

CONSTRUCTING ASYMPTOTIC SERIES
FOR PROBABILITY DISTRIBUTIONS
OF MARKOV CHAINS
WITH WEAK AND STRONG INTERACTIONS

BY

R. Z. KHASHMINSKII (*Department of Mathematics, Wayne State University, Detroit, MI*),

G. YIN (*Department of Mathematics, Wayne State University, Detroit, MI*),

AND

Q. ZHANG (*Department of Mathematics, University of Georgia, Athens, GA*)

Abstract. Many applications arise in manufacturing systems, and queueing network problems involve Markov chains having slow and fast components. These components are coupled through weak and strong interactions. The main goal of this work is to study asymptotic properties for the probability distribution of the aforementioned Markov chains. Explicit construction of series expansions, consisting of regular part and boundary layer part or singular part, are developed by means of singular perturbation methods. The regular part is obtained by solving algebraic-differential equations, and the singular part is derived via solution of differential equations. One of the key points in the constructions is to select appropriate initial conditions. This is done by taking into consideration the regular part and the singular part together with their interactions. It is shown that the singular part decays exponentially fast. Analysis of residue is carried out, and the error bound for the remainder terms is ascertained.

1. Introduction. The preparation of this work was promoted by a wide range of applications involving a singularly perturbed Markov chain consisting of slow and fast motions with weak and strong interactions. Our main interest is to study the asymptotic behavior of the underlying Markov chains. In a recent paper [6], the authors have proved that the probability distribution of a class of singularly perturbed Markov chains with rapidly fluctuating motion, admits an asymptotic series expansion. This paper is a

Received July 13, 1994.

1991 *Mathematics Subject Classification.* Primary 34E05, 34E15, 60J27.

Key words and phrases. Asymptotic expansion, slow and fast motions, singular perturbation, Markovian generator.

Research of the first author was supported in part by the Office of Naval Research under grant N00014-93-1-0936, and N00014-95-1-0793.

Research of the second author was supported in part by the National Science Foundation under grant DMS-9529738, and in part by Wayne State University.

Research of the third author was supported in part by the Office of Naval Research Grant N00014-96-1-0263 and in part by the University of Georgia Faculty Research Grant.

substantial generalization of [6]. We address the issue of asymptotic properties when a system is governed by a Markov chain that consists of both slow and fast processes. An immediate question is: can we still develop an asymptotic series expansion? Our findings indicate that the answer to the question above is affirmative.

To model Markov chains involving slow and fast motions with weak and strong interactions, let $\varepsilon > 0$ be a small parameter, and $\alpha^\varepsilon(t)$ be a Markov chain with generator

$$Q^\varepsilon(t) = \frac{1}{\varepsilon}A(t) + B(t),$$

where $A(t)$ and $B(t)$ are $N \times N$ measurable matrix-valued functions, for all $t \in [0, T]$ and some $0 < T < \infty$. We assume that $A(t)$ has the diagonal form

$$A(t) = \begin{pmatrix} A^1(t) & & & \\ & A^2(t) & & \\ & & \ddots & \\ & & & A^l(t) \end{pmatrix} \text{ and } B(t) = \begin{pmatrix} B^{11}(t) & B^{12}(t) & \cdots & B^{1l}(t) \\ B^{21}(t) & B^{22}(t) & \cdots & B^{2l}(t) \\ \vdots & \vdots & \cdots & \vdots \\ B^{l1}(t) & B^{l2}(t) & \cdots & B^{ll}(t) \end{pmatrix},$$

where $A^\nu(t) \in \mathbb{R}^{n_\nu \times n_\nu}$ for $\nu = 1, 2, \dots, l$ are themselves generators with $\sum_{\nu=1}^l n_\nu = N$. Similarly, $B^{ij}(t)$ for $i, j = 1, \dots, l$ are blocks of submatrices that have comparable dimensions. Since $A^\nu(t)$ with $\nu = 1, 2, \dots, l$, $A(t)$ and $B(t)$ are generators of Markov chains,

$$\sum_{j=1}^{n_\nu} A_{ij}^\nu(t) = 0 \text{ for } i = 1, \dots, \nu, \sum_{j=1}^N A_{ij}(t) = 0 \text{ and } \sum_{j=1}^N B_{ij}(t) = 0 \text{ for } i = 1, \dots, N.$$

In what follows, we use the convention that for integers, superscripts denote the partitioned submatrices and subscripts denote the components of a vector and/or entries of a matrix or a submatrix.

The study of asymptotic properties of the underlying Markov chains is facilitated by considering the following ordinary differential equation:

$$\begin{aligned} \dot{y}^\varepsilon(t) &= y^\varepsilon(t) \left(\frac{1}{\varepsilon}A(t) + B(t) \right), \\ y^\varepsilon(0) &= y(0), \quad y_i(0) \geq 0, \quad i = 1, 2, \dots, N, \quad \sum_{i=1}^N y_i(0) = 1. \end{aligned} \tag{1}$$

Our main objective is to develop series expansions for the solution $y^\varepsilon(\cdot)$ with small $\varepsilon > 0$.

It is well known that (see [2] and [3]) the solution of (1) defines the probability distribution of the Markov chain $\alpha^\varepsilon(t)$ at time t with initial distribution $y^\varepsilon(0)$, i.e.,

$$\begin{aligned} (P(\alpha^\varepsilon(t) = 1), \dots, P(\alpha^\varepsilon(t) = N)) &= y^\varepsilon(t) \\ \text{with initial distribution } (P(\alpha^\varepsilon(0) = 1), \dots, P(\alpha^\varepsilon(0) = N)) &= y^\varepsilon(0). \end{aligned}$$

Since the differential equation is linear, (1) has a unique solution. Consequently,

$$0 \leq y^\varepsilon(t) \leq 1 \quad \text{and} \quad \sum_{i=1}^N y_i^\varepsilon(t) = 1. \tag{2}$$

In many problems arising in manufacturing systems, queueing networks, random fatigue analysis, and system reliability, finite state Markov chains with singular perturbation play an important role. To illustrate, consider the following problem. Suppose that two unreliable machines are lined up in a cascade form. The machines are subject to random breakdown and repair. However, the rates of breakdown for the two machines are not identical. In fact, one machine breaks down more often than the other. One way to formulate the problem is to assume the system capacity is a finite state Markov chain $\alpha^\varepsilon(t)$ having a generator $(1/\varepsilon)A(t) + B(t)$, with $A(t)$ and $B(t)$ given above. In order to design optimal controls for the manufacturing systems, the first and foremost important task is to investigate the asymptotic properties of the Markov chain involved. Heuristically, as ε gets smaller and smaller, the Markov chain will approach its “average” in an appropriate sense. Our results in this paper show that if $A(t) \equiv A$ and $B(t) \equiv B$ are constant matrices, $y^\varepsilon(t)$ converges to the equilibrium distribution (see [2]) $\bar{\pi}$ such that $\sum_{i=1}^N \bar{\pi}_i = 1$. In the more general cases, with time-dependent generators, $\bar{\pi}$ is replaced by $\bar{\pi}(t)$, a quasi-equilibrium distribution (to be defined in Sec. 2). In addition, a full asymptotic development is obtained, which shades more light in studying many properties of the Markov chains as well as related optimal control problems.

Notice that the formulation above is also frequently encountered in queueing network problems. For instance, consider two service stations in cascade form such that the service rate of one station is much faster than the other.

Using singular perturbation methods, we proved in [6] that when $A(t)$ has only one block, i.e., for $l = 1$, and $B(t) = 0$, $y^\varepsilon(t)$ has an asymptotic series expansion. Such cases correspond to a Markov chain with rapidly varying motions. As was pointed out in [6], although singular perturbation has been the focus of many researchers for years, the existing results (see [1], [7], [9], [10] and the references therein) are not applicable to the system we wish to study due to the singularity of the matrix $A(t)$. (Note that 0 is an eigenvalue for the generator with corresponding right eigenvector $\mathbb{1} = (1, 1, \dots, 1)' \in \mathbb{R}^N$, where \prime denotes the transpose.)

In the current paper, we treat more complicated situations in which the Markov chains have both slow and fast components. The fast components are governed by each of the blocks in $A(t)$, and the slow components are governed mainly by the matrices $B(t)$. The slow and fast components are coupled through weak and strong interactions. The states corresponding to $A^\nu(t)$, $\nu = 1, \dots, l$, are not isolated or independent of each other. More precisely, if we group the states corresponding to $A^i(t)$ as S_i , $i = 1, 2, \dots, l$, then these groups are coupled through the matrix $B(t)$, and transitions from S_i to S_j , $i \neq j$, are possible.

Under close scrutiny, we establish in this paper that an asymptotic expansion can be constructed. Explicit forms of the “coefficients of ε^n ’s” (functions of t) in the expansion are given. Their desired properties are obtained. The analysis of residue or remainder of the error terms is provided. Owing to the complexity of introducing diagonal blocks in $A(t)$, the derivation of the asymptotic expansion is much more involved than that of [6]. In fact, the regular part of the expansion is obtained through solutions of algebraic-differential equations. The solutions of these equations highly depends on a number of “multipliers”, which represent the interactions among different blocks. The multipliers,

in turn, depend on the general solution of the boundary layers. On the other hand, the choice of the initial conditions for the boundary layer terms relies on the solution to the regular part of the expansion. As a result, the selection of the initial conditions is a rather subtle and delicate issue.

The rest of the paper is arranged as follows. Section 2 gives precise formulation of the problem. We then begin the construction of the asymptotic series. The regular part of the expansion is constructed in Sec. 3, and boundary layer corrections are dealt with in Sec. 4. To establish the desired asymptotic properties, we need to make sure that the singular part of the expansion decays sufficiently fast. This is treated in Sec. 5, which exploits the key relation between the regular and singular parts. Among other things, the initial conditions are chosen to ensure matched asymptotic expansions. In Sec. 6, we analyze the remainder term in detail. The main results of the paper are summarized in Sec. 7. In addition, an outline of the computation procedure is also presented. Finally, we close the paper with an illustrative example.

2. Problem formulation. In this paper, we only consider finite state Markov chains. The qualifier “finite state” will be dropped in the sequel.

We first recall some definitions of Markov chains. For each $t \in [0, T]$, let a Markov chain have a generator $\tilde{Q}(t) \in \mathbb{R}^{r \times r}$. The matrix $\tilde{Q}(t)$ has the following form:

$$\tilde{Q}(t) = \begin{pmatrix} -\tilde{q}_1(t) & \tilde{q}_{12}(t) & \cdots & \tilde{q}_{1r}(t) \\ \tilde{q}_{21}(t) & -\tilde{q}_2(t) & \cdots & \tilde{q}_{2r}(t) \\ \vdots & \vdots & \cdots & \vdots \\ \tilde{q}_{r1}(t) & \tilde{q}_{r2}(t) & \cdots & -\tilde{q}_r(t) \end{pmatrix}$$

where $\tilde{q}_i(t) = \sum_{j \neq i} \tilde{q}_{ij}(t)$.

As defined in [6], a vector $y(t)$ is said to be a quasi-equilibrium (or quasi-stationary) distribution if it satisfies

$$\begin{aligned} y_i(t) &\geq 0, \quad \sum_{i=1}^r y_i(t) = 1, \\ y(t)\tilde{Q}(t) &= 0. \end{aligned} \tag{3}$$

We say that a Markov chain generated by $\tilde{Q}(t) \in \mathbb{R}^{r \times r}$ is weakly irreducible if the system

$$\begin{aligned} y(t)\tilde{Q}(t) &= 0, \\ \sum_{i=1}^r y_i(t) &= 0 \end{aligned} \tag{4}$$

has a unique solution.

With a slight abuse of notation, if a Markov chain is irreducible, we also say the corresponding generator $\tilde{Q}(t)$ is irreducible. Notice that as a consequence of the definition above, if a chain is weakly irreducible, there exists a unique quasi-equilibrium distribution. In fact, this can be used as an alternative definition of weak irreducibility. The latter emphasizes the probabilistic aspects of weak irreducibility, whereas the former stresses

the algebraic structure. Observe that if $\tilde{Q}(t)$ is weakly irreducible, $\text{rank } \tilde{Q}(t) = r - 1$ and there is a minor $M(t) \in \mathbb{R}^{(r-1) \times (r-1)}$ such that $\det M(t) \neq 0$.

We remark that the notion of weak irreducibility, which was defined in [6], is different from the usual definition of irreducibility (see [2] and [3]). To illustrate, consider the stationary case. That is, $\tilde{Q}(t) = \tilde{Q}$, a constant matrix. The usual notion of irreducibility requires that the equilibrium distribution be strictly positive, i.e., $\tilde{\pi}_i > 0$, $i = 1, \dots, r$, whereas the irreducibility in the weak sense means that some of the $\tilde{\pi}_i$'s are allowed to be zero. As for the time-varying case, for example $\tilde{Q}(t) = \begin{pmatrix} -\lambda(t) & \lambda(t) \\ \mu(t) & -\mu(t) \end{pmatrix}$, the usual irreducibility requires that $\lambda(t) > 0$ and $\mu(t) > 0$ for all t . However, only $\lambda(t) + \mu(t) > 0$ is needed in the weak formulation. This is particularly useful in studying failure-prone manufacturing systems. For example, in a manufacturing model, $\mu(\cdot)$ may denote the repairing rate of a machine. The case $\mu(t) = 0$ corresponds to, for example, breaks of the repairing workers are allowed or a waiting period is needed to get the required parts.

To proceed, we make the following assumptions.

(A1) For each $t \in [0, T]$, $A^\nu(t)$, $\nu = 1, 2, \dots, l$, are weakly irreducible.

(A2) For some n , $A(\cdot)$, $B(\cdot) \in C^{n+2}[0, T]$, i.e., they are $n+2$ continuously differentiable in $[0, T]$.

Notice that, in view of the definition of weak irreducibility, conditions (A1) and (A2) imply that $A^\nu(\cdot)$, $\nu = 1, 2, \dots, l$ are uniformly weakly irreducible in $[0, T]$, i.e., for each $t \in [0, T]$, there exists $M^\nu(t)$, a minor of $A^\nu(t)$ such that the determinant of $M^\nu(t) \neq 0$. In the sequel, superscripts with curly bracket denote indices of a sequence. Based on the framework of singular perturbation methods, we seek solutions of (1) of the form

$$\begin{aligned} Y^\varepsilon(t) &= \tilde{p}^\varepsilon(t) + \tilde{q}^\varepsilon(t/\varepsilon) \\ &= p^{\{0\}}(t) + \varepsilon p^{\{1\}}(t) + \varepsilon^2 p^{\{2\}}(t) + \dots \\ &\quad + q^{\{0\}}(t/\varepsilon) + \varepsilon q^{\{1\}}(t/\varepsilon) + \varepsilon^2 q^{\{2\}}(t/\varepsilon) + \dots \end{aligned} \quad (5)$$

We call $p^{\{i\}}(\cdot)$, $i = 0, 1, 2, \dots$ the regular part of the expansion, and $q^{\{i\}}(\cdot)$, $i = 0, 1, 2, \dots$ the boundary layer correction or the singular part of the expansion. It will be shown in the sequel that $p^{\{0\}}(\cdot)$ is nothing but the quasi-equilibrium distribution. As was mentioned in the introduction, classical results in singular perturbation are not applicable due to the highly degenerate nature of the matrix $A(t)$.

Many applications in manufacturing systems and in production planning consist of a Markov chain $\alpha^\varepsilon(t) = (\tilde{\alpha}^\varepsilon(t), \tilde{\beta}(t))$ such that $\tilde{\alpha}^\varepsilon(\cdot)$ is a fast-changing component and $\tilde{\beta}(\cdot)$ is a slowly-varying one. Very often, the generators of the two components of the Markov chain commute. See the example in Sec. 6 for a typical case. In this case, the proof in the sequel can be much simplified. Our aim, however, is to obtain the desired results without further restrictions (other than the weak irreducibility and the smoothness conditions (A1) and (A2)) on the generators.

For each $k = 1, 2, \dots$, define

$$Y_k^\varepsilon(t) = \sum_{i=0}^k \varepsilon^i p^{\{i\}}(t) + \sum_{i=0}^k \varepsilon^i q^{\{i\}}(t/\varepsilon) = \tilde{p}_k^\varepsilon(t) + \tilde{q}_k^\varepsilon(t/\varepsilon). \quad (6)$$

We show in this paper that $Y_n^\varepsilon(t)$ can be constructed such that

$$\sup_{t \in [0, T]} |y^\varepsilon(t) - Y_n^\varepsilon(t)| = O(\varepsilon^{n+1}).$$

The main difficulty lies on the interactions between different block matrices. In our early paper [6], the functions $p^{\{i\}}(\cdot)$ for $i = 1, \dots, n$, were constructed independently of $q^{\{i\}}(\cdot)$; the functions $q^{\{i\}}(\cdot)$ were constructed essentially independently of $p^{\{i\}}(\cdot)$ except for the initial conditions $p^{\{i\}}(0)$. Now for the harder part; $p^{\{i\}}(\cdot)$ and $q^{\{i\}}(\cdot)$ are highly intertwined and tangled together. The trick of the methods in the sequel is to find $p^{\{i\}}(\cdot)$ and $q^{\{i\}}(\cdot)$ jointly and recursively. In the process of construction, we bare in mind to make $q^{\{i\}}(\cdot)$ decay sufficiently fast. One of the crucial and delicate points is to select the “right” initial conditions.

For future use, define a differential operator \mathcal{L}^ε on the space of $\mathbb{R}^{1 \times N}$ -valued functions by

$$\mathcal{L}^\varepsilon f = \varepsilon \frac{df}{dt} - f(A + \varepsilon B). \tag{7}$$

The formulation is completed. We are now in a position to derive the asymptotic expansion.

3. Regular part in the asymptotic expansion. To construct the regular part of the expansion, setting $\mathcal{L}^\varepsilon \bar{p}^\varepsilon(t) = 0$, and equating like powers of ε^i , $i = 0, 1, 2, \dots$, we obtain

$$\begin{aligned} \varepsilon^0 : p^{\{0\}}(t)A(t) &= 0 \\ \varepsilon^1 : \dot{p}^{\{0\}}(t) &= p^{\{1\}}(t)A(t) + p^{\{0\}}(t)B(t) \\ \varepsilon^2 : \dot{p}^{\{1\}}(t) &= p^{\{2\}}(t)A(t) + p^{\{1\}}(t)B(t) \\ &\dots \\ \varepsilon^n : \dot{p}^{\{n\}}(t) &= p^{\{n+1\}}(t)A(t) + p^{\{n\}}(t)B(t) \\ &\dots \end{aligned} \tag{8}$$

In the case of absence of $B(t)$ and without subdivision of $A(t)$, i.e., $A(t)$ consists of only one block of weakly irreducible matrices, the problem was solved in [6] by computing the solution of the first equation in (8) together with $\sum_{i=1}^N p_i^{\{0\}}(t) = 1$, then substituting the resulting $p^{\{0\}}(t)$ into the second equation, together with $\sum_{i=1}^N p_i^{\{1\}}(t) = 0$, to obtain $p^{\{1\}}(t)$ and so on. That is, all the $p^{\{i\}}(\cdot)$, $i = 0, 1, 2, \dots$ are obtained by solving a set of algebraic equations. However, in this paper the picture is quite different. To begin with, note that the system

$$\begin{aligned} p^{\{0\}}(t)A(t) &= 0, \\ \sum_{i=1}^N p_i^{\{0\}}(t) &= 1 \end{aligned}$$

has infinitely many solutions due to the fact that $\text{rank } A(t) < N - 1$ if $l > 1$. More information is needed to determine $p^{\{0\}}(t)$ uniquely.

Since $A^\nu(t)$, $\nu = 1, 2, \dots, l$ are weakly irreducible, $\text{rank } A^\nu(t) = n_\nu - 1$ for each $t \in [0, T]$ and $\nu = 1, 2, \dots, l$. As a result, $\text{rank } A(t) = N - l$ for each t . This indicates

that in order to get a unique solution, we need to supply l auxiliary equations. Where can we find these equations? Loosely, the idea is to apply the Fredholm alternative, and use the orthogonality condition to choose l additional equations to replace l equations in the system represented by the first equation in (8). Consider the conjugate equation

$$A(t)z(t) = 0, \quad z(t) \in \mathbb{R}^N.$$

It is easy to see that there exist l linearly independent solutions $z_1(t), \dots, z_l(t)$ such that

$$\begin{aligned} z_1(t) &= (\underbrace{1, \dots, 1}_{n_1}, 0, \dots, 0)', \\ z_2(t) &= (\underbrace{0, \dots, 0}_{n_1}, \underbrace{1, \dots, 1}_{n_2}, 0, \dots, 0)', \\ &\dots \\ z_l(t) &= (0, \dots, 0, \underbrace{1, \dots, 1}_{n_l})'. \end{aligned} \tag{9}$$

In view of (9), the Fredholm alternative implies that the second equation in (8) has a solution if and only if

$$\tilde{p}^{\{0\}}(t) - p^{\{0\}}(t)B(t) \text{ is orthogonal to } z_\nu(t) \tag{10}$$

for $\nu = 1, 2, \dots, l$. Thus we have the equations:

$$\begin{aligned} p^{\{0\}}(t)A(t) &= 0, \\ \sum_{i=1}^N p_i^{\{0\}}(t) &= 1, \\ \sum_{i=\tilde{c}_{\nu-1}+1}^{\tilde{c}_{\nu-1}+n_\nu} \tilde{p}_i^{\{0\}}(t) - \sum_{i=\tilde{c}_{\nu-1}+1}^{\tilde{c}_{\nu-1}+n_\nu} (p^{\{0\}}(t)B(t))_i &= 0, \quad \text{for } \nu = 1, 2, \dots, l, \end{aligned} \tag{11}$$

where $\tilde{c}_{\nu-1} = \sum_{j=1}^{\nu-1} n_j$. The last equation in (11) is a direct consequence of (10).

To formalize what was mentioned above, we introduce some notation and conveniently write the above equations in their equivalent forms as follows. Partition the vector $p^{\{0\}}(t)$ as $(p^{\{0\},1}(t), \dots, p^{\{0\},l}(t))$ such that $p^{\{0\},\nu}(t) \in \mathbb{R}^{1 \times n_\nu}$. Corresponding to this partition, the first equation in (8) can be written as

$$p^{\{0\},\nu}(t)A^\nu(t) = 0, \quad \nu = 1, \dots, l. \tag{12}$$

Since $\text{rank} A^\nu(t) = n_\nu - 1$, for each $t \in [0, T]$, the null space of $A^\nu(t)$ is one dimensional. Notice that $\sum_{i=1}^{n_\nu} p_i^{\{0\},\nu}(t) \neq 1$ in general. However, $\sum_{i=1}^N p_i^{\{0\}}(t) = 1$. Let

$$\sum_{i=1}^{n_\nu} p_i^{\{0\},\nu}(t) = f^{\{0\},\nu}(t).$$

We establish the following:

LEMMA 3.1. Assume that the conditions (A1) and (A2) are satisfied. Then for each $\nu = 1, \dots, l$, the solution of

$$\begin{aligned}
 p^{\{0\},\nu}(t)A^\nu(t) &= 0, \\
 \sum_{i=1}^{n_\nu} p_i^{\{0\},\nu}(t) &= f^{\{0\},\nu}(t)
 \end{aligned} \tag{13}$$

can be uniquely expressed as $p^{\{0\},\nu}(t) = \mu^\nu(t)f^{\{0\},\nu}(t)$, where $\mu^\nu(\cdot) \in C^{n+2}[0, T]$, with

$$\mu_i^\nu(t) \geq 0, \quad i = 1, \dots, n_\nu, \quad \sum_{i=1}^{n_\nu} \mu_i^\nu(t) = 1, \quad \mu^\nu(t)A^\nu(t) = 0.$$

REMARK. Although it is unique, $f^{\{0\},\nu}(\cdot)$ needs to be determined later. Observe that $f^{\{0\},\nu}(\cdot)$ satisfies $\sum_{\nu=1}^l f^{\{0\},\nu}(t) = 1$. For $\nu = 1, 2, \dots, l$, $\mu^\nu(t)$ are the quasi-equilibrium distributions for the sub-chain generated by $A^\nu(t)$. Nevertheless, between different block matrices, there are weak interactions. As a consequence, the individual quasi-equilibrium distribution is no longer a quasi-equilibrium distribution for the entire system. The interactions are embedded in the scalar-valued function $f^{\{0\},\nu}(t)$. Roughly, the lemma states that the quasi-equilibrium distribution for the Markov chain generated by $A(t)$ is proportional to or a “multiple” of the quasi-equilibrium distributions of the Markov chains generated by $A^\nu(t)$, for $\nu = 1, \dots, l$. The multiplier $f^{\{0\},\nu}(t)$ reflects the interactions for different sub-chains.

Proof. Since $\text{rank}A^\nu(t) = n_\nu - 1$, there are $n_\nu - 1$ linearly independent equations in the system of equations in (13). For fixed $t \in [0, T]$, without loss of generality, suppose the first $n_\nu - 1$ equations are linearly independent. Replace the n_ν th equation by the last equation in (13). The resulting coefficient matrix then has full rank. By virtue of the weak irreducibility,

$$\Delta(t) = \begin{vmatrix} A_{11}^\nu(t) & \cdots & A_{n_\nu,1}^\nu(t) \\ & \cdots & \\ A_{1,n_\nu-1}^\nu(t) & \cdots & A_{n_\nu,n_\nu-1}^\nu(t) \\ 1 & \cdots & 1 \end{vmatrix} \neq 0.$$

Owing to Cramer’s rule, for each $i = 1, \dots, n_\nu$,

$$\begin{aligned}
 p_i^{\{0\},\nu}(t) &= \frac{1}{\Delta(t)} \begin{vmatrix} A_{11}^\nu(t) & \cdots & 0 & \cdots & A_{n_\nu,1}^\nu(t) \\ & \cdots & \cdots & \cdots & \\ A_{1,n_\nu-1}^\nu(t) & \cdots & \cdots & \cdots & A_{n_\nu,n_\nu-1}^\nu(t) \\ 1 & \cdots & \underbrace{f^{\{0\},\nu}(t)}_{i\text{th column}} & \cdots & 1 \end{vmatrix} \\
 &= \frac{f^{\{0\},\nu}(t)}{\Delta(t)} \begin{vmatrix} A_{11}^\nu(t) & \cdots & 0 & \cdots & A_{n_\nu,1}^\nu(t) \\ & \cdots & \cdots & \cdots & \\ A_{1,n_\nu-1}^\nu(t) & \cdots & \cdots & \cdots & A_{n_\nu,n_\nu-1}^\nu(t) \\ 1 & \cdots & \underbrace{1}_{i\text{th column}} & \cdots & 1 \end{vmatrix}.
 \end{aligned}$$

Notice that

$$\frac{1}{\Delta(t)} \begin{vmatrix} A_{11}^\nu(t) & \cdots & 0 & \cdots & A_{n_\nu,1}^\nu(t) \\ & \cdots & \cdots & \cdots & \\ A_{1,n_\nu-1}^\nu(t) & \cdots & \cdots & \cdots & A_{n_\nu,n_\nu-1}^\nu(t) \\ & 1 & \cdots & \underbrace{1}_{i\text{th column}} & \cdots \\ & & & & 1 \end{vmatrix} = \mu_i^\nu(t)$$

is the unique solution of the system (13) with $f^{\{0\},\nu}(t) \equiv 1$, and it is C^{n+2} smooth. The lemma is thus concluded.

Our next task is to determine these $f^{\{0\},\nu}(\cdot)$'s. Define an $N \times l$ matrix

$$\tilde{\mathbb{1}} = \begin{pmatrix} \mathbb{1}_{n_1} & & & \\ & \mathbb{1}_{n_2} & & \\ & & \ddots & \\ & & & \mathbb{1}_{n_l} \end{pmatrix}$$

where $\mathbb{1}_{n_i} = (1, \dots, 1)' \in \mathbb{R}^{n_i}$. Then $\mathcal{L}^\varepsilon \tilde{p}^\varepsilon(t) \tilde{\mathbb{1}} = 0$. Noticing $A(t) \tilde{\mathbb{1}} = 0$ and rewriting $p^{\{0\}}(t)$ as

$$p^{\{0\}}(t) = (f^{\{0\},1}(t), \dots, f^{\{0\},l}(t)) \begin{pmatrix} \mu^1(t) & & & \\ & \mu^2(t) & & \\ & & \ddots & \\ & & & \mu^l(t) \end{pmatrix},$$

by equating the first power of ε in $\mathcal{L}^\varepsilon \tilde{p}^\varepsilon(t) \tilde{\mathbb{1}} = 0$, we arrive at

$$\begin{aligned} \frac{d}{dt} (f^{\{0\},1}(t), \dots, f^{\{0\},l}(t)) &= (f^{\{0\},1}(t), \dots, f^{\{0\},l}(t)) G(t), \\ (f^{\{0\},1}(0), \dots, f^{\{0\},l}(0)) &= \left(\sum_{i=1}^{n_1} y_i(0), \dots, \sum_{i=N-n_l+1}^N y_i(0) \right), \end{aligned} \quad (14)$$

where

$$G(t) = \begin{pmatrix} \mu^1(t) & & & \\ & \mu^2(t) & & \\ & & \ddots & \\ & & & \mu^l(t) \end{pmatrix} B(t) \tilde{\mathbb{1}}.$$

Equation (14) is a linear system of differential equations. Solving this system yields that

$$(f^{\{0\},1}(t), \dots, f^{\{0\},l}(t)) = \left(\sum_{i=1}^{n_1} y_i(0), \dots, \sum_{i=N-n_l+1}^N y_i(0) \right) X(t, 0),$$

where $X(t, \tau)$ is the principal matrix solution (see [5]) of (14). Since $X(\cdot, \cdot)$ purely depends on the properties of $A(t)$ and $B(t)$, $(f^{\{0\},1}(\cdot), \dots, f^{\{0\},l}(\cdot)) \in C^{n+2}[0, T]$. Up to now, we have shown that $p^{\{0\}}(\cdot)$ can be constructed such that $p^{\{0\}}(\cdot) \in C^{n+2}[0, T]$. We summarize the discussion above as:

PROPOSITION 3.2. Denote $f^{\{0\}}(t) = (f^{\{0\},1}(t), \dots, f^{\{1\},l}(t))$. For $t \in [0, T]$, $p^{\{0\}}(t)$ can be obtained uniquely by solving the system

$$\begin{aligned}
 & p^{\{0\},\nu}(t)A^\nu(t) = 0, \\
 & \sum_{i=1}^{n_\nu} p_i^{\{0\},\nu}(t) = f^{\{0\},\nu}(t), \\
 & \frac{d}{dt}f^{\{0\}}(t) = f^{\{0\}}(t)G(t) \tag{15}
 \end{aligned}$$

with $f^{\{0\}}(0) = \left(\sum_{i=1}^{n_1} y_i(0), \dots, \sum_{i=N-n_l+1}^N y_i(0) \right)$

such that $p^{\{0\}}(\cdot) \in C^{n+2}[0, T]$.

Since $p^{\{0\}}(t)$ is obtained, the vector

$$b^{\{0\}}(t) = \dot{p}^{\{0\}}(t) - p^{\{0\}}(t)B(t)$$

becomes known. To proceed, consider the system of equations

$$\begin{aligned}
 & p^{\{1\}}(t)A(t) = b^{\{0\}}(t), \\
 & \sum_{i=1}^N p_i^{\{1\}}(t) = 0.
 \end{aligned}$$

Partitioning the matrices yields

$$(p^{\{1\},1}(t), \dots, p^{\{1\},l}(t)) \begin{pmatrix} A^1(t) & & \\ & \ddots & \\ & & A^l(t) \end{pmatrix} = (b^{\{0\},1}(t), \dots, b^{\{0\},l}(t)) = b^{\{0\}}(t).$$

Again, by virtue of the Fredholm alternative, the equation $p^{\{1\},\nu}(t)A^\nu(t) = b^{\{0\},\nu}(t)$ has a solution if and only if $b^{\{0\},\nu}(t)$ is orthogonal to $z_\nu(t)$. This condition is fulfilled (see the last equation in (11)). To obtain the unique solution, we add the normalizing condition $\sum_{i=1}^{n_\nu} p_i^{\{1\},\nu}(t) = f^{\{1\},\nu}(t)$. Notice that $\sum_{\nu=1}^l f^{\{1\},\nu}(t) = 0$. Consequently, we have the system of equations:

$$\begin{aligned}
 & p^{\{1\},\nu}(t)A^\nu(t) = b^{\{0\},\nu}(t), \quad \nu = 1, \dots, l, \\
 & \sum_{i=1}^{n_\nu} p_i^{\{1\},\nu}(t) = f^{\{1\},\nu}(t). \tag{16}
 \end{aligned}$$

Then the solution can be written as

$$p^{\{1\},\nu}(t) = \tilde{b}^{\{0\},\nu}(t) + f^{\{1\},\nu}(t)\mu^\nu(t), \tag{17}$$

where $f^{\{1\},\nu}(t)\mu^\nu(t)$ is a solution of the homogeneous equation $p^{\{1\},\nu}(t)A^\nu(t) = 0$ and $\tilde{b}^{\{0\},\nu}(t)$ is a solution of the nonhomogeneous equation with the right-hand side of (16), namely, $b^{\{0\},\nu}(t)$ being orthogonal to z_ν .

Similar to the previous development, we proceed to determine $f^{\{1\},\nu}(t)$ by solving the system of equations in $\mathcal{L}^\varepsilon \tilde{p}^\varepsilon(t) \tilde{\mathbb{I}} = 0$.

Using $\sum_{i=1}^{n_\nu} \tilde{b}_i^{\{0\},\nu}(t) = 0$ and $\sum_{i=1}^{n_\nu} \mu_i^\nu(t) = 1$, we have

$$p^{\{1\}}(t) \tilde{\mathbb{I}} = (f^{\{1\},1}(t), \dots, f^{\{1\},l}(t))$$

and

$$p^{\{1\}}(t)B(t) \tilde{\mathbb{I}} = (f^{\{1\},1}(t), \dots, f^{\{1\},l}(t))G(t),$$

where $G(t)$ is the same as before.

By equating the second power of ε in $\mathcal{L}^\varepsilon \tilde{p}^\varepsilon(t) \tilde{\mathbb{I}} = 0$, we obtain a linear nonhomogeneous equation

$$\begin{aligned} \frac{d}{dt}(f^{\{1\},1}(t), \dots, f^{\{1\},l}(t)) &= (f^{\{1\},1}(t), \dots, f^{\{1\},l}(t))G(t) \\ &+ (\tilde{b}^{\{0\},1}(t), \dots, \tilde{b}^{\{0\},l}(t))B(t) \tilde{\mathbb{I}}, \end{aligned} \tag{18}$$

with initial conditions $f^{\{1\},\nu}(0)$, for $\nu = 1, 2, \dots, l$ such that $\sum_{\nu=1}^l f^{\{1\},\nu}(0) = 0$.

The initial conditions $f^{\{1\},\nu}(0)$ for $\nu = 1, \dots, l$, have not yet been specified. We do this later to ensure the matched asymptotic expansion. Once the $f^{\{1\},\nu}(0)$'s are given, the solution of the above equation is:

$$\begin{aligned} (f^{\{1\},1}(t), \dots, f^{\{1\},l}(t)) &= (f^{\{1\},1}(0), \dots, f^{\{1\},l}(0))X(t, 0) \\ &+ \int_0^t (\tilde{b}^{\{0\},1}(\tau), \dots, \tilde{b}^{\{0\},l}(\tau))B(\tau) \tilde{\mathbb{I}}X(t, \tau) d\tau. \end{aligned}$$

Therefore $f^{\{1\},\nu}(\cdot)$, $\nu = 1, \dots, l$ are found, and so is $p^{\{1\}}(\cdot)$. Moreover, $p^{\{1\}}(\cdot) \in C^{n+1}[0, T]$. In fact, $p^{\{1\}}(\cdot)$ is obtained by solving a system of equations similar to (15), except the initial conditions of $f^{\{1\}}(0)$ have not been given at this point.

Similar to $\tilde{b}^{\{0\},\nu}(\cdot)$, define $\tilde{b}^{\{i\},\nu}(\cdot)$ and write

$$\tilde{b}^{\{i\}}(t) = (\tilde{b}^{\{i\},1}(t), \dots, \tilde{b}^{\{i\},l}(t)).$$

Proceeding inductively, we obtain:

PROPOSITION 3.3. Under the conditions (A1) and (A2), sequences of row vector-valued functions $p^{\{i\}}(\cdot)$, and $f^{\{i\}}(\cdot)$ for $i = 1, 2, \dots, n$, can be obtained by solving the system of algebraic-differential equations

$$\begin{aligned} p^{\{i\}}(t)A(t) &= \dot{p}^{\{i-1\}}(t) - p^{\{i-1\}}(t)B(t), \\ \sum_{j=1}^{n_\nu} p_j^{\{i\},\nu}(t) &= f^{\{i\},\nu}(t), \\ \frac{d}{dt}f^{\{i\}}(t) &= f^{\{i\}}(t)G(t) + \tilde{b}^{\{i-1\}}(t)B(t) \tilde{\mathbb{I}} \end{aligned} \tag{19}$$

such that $p^{\{i\}}(\cdot) \in C^{n+2-i}[0, T]$, provided the initial conditions $f^{\{i\}}(0)$, for $i = 1, 2, \dots, l$ are given.

So far, with the proviso of specified initial conditions, the construction of the regular part has been completed; the smoothness of $p^{\{i\}}(\cdot)$ has been established, and the section is concluded.

4. Construction of the singular part. The regular part alone gives a good approximation of $y^\varepsilon(\cdot)$ when t is away from 0. Nevertheless, when t is sufficiently near 0, this part no longer gives a reasonable approximation due to the apparent nonuniformity related to the $1/\varepsilon$ term in (1). What happens is that a boundary layer of thickness ε is developed. To circumvent the difficulty, we enlarge the picture near 0 by defining a stretched variable

$$s = \frac{t}{\varepsilon}, \tag{20}$$

and considering a sequence of boundary layer correction terms. In what follows, we use $z^{(n)}$ to denote $d^n z/ds^n$, and \dot{z} to denote dz/ds .

Substituting $\tilde{q}^\varepsilon(s)$ into $\mathcal{L}^\varepsilon f = 0$, and noting the change of variable (20),

$$\frac{d}{ds}(q^{\{0\}}(s) + \varepsilon q^{\{1\}}(s) + \dots) = (q^{\{0\}}(s) + \varepsilon q^{\{1\}}(s) + \dots)(A(\varepsilon s) + \varepsilon B(\varepsilon s)).$$

By means of the Taylor expansion,

$$\begin{aligned} A(\varepsilon s) + \varepsilon B(\varepsilon s) &= A(0) + \varepsilon s A^{(1)}(0) + \dots + \frac{(\varepsilon s)^n}{n!} A^{(n)}(0) + \dots \\ &+ \varepsilon B(0) + \varepsilon^2 s B^{(1)}(0) + \dots + \frac{\varepsilon(\varepsilon s)^n}{n!} B^{(n)}(0) + \dots \end{aligned}$$

Using this expansion and equating coefficients of like powers of ε^i , we obtain

$$\begin{aligned} \varepsilon^0 : \dot{q}^{\{0\}}(s) &= q^{\{0\}}(s)A(0) \\ \varepsilon^1 : \dot{q}^{\{1\}}(s) &= q^{\{1\}}(s)A(0) + q^{\{0\}}(s)(B(0) + sA^{(1)}(0)) \\ &\dots \\ \varepsilon^n : \dot{q}^{\{n\}}(s) &= q^{\{n\}}(s)A(0) + \sum_{i=0}^{n-1} q^{\{n-i-1\}}(s) \left(\frac{s^i B^{(i)}(0)}{i!} + \frac{s^{i+1} A^{(i+1)}(0)}{(i+1)!} \right). \end{aligned} \tag{21}$$

Our next task is to solve these equations. The initial conditions are at our disposal. We will choose them so that the expansion becomes a matched one. The discussion on the choice of initial conditions for $q^{\{i\}}(\cdot)$ is deferred until the next section. Suppose that $q^{\{0\}}(0), q^{\{1\}}(0), \dots, q^{\{n\}}(0)$ are given. Then the solution is:

$$\begin{aligned} q^{\{0\}}(s) &= q^{\{0\}}(0) \exp(A(0)s) \\ &\dots \\ q^{\{n\}}(s) &= q^{\{n\}}(0) \exp(A(0)s) \\ &+ \sum_{i=0}^{n-1} \int_0^s q^{\{n-i-1\}}(\tau) \left(\frac{\tau^i B^{(i)}(0)}{i!} + \frac{\tau^{i+1} A^{(i+1)}(0)}{(i+1)!} \right) \exp(A(0)(s-\tau)) d\tau. \end{aligned} \tag{22}$$

We summarize this into the following proposition.

PROPOSITION 4.1. Under the conditions (A1) and (A2), $q^{\{0\}}(\cdot)$ is uniquely determined by choosing $q^{\{0\}}(0) = y(0) - p^{\{0\}}(0)$. For each $i = 1, 2, \dots, n$, if $q^{\{i\}}(0)$ is chosen, a sequence of boundary layer terms can be obtained by (22).

REMARK. Since $f^{\{i\},\nu}(0)$ have not been specified, neither have $q^{\{i\},\nu}(0)$. Our next task is to select these initial conditions in a clever way so that $q^{\{i\}}(\cdot)$ decay exponentially fast. In addition,

$$\sum_{i=0}^n \varepsilon^i [p^{\{i\}}(0) + q^{\{i\}}(0)] = y(0).$$

That is, the asymptotic expansion is a matched one.

5. Exponential decay and choice of initial conditions. To ensure the boundary layer terms act properly, we need to verify that when s is sufficiently large, $q^{\{i\}}(\cdot)$ is negligibly small. To be more precise, it is shown in the sequel that $q^{\{i\}}(\cdot)$ for $i \leq n$, decay exponentially. For each $\nu = 1, \dots, l$, denote the equilibrium distribution corresponding to the generator $A^\nu(0)$ by $\bar{\mu}^\nu$. Define

$$\pi = \tilde{\mathbb{1}} \bar{\mu} = \tilde{\mathbb{1}} \begin{pmatrix} \bar{\mu}^1 & & & \\ & \bar{\mu}^2 & & \\ & & \ddots & \\ & & & \bar{\mu}^l \end{pmatrix} = \begin{pmatrix} \mathbb{1}_{n_1} \bar{\mu}^1 & & & \\ & \mathbb{1}_{n_2} \bar{\mu}^2 & & \\ & & \ddots & \\ & & & \mathbb{1}_{n_l} \bar{\mu}^l \end{pmatrix}.$$

In the rest of the paper, K denotes a generic positive constant. Its value may be different at each appearance. Thus $K + K = K$ and $KK = K$, etc.

PROPOSITION 5.1. Assume the conditions of Proposition 4.1 are fulfilled.

1. For $i = 1, \dots, n$, the initial conditions are selected as follows:

(a) For $\nu = 1, 2, \dots, l$, find $\sum_{j=1}^{n_\nu} q_j^{\{i\},\nu}(0)$ from the equation

$$q^{\{i\}}(0)\pi = - \left(\sum_{j=0}^{i-1} \int_0^\infty \frac{\tau^j}{j!} q^{\{i-j-1\}}(\tau) d\tau B^{(j)}(0) \right) \pi. \tag{23}$$

(b) Choose $f^{\{i\},\nu}(0) = - \sum_{j=1}^{n_\nu} q_j^{\{i\}}(0)$, for $\nu = 1, \dots, l$.

(c) Choose $q^{\{i\}}(0) = -p^{\{i\}}(0)$.

2. There is a positive real number $\gamma = \gamma_n$ such that

$$|q^{\{n\}}(s)| \leq K \exp(-\gamma s).$$

REMARK. By Proposition 5.1, the initial conditions for $f^{\{i\},\nu}(0)$ are specified. As a consequence, $p^{\{i\}}(\cdot)$ are uniquely determined by Proposition 3.3. The solutions of $p^{\{i\}}(\cdot)$ give specified initial conditions for $q^{\{i\}}(0)$. Thus, $q^{\{i\}}(\cdot)$ are uniquely determined by Proposition 4.1. We emphasize that owing to the structure of the matrix π , only l unknowns, namely $\sum_{j=1}^{n_\nu} q_j^{\{i\},\nu}(0)$, for $\nu = 1, \dots, l$, are involved in (23). In fact, for a vector $V \in \mathbb{R}^{1 \times N}$, let V^1, \dots, V^l denote a partition of the vector with $V^\nu \in \mathbb{R}^{1 \times n_\nu}$ for $\nu = 1, 2, \dots, l$. We have

$$V\pi = (V^1 \mathbb{1}_{n_1} \bar{\mu}^1, \dots, V^l \mathbb{1}_{n_l} \bar{\mu}^l).$$

Since

$$\mathbb{1}_{n_\nu} \bar{\mu}^\nu = \begin{pmatrix} \bar{\mu}_1^\nu & \cdots & \bar{\mu}_{n_\nu}^\nu \\ & \cdots & \\ \bar{\mu}_1^\nu & \cdots & \bar{\mu}_{n_\nu}^\nu \end{pmatrix},$$

$$V\pi = \left(\left(\sum_{j=1}^{n_1} V_j^1 \right) \bar{\mu}^1, \dots, \left(\sum_{j=1}^{n_l} V_j^l \right) \bar{\mu}^l \right).$$

Proof. First notice that owing to the block-diagonal structure of $A(0)$,

$$\exp(A(0)s) = \begin{pmatrix} \exp(A^1(0)s) & & & \\ & \exp(A^2(0)s) & & \\ & & \ddots & \\ & & & \exp(A^l(0)s) \end{pmatrix}.$$

For each $\nu = 1, 2, \dots, l$, consider the transition probability matrix $P^\nu(s)$ corresponding to the generator $A^\nu(0)$. It satisfies the differential equation

$$\begin{aligned} \dot{P}^\nu(s) &= P^\nu(s)A^\nu(0), \\ P^\nu(0) &= I. \end{aligned}$$

The solution is $P^\nu(s) = \exp(A^\nu(0)s)$. By the well-known ergodic theorem (see [2], Theorem II.10.1), $\lim_{s \rightarrow \infty} P^\nu(s)$ exists. Clearly 0 is an eigenvalue of $A^\nu(0)$. In view of the weak irreducibility, the multiplicity of the eigenvalue $\lambda_1^\nu = 0$ is 1. Since the limit $P^\nu(s)$ exists, the rest of the eigenvalues of $A^\nu(0)$ must have negative real parts. By virtue of [2], Theorem II.12.8, $P^\nu(s) \xrightarrow{s \rightarrow \infty} \mathbb{1}_{n_\nu} \bar{\mu}^\nu$, where as noted $\bar{\mu}^\nu$ is a row vector and is precisely the equilibrium distribution corresponding to the generator $A^\nu(0)$. Consequently, for some $\gamma_\nu > 0$,

$$|\exp(A^\nu(0)s) - \mathbb{1}_{n_\nu} \bar{\mu}^\nu| \leq K \exp(-\gamma_\nu s).$$

The discussion above implies that

$$\exp(A(0)s) \rightarrow \pi \text{ as } s \rightarrow \infty.$$

Since $(y(0) - p^{\{0\}}(0))\tilde{\mathbb{1}} = 0$, $(y(0) - p^{\{0\}}(0))\pi = 0$. It in turn yields

$$\begin{aligned} q^{\{0\}}(s) &= (y(0) - p^{\{0\}}(0))\pi + (y(0) - p^{\{0\}}(0))(\exp(A(0)s) - \pi) \\ &= (y(0) - p^{\{0\}}(0))(\exp(A(0)s) - \pi). \end{aligned}$$

As a result,

$$\begin{aligned} |q^{\{0\}}(s)| &\leq K \sum_{\nu=1}^l \left| (y^\nu(0) - p^{\{0\}, \nu}(0))(\exp(A^\nu(0)s) - \mathbb{1}_{n_\nu} \bar{\mu}^\nu) \right| \\ &\leq K \sum_{\nu=1}^l \exp(-\gamma_\nu s) \\ &\leq K \exp(-\tilde{\gamma}s), \end{aligned}$$

where $\tilde{\gamma} = \min_{\nu} \gamma_{\nu}$. In fact, we have shown that, for any vector $v \in \mathbb{R}^{1 \times N}$,

$$|v (\exp(A(0)s) - \pi)| \leq K \exp(-\tilde{\gamma}s).$$

We continue our estimate for the next term, $q^{\{1\}}(s)$, and show it decays exponentially. In this process, the crucial point is to choose the initial condition in a suitable way.

In view of (22),

$$\begin{aligned} q^{\{1\}}(s) &= q^{\{1\}}(0) \exp(A(0)s) + \int_0^s q^{\{0\}}(\tau) \exp(A(0)\tau) B(0) \exp(A(0)(s - \tau)) d\tau \\ &\quad + \int_0^s \tau q^{\{0\}}(\tau) \exp(A(0)\tau) A^{(1)}(0) \exp(A(0)(s - \tau)) d\tau. \end{aligned} \quad (24)$$

Recall that $q^{\{1\}}(0)$ has not been specified as yet.

Similar to [6], for each $t \in [0, T]$, $A(t)\tilde{\mathbb{I}} = 0$. Differentiating the equation above with respect to t , $d^i/dt^i(A(t)\tilde{\mathbb{I}}) = 0$. Therefore, $A^{(i)}(0)\pi = 0$ for $i = 1, \dots$. This together with $q^{\{0\}}(\tau)\pi = 0$ yields that

$$\begin{aligned} &\left| \int_0^s \tau q^{\{0\}}(\tau) \exp(A(0)\tau) A^{(1)}(0) \exp(A(0)(s - \tau)) d\tau \right| \\ &\leq \int_0^s \tau |q^{\{0\}}(\tau) [\exp(A(0)\tau) - \pi]| |A^{(1)}(0) [\exp(A(0)(s - \tau)) - \pi]| d\tau \\ &\leq K s^2 \exp(-\tilde{\gamma}s). \end{aligned} \quad (25)$$

To obtain the desired property, we need only work with the first two terms on the right side of the equality sign of (24). Noticing the exponential decay property of $q^{\{0\}}(s)$,

$$\int_0^{\infty} |q^{\{0\}}(\tau) \exp(A(0)\tau)| d\tau < \infty,$$

i.e., the improper integral converges absolutely.

By virtue of the same argument as in the first paragraph of the proof of this lemma,

$$\begin{aligned} \lim_{s \rightarrow \infty} q^{\{1\}}(0) \exp(A(0)s) &= q^{\{1\}}(0)\pi, \\ \lim_{s \rightarrow \infty} \int_0^s q^{\{0\}}(\tau) \exp(A(0)\tau) B(0) \exp(A(0)(s - \tau)) d\tau \\ &= \int_0^{\infty} q^{\{0\}}(\tau) \exp(A(0)\tau) d\tau B(0)\pi \\ &= \bar{q}^{\{0\}}\pi, \end{aligned} \quad (26)$$

where $\bar{q}^{\{0\}} \in \mathbb{R}^{1 \times N}$ is a constant vector.

By virtue of the structure of π , partitioning the vector $\bar{q}^{\{0\}}$ as $(\bar{q}^{\{0\},1}, \dots, \bar{q}^{\{0\},l})$, for $\nu = 1, \dots, l$, we have

$$\begin{aligned} q^{\{1\}}(0)\pi &= (q^{\{1\},1}(0) \mathbb{1}_{n_1} \bar{\mu}^1, \dots, q^{\{1\},l}(0) \mathbb{1}_{n_l} \bar{\mu}^l), \\ \bar{q}^{\{0\}}\pi &= (\bar{q}^{\{0\},1} \mathbb{1}_{n_1} \bar{\mu}^1, \dots, \bar{q}^{\{0\},l} \mathbb{1}_{n_l} \bar{\mu}^l). \end{aligned} \quad (27)$$

Note that

$$\mathbb{1}_{n_\nu} \bar{\mu}^\nu = \begin{pmatrix} \bar{\mu}_1^\nu & \cdots & \bar{\mu}_{n_\nu}^\nu \\ & \cdots & \\ \bar{\mu}_1^\nu & \cdots & \bar{\mu}_{n_\nu}^\nu \end{pmatrix}.$$

We thus have

$$\begin{aligned} q^{\{1\},\mu}(0) \mathbb{1}_{n_\nu} \bar{\mu}^\nu &= \left(\sum_{j=1}^{n_\nu} q_j^{\{1\},\nu}(0) \right) \bar{\mu}^\nu, \\ \bar{q}^{\{0\},\mu} \mathbb{1}_{n_\nu} \bar{\mu}^\nu &= \left(\sum_{j=1}^{n_\nu} \bar{q}_j^{\{0\},\nu} \right) \bar{\mu}^\nu. \end{aligned} \tag{28}$$

Set

$$q^{\{1\}}(0)\pi + \bar{q}^{\{0\}}\pi = 0. \tag{29}$$

In view of (27) and (28), although there are N equations in (29), only l of them are unknowns, namely,

$$\sum_{j=1}^{n_\nu} q_j^{\{1\},\nu}(0) \quad \text{for } \nu = 1, \dots, l.$$

This linear system is easily solved. Choose

$$f^{\{1\},\nu}(0) = - \sum_{i=1}^{n_\nu} q_i^{\{1\},\nu}(0) \quad \text{for } \nu = 1, \dots, l.$$

Then (18) is uniquely solvable. Substituting the results into (17), we obtain $p^{\{1\}}(\cdot)$. Finally, choose $q^{\{1\}}(0) = -p^{\{1\}}(0)$. The process of choosing initial conditions for $p^{\{1\}}(\cdot)$ and $q^{\{1\}}(\cdot)$ is completed.

Next, we have to verify that the first two terms on the right side of the equality sign in (24) decay exponentially. To this end,

$$\begin{aligned} &\left| q^{\{1\}}(0) \exp(A(0)s) + \int_0^s q^{\{0\}}(\tau) \exp(A(0)\tau) B(0) \exp(A(0)(s - \tau)) d\tau \right| \\ &\leq |q^{\{1\}}(0)(\exp(A(0)s) - \pi)| + \left| q^{\{1\}}(0)\pi + \int_0^\infty q^{\{0\}}(\tau) \exp(A(0)\tau) d\tau B(0)\pi \right| \\ &\quad + \int_0^s |q^{\{0\}}(\tau) \exp(A(0)\tau) B(0) [\exp(A(0)(s - \tau)) - \pi]| d\tau \\ &\quad + \int_s^\infty |q^{\{0\}}(\tau) \exp(A(0)\tau)| d\tau |B(0)\pi| \\ &\leq K \left(\exp(-\tilde{\gamma}s) + \int_0^s \exp(-\tilde{\gamma}\tau) \exp(-\tilde{\gamma}(s - \tau)) d\tau |B(0)| - \frac{1}{\tilde{\gamma}} |B(0)\pi| \exp(-\tilde{\gamma}\tau)|_s^\infty \right) \\ &\leq K \left(\exp(-\tilde{\gamma}s) + \int_0^s \exp(-\tilde{\gamma}\tau) d\tau |B(0)| \right) \\ &\leq K(1 + s) \exp(-\tilde{\gamma}s). \end{aligned}$$

Combining the estimate above with (25), we arrive at

$$|q^{\{1\}}(s)| \leq K(1 + s + s^2) \exp(-\tilde{\gamma}s) \leq K \exp(-\gamma_1 s),$$

where $0 < \gamma_1 < \tilde{\gamma}$.

Proceeding with the next term,

$$\begin{aligned}
 q^{\{2\}}(s) &= q^{\{2\}}(0) \exp(A(0)s) + \int_0^s q^{\{1\}}(\tau) B(0) \exp(A(0)(s-\tau)) d\tau \\
 &\quad + \int_0^s \tau q^{\{0\}}(\tau) B^{(1)}(0) \exp(A(0)(s-\tau)) d\tau \\
 &\quad + \int_0^s \tau q^{\{1\}}(\tau) A^{(1)}(0) \exp(A(0)(s-\tau)) d\tau \\
 &\quad + \int_0^s \tau^2 q^{\{0\}}(\tau) A^{(2)}(0) \exp(A(0)(s-\tau)) d\tau.
 \end{aligned} \tag{30}$$

Similar to the previous case, choose

$$f^{\{2\},\nu}(0) = - \sum_{j=1}^{n_\nu} q_j^{\{2\},\nu}(0) \quad \text{for } \nu = 1, \dots, l$$

where $\sum_{j=1}^{n_\nu} q_j^{\{2\},\nu}(0)$ are the solutions of

$$q^{\{2\}}(0)\pi + \left(\int_0^\infty q^{\{1\}}(\tau) d\tau B(0) + \int_0^\infty \tau q^{\{0\}}(\tau) d\tau B^{(1)}(0) \right) \pi = 0.$$

Then $p^{\{2\}}(\cdot)$ is uniquely determined. Next, choose $q^{\{2\}}(0) = -p^{\{2\}}(0)$. Similar to the estimate of $q^{\{1\}}(s)$,

$$|q^{\{2\}}(s)| \leq K \exp(-\gamma_2 s) \quad \text{for some } 0 < \gamma_2 \leq \tilde{\gamma}.$$

Suppose it has been shown that

$$|q^{\{i\}}(s)| \leq K \exp(-\gamma_i s), \quad i < n$$

for some $0 < \gamma_i < \tilde{\gamma}$. Solve

$$q^{\{n\}}(0)\pi = - \left(\sum_{i=0}^{n-1} \int_0^\infty \frac{\tau^i}{i!} q^{\{n-i-1\}}(\tau) d\tau B^{(i)}(0) \right) \pi$$

to obtain $\sum_{j=1}^{n_\nu} q_j^{\{n\},\nu}(0)$. Set

$$f^{\{n\},\nu}(0) = - \sum_{j=1}^{n_\nu} q_j^{\{n\},\nu}(0) \quad \text{for } \nu = 1, \dots, l.$$

Finally choose $q^{\{n\}}(0) = -p^{\{n\}}(0)$. Exactly the same argument leads to

$$|q^{\{n\}}(s)| \leq K \exp(-\gamma_n s) \quad \text{for some } 0 < \gamma_n < \tilde{\gamma}.$$

Thus the proposition follows.

6. Analysis of remainder. In the previous sections, regular and singular parts of the approximating sequence are constructed, and their desired properties are proved. The objective of this section is to carry out error analysis.

To assist us for subsequent study, we derive the following lemma.

LEMMA 6.1. Let $u^\varepsilon(\cdot)$ be the solution (on $[0, T]$) of the initial-value problem

$$\begin{aligned} \mathcal{L}^\varepsilon u^\varepsilon(t) &= \eta^\varepsilon(t) \text{ with } \sup_{t \in [0, T]} |\eta^\varepsilon(t)| = O(\varepsilon^{i+1}) \text{ for } i \leq n, \\ u^\varepsilon(0) &= 0. \end{aligned} \tag{31}$$

Then it satisfies

$$u^\varepsilon(t) = O(\varepsilon^i) \text{ uniformly in } t \in [0, T].$$

Proof. Using the notation $Q^\varepsilon(t) = \varepsilon^{-1}A(t) + B(t)$, the differential equation in (31) can be written as

$$\frac{du^\varepsilon(t)}{dt} = u^\varepsilon(t)Q^\varepsilon(t) + \frac{\eta^\varepsilon(t)}{\varepsilon}.$$

Denoting the principal matrix solution by $\Phi^\varepsilon(t, \tau)$, then

$$u^\varepsilon(t) = \frac{1}{\varepsilon} \int_0^t \eta^\varepsilon(\tau)\Phi^\varepsilon(t, \tau)d\tau.$$

By virtue of (2), $\Phi^\varepsilon(t)$ is bounded on $[0, T]$. Hence

$$|u^\varepsilon(t)| \leq \varepsilon^i \int_0^t |\Phi^\varepsilon(t, \tau)|d\tau \leq K\varepsilon^i.$$

With Lemma 6.1, we are now ready to derive:

PROPOSITION 6.2. For each $i = 0, 1, 2, \dots$, define

$$e^{\{i\}, \varepsilon}(t) = y^\varepsilon(t) - (\tilde{p}_i^\varepsilon(t) + \tilde{q}_i^\varepsilon(t/\varepsilon)). \tag{32}$$

Then

$$\sup_{0 \leq t \leq T} |e^{\{n\}, \varepsilon}(t)| = O(\varepsilon^{n+1}).$$

REMARK. In view of Lemma 6.1, it seems that the order in terms of ε for the solution of (31) is lower than $\mathcal{L}^\varepsilon u^\varepsilon$. If we wish to get (32), a direct estimate seems to be rather difficult. Thus we use the technique “back up one step” as in our work [6]. For notational simplicity, in what follows, suppress the ε -dependence in $e^{\{i\}, \varepsilon}(\cdot)$ and write it as $e^{\{i\}}(\cdot)$ instead.

Proof. For $i = 0, 1, 2, \dots, n$, it is easily seen that $e^{\{i\}}(0) = 0$ by the initial-value problem (1) and the initial data selected for $p^{\{i\}}(\cdot)$ and $q^{\{i\}}(\cdot)$. Since $\mathcal{L}^\varepsilon y^\varepsilon(t) = 0$, for each $i = 0, 1, 2, \dots, n$,

$$\mathcal{L}^\varepsilon e^{\{i\}}(t) = -\mathcal{L}^\varepsilon \sum_{j=0}^i \varepsilon^j p^{\{j\}}(t) - \mathcal{L}^\varepsilon \sum_{j=0}^i \varepsilon^j q^{\{j\}}(t/\varepsilon).$$

First consider $e^{\{1\}}(\cdot)$. Noticing that

$$p^{\{0\}}(t)A(t) = 0 \quad \text{and} \quad \dot{p}^{\{1\}}(t) = p^{\{2\}}(t)A(t) + p^{\{1\}}(t)B(t),$$

and using (8) and (21), upon cancellation,

$$\begin{aligned} \mathcal{L}^\varepsilon e^{\{1\}}(t) &= -\varepsilon \frac{dp^{\{0\}}(t)}{dt} + p^{\{0\}}(t)A(t) + \varepsilon p^{\{0\}}(t)B(t) \\ &\quad - \varepsilon^2 \frac{dp^{\{1\}}(t)}{dt} + \varepsilon p^{\{1\}}(t)A(t) + \varepsilon^2 p^{\{1\}}(t)B(t) \\ &\quad - \dot{q}^{\{0\}}(t/\varepsilon) + q^{\{0\}}(t/\varepsilon)A(t) + \varepsilon q^{\{0\}}(t/\varepsilon)B(t) \\ &\quad - \varepsilon \dot{q}^{\{1\}}(t/\varepsilon) + \varepsilon q^{\{1\}}(t/\varepsilon)A(t) + \varepsilon^2 q^{\{1\}}(t/\varepsilon)B(t) \\ &= -\varepsilon^2 p^{\{2\}}(t)A(t) + q^{\{0\}}(t/\varepsilon)[A(t) - (A(0) + tA^{(1)}(0)) + \varepsilon(B(t) - B(0))] \\ &\quad + \varepsilon q^{\{1\}}(t/\varepsilon)(A(t) - A(0)) + \varepsilon^2 q^{\{1\}}(t/\varepsilon)B(t) \\ &= O(\varepsilon^2) + q^{\{0\}}(t/\varepsilon)O(t^2 + \varepsilon t) + \varepsilon q^{\{1\}}(t/\varepsilon)O(t) + O(\varepsilon^2) \\ &= O(\varepsilon^2), \end{aligned} \tag{33}$$

where $\dot{q}^{\{i\}}(s) = \frac{d}{ds}q^{\{i\}}(s)$. In the above, we have used the Taylor expansions for $A(t)$ and $B(t)$ as well as Proposition 5.1, and noted that

$$|t^2 q^{\{0\}}(t/\varepsilon)| \leq K\varepsilon^2, \quad |t q^{\{0\}}(t/\varepsilon)| \leq K\varepsilon, \quad \text{and} \quad |t q^{\{1\}}(t/\varepsilon)| \leq K\varepsilon.$$

Applying Lemma 6.1, $e^{\{1\}}(t) = O(\varepsilon)$. However,

$$e^{\{1\}}(t) = e^{\{0\}}(t) - \varepsilon p^{\{1\}}(t) - \varepsilon q^{\{1\}}(t/\varepsilon).$$

Since $\varepsilon p^{\{1\}}(t) + \varepsilon q^{\{1\}}(t/\varepsilon) = O(\varepsilon)$, $e^{\{0\}}(t) = O(\varepsilon)$ as well. Moreover, the bound is uniform in $t \in [0, T]$.

Likewise, similar estimates lead to

$$\begin{aligned} \mathcal{L}^\varepsilon e^{\{n+1\}}(t) &= -\varepsilon \frac{d}{dt} \left(\sum_{i=0}^{n+1} \varepsilon^i p^{\{i\}}(t) + \sum_{i=0}^{n+1} \varepsilon^i q^{\{i\}}(t/\varepsilon) \right) \\ &\quad + \left(\sum_{i=0}^{n+1} \varepsilon^i p^{\{i\}}(t) + \sum_{i=0}^{n+1} \varepsilon^i q^{\{i\}}(t/\varepsilon) \right) (A(t) + \varepsilon B(t)) \\ &= -\varepsilon^{n+2} p^{\{n+2\}}(t)A(t) \\ &\quad + \sum_{i=0}^n \varepsilon^i q^{\{i\}}(t/\varepsilon) \left\{ \left[A(t) - \sum_{j=0}^{n-i+1} \frac{t^j A^{(j)}(0)}{j!} \right] + \varepsilon \left[B(t) - \sum_{j=0}^{n-i} \frac{t^j B^{(j)}(0)}{j!} \right] \right\} \\ &\quad + \varepsilon^{n+1} q^{\{n+1\}}(t/\varepsilon)(A(t) - A(0)) + \varepsilon^{n+2} q^{\{n+1\}}(t/\varepsilon)B(t). \end{aligned} \tag{34}$$

Using the Taylor expansions for $A(t)$ and $B(t)$, we can show that the above expression has an upper bound

$$\begin{aligned} &O(\varepsilon^{n+2}) + \sum_{i=0}^n \varepsilon^i q^{\{i\}}(t/\varepsilon)O(t^{n-i+2} + \varepsilon t^{n-i+1}) + \varepsilon^{n+1} q^{\{n+1\}}(t/\varepsilon)O(t) + O(\varepsilon^{n+2}) \\ &= O(\varepsilon^{n+2}) + q^{\{0\}}(t/\varepsilon)O(t^{n+2} + \varepsilon t^{n+1}) + \varepsilon q^{\{1\}}(t/\varepsilon)O(t^{n+1} + \varepsilon t^n) \\ &\quad + \dots + \varepsilon^{n+1} q^{\{n+1\}}(t/\varepsilon)O(t + \varepsilon), \end{aligned}$$

which is further bounded by

$$K(\varepsilon^{n+2} + (t^{n+2} + \varepsilon t^{n+1} + \dots + \varepsilon^{n+1}t)e^{-\gamma t/\varepsilon}) \leq K\varepsilon^{n+2}.$$

By virtue of Lemma 6.1, $e^{\{n+1\}}(t) = O(\varepsilon^{n+1})$ uniformly in $t \in [0, T]$. However,

$$\begin{aligned} e^{\{n+1\}}(t) &= e^{\{n\}}(t) - \varepsilon^{n+1}p^{\{n+1\}}(t) - \varepsilon^{n+1}q^{\{n+1\}}(t/\varepsilon) \\ &= e^{\{n\}}(t) + O(\varepsilon^{n+1}) \text{ uniformly in } t \in [0, T]. \end{aligned}$$

Therefore, $e^{\{n\}}(t) = O(\varepsilon^{n+1})$ uniformly in $t \in [0, T]$ as desired.

7. Summary of results and procedures. This section consists of two parts. First, the main results of this paper are summarized in Theorem 7.1. Then the second subsection gives an outline of the asymptotic expansion procedure.

7.1. *Summary of results.* The main results are given below.

THEOREM 7.1. Suppose the conditions (A1) and (A2) are satisfied. To approximate the solution $y^\varepsilon(\cdot)$ of (1), an asymptotic series expansion

$$Y_n^\varepsilon(\cdot) = \sum_{i=0}^n (\varepsilon^i p^{\{i\}}(t) + \varepsilon^i q^{\{i\}}(t/\varepsilon))$$

can be constructed in the following way: For each $i = 0, 1, \dots, n$,

1. $p^{\{0\}}(\cdot)$ is obtained uniquely from Proposition 3.2, and $q^{\{0\}}(\cdot)$ is obtained from Proposition 4.1 with the initial condition $q^{\{0\}}(0) = y(0) - p^{\{0\}}(0)$.
2. With temporarily unspecified initial data $f^{\{i\},\nu}(0)$, $p^{\{i\}}(\cdot)$ are obtained from Proposition 3.3.
3. With temporarily unspecified initial data $q^{\{i\}}(0)$, $q^{\{i\}}(\cdot)$ are obtained from (22).
4. Use Eq. (23) to determine $\sum_{j=1}^{n_\nu} q_j^{\{i\},\nu}(0)$.
5. For $i = 1, \dots, n$, $f^{\{i\}}(0)$ are selected so that

$$f^{\{i\},\nu}(0) = - \sum_{j=1}^{n_\nu} q_j^{\{i\},\nu}(0).$$

This together with 2 above uniquely determines $p^{\{i\}}(\cdot)$.

6. Choose $q^{\{i\}}(0) = -p^{\{i\}}(0)$. This together with 3 above uniquely determines $q^{\{i\}}(\cdot)$.
7. $p^{\{i\}}(\cdot) \in C^{n+2-i}[0, T]$.
8. $|q^{\{i\}}(t)| \leq K \exp(-\gamma t)$ for some $\gamma > 0$.
9. $y^\varepsilon(t) - Y_n^\varepsilon(t) = O(\varepsilon^{n+1})$ uniformly in $t \in [0, T]$.

REMARK. If $B(t)$ is a block diagonal matrix compatible with $A(t)$, i.e., it has the same kind of partition $B^\nu(t)$ with the same dimension as that of $A^\nu(t)$, then the results of Theorem 7.1 can be obtained as in [6].

We point out once more that (23) contains only l unknowns $\sum_{j=1}^{n_\nu} q_j^{\{i\}}(0)$. Thus using this equation, we do not determine all $q_j^{\{i\}}(0)$, for $j = 1, \dots, N$, but only the sum of the

components corresponding to each block partition. This is easily seen by noticing the structure of the matrices $\mathbb{1}_{n_\nu} \bar{\mu}^\nu$ and the structure of $\bar{q}^{\{i\}} \pi$ and $q^{\{i\}}(0) \pi$. In fact, denote

$$\bar{q}^{\{i\}} = \sum_{j=0}^{i-1} \int_0^\infty \frac{\tau^j}{j!} q^{\{i-j-1\}}(\tau) d\tau B^{(j)}(0).$$

Since

$$\mathbb{1}_{n_\nu} \bar{\mu}^\nu = \begin{pmatrix} \bar{\mu}_1^\nu & \cdots & \bar{\mu}_{n_\nu}^\nu \\ \vdots & \ddots & \vdots \\ \bar{\mu}_1^\nu & \cdots & \bar{\mu}_{n_\nu}^\nu \end{pmatrix},$$

for $\nu = 1, \dots, l$, we have

$$\begin{aligned} q^{\{i\}}(0) \pi &= (q^{\{i\},1}(0) \mathbb{1}_{n_1} \bar{\mu}^1, \dots, q^{\{i\},l}(0) \mathbb{1}_{n_l} \bar{\mu}^l) \\ &= \left(\left(\sum_{j=1}^{n_1} q_j^{\{i\},1}(0) \right) \bar{\mu}^1, \dots, \left(\sum_{j=1}^{n_l} q_j^{\{i\},l}(0) \right) \bar{\mu}^l \right) \end{aligned}$$

and

$$\begin{aligned} \bar{q}^{\{i\}} \pi &= (\bar{q}^{\{i\},1} \mathbb{1}_{n_1} \bar{\mu}^1, \dots, \bar{q}^{\{i\},l} \mathbb{1}_{n_l} \bar{\mu}^l) \\ &= \left(\left(\sum_{j=1}^{n_1} \bar{q}_j^{\{i\},1} \right) \bar{\mu}^1, \dots, \left(\sum_{j=1}^{n_l} \bar{q}_j^{\{i\},l} \right) \bar{\mu}^l \right). \end{aligned}$$

7.2. The computation procedure. Since the construction of $p^{\{i\}}(\cdot)$ and $q^{\{i\}}(\cdot)$ is rather involved, and the choice of initial conditions is very tricky, we summarize the procedure below. This procedure can certainly be used as a user's guide for developing the asymptotic expansion. The procedure is divided into two main stages.

Step 1: Initialization—finding $p^{\{0\}}(\cdot)$ and $q^{\{0\}}(\cdot)$.

1. Obtain the unique solution $p^{\{0\}}(\cdot)$ via (15).
2. Obtain the unique solution $q^{\{0\}}(\cdot)$ via (22) and the initial condition $q^{\{0\}}(0) = y(0) - p^{\{0\}}(0)$.

Step 2. Iteration—finding $p^{\{i\}}(\cdot)$ and $q^{\{i\}}(\cdot)$ for $i \leq n$.

While $i \leq n$, do the following:

1. Find $p^{\{i\}}(\cdot)$, the solution of (19) with temporarily unspecified $f^{\{i\},\nu}(0)$ for $\nu = 1, \dots, l$.
2. Obtain $q^{\{i\}}(\cdot)$ from (22) with temporarily unspecified $q^{\{i\}}(0)$.
3. Use the equation

$$q^{\{i\}}(0) \pi = - \left(\sum_{j=0}^{i-1} \int_0^\infty \frac{\tau^j}{j!} q^{i-j-1}(\tau) d\tau B^{(j)}(0) \right) \pi$$

to determine $\sum_{j=1}^{n_\nu} q_j^{\{i\},\nu}(0)$.

4. Set $f^{\{i\},\nu}(0) = - \sum_{j=1}^{n_\nu} q_j^{\{i\},\nu}(0)$. Up to now, $p^{\{i\}}(\cdot)$ are determined uniquely.
5. Choose $q^{\{i\}}(0) = -p^{\{i\}}(0)$. By now, $q^{\{i\}}(\cdot)$ have also been determined uniquely.
6. Set $i = i + 1$.
7. If $i > n$, stop.

8. An example. Let us consider a failure-prone manufacturing system that consists of a two-machine flowshop. Each of the two machines has two states up, denoted by 1, and down, denoted by 0. Then, the system has four states, represented by $\{(1, 1), (0, 1), (1, 0), (0, 0)\}$. Consider further that the state of the first machine is changing more frequently than the second one. Moreover, these two machines are probabilistic independent. One way of modeling the scenario is to formulate the state process (see, for example, [8] and the references therein) as a Markov chain with generator

$$Q^\varepsilon(t) = \frac{1}{\varepsilon} \begin{pmatrix} -\lambda(t) & \lambda(t) & 0 & 0 \\ \tilde{\mu}(t) & -\tilde{\mu}(t) & 0 & 0 \\ 0 & 0 & -\lambda(t) & \lambda(t) \\ 0 & 0 & \tilde{\mu}(t) & -\tilde{\mu}(t) \end{pmatrix} + \begin{pmatrix} -\lambda_1(t) & 0 & \lambda_1(t) & 0 \\ 0 & -\lambda_1(t) & 0 & \lambda_1(t) \\ \tilde{\mu}_1(t) & 0 & -\tilde{\mu}_1(t) & 0 \\ 0 & \tilde{\mu}_1(t) & 0 & -\tilde{\mu}_1(t) \end{pmatrix}.$$

In this model, the breakdown rate and repair rate for the first machine (or the second machine) are $\lambda(t)/\varepsilon$ and $\tilde{\mu}(t)/\varepsilon$ (or $\lambda_1(t)$ and $\tilde{\mu}_1(t)$), respectively. The probability of the state process is given by the $y^\varepsilon(t)$ that satisfies

$$\frac{dy^\varepsilon(t)}{dt} = y^\varepsilon(t)Q^\varepsilon(t),$$

$$y^\varepsilon(0) = y(0) = (y_1(0), y_2(0), y_3(0), y_4(0)), \text{ such that } y_i(0) \geq 0 \text{ and } \sum_{i=1}^4 y_i(0) = 1.$$

To solve this set of equations, note that

$$\begin{aligned} \frac{d}{dt}(y_1^\varepsilon(t) + y_3^\varepsilon(t)) &= -\frac{\lambda(t)}{\varepsilon}(y_1^\varepsilon(t) + y_3^\varepsilon(t)) + \frac{\tilde{\mu}(t)}{\varepsilon}(y_2^\varepsilon(t) + y_4^\varepsilon(t)), \\ \frac{d}{dt}(y_2^\varepsilon(t) + y_4^\varepsilon(t)) &= \frac{\lambda(t)}{\varepsilon}(y_1^\varepsilon(t) + y_3^\varepsilon(t)) - \frac{\tilde{\mu}(t)}{\varepsilon}(y_2^\varepsilon(t) + y_4^\varepsilon(t)), \\ \frac{d}{dt}(y_1^\varepsilon(t) + y_2^\varepsilon(t)) &= -\lambda_1(t)(y_1^\varepsilon(t) + y_2^\varepsilon(t)) + \tilde{\mu}_1(t)(y_3^\varepsilon(t) + y_4^\varepsilon(t)), \\ \frac{d}{dt}(y_3^\varepsilon(t) + y_4^\varepsilon(t)) &= \lambda_1(t)(y_1^\varepsilon(t) + y_2^\varepsilon(t)) - \tilde{\mu}_1(t)(y_3^\varepsilon(t) + y_4^\varepsilon(t)). \end{aligned}$$

To proceed, let

$$\begin{aligned} a_{13}(t) &= (y_1(0) + y_3(0)) \exp\left(-\frac{1}{\varepsilon} \int_0^t (\lambda(\tau) + \tilde{\mu}(\tau)) d\tau\right) \\ &\quad + \int_0^t \frac{\tilde{\mu}(u)}{\varepsilon} \exp\left(-\frac{1}{\varepsilon} \int_0^t (\lambda(\tau) + \tilde{\mu}(\tau)) d\tau\right) du, \\ a_{24}(t) &= (y_2(0) + y_4(0)) \exp\left(-\frac{1}{\varepsilon} \int_0^t (\lambda(\tau) + \tilde{\mu}(\tau)) d\tau\right) \\ &\quad + \int_0^t \frac{\lambda(u)}{\varepsilon} \exp\left(-\frac{1}{\varepsilon} \int_0^t (\lambda(\tau) + \tilde{\mu}(\tau)) d\tau\right) du, \\ a_{12}(t) &= (y_1(0) + y_2(0)) \exp\left(-\int_0^t (\lambda_1(\tau) + \tilde{\mu}_1(\tau)) d\tau\right) \\ &\quad + \int_0^t \tilde{\mu}_1(u) \exp\left(-\int_0^t (\lambda_1(\tau) + \tilde{\mu}_1(\tau)) d\tau\right) du, \end{aligned}$$

$$a_{34}(t) = (y_3(0) + y_4(0)) \exp\left(-\int_0^t (\lambda_1(\tau) + \tilde{\mu}_1(\tau)) d\tau\right) \\ + \int_0^t \lambda_1(u) \exp\left(-\int_0^t (\lambda_1(\tau) + \tilde{\mu}_1(\tau)) d\tau\right) du.$$

Then

$$\begin{aligned} y_1^\varepsilon(t) + y_3^\varepsilon(t) &= a_{13}(t), \\ y_2^\varepsilon(t) + y_4^\varepsilon(t) &= a_{24}(t), \\ y_1^\varepsilon(t) + y_2^\varepsilon(t) &= a_{12}(t), \\ y_3^\varepsilon(t) + y_4^\varepsilon(t) &= a_{34}(t). \end{aligned}$$

Note also that

$$\frac{d}{dt} y_1^\varepsilon(t) = -\left(\frac{\lambda(t)}{\varepsilon} + \frac{\tilde{\mu}(t)}{\varepsilon} + \lambda_1(t) + \tilde{\mu}_1(t)\right) y_1^\varepsilon(t) + \frac{\tilde{\mu}(t)}{\varepsilon} a_{12}(t) + \tilde{\mu}_1(t) a_{13}(t).$$

Thus,

$$y_1^\varepsilon(t) = y_1(0) \exp\left(-\int_0^t \left(\frac{\lambda(\tau) + \tilde{\mu}(\tau)}{\varepsilon} + \lambda_1(\tau) + \tilde{\mu}_1(\tau)\right) d\tau\right) \\ + \int_0^t \left(\frac{\tilde{\mu}(u)}{\varepsilon} a_{12}(u) + \tilde{\mu}_1(u) a_{13}(u)\right) \exp\left(-\int_0^t \left(\frac{\lambda(\tau) + \tilde{\mu}(\tau)}{\varepsilon} + \lambda_1(\tau) + \tilde{\mu}_1(\tau)\right) d\tau\right) du.$$

Consequently,

$$\begin{aligned} y_2^\varepsilon(t) &= a_{12}(t) - y_1^\varepsilon(t), \\ y_3^\varepsilon(t) &= a_{13}(t) - y_1^\varepsilon(t), \\ y_4^\varepsilon(t) &= a_{24}(t) - y_2^\varepsilon(t). \end{aligned}$$

In this example,

$$p^{\{0\}}(t) = (p^{\{0\},1}(t) f^{\{0\},1}(t), p^{\{0\},2}(t) f^{\{0\},2}(t)),$$

where

$$p^{\{0\},1}(t) = p^{\{0\},2}(t) = \left(\frac{\tilde{\mu}(t)}{\lambda(t) + \tilde{\mu}(t)}, \frac{\lambda(t)}{\lambda(t) + \tilde{\mu}(t)}\right)$$

and

$$\frac{d}{dt} (f^{\{0\},1}(t), f^{\{0\},2}(t)) = (f^{\{0\},1}(t), f^{\{0\},2}(t)) \begin{pmatrix} -\lambda_1(t) & \lambda_1(t) \\ \tilde{\mu}_1(t) & -\tilde{\mu}_1(t) \end{pmatrix},$$

with initial value $(f^{\{0\},1}(0), f^{\{0\},2}(0)) = (y_1(0) + y_2(0), y_3(0) + y_4(0))$.

$q^{\{0\}}(s)$ is given by

$$\dot{q}^{\{0\}}(s) = q^{\{0\}}(s) A(0), \quad q^{\{0\}}(0) = y^\varepsilon(0) - p^{\{0\}}(0).$$

By virtue of Theorem 7.1,

$$y^\varepsilon(t) - p^{\{0\}}(t) - q^{\{0\}}(t) = O(\varepsilon),$$

provided that $Q^\varepsilon(t) \in C^2[0, T]$. Noticing the exponential decay of $q^{\{0\}}(t)$, we have

$$y^\varepsilon(t) = p^{\{0\}}(t) + O(\varepsilon + e^{-\gamma t/\varepsilon}).$$

This implies, in particular, that for any $t > 0$,

$$\lim_{\varepsilon \rightarrow 0} y^\varepsilon(t) = p^{\{0\}}(t).$$

Namely, $p^{\{0\}}(t)$ is the "average" distribution of the Markov chain generated by $Q^\varepsilon(t)$.

REFERENCES

- [1] N. N. Bogoliubov and Y. A. Mitropolskii, *Asymptotic Methods in the Theory of Nonlinear Oscillations*, Gordon and Breach Science Publishers, New York, 1961
- [2] K. L. Chung, *Markov Chains with Stationary Transition Probabilities, Second Edition*, Springer-Verlag, New York, 1967
- [3] E. B. Dynkin, *Markov Processes*, Springer-Verlag, Berlin, 1965
- [4] S. N. Ethier and T. G. Kurtz, *Markov Processes, Characterization and Convergence*, J. Wiley, New York, 1986
- [5] J. K. Hale, *Ordinary Differential Equations, Second Edition*, R. E. Krieger Publishing Co., Malabar, FL, 1980
- [6] R. Z. Khasminskii, G. Yin, and Q. Zhang, *Asymptotic expansions of singularly perturbed systems involving rapidly fluctuating Markov chains*, SIAM J. Appl. Math. **56**, 277–293 (1996)
- [7] R. E. O'Malley, Jr., *Singular Perturbation Methods for Ordinary Differential Equations*, Springer-Verlag, New York, 1991
- [8] S. P. Sethi and Q. Zhang, *Hierarchical Decision Making in Stochastic Manufacturing Systems*, Birkhäuser, Boston, Cambridge, MA, 1994
- [9] A. B. Vasil'eva and V. F. Butuzov, *Asymptotic expansions of the solutions of singularly perturbed equations*, Nauka, Moscow, 1973
- [10] W. Wasow, *Asymptotic Expansions for Ordinary Differential Equations*, Wiley, New York, 1965