidelberg (2016)
37. Petersson, M.: Perturbed discrete time stochastic models. Doctoral dissertation, Stockholm University (2016)
38. Silvestrov, D., Manca, R.: Reward algorithms for semi-Markov processes. Methodol. Comput. Appl. Probab. **19**(4), 1191–1209 (2017)
39. Silvestrov, D., Silvestrov, S.: 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. Springer Proceedings in Mathematics and Statistics, vol. 179, pp. 151–222. Springer, Cham (2016)
40. Silvestrov, D., Silvestrov, S.: Asymptotic expansions for stationary distributions of nonlinearly perturbed semi-Markov processes 1, 2. Methodol. Comput. Appl. Probab. 20 (2017). Part 1: https://doi.org/10.1007/s11009-017-9605-0, Part 2: https://doi.org/10.1007/s11009-017-9607-y
41. Silvestrov, D., Silvestrov, S.: Nonlinearly Perturbed Semi-Markov Processes. Springer Briefs in Probability and Mathematical Statistics, xiv+143 pp. Springer, Berlin (2017)
42. Stewart, G.W.: Matrix Algorithms. Vol. I. Basic Decompositions, xx+458 pp. SIAM, Philadelphia (1998)

43. Stewart, G.W.: Matrix Algorithms. Vol. II. Eigensystems, xx+469 pp. SIAM, Philadelphia (2001)
44. Tsoularis, A., Wallace, J.: Analysis of logistic growth models. Math. Biosci. **179**, 21–55 (2002)
45. Verhulst, P.F.: Notice sur la loi que la population suit dans son accroissement. Corr. Math. Phys. **10**, 113–121 (1838)
46. Weiss, G.H., Dishon, M.: On the asymptotic behavior of the stochastic and deterministic models of an epidemic. Math. Biosci. **11**, 261–265 (1971)
47. Whittle, P.: On the use of the normal approximation in the treatment of stochastic processes. J. R. Stat. Soc. Ser. B **19**, 268–281 (1957)
48. Yin, G.G., Zhang, Q.: Discrete-Time Markov Chains. Two-Time-Scale Methods and Applications. Stochastic Modelling and Applied Probability, vol. 55, xix+348 pp. Springer, New York (2005)
49. Yin, G.G., Zhang, Q.: Continuous-Time Markov Chains and Applications. A Two-Time-Scale Approach. Stochastic Modelling and Applied Probability, vol. 37, 2nd edn, xxii+427 pp. Springer, New York (2013) (An extended variant of the first 1998 edition)