Singularly Perturbed Population Models with Reducible Migration Matrix: 2. Asymptotic Analysis and Numerical Simulations

Jacek Banasiak, Amartya Goswami, and Sergey Shindin

Abstract. The paper is concerned with asymptotic analysis of a singularly perturbed system of McKendrick equations of population with age and geographical structure. It is assumed that the migration between geographical patches occurs on a much faster time scale than the demographic processes and is described by a reducible Kolmogorov matrix. We apply a novel regularizing technique which makes the error estimates easier than that in previous papers and provide a numerical illustration of theoretical results.

Mathematics Subject Classification (2010). 92D25, 35B25, 35Q80, 47D06, 47N20.

Keywords. Structured population models, aggregation, singular perturbation, asymptotic analysis, semigroups, multiple time scales, Chapman–Enskog asymptotic expansion, initial layer, Reducible matrices.

1. Introduction

The present paper is an extension of [3–6] to reducible migration matrices and it can be treated as a (self-consistent) companion paper to [8]. All three papers are concerned with asymptotic analysis of McKendrick type equations with age and spatial structure, modelling fast migrations between the spatial patches and showing that for large migration rates the solutions of the original problem can be approximated by solutions to a much simpler aggregated model. In [8] and here we consider a more general mechanism of migrations than in [3–6], which is described in detail in Sect. 2. However, in
we focused on the functional analytic aspects of analysis by showing how the model fits into the general framework of Trotter–Kato and Sova–Kurtz theorems, see [9], and how the practical implementation of these theorems can be facilitated by a careful use of classical asymptotic expansions. However, we stopped short of performing the full analysis of the layers which is necessary for the convergence to occur uniformly in time and for all initial conditions. In this paper we complete the analysis by considering all necessary layers; that is, the initial, the boundary and the corner layers provide biological interpretation of the obtained results and illustrate them on a worked example. Using this example we also numerically demonstrate the accuracy of the approximations obtained by the asymptotic analysis. We emphasize that the asymptotic analysis used in this paper yields a more detailed error analysis than that yielded by [8] but at a much higher computational cost.

To describe the results in more detail, we have to recall the model.

1.1. The Model and Some Heuristic Considerations

In recent years there has been an interest in structured population models describing populations with various levels of organization such as individual, population, community or ecosystem levels. In many cases we observe different time scales at which each level evolves. For instance, considering demography of a population together with migratory processes between various geographical patches its individuals occupy, we often find that the latter, consisting, e.g. of moving between feeding and resting locations occurs at a much faster pace than the former (which includes e.g. breeding and deaths). To account for this, in a series of papers, [1, 6, 8, 16], the authors consider the following version of the McKendrick system:

\[ \partial_t u_{\epsilon} = Su_{\epsilon} - Mu_{\epsilon} + \frac{1}{\epsilon} Cu_{\epsilon}, \quad (1) \]

\[ (\gamma u_{\epsilon})(t) := u_{\epsilon}(0, t) = (B u_{\epsilon})(t) = \int_0^{\infty} B(a) u_{\epsilon}(a, t) \, da, \quad (2) \]

\[ u_{\epsilon}(a, 0) = w(a). \quad (3) \]

Here, as in [6, 8], \( u_{\epsilon}(a, t) = (u_{1,\epsilon}(a, t), \ldots, u_{n,\epsilon}(a, t)) \), where \( u_{i,\epsilon}(a, t) \) is the population density at time \( t \) of individuals residing in patch \( i \) and being of age \( a \). Further, the symbol \( \partial \) denotes the differentiation with respect to the variable appearing as the subscript and, accordingly, \( Su := -\partial_a u \) is the operator describing the ageing, \( M(a) = \{\mu_{ij}(a)\}_{1 \leq i, j \leq n} \) is the matrix describing mortality as well as possible slow migrations. The matrix \( C(a) = \{c_{ij}(a)\}_{1 \leq i, j \leq n} \) is a Kolmogorov matrix (whose columns sum to 0) which describes the migration of individuals between patches and \( 1/\epsilon \), with small \( \epsilon > 0 \), accounts for the fact that the rate of migration is much higher than the demographic rates. In the boundary condition, \( \gamma \) denotes the operator of taking the trace at \( a = 0 \) and \( B(a) = \{\beta_{ij}(a)\}_{1 \leq i, j \leq n} \) is the maternity matrix. The problem (1)–(3) is considered as an abstract Cauchy problem in the space \( X_n := L_1(\mathbb{R}^+, \mathbb{R}^n) = (L_1(\mathbb{R}^+, \mathbb{R}))^n \). Accordingly, the domain of \( S \) is the subspace of the Sobolev space \( Y_n := (W_1^1(\mathbb{R}^+))^n \).
If the rate of migration is much faster than the demographical processes, and the number of patches large, then it is of interest to provide an ‘aggregated’ description of the system which ‘averages’ over the geographical structure. Assuming for simplicity that $\mathcal{M}$ and $B$ are diagonal matrices, we can use the Kolmogorov structure of $\mathcal{C}$ to remove $\epsilon$ from (1) by adding the equations to get

$$\partial_t u = -\partial_a u - \sum_{i=1}^{n} \mu_i u_i,$$

where $u = \sum_{i=1}^{n} u_i$ is the density of the total population across all patches. However, Eq. (4) is not closed unless $\mu_1 = \cdots = \mu_n$. To approximately close it, in applications typically it is assumed that the ratios $u_i/u$ stabilize in time and thus can be replaced by the components of a constant vector $k = (k_1, \ldots, k_n)$ with $k_i \approx u_i/u$ for $i = 1, \ldots, n$, called the stable patch distribution. This leads to the approximating aggregated equation

$$\partial_t u \approx -\partial_a u - \mu^* u,$$

where $\mu^* = \sum_{i=1}^{n} \mu_i k_i$ is the aggregated mortality. The aggregated maternity rate $\beta^*$ is defined in the analogous way.

The results of [1,6,16] confirm this biological heuristics but only if the migration matrix $\mathcal{C}$ is irreducible, see, e.g. [20]. This ensures that $\lambda = 0$ is the dominant simple eigenvalue of $\mathcal{C}$ with corresponding positive right eigenvector $k$, and a left eigenvector $1 = (1, 1, \ldots, 1)$. These vectors are normalized to satisfy $1 \cdot k = 1$ so that $k$ indeed gives the stable patch distribution mentioned above. Then, introducing the spectral decomposition of $\mathbb{R}^n$ corresponding to $\lambda = 0$, by $\mathbb{R}^n = \text{Span}\{k, w\}$, we can write $u_\epsilon(t) = u_\epsilon(t)k + w_\epsilon(t)$ and $\tilde{u} = \tilde{u}k + \tilde{w}$. We denote by $\tilde{u}$ the solution of the full aggregated model corresponding to (5),

$$\partial_t \tilde{u} = -\partial_a \tilde{u} - \mu^* \tilde{u},$$

$$\tilde{u}(0, t) = \int_{0}^{\infty} \beta^*(a) \tilde{u}(a, t) \ da, \quad \tilde{u}(a, 0) = \sum_{i=1}^{n} \tilde{u}_i.$$ 

Actually, the analysis can be done for general $\mathcal{M}$ and $B$. Then, see [6],

$$\|u_\epsilon(\cdot, t) - \tilde{u}(\cdot, t)\|_{L_1(\mathbb{R}^+)} = O(\epsilon), \quad \|w_\epsilon(\cdot, t) - \epsilon^{\frac{1}{2}} \mathcal{C}(\cdot) \tilde{w}(\cdot)\|_{(L_1(\mathbb{R}^n))^n} = O(\epsilon),$$

uniformly in $t$ on finite intervals of $[0, \infty)$. Since $\lambda = 0$ is the strictly dominant eigenvalue, we recognize $\epsilon^{\frac{1}{2}} \mathcal{C}(\cdot) \tilde{w}(\cdot)$ as the initial layer contribution to the approximation, whose influence rapidly decreases for $t > 0$ and small $\epsilon$.

It is to be emphasized that the above result only is valid for irreducible migration matrices. This assumption often proves to be too restrictive. For instance, in population theory it amounts to requiring that any geographical patch is accessible from any other, see e.g. [10]. In many cases, however, one has to consider patches that are either isolated or only admitting emigration (or only immigration). A typical example is a population with a breeding refuge which typically is inaccessible to males living in other patches but
supplies them with offspring, or a pool above a waterfall which forms a barrier for fish living downstream. Such models result in systems of the form (1), but with a reducible matrix $C$. In such cases the total population $u$ cannot be approximated by the solution of a scalar equation, such as Eq. (6), but instead one must use a system of equations for the population density in each of the so-called ergodic components of the system.

The appropriate aggregated system, that replaces (6) if $C$ is reducible, was constructed in [8,12], where we also used the Sova–Kurtz theory to prove the convergence of $u$, to the solutions of this system as $\epsilon \to 0$, but only for the initial conditions from the ergodic states. The main aim of this paper was to provide a detailed description of the decomposition of the state space into ergodic and the complementary, transient, states and, by a careful analysis of the appropriate layers, to extend the convergence results to the solutions emanating from arbitrary initial conditions. The paper is concluded by a worked example and numerical simulations illustrating the accuracy of the constructed approximations.

2. The Migration Matrix
As we know from [6,8], the key to asymptotic analysis of (1)–(3) is understanding the large time behaviour of the solutions to the problem

$$\partial_t u = Cu, \quad u(0) = u_0,$$

where $u = (u_1, \ldots, u_n) \in \mathbb{R}^n$ and $u_0$ is an initial condition, which describes the so-called fast dynamics of (1). Here the age variable $a$ is a parameter and thus it will be suppressed in this section.

The general theory explicitly describing behaviour of systems with reducible transition matrices seems to be hard to achieve due to a complicated internal structure they may have, see, e.g. [7]. However, in the structured population models considered here, we only consider matrices $C$ that describe migration and, as mentioned before, the principle of conservation of the total population yields that they are Kolmogorov matrices. A Kolmogorov matrix is a particular case of an ML matrix, see [20, Section 2.3], or a Metzler matrix, [17, Section 6.5], in which columns sum up to zero. In other words, a Kolmogorov matrix is the (transpose) of a $Q$ matrix describing a continuous time Markov process with finitely many states. The theory of the large time behaviour of such processes, even in the reducible case, is relatively well developed in the probabilistic context, especially in the discrete time, see, e.g. [18, Section 8.4] or [2, Chapter 5]. Here we shall look at it from the dynamical systems point of view which seems to be less available.

Before we discuss the relevant properties of Kolmogorov matrices, let us introduce basic notation. The spectrum of $C$ will be denoted by $\sigma(C)$. The spectral bound of $C$ is defined as $s(C) := \max\{\Re \lambda; \lambda \in \sigma(C)\}$. It is easy to see that $1 := (1, \ldots, 1) \in \mathbb{R}^n$ is a left eigenvector of any Kolmogorov matrix $C$, corresponding to the eigenvalue 0 which is dominant, that is, all other eigenvalues $\lambda \in \sigma(C)$ satisfy $\Re \lambda < 0$, see, e.g. [17, Section 6.5, Theorem 1].
Let us recall, [11, p. 90] (note, however, that due to the particular application we have in mind, we use the notation of [11] in the transposed form), that a reducible Kolmogorov matrix $C$ can be written, by simultaneous permutation of rows and columns, in the following normal form:

$$
C := \begin{bmatrix}
C_1 & \ldots & \mathcal{O} & A_{1,m+1} & \ldots & A_{1,r} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathcal{O} & \ldots & C_m & A_{m,m+1} & \ldots & A_{m,r} \\
\mathcal{O} & \ldots & \mathcal{O} & T_{m+1} & \ldots & T_{m+1,r} \\
\vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\
\mathcal{O} & \ldots & \mathcal{O} & \mathcal{O} & \ldots & T_r
\end{bmatrix},
$$

where $C_l = [c_{ij}]_{n_l \leq i,j \leq N_l}$ for $l \leq m$ and $T_l = [c_{ij}]_{n_l \leq i,j \leq N_l}$ for $m + 1 \leq l \leq r$, where $n_1 = 1, n_{l+1} = N_l + 1$ and $N_r = n$, are square irreducible matrices (with some possibly being the scalar 0). Furthermore, if $r > m$, then in each column of (9) at least one of the off-diagonal matrices is not equal to $\mathcal{O}$. Clearly, the matrices $A_{i,j}$, $1 \leq i \leq m, m + 1 \leq j \leq r$ and $T_{i,j}$, $m + 1 \leq i \leq r, i + 1 \leq j \leq r$ are nonnegative. The structure of (9) emphasizes the division of the space $\mathbb{R}^n$ of states into two sets, namely the closed, or ergodic, states from which no transition outside is possible ($1 \leq j \leq N_m$) and transient states from which a transition to other states is possible ($n_{m+1} \leq j \leq N_r$).

Using the standard equivalence of positive off-diagonal matrices and nonnegative matrices, see, e.g. [20, Section 2.3], and our convention for writing the transpose of the normal form, [11, Chapter 3, Theorems 6 & 7] yields

**Theorem 2.1.** Let $C$ be a reducible Kolmogorov matrix written in the form (9). Then

1. $\lambda = s(C) = 0$ is the dominant eigenvalue of $C$, semi-simple with multiplicity $m$, and $1 = (1, \ldots, 1)$ is a strictly positive left eigenvector corresponding to $\lambda = 0$.
2. For each $1 \leq l \leq m$, each matrix $C_l$ is irreducible and $s(C_l) = 0$ is its dominant eigenvalue.
3. For each $m + 1 \leq l \leq r$, the matrices $T_l$, $m + 1 \leq l \leq r$ are irreducible and their dominant eigenvalues satisfy $s(T_l) < 0$.
4. There is a nonnegative basis of $N_r(C)$, the right null space of $C$, consisting of right eigenvectors $\{e_1, \ldots, e_m\}$, and if $e_l = (e_{l,j})_{1 \leq j \leq n_l}, l = 1, \ldots, m$, then the entries can be selected to satisfy $e_{l,j} > 0$ for $n_l \leq j \leq N_l$, $e_{l,j} = 0$ for all other $j$ and $\sum_{j=n_l}^{N_l} e_{l,j} = 1$ for each $l, 1 \leq l \leq m$.
5. There is a nonnegative basis of $N_l(C)$, the left null space of $C$, consisting of left eigenvectors $\{x_1, \ldots, x_m\}, x_l = (x_{l,j})_{1 \leq j \leq n_l}, l = 1, \ldots, m$, which can be selected to satisfy $x_{l,j} = 1$ for $n_l \leq j \leq N_l, x_{l,j} = 0$ for $1 \leq j \leq N_m, j \neq n_l, \ldots, N_l$ and $x_1 + \cdots + x_m = 1$.
6. There is a strictly positive right eigenvector of $C$ corresponding to $\lambda = 0$ if and only if $r = m$, that is, when there are only ergodic states.

We call a left (resp. right) eigenvector corresponding to the dominant eigenvalue $\lambda = 0$, a left (resp. right) Perron eigenvector.
Remark 2.1. Statements 2 and 3 of the theorem are equivalent to the existence of a strictly positive left Perron eigenvector which, in the present case, is $1$. Considering, however, only Kolmogorov matrices does not compromise generality of the considerations since any matrix $A$ with strictly positive left Perron eigenvector and $s(A) = 0$ is similar to a Kolmogorov matrix. Indeed, let $p = (p_1, \ldots, p_n)$ be a strictly positive left Perron eigenvector. Denoting $P = \text{diag}\{ p_1, \ldots, p_n \}$ and $Z = \{ z_{ij} \}_{1 \leq i, j \leq n} = PA^{-1}$ we have
\[
\sum_{i=1}^{n} z_{i,j} = p_j^{-1} \sum_{i=1}^{n} p_i a_{i,j} = 0
\]
for any $1 \leq j \leq n$.

A discussion of Theorem 2.1 and proofs of its less known statements is given in [8]. Here we recall the formulae which are necessary for the asymptotic analysis carried out in this paper and provide an interpretation of them in the context of population theory.

Let us denote by $\nu_i$ the dimension of the matrices $C_i, 1 \leq i \leq m$, and $T_i, m + 1 \leq i \leq r$, introduced in (9); that is $\nu_i = N_i - n_i + 1, 1 \leq i \leq r$. Any right Perron eigenvector $k := (k_1, k_2, \ldots, k_r)$, (10) with $k_i := (k_{i,n_i}, \ldots, k_{i,N_i})$, $1 \leq i \leq r$ can be obtained by solving the system
\[
C_i k_i + \sum_{j=m+1}^{r} A_{i,j} k_j = 0, \quad 1 \leq i \leq m,
\]
\[
T_i k_i + \sum_{j=i+1}^{r} T_{i,j} k_j = 0, \quad m + 1 \leq i \leq r,
\]
where hereafter we adopt the convention that any summation over an empty set of indices is 0. Only $m$ first entries in (10) are nonzero (even strictly positive) and the eigenvectors $e_i$, $i = 1, \ldots, m$, of Theorem 2.1, point 4., can be obtained by normalizing them and extending by 0 to $n$ dimensional vectors.

For a left Perron eigenvector $y = (y_1, \ldots, y_r)$ of $C$ [notation as in (10)] we clearly get $y_i = 1_{\nu_i} = (1, \ldots, 1)$ and, since $T_i, m + 1 \leq i \leq r$ are invertible,
\[
y_i = \left( \sum_{j=1}^{m} y_j A_{j,i} + \sum_{j=m+1}^{i-1} y_j T_{j,i} \right) (-T_i)^{-1}, \quad m + 1 \leq i \leq r.
\]
This formula recursively defines $y_i, m + 1 \leq i \leq r$, since the first term is given explicitly by
\[
y_{m+1} = \left( \sum_{j=1}^{m} y_j A_{j,m+1} \right) (-T_{m+1})^{-1}.
\]
The explicit formula for $y_l, l \geq m + 1$, is related to the characterization of absorbing states in reducible Markov chains (see, e.g. [18, Section 8.4] in the discrete case). As the continuous case seems to be less known, we provide an
elementary algebraic discussion of it below. For this, we have to introduce some notation. For given \( m + 1 \leq j, l \leq r \), \( V_{j,l} \) will denote the set of all sequences of indices \( s := \{ j, s_1, \ldots, s_k, l \} \) with \( j < s_1 < \cdots < s_k < l \). In other words, \( V_{j,l} \) is the set of all potential paths from \( j \) to \( l \). Let \( n_s \) denote the length of \( s \). Then, identifying \( s_0 = j, s_{n_s} = l \), we write \( s = \{ s_i \}_{0 \leq i \leq n_s} \).

**Proposition 2.1.** For \( m + 1 \leq l \leq r \)

\[
y_l = \left( \sum_{j=m+1}^{l} \sum_{i=1}^{m} y_i A_{ij} \left( \sum_{s \in V_{j,l}, s_k \in s \setminus \{ l \}} \prod_{k} (-T_{s_k})^{-1}T_{s_k s_{k+1}} \right) \right) (-T_l)^{-1}. \tag{13}
\]

Here, we adopt convention that any summation over an empty set of indices is 0 and the product over such set of indices is the identity matrix. Further, in the products of matrices the indices increase from left to right.

**Proof.** To shorten notation we denote

\[
T_{j,l} = \sum_{s \in V_{j,l}, s_k \in s \setminus \{ l \}} (-T_{s_k})^{-1}T_{s_k s_{k+1}}.
\]

To prove (13), we use induction. The formula is true for \( l = m + 1 \). Indeed, then \( V_{m+1,m+1} = \{ m + 1 \} \), the set over which we take the product is empty and the formula reduces to (12). Assuming now that (13) is valid for all \( i = m + 1, \ldots, l \), from (11) we have

\[
y_{l+1} = \left( \sum_{i=1}^{m} y_i A_{il+1} + \sum_{i=m+1}^{l} y_i T_{i,l+1} \right) (-T_{l+1})^{-1}
\]

\[
= \left( \sum_{i=1}^{m} y_i A_{il+1} + \sum_{i=m+1}^{l} \left( \sum_{j=m+1}^{l} \sum_{p=1}^{m} y_p A_{pj} T_{j,i} \right) (-T_i)^{-1}T_{i,l+1} \right) (-T_{l+1})^{-1}.
\]

For \( m + 1 \leq j \leq l \), the coefficient at \( \sum_{p=1}^{m} y_p A_{pj} \) is \( \sum_{i=j}^{l} T_{j,i}(-T_i)^{-1}T_{i,l+1} = T_{j,l+1} \), which gives

\[
y_{l+1} = \left( \sum_{j=m+1}^{l+1} \sum_{i=1}^{m} y_i A_{ij} T_{j,l+1} \right) (-T_{l+1})^{-1}.
\]

This ends the proof of (13). \( \square \)

Defining \( y_l^{(i)} = y_i \left( \sum_{j=m+1}^{l} A_{ij} T_{j,l}(-T_l)^{-1} \right) \) for a fixed \( 1 \leq i \leq m \) and \( m + 1 \leq l \leq r \), we obtain the basis of the left null space of \( C \) as the span of

\[
x_1 = \left( y_1, 0, \ldots, 0, y_{m+1}^{(1)}, \ldots, y_r^{(1)} \right),
\]

\[
\vdots
\]

\[
x_m = \left( 0, \ldots, 0, y_m, y_{m+1}^{(m)}, \ldots, y_r^{(m)} \right), \tag{14}
\]

where \( y_1 = (1, \ldots, 1) \), \( \ldots, y_m = (1, \ldots, 1) \) and \( y_{l}^{(i)} = (f_{i,n_1}, \ldots, f_{i,N_1}) \).
We can strengthen, [8], the statement 5 of Theorem 2.1 to
\[ \sum_{i=1}^{m} f_{i,p} = 1 \quad n_{m+1} \leq p \leq n. \tag{15} \]

Furthermore, since \( T_l, m + 1 \leq l \leq r \), are positive off diagonal matrices with \( s(T_l) < 0 \), they are resolvent positive so that \( (-T_l)^{-1} \geq 0 \). Hence, \( x_i \geq 0 \), \( i = 1, \ldots, m \).

Using the constructed eigenvectors, we obtain the spectral decomposition of the state space as
\[ \mathbb{R}^n = \text{Span}\{ e_1, \ldots, e_m \} \oplus W =: V \oplus W, \tag{16} \]
with the spectral projections onto the eigenvectors defined by
\[ P_k u = (x_k \cdot u) e_k = \left( \sum_{i=n_k}^{N_k} u_i + \sum_{i=n_{m+1}}^{n} f_{k,i} u_i \right) e_k =: u^k e_k, \quad 1 \leq k \leq m. \tag{17} \]
Thus, the projection of \( u \) onto the null space \( V \) of \( C \) is given by
\[ v = P u := \sum_{k=1}^{m} P_k u = \sum_{k=1}^{m} u^k e_k. \tag{18} \]

We also will use the \( m \)-dimensional vector of coordinates \( v = (u^1, \ldots, u^m) \) which satisfies
\[ v = \mathcal{E}_v u := \begin{pmatrix} | & \cdots & | \\ e_1 & \cdots & e_m \end{pmatrix} u. \tag{19} \]

The complementary space \( W \) defined by
\[ W = \{ x \in \mathbb{R}^n; P_k x = 0, k = 1, \ldots, m \} \tag{20} \]
with the complementary projection
\[ Q u := [I - (P_1 + \cdots + P_m)] u =: w. \]

However, in contrast to [8], there is no need to construct the full spectral basis in \( \mathbb{R}^n \) as we use another approach. Hence, we leave elements of \( W \) in coordinate free form.

**Remark 2.2.** In the considered meta-population it is clear that the states \( j = n_{m+1}, \ldots, n \), eventually will become depleted with the population moving towards states \( j = 1, \ldots, N_m \). However, according to (17), the population in any final state will contain a contribution from the transient states that communicate with that state (but not from any other ergodic state). It is interesting to quantify these contributions. First, let \( v_l \) be a solution of
\[ \partial_t v_l = T_l v_l, \quad v_l(0) = v_{l,0} \in \mathbb{R}^{n_l}, \quad m + 1 \leq l \leq r. \tag{21} \]

Using the fact that \( s(T_l) < 0 \), \( v_l(t) = e^{T_l t} v_{l,0} \) is integrable on \( \mathbb{R}^+ \) and
\[ \int_0^\infty v_l(t) \, dt = -(-T_l)^{-1} \int_0^\infty T_l e^{T_l t} v_{l,0} \, dt = -(-T_l)^{-1} \int_0^\infty \partial_t v_l(t) \, dt = -(-T_l)^{-1} v_{l,0}. \]
Hence, \((-T_l)^{-1}v_{l,0}\) is the vector of all individuals that originated from the initial state \(v_{l,0}\) and existed at any time at any state \(n_l \leq j \leq N_l\). This interpretation is clear if \(T_l = t_l\) is one-dimensional. Then \(1/t_l\) is the average lifespan of an individual and thus \(v_{l,0}/t_l\) is the average number of individuals from the initial cohort \(v_{l,0}\) that will be in the system over the whole time.

Recall that
\[
y_l^{(i)} = y_t \left( \sum_{j=m+1}^{l} A_{i,j} T_{j,l} (-T_l)^{-1} \right).
\] (22)

Now, since \(T_{j,l}\) is the matrix covering all possible transitions from the block of states \(l\) to the block \(j\) with \(m+1 \leq j, l \leq r\), and \(A_{i,j}\) is the matrix of rates, with respect to the chosen unit of time, of transfers from states \(m+1 \leq j \leq r\) to the states \(1 \leq i \leq m\), \(A_{i,j} T_{j,l} (-T_l)^{-1} v_{l,0}\) is the vector giving all individuals from the cohort \(v_{l,0}\) who moved to the states in the block \(i\) over the whole time. Now, remembering that \(y_i = 1\), we see, by multiplying both sides of (22) with \(v_{l,0}\), that \(y_l^{(i)} v_{l,0}\) is the total contribution to the states in the block \(i\) with \(1 \leq i \leq m\) from all states in the block \(l\) with \(m+1 \leq l \leq r\).

3. Model with General \(\mathcal{M}(a), \mathcal{B}(a)\) and Reducible \(\mathcal{C}(a)\)

In the previous section we suppressed the age variable \(a\) in the notation since in the analysis of (8) it has played the role of a parameter. Hereafter we return to the analysis of the full problem (1)-(3). Let, for \(j \in \mathbb{N}, X_j := L^1(\mathbb{R}^+, \mathbb{R}^j)\) and \(Y_j := W^1_1(\mathbb{R}^+, \mathbb{R}^j)\). We recall that \(X_N\) is our state space. We denote by \(|\cdot|_j\) the norm in the \(l^1_j\) equipped with \(l^1\) norm). By \(|\cdot|_j\) we denote the norm of a function \(u : X \rightarrow l^1_j\),
\[
|u|_j = \int_0^\infty |u(a)|_j \, da.
\]

We drop the index \(j\) in the notation of the space and the respective norm if the dimension of the space is clear from the context. For any other norm in a Banach space \(Z\) we shall write \(|\cdot|_Z\). By \(\mathcal{B}(Z_1, Z_2)\) we denote the space of linear bounded operators from \(Z_1\) to \(Z_2\), equipped with the standard operator norm \(|\cdot|_\mathcal{B}(Z_1, Z_2)\) (with short notation \(|\cdot|_\mathcal{B}\) if no ambiguity arises).

We consider the problem (1)-(3) with general \(a\) dependent matrices \(\mathcal{M}(a)\) and \(\mathcal{B}(a)\geq 0\) discussed in Introduction. We assume that their entries are bounded. The matrix \(-\mathcal{M}(a)\) is assumed to be sub-Kolmogorov; that is, \(-\mathcal{M}(a)\) is positive off-diagonal and satisfies \(-\sum_{i=1}^n \mu_{ij}(a) \leq 0\) for any \(1 \leq j \leq n\) and a.e. \(a \in \mathbb{R}_+\). Furthermore, we assume that \(a \mapsto \mathcal{M}(a) \in C^1_b(\mathbb{R}_+, \mathbb{R}^{n^2})\) and \(a \mapsto \mathcal{C}(a) \in C^2_b(\mathbb{R}_+, \mathbb{R}^{n^2})\) (bounded functions with bounded derivatives), that the structure (9) is independent of \(a\) and all eigenvalues of \(\mathcal{C}_W := \mathcal{C}|_W\) have negative real parts with
\[
\max_{\lambda \in \sigma(\mathcal{C}_W(a)), a \in \mathbb{R}_+} \Re \lambda =: \zeta < 0.
\] (23)
Lemma 3.1. [8, Lemma 4.1] Under the above assumptions, $C_{W}^{-1} \in C_{b}^{2}(\mathbb{R}^{+}, \mathbb{R}^{n})$ and $e_{j} \in C_{b}^{2}(\mathbb{R}^{+}, \mathbb{R}^{n})$ for $1 \leq j \leq m$.

In what follows, slightly abusing the notation, we denote the operators in $X$, generated by pointwise multiplication of elements of $X$ by matrices $M(a), C(a), P(a), \text{etc}$; that is, e.g. $X \ni f(\cdot) \to C(\cdot)f(\cdot) \in X$, with the same letters. Due to the adopted assumptions, these operators are bounded in $X$.

We further denote $Bf = \int_{0}^{\infty} B(a) f(a) \, da$ and, finally, $A_{\epsilon}$ is the operator given by the differential expression $A_{\epsilon} = S - M + \epsilon^{-1}C$ restricted to the domain $D(A_{\epsilon}) := \{ u \in Y; u(0) = Bu \}$. It generates a $C_{0}$-semigroup, say $(T_{A_{\epsilon}}(t))_{t \geq 0}$, satisfying

$$|||T_{A_{\epsilon}}(t)||| \leq e^{\omega t},$$

(24)

where $\omega$ is a constant depending on $B$ and $M$ but independent of $\epsilon$, see [8,15,21].

For the error estimates one needs to deal with inhomogeneous boundary conditions in problems of the form

$$u_{t} = -Mu + Su + \frac{1}{\epsilon}Cu + h, \quad \gamma u = Bu + f, \quad u|_{t=0} = \bar{u},$$

(25)

where $f$ is a function differentiable with respect to $t$. If $f \neq 0$, the problem cannot be approached by only using semigroup theory but some version of the lifting technique must be used. By linearity, we can consider the case with $\bar{u} = 0$ and $h = 0$.

Let $L_{\epsilon}(a_{0}, a)$ be the fundamental solution matrix of

$$z_{a}' = \left( -M + \frac{1}{\epsilon}C \right) z, \quad z(a_{0}) = z_{0};$$

(26)

that is, $z(a) = L_{\epsilon}(a_{0}, a)z_{0}$ is the solution of this problem. Then, by [6, Lemma 4.2], any solution to (25) with $\bar{u} = 0$ and $h = 0$ is given by

$$u(t, a) = L_{\epsilon}(0, a)\omega(t, a),$$

(27)

where

$$\omega(t, a) = \begin{cases} 0, & a > t, \\ ((I - BL_{\epsilon}(0, \cdot))^{-1}f)(t - a), & a < t. \end{cases}$$

One of the problems with classical asymptotic approach to (1)–(3) is that the solution to it as well as the solution to the aggregated system (6) are discontinuous unless the initial value satisfies compatibility conditions described in the definition of the generator of the semigroup. Thus, in general, using the differential form of the equation is in both cases incorrect. One can deal with this difficulty as in [6] and write the problem in the integral form. Such an approach gives optimal error estimates in $\epsilon$, but it requires cumbersome calculations and depends on the fact that the differentiation operator $S$ is an $n$-fold copy of $-\partial_{a}$, which restricts its applicability. Here we suggest an alternative approach which allows for working with the original differential formulation and thus also can be applied to systems with different growth rates, such as occur in size structured populations. The method depends on the following lemma:
Lemma 3.2. Let $\psi \in Y$. For any sufficiently small $\delta > 0$ there exists $\psi_\delta \in D(A_\epsilon)$ such that $\|\psi - \psi_\delta\|_X \leq C\delta\|\psi\|_Y$ and $\|\partial_a \psi - \partial_a \psi_\delta\|_X \leq C\|\psi\|_Y$, where $C$ is a constant independent of $\delta$ and $\psi$.

Proof. Consider $\psi_\delta(a) = \psi(a) + \phi e^{-a/\delta}$, where $\phi$ is an undetermined vector. We require

$$\psi_\delta(0) = \psi(0) + \phi = \int_0^\infty B(a)\psi(a) \, da + \left( \int_0^\infty e^{-a/\delta} B(a) \, da \right) \phi. \quad (28)$$

Now, $||| \int_0^\infty e^{-a/\delta} B(a) \, da |||_{B(1)} \leq b\delta$ for some $b$ and, therefore, for sufficiently small $\delta$, (28) can be solved for $\phi$ with $\phi = H_\delta \left( \int_0^\infty B(a)\psi(a) \, da - \psi(0) \right)$, where $||| H_\delta |||_{B(1)} \leq (1 - \delta b)^{-1}$ and, consequently, for sufficiently small $\delta < \delta_0 \leq 1/2b$,

$$|\phi| \leq \frac{1}{1 - \delta b} \left( b\|\psi\|_X + C'\|\psi\|_Y \right) \leq C\|\psi\|_Y,$$

where we used the continuity of the trace operator $\psi \to \psi(0)$ from $Y = (W_1'(\mathbb{R}_+))^n$ to $\mathbb{R}^n$ and $C$ only depends on $\delta_0, b$ and the norm of the trace operator. Hence the lemma follows by

$$\|\psi - \psi_\delta\|_X \leq |\phi| \int_0^\infty e^{-a/\delta} \, da = \delta|\phi| \leq C\delta\|\psi\|_Y,$$

$$\|\partial_a \psi - \partial_a \psi_\delta\|_X \leq \frac{|\phi|}{\delta} \int_0^\infty e^{-a/\delta} \, da = |\phi| \leq C\|\psi\|_Y.$$

4. The Asymptotic Expansion and Its Convergence

Decomposition (16) was constructed for a fixed $a \in \mathbb{R}_+$ but it induces a corresponding decomposition of the space $X$ into $V$ and $W$, where $V$ is the subspace $V = N(C) = \{u \in X; C(a)u(a) = 0, a \in \mathbb{R}_+\}$ and $W$ is defined analogously using $W$. Then, according to the convention, the operators $P_j$, defined by (17), are spectral projections onto $V$ and thus $P_jCu = 0$ for $1 \leq j \leq m$.

The following calculations are similar to those carried out in [8, Section 5]. The only difference is that in the approach of this paper there is no need to write the complementary, kinetic, part $w$ of the solution in the spectral coordinates. Thus, we will skip details of some calculations.

We have, for $1 \leq j \leq m$,

$$P_j\mathcal{M}u = P_j\mathcal{M}\sum_{k=1}^m P_ku + P_j\mathcal{M}Qu = \left( \sum_{k=1}^m \mu_{jk}^* u_k \right) e_j + (x_j \cdot \mathcal{M}w)e_j,$$

where $u_k$ is defined (17) and $\mu_{jk}^* = x_j \cdot \mathcal{M}e_k$. We define $\mathcal{M}^* := \{\mu_{jk}^*\}_{1 \leq j, k \leq m}$. It follows, [8], that $-\mathcal{M}^* = -\{\mu_{ij}^*\}_{1 \leq i, j \leq m}$ is a sub-Kolmogorov matrix.
Similarly, with $B^* := \{\beta^*_j\}_{1 \leq j, k \leq m}$, where $\beta^*_j := x_j \cdot Be_k$, we have

$$\mathcal{P}_jBu = \left( \sum_{k=1}^{m} \beta^*_j u^k \right) e_j + (x_j \cdot Bw)e_j.$$  

Writing the solution as $u = v + w = \sum_{j=1}^{m} w^j e_j + w$, we obtain

$$Su = -\sum_{j=1}^{m} \partial_a w^j e_j - \sum_{j=1}^{m} u^j \partial_a e_j - \partial_a w.$$  

Properties of $e_j$ yield $\sum_{j=1}^{m} w^j \partial_a e_j \in W$ and, since $x_j \cdot w = 0, 1 \leq j \leq m$, we obtain

$$\sum_{i=n_j}^{N_j} \partial_a w_i + \sum_{i=n_{m+1}}^{n} f_j, \partial_a w_i = -\sum_{i=n_{m+1}}^{n} \partial_a f_j, \partial_a w_i,$$

where $w = (w_1, \ldots, w_n)$; that is, $w_i, i = 1, \ldots, n$ are coefficients of the expansion of $w$ in the canonical basis in $\mathbb{R}^n$. Hence

$$\mathcal{P}_j SQu = -\mathcal{P}_j \partial_a w = -(x_j \cdot \partial_a w)e_j = (\partial_a f_j \cdot w)e_j,$$

where $f_j = (0, f_j, n_{m+1}, \ldots, f_j, n)$. Similarly,

$$QSQw = -\partial_a w + \sum_{j=1}^{m} \mathcal{P}_j \partial_a w = -\partial_a w - \sum_{i=n_{m+1}}^{n} w_i \left( \sum_{j=1}^{m} \partial_a f_j, i e_j \right).$$

Now, applying projections $\mathcal{P}_j, 1 \leq j \leq m$, and $Q$ to (1) we obtain the projected system in a compact form as

$$\partial_t v_\epsilon = -\partial_a v_\epsilon - M^* v_\epsilon + \mathcal{F} w_\epsilon,$$

$$\partial_t w_\epsilon = -\partial_a w_\epsilon + G v_\epsilon + \mathcal{H} w_\epsilon + \frac{1}{\epsilon^2} \mathcal{C} w_\epsilon,$$

$$v_\epsilon(0, t) = B^*[v_\epsilon(\cdot, t)] + J[w_\epsilon(\cdot, t)],$$

$$w_\epsilon(0, t) = B[\mathcal{E}_\epsilon v_\epsilon(\cdot, t)] + Bw_\epsilon(\cdot, t) - \mathcal{E}_\epsilon(0)v_\epsilon(0, t),$$

$$v_\epsilon(a, 0) = \hat{v}(a), \quad w_\epsilon(a, 0) = \hat{w}(a), \quad (29)$$

where $\mathcal{E}_\epsilon$ and $v_\epsilon = (v_1^\epsilon, \ldots, v_m^\epsilon)$ were defined in (19), $\mathcal{F}, G, \mathcal{H}, J$ are linear bounded operators on respective subspaces of $X$, related to the projections of the operator $\mathcal{M}$ and derivatives of the spectral projections, whose particular form does not have any bearing on the calculations, $B^* = \int_0^\infty B^*$, $\mathcal{P} \hat{u} = \hat{v} = \mathcal{E}_\epsilon \hat{v}$ and $Q \hat{u} = \hat{w}$. If $\hat{u} = \mathcal{E}_\epsilon \hat{v} + \hat{w} \in D(\mathcal{A}_\epsilon)$, then

$$u_\epsilon(t) = T_{\mathcal{A}_\epsilon}(t) \hat{u} = \mathcal{E}_\epsilon v(t) + w_\epsilon(t) \in D(\mathcal{A}_\epsilon)$$

is a solution to (1)–(3) and, therefore,

$$(v_\epsilon(t), w_\epsilon(t)) = (\mathcal{E}_\epsilon^{-1} \mathcal{P} T_{\mathcal{A}_\epsilon}(t)(\mathcal{E}_\epsilon \hat{v} + \hat{w}), QT_{\mathcal{A}_\epsilon}(t)(\mathcal{E}_\epsilon \hat{v} + \hat{w})) \quad (30)$$

is a solution to (29). Note that the above formula can be used to define projections $(v_\epsilon(t), w_\epsilon(t))$ of $T_{\mathcal{A}_\epsilon}(t) \hat{u}$ with arbitrary $\hat{u} \in X$; that is, for mild solutions to (1)–(3), but such $(v_\epsilon(t), w_\epsilon(t))$ no longer are (classical) solutions to (29).
4.1. Deriving the Aggregated System

We obtain the aggregated system corresponding to (6) by performing an asymptotic expansion with respect to the small parameter \( \epsilon \) of the solution and approximately closing the obtained hierarchy at the zeroth level. Since our aim is to obtain an approximate equation for \( v_\epsilon \) (or \( u_\epsilon \)), the best suited for this is the Chapman–Enskog procedure in the version introduced in [19]. According to this approach, we leave \( v_{1,\epsilon}, \ldots, v_{m,\epsilon} \) unexpanded so that

\[
\mathbf{u}_\epsilon = (v_{1,\epsilon}, \ldots, v_{m,\epsilon}, w_\epsilon) = (v_{1,\epsilon}, \ldots, v_{m,\epsilon}, w_0 + \epsilon w_1 + \cdots),
\]

where \( v_{k,\epsilon} = u_k^0 e_k, k = 1, \ldots, m \). Recall also that \( v_\epsilon = \mathcal{E}_\epsilon v_\epsilon \), see (19). Since the formal procedure is analogous to that in [6,8], we skip the details. Inserting (31) into (29) and retaining only terms of zeroth and first order in \( \epsilon \), we obtain

\[
\begin{align*}
\partial_t v_\epsilon &= -\partial_a v_\epsilon - M^* v_\epsilon + \mathcal{F}(w_0 + \epsilon w_1) + \cdots, \\
\epsilon \tilde{w}_{0,t} &= -\epsilon \tilde{w}_{0,a} + \epsilon \mathcal{G} v_\epsilon + \epsilon \mathcal{H} w_0 + \mathcal{C}_W(w_0 + \epsilon w_1) + \cdots, \\
\gamma v_\epsilon(t) &= B^* v_\epsilon(t) + \mathcal{J}(w_0(t) + \epsilon w_1(t)) + \cdots, \\
\gamma(w_0(t) + \epsilon w_1(t)) &= B[\mathcal{E}_\epsilon v_\epsilon(t)] + B(w_0(t) + \epsilon w_1(t)) - \mathcal{E}_\epsilon(0)\gamma v_\epsilon(t) + \cdots, \\
v_\epsilon(0) &= \tilde{v}, \quad \tilde{w}_0(0) + \epsilon \tilde{w}_1(0) + \cdots = \tilde{w}.
\end{align*}
\]

The invertibility of \( \mathcal{C}_W \) on \( \mathcal{W} \) yields \( \tilde{w}_0 = 0 \) and this allows for the approximate closure of (32) by only retaining the zeroth order terms. The solution of the obtained system will be denoted by \( \tilde{v} \); that is,

\[
\begin{align*}
\partial_t \tilde{v} &= -\partial_a \tilde{v} - M^* \tilde{v}, \\
\gamma \tilde{v}(t) &= B^* \tilde{v}(t), \quad \tilde{v}(0) = \tilde{v}.
\end{align*}
\]

The first order corrector in \( \mathcal{W} \),

\[
\tilde{w}_1 = -\mathcal{C}_W^{-1} \mathcal{G} \tilde{v},
\]

will be useful in the error estimates.

System (33) has the same structure as (1)–(3). Thus, similarly, we define by \( \mathcal{A} \) the realization of the differential expression \( \mathcal{A} = -\text{diag}\{\partial_a\} - M^* \) on the domain \( D(\mathcal{A}) = \{v \in \mathcal{Y}_m; \; v(0) = B^* v\} \) where, recall, \( \mathcal{Y}_m = (W^{1,1}(\mathbb{R}^+))^m \). Then \( \mathcal{A} \) generates a strongly continuous semigroup \( \{\mathcal{T}_\mathcal{A}(t)\}_{t \geq 0} \) such that \( \|\mathcal{T}_\mathcal{A}(t)\| \leq e^{\omega^* t} \), where \( \omega^* \) is a constant depending on \( B^* \) and \( M^* \). Furthermore, if \( \tilde{v} \in D(\mathcal{A}) \), then for any \( T > 0 \) there is \( C_T \) such that

\[
\text{ess sup}_{t \in [0,T]} \|v(t)\|_{\mathcal{Y}_m} \leq C_T \|\tilde{v}\|_{\mathcal{Y}_m}.
\]

This estimate follows since then \( v(t) \) is a classical solution of (33) and hence \( \partial_t v(t) = \mathcal{A} v(t) = \mathcal{T}_\mathcal{A}(t) \mathcal{A} \tilde{v} \). With this

\[
\|\partial_t v(t)\| \leq \|\mathcal{T}_\mathcal{A}(t) \mathcal{A} \tilde{v}\| \leq e^{\omega^* T} \|\mathcal{A} \tilde{v}\| \leq C_T \|\tilde{v}\|_{\mathcal{Y}_m}
\]

for some constant \( C_T' \) (remember that here \( \|\cdot\| = \|\cdot\|_{\mathcal{X}_m} \)). Then (35) follows from (33) since the coefficients of \( M^* \) are bounded.

Now, instead of using the solution \( \tilde{v} \) to (33) with the initial condition \( \tilde{v} = \mathcal{P} \hat{u} \) which, in general, is not differentiable, we use \( \tilde{v}_\delta \) which
solves (33), but with the initial condition \( \tilde{v}_\delta \) approximating \( \tilde{v} \), as constructed in Lemma 3.2. Accordingly, the \( \epsilon \)-order corrector of \( w \) is given as \( \tilde{w}_{1,\delta} = -C_W^{-1}G\tilde{v}_\delta \). We note that in the final estimates we shall set \( \epsilon = \delta \) here; however, we use the parameter \( \delta \) to indicate approximate terms which can be controlled independently of the original small parameter \( \epsilon \).

4.2. Construction of the Layers and the Error Equation

The construction of the initial, boundary and corner layers is carried out as for irreducible \( C \), [6], and thus we just list the final results.

**Initial Layer.** To construct the initial layer we rescale \( t \) according to \( \tau := t/\epsilon \) and denote \( \tilde{u}(a, \tau) = u(a, t/\epsilon) \). Then we write \( \tilde{u} := (\tilde{v} + \cdots, \tilde{w}_0 + \cdots) \) and repeat the asymptotic expansion and truncate it at the zeroth order level, obtaining \( \tilde{v}_0 = 0 \) and

\[
\tilde{w}_0(a, \tau) = e^{\tau C_W(a)} \tilde{w}(a).
\]  

(36)

Since, by (23), all eigenvalues of \( C_W \) have negative real parts, \( \tilde{w}_0 \) exponentially decays to 0 with \( \tau \to \infty \).

**Boundary Layer.** The boundary layer is constructed is a similar way but by rescaling the age variable \( a \) as \( \alpha := a/\epsilon \). Accordingly, we define \( \tilde{u}(a, \tau) := (\tilde{v}(\alpha, \tau), \tilde{w}(\alpha, \tau)) \) and, expanding the components as before, at the \( \epsilon^0 \) order we get

\[
\begin{align*}
\partial_\alpha \tilde{v}_0 &= 0, & \gamma \tilde{v}_0 &= 0, \\
\partial_\alpha \tilde{w}_0,\delta &= C_W(0)\tilde{w}_0,\delta, & \gamma \tilde{w}_0,\delta(t) := B[\mathcal{E}_\nu \tilde{v}_\delta(t)] - \mathcal{E}_\nu(0) \gamma \tilde{v}_\delta(t),
\end{align*}
\]

(37)

where the subscript \( \delta \) indicates that the kinetic part of the boundary layer is constructed for the approximation of \( \tilde{v} \) introduced in Lemma 3.2.

**Corner Layer.** The corner layer is constructed by simultaneous rescaling of \( t \) and \( a \): \( \tau := t/\epsilon \) and \( \alpha := a/\epsilon \). Proceeding as before, we get

\[
\begin{align*}
\partial_\tau \tilde{v}_0 &= -\partial_\alpha \tilde{v}_0, \\
\partial_\tau \tilde{w}_0 &= -\partial_\alpha \tilde{w}_0 + C_W(0)\tilde{w}_0,
\end{align*}
\]

(38)

with the following side conditions:

\[
\begin{align*}
\gamma \tilde{v}_0(\tau) &= \mathcal{J} \tilde{w}_0(\tau), & \gamma \tilde{w}_0(\tau) &= \mathcal{B} \tilde{w}_0(\tau) - \mathcal{E}(0) \gamma \tilde{v}_0(\tau), \\
\tilde{v}_0(a, 0) &= 0, & \tilde{w}_0(a, 0) &= 0.
\end{align*}
\]

The Error Equation. Taking all layers into account, we use here the following approximation:

\[
(v_\epsilon(a, t), w_\epsilon(a, t)) \approx (\tilde{v}_\delta(a, t) + \tilde{v}(a/\epsilon, t/\epsilon),
\]

\[
\epsilon w_{1,\delta}(a, t) + \tilde{w}_0(a, t/\epsilon) + \tilde{w}_{0,\delta}(a/\epsilon, t) + \tilde{w}_0(a/\epsilon, t/\epsilon),
\]

with the error of this approximation given by

\[
\begin{align*}
E(a, t) := (g(a, t), h(a, t)) \\
&= (v_\epsilon(a, t) - \tilde{v}_\delta(a, t) - \tilde{v}_0(a/\epsilon, t/\epsilon), \\
&\quad w_\epsilon(a, t) - \epsilon \tilde{w}_{1,\delta}(a, t) - \tilde{w}_0(a, t/\epsilon) - \tilde{w}_{0,\delta}(a/\epsilon, t) - \tilde{w}_0(a/\epsilon, t/\epsilon)).
\end{align*}
\]
If we assume for a moment that \( \hat{u} \in D(A_\epsilon) \) (which will be justified in the proof of Theorem 4.1) then, by construction, all terms of the error are differentiable (a.e.) and we find that the error satisfies

\[
\partial_t \mathbf{E} = S \mathbf{E} - \mathcal{M} \mathbf{E} + \frac{1}{\epsilon} C \mathbf{E} + \mathbf{F},
\]

\[
\gamma \mathbf{E} = B \mathbf{E} + \mathbf{H}, \quad \mathbf{E}(0) = \mathbf{E}_0,
\]

where

\[
\mathbf{E}_0(a, 0) = \left[ -\phi e^{-a/\delta}, -\hat{w}_{0, \delta}(a/\epsilon, 0) + \epsilon \hat{w}_{1, \delta}(a, 0) \right],
\]

the inhomogeneity \( \mathbf{F} \) in the equation is given by

\[
\mathbf{F} = \left[ \frac{F\hat{w}_0}{\mathcal{H}\hat{w}_0 - \partial_a \hat{w}_0} + \epsilon \left[ -\partial_a \hat{w}_{1, \delta} - \partial_t \hat{w}_{0, \delta} + \mathcal{H} \hat{w}_{1, \delta} \right] + \frac{F\hat{w}_{0, \delta}}{\mathcal{H}\hat{w}_0 - \partial_t \hat{w}_{0, \delta} + \frac{1}{\epsilon} (C_W - C_W(0)) \hat{w}_{0, \delta}} \right]
\]

and the inhomogeneity \( \mathbf{H} \) in the boundary conditions is given by

\[
\mathbf{H} = \left[ \frac{B^* \partial_0 \mathbf{v} + F[\epsilon \hat{w}_{1, \delta} + \hat{w}_{0, \delta} + \hat{w}_0]}{B \mathbf{E} \partial_0 \mathbf{v} - \epsilon \gamma \hat{w}_{1, \delta} + B[\epsilon \hat{w}_{1, \delta} + \hat{w}_{0, \delta} + \hat{w}_0]} \right].
\]

**Theorem 4.1.** Let us assume that \( C, B, \) and \( \mathcal{M} \) satisfy assumptions introduced in Sect. 3 and \( u_t(t) = T_\epsilon(t) \hat{u} = \mathcal{C}_\epsilon \mathbf{v}_\epsilon(t) + \mathbf{w}_\epsilon(t) \) be the (mild) solution to (1)–(3) (where \( (\mathbf{v}_\epsilon, \mathbf{w}_\epsilon) \) are defined by (30)) with \( \hat{u} \in Y_n \). Then, for each \( T < \infty \), there exists a constant \( K(T, M, B, C) \) such that

\[
\| \mathbf{v}_\epsilon(\cdot, t) - \mathbf{v}(t) \|_{X_m} \leq \epsilon K(T, M, B, C) \| \hat{u} \|_{Y_n},
\]

and

\[
\| \mathbf{w}_\epsilon(t) - e^{\frac{1}{\epsilon} C} \hat{w} \|_{X_n} \leq \epsilon K(T, M, B, C) \| \hat{u} \|_{Y_n},
\]

uniformly on \([0, T]\), where \( \mathbf{v} \) is the solution to (33).

**Proof.** Let us fix an arbitrary \( T > 0 \).

First, we note that the assumption \( \hat{u} \in W^1_1(\mathbb{R}_+, \mathbb{R}^n) \) is not sufficient for \( u_t \) to be differentiable so that the error equation (39) cannot be directly used. Thus, we begin by modifying the solution \( u_\epsilon \) to (1)–(3) by the first part of Lemma 3.2 (note that here the estimates of the derivatives of the solution are not needed). Hence, let \( \hat{u}_\eta \in D(A_\epsilon) \) satisfy

\[
\| \hat{u}_\eta - \hat{u} \|_{X_n} \leq \eta C \| \hat{u} \|_{Y_n}, \quad \| \partial_a \hat{u}_\eta - \partial_a \hat{u} \|_{X_n} \leq C \| \hat{u} \|_{Y_n},
\]

for sufficiently small \( \eta > 0 \). If \( \mathbf{u}_\eta \) is the solution to (1)–(3) with the initial condition \( \mathbf{u}_\eta(0) = \hat{u}_\eta \) then, by (24), we have
for \( t \in [0, T] \), where \( C_T \) is a constant independent of \( \epsilon \) and \( \eta \).

Next, let \( \bar{v}^{\eta} \) denote the (mild) solution to (33) with the initial condition \( \bar{v}^{\eta}(0) = v^{\eta} = \mathcal{E}^{-1}\mathcal{P}u^{\eta} \); similarly, \( \bar{w}_0^{\eta}(\tau) = e^{\tau C}\bar{w}^{\eta} \), where \( \bar{w}^{\eta} = Qu^{\eta} \). Since the projections are continuous, we obtain

\[
\|\bar{v}(t) - \bar{v}^{\eta}(t)\|_{X_m} \leq C_T\eta\|\bar{u}\|_{Y_n}, \quad \|\bar{w}_0(\tau) - \bar{w}_0^{\eta}(\tau)\|_{X_m} \leq C\eta\|\bar{u}\|_{Y_n} \tag{47}
\]

on \([0, T]\), for some, possibly different, constants \( C_T, C \); note that the second estimate is independent of \( t \) on account of the negative type of \( e^{\tau C} \) on \( QX_n \).

Since, in general, \( \bar{v}^{\eta} \notin D(A) \), we use again Lemma 3.2 to approximate it by \( \bar{v}_{\delta}^{\eta} \in D(A) \) satisfying

\[
\|\bar{v}_{\delta}^{\eta} - v^{\eta}\|_{X_m} \leq \delta C'\|\bar{v}^{\eta}\|_{Y_m} \leq \delta C\|\bar{u}\|_{Y_n}, \quad \|\partial_a\bar{v}_{\delta}^{\eta} - \partial_a v^{\eta}\|_{X_m} \leq C'\|\bar{v}^{\eta}\|_{Y_m} \leq C\|\bar{u}\|_{Y_n}, \tag{48}
\]

and consequently

\[
\|\bar{v}_{\delta}^{\eta}(t) - \bar{v}^{\eta}(t)\|_{X_m} \leq C_T\delta\|\bar{v}^{\eta}\|_{Y_m} \leq C_T\delta\|\bar{u}\|_{Y_n} \tag{49}
\]

with constants depending neither on \( \delta \) nor \( \eta \) (for sufficiently small \( \eta \) and \( \delta \)).

Now, we can use the error equation (39) for \( u_\xi^{\eta} \) and the (classical) solution \( v_{\delta}^{\eta} \) of (33) with the initial condition \( \bar{v}_{\delta}^{\eta} \) (together with other elements of the asymptotic expansion constructed for \( v_{\delta}^{\eta} \) and \( \bar{w}^{\eta} \)). Thus, each term in (40)–(41) depends on \( \eta \) and those with the subscript \( \delta \) additionally depend on this parameter.

We note that to be completely correct we should also regularize the corner layer since the compatibility conditions are not necessarily satisfied. Since, however, the error equations do not contain derivatives of the corner layer, we skip this step for simplicity.

By linearity, we can split the error into the part \( E_1 \), coming from \( F \) and the initial condition \( E_0 \) (with \( H = 0 \)), and \( E_2 \) which is due to the boundary inhomogeneity \( H \). The error estimates of \( E_2 \) use the representation (27) and thus they are practically identical to those for the reducible matrix \( C \) carried out in [6]. Hence, they are only briefly sketched. On the other hand, the estimates for \( E_1 \) use the differential formulation (39) of the error equation which is simpler to handle than the integrated version used in [6] but requires some additional estimates described below.

Let us recall that the semigroup \( (T_A(t))_{t \geq 0} \) is equibounded in \( \epsilon \), see (24). Setting \( H = 0, E_1 \) can be written using the Duhamel formula and consequently, given \( T < \infty \), for any \( t \in [0, T] \) we have

\[
\|E_1(t)\| \leq e^{\omega T} \left( \|E_0\| + \int_0^t \|F(s)\| \, ds \right). \tag{46}
\]

In what follows, we use a generic constant \( c \) only depending on the coefficients of the problem and \( T \), but not on the initial data or the parameters \( \eta \) or \( \delta \).

To estimate \( \|E_0\| \) we see that for the first row of \( E_0 \) we have
where we used the first equation of (45). For the term $\epsilon \tilde{w}_{1,\delta}(a,0)$ of the second row, we use (34), the assumptions on $M$ and the first equation of (48) to get as above

$$\|\epsilon C_W^{-1}G \tilde{v}_\delta^\eta\|_{X_m} \leq \epsilon c\|\tilde{v}_\delta^\eta\|_{Y_m} \leq \epsilon c\|\tilde{u}\|_{Y_n},$$

For the last term in $E_0$, we observe that $\tilde{v}_\delta^\eta$ is a classical solution; hence it is continuous up to the boundaries and thus $\gamma \tilde{v}_\delta^\eta|_{t=0} = \tilde{v}_\delta^\eta(0)$. Next, due to (23) and the assumptions on $B$, for some $0 < \sigma < -\zeta$, we have

$$\left\| e^{\epsilon C_W(0)} \left( B[\mathcal{E} \tilde{v}_\delta^\eta] - \mathcal{E}(0)\tilde{v}_\delta^\eta(0) \right) \right\|_{X_m} \leq c \|\tilde{v}_\delta^\eta\|_{Y_m} \int_0^\infty e^{-\sigma s} \delta c \eta \|\tilde{u}\|_{Y_n},$$

where we used the fact that $\tilde{v}_\delta^\eta \in Y_m$ and thus the trace operator at $0$ can be estimated by the $Y_m$-norm. We also used second inequality in (48).

Next, let us consider $\tilde{F}$. The term $\partial_a \tilde{w}_{1,\delta}^\eta$ is well defined due to the assumption that $\tilde{u}$ (and thus $\tilde{u}^\eta$) are in $Y_n$ and differentiability of $\mathcal{C}$ with respect to $a$. Thus, $\tilde{F}$ is bounded by $c e^{-\sigma t}$, where $\sigma$ is as above, and

$$\left\| \int_0^t T_{A_\epsilon}(t-s)\tilde{F}(s) \delta c \eta \|\tilde{w}\|_{Y_n} \leq \epsilon c \|\tilde{u}\|_{Y_n}. \tag{50}$$

To get the contribution of $\tilde{F}$ to the error, we begin with the term

$$\partial_t \tilde{w}_{1,\delta}^\eta + \partial_a \tilde{w}_{1,\delta}^\eta = -C_W^{-1}G(\partial_t \tilde{v}_\delta^\eta + \partial_a \tilde{v}_\delta^\eta) - \partial_a[C_W^{-1}G]\tilde{v}_\delta^\eta \tag{51}$$

$$= C_W^{-1}G \mathcal{M}^* - \partial_a[C_W^{-1}G]\tilde{v}_\delta^\eta.$$  

We note that this calculation is only possible since $\tilde{w}_{1,\delta}^\eta$ is differentiable with respect to both $a$ and $t$ separately which is obtained by regularization of the initial conditions. Without it, in [6] we had to resort to the cumbersome integral formulation to achieve essentially the same result. Now, using regularity of $C_W^{-1}$ and $\mathcal{M}$, we get as above

$$\epsilon \left\| \int_0^t T_{A_\epsilon}(t-s)\tilde{F}(s) \delta c \eta \|\tilde{u}\|_{Y_n} \leq \epsilon c \|\tilde{u}\|_{Y_n}. \tag{51}$$

All the other estimates are the same as in [6] (compare the error terms [6, (5.17), (5.18)] with (41) and (42)) with the only difference that here the aggregated part is an $m$-dimensional vector and not a scalar quantity. This, however, does not have any bearing on the estimates. Further, the estimates here naturally involve $\|\tilde{u}\|_{Y_n}$ rather than $\|\tilde{u}\|_{Y_n}$ but this is easily overcome by (45).
Summarizing, following the proof of [6, Theorem 5.1], we can prove (43) and (44) for the regularized initial conditions; that is
\[
\|v_\epsilon^n(\cdot, t) - \bar{v}_\epsilon^n(t)\|_{X_m} \leq \epsilon K(T, M, B, C) \|\bar{u}\|_{Y_n},
\]
and
\[
\|\bar{w}_\epsilon^n(t) - \epsilon^j C \bar{w}_\epsilon^n\|_{X_n} \leq \epsilon K(T, M, B, C) \|\bar{u}\|_{Y_n}.
\]
However, using the estimates (46), (47) and (49), we immediately obtain (43) and (44) for arbitrary $W_1^1$ initial conditions. \hfill \Box

Remark 4.1. Let us summarize and introduce a more compact notation. In the construction of the approximation of $u$, we use the following terms: the bulk approximation $\bar{u} + \epsilon \bar{w}_0 = \sum_{j=1}^{m} \bar{u}^j e_j + \epsilon \bar{w}_0$ (where $\bar{v} = (\bar{u}, \ldots, \bar{u}^m)$), the initial layer $\hat{u} = \bar{w}_0$, the boundary layer $\hat{u} = \bar{w}_0$, and the corner layer $\hat{u} = E \bar{v}_0 + \bar{w}_0$. However, it turns out that the $X_n = (L_1(R_+))^n$ norms of $\epsilon \bar{w}_1, \bar{u}$ and $\hat{u}$ are of order of $\epsilon$ themselves so that, finally,
\[
\|u_\epsilon(t) - \bar{u}(t) - \hat{u}(t/\epsilon)\|_{X_n} = O(\epsilon)
\]
uniformly on finite time intervals, provided $\bar{u} \in (W_1^1(R_+))^n$. We emphasize here that there is no need for $\bar{u} \in D(A_\epsilon)$ or $P \bar{u} \in D(A)$; these regularity requirements were necessary only for interim calculations and are not necessary for validity of (54).

5. A Worked Example

In this section, we will illustrate the asymptotic procedure as applied to the McKendrick models with 5 patches. To make calculations simple, we consider age-independent matrices $M, B$ and $C$ with
\[
C := \begin{bmatrix}
-1 & 1 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 1 \\
0 & 0 & -1 & 1 & 0 \\
0 & 0 & 1 & -1 & 1 \\
0 & 0 & 0 & 0 & -2
\end{bmatrix}.
\]
It is a continuation of Example 3.1 of [8], so the initial section is only briefly sketched. The matrix $C$ is reducible, with $\sigma(C) = \{0, -2\}$. The eigenvalue $\lambda = 0$ is dominant, semisimple with multiplicity 2, while $\lambda = -2$ is of multiplicity 3. Accordingly, the null space of $C$ is also two dimensional.

The nonnegative basis of (normalized) vectors for the null space of $C$, described in Theorem 2.1, is given by $e_1 := (\frac{1}{2}, \frac{1}{2}, 0, 0, 0)$ and $e_2 := (0, 0, \frac{1}{2}, -\frac{1}{2}, 0)$, while in $W$ we can select the basis $e_3 := (1, -1, 0, 0, 0), e_4 := (0, 0, 1, -1, 0)$ and $e_5 := (0, -\frac{1}{2}, 0, -\frac{1}{2}, 1)$. Similarly, the bi-orthogonal basis (14) of the left null space is $x_1 := (1, 1, 0, 0, \frac{1}{2})$ and $x_2 := (0, 0, 1, 1, \frac{1}{2})$. Let $u := (u_1, u_2, u_3, u_4, u_5)$. Accordingly, for $u = v + w \in V \oplus W$ we have $v = u^1 e_1 + u^2 e_2$ with
\[
u^1 := u_1 + u_2 + \frac{1}{2} u_5 \quad \text{and} \quad u^2 := u_3 + u_4 + \frac{1}{2} u_5,
\]
while \( w = w^3e_3 + w^4e_4 + w^5e_5 \), where
\[
w^3 = \frac{1}{2}u_1 - \frac{1}{2}u_2 - \frac{1}{4}u_5, \quad w^4 = \frac{1}{2}u_3 - \frac{1}{2}u_4 - \frac{1}{4}u_5, \quad \text{and} \quad w^5 = u_5.
\] (56)
It is easy to see that this transformation is invertible.

We take general constant mortality and birth matrices, respectively, \( \mathcal{M} := \{\mu_{ij}\}_{1 \leq i, j \leq 5} \) and \( B := \{\beta_{ij}\}_{1 \leq i, j \leq 5} \). Applying projections \( P_j, 1 \leq j \leq 2 \), and \( Q \), we get the hydrodynamics part of (29) as
\[
\partial_t u^1 = -\partial_a u^1 - \mu_{11}^* u^1 + \mu_{12}^* u^2 - \mu_{13}^* w^3 + \mu_{14}^* w^4 + \mu_{15}^* w^5,
\]
\[
\partial_t u^2 = -\partial_a u^2 - \mu_{21}^* u^1 + \mu_{22}^* u^2 - \mu_{23}^* w^3 + \mu_{24}^* w^4 + \mu_{25}^* w^5,
\]
where the explicit values of the coefficients \( \mu_{ij}^* \) are found as follows. First,
\[
\mathcal{P}_1 \mathcal{M}(\mathcal{P}_1 u + \mathcal{P}_2 u) = (x_1 \cdot \mathcal{M}(u^1 e_1 + u^2 e_2)) e_1 = (\mu_{11}^* u^1 + \mu_{12}^* u^2) e_1,
\] (57)
so that
\[
\mu_{11}^* = \frac{1}{2} \left( \mu_{11} + \mu_{12} + \mu_{21} + \mu_{22} + \frac{1}{2} \mu_{51} + \frac{1}{2} \mu_{52} \right),
\]
\[
\mu_{12}^* = \frac{1}{2} \left( \mu_{13} + \mu_{14} + \mu_{23} + \mu_{24} + \frac{1}{2} \mu_{53} + \frac{1}{2} \mu_{54} \right).
\]
Similarly,
\[
\mu_{21}^* = \frac{1}{2} \left( \mu_{31} + \mu_{32} + \mu_{41} + \mu_{42} + \frac{1}{2} \mu_{51} + \frac{1}{2} \mu_{52} \right),
\]
\[
\mu_{22}^* = \frac{1}{2} \left( \mu_{33} + \mu_{34} + \mu_{43} + \mu_{44} + \frac{1}{2} \mu_{53} + \frac{1}{2} \mu_{54} \right).
\]
The other components of \( \{\mu_{i,j}\}_{1 \leq i \leq 2, 1 \leq j \leq 5} \) can be calculated in the same way, but as they do not play any role in the final approximation, we skip them here, referring the reader to [12]. The other essential term in the expansion is the operator \( C_W = QCQ \) which is given by
\[
C_W w = \left(-2w^3 - \frac{1}{2}w^5\right) e_3 + \left(-2w^4 - \frac{1}{2}w^5\right) e_4 - 2w^5 e_5.
\] (58)
Using projections \( P_j, 1 \leq j \leq 2 \), and \( Q \), we can also derive the corresponding expressions of the boundary equations, see [12].

Hence, system (33) takes the following form, where \( \bar{\nu} = (\bar{u}^1, \bar{u}^2) \):
\[
\partial_t \bar{u}^1 = -\partial_a \bar{u}^1 - \mu_{11}^* \bar{u}^1 - \mu_{12}^* \bar{u}^2,
\]
\[
\partial_t \bar{u}^2 = -\partial_a \bar{u}^2 - \mu_{21}^* \bar{u}^1 - \mu_{22}^* \bar{u}^2,
\]
\[
\bar{u}^1(0, t) = \int_0^\infty (\beta_{11}^* \bar{u}^1(a, t) + \beta_{12}^* \bar{u}^2(a, t)) \, da,
\]
\[
\bar{u}^2(0, t) = \int_0^\infty (\beta_{21}^* \bar{u}^1(a, t) + \beta_{22}^* \bar{u}^2(a, t)) \, da,
\]
\[ \bar{u}^1(a,0) = \circ u_1(a) + \circ u_2(a) + \frac{1}{2} \circ u_5(a), \]
\[ \bar{u}^2(a,0) = \circ u_3(a) + \circ u_4(a) + \frac{1}{2} \circ u_5(a), \]

where \( \{\beta^j_{ij}\}_{1 \leq i \leq 2, 1 \leq j \leq 2} \) is given by formulæ analogous to \( \{\mu^j_{ij}\}_{1 \leq i \leq 2, 1 \leq j \leq 2} \). According to Theorem 4.1, to write down the approximation we need the initial layer; that is, the solution to \( \partial_t \tilde{w}_0 = C_W \tilde{w}_0 \). Using the coordinate form (58), we find the solution as

\[ \tilde{w}^3_0 = e^{-2\tau} \left( \frac{\tau^2}{8} \tilde{w}^5 - \frac{\tau}{2} \tilde{w}^4 + \tilde{w}^3 \right), \]
\[ \tilde{w}^4_0 = e^{-2\tau} \left( -\frac{\tau}{2} \tilde{w}^5 + \tilde{w}^4 \right), \]
\[ \tilde{w}^5_0 = e^{-2\tau} \tilde{w}^5, \]

where, by (56),

\[ \tilde{w}^3 = \frac{1}{2} u_1 - \frac{1}{2} u_2 - \frac{1}{4} u_5, \quad \tilde{w}^4 = \frac{1}{2} u_3 - \frac{1}{2} u_4 - \frac{1}{4} u_5 \quad \text{and} \quad \tilde{w}^5 = u_5. \]

Thus, as in (54), we can write the combined estimates (43) and (44) as

\[ \mathbf{u}(t) = \mathbf{u}(t) + \mathbf{u}(\tau) + O(\epsilon) = \bar{u}^1(t)e_1 + \bar{u}^2(t)e_2 + e^{-2\tau} \left( \frac{\tau^2}{8} \bar{w}^5 - \frac{\tau}{2} \bar{w}^4 + \bar{w}^3 \right)e_3 + e^{-2\tau} \left( -\frac{\tau}{2} \bar{w}^5 + \bar{w}^4 \right)e_4 + e^{-2\tau} \bar{w}^5 e_5 + O(\epsilon) \]

in \( \mathbf{X}_5 = (L_1(\mathbb{R}^+))^5 \) norm, uniformly in \( t \) on \([0,T]\). Using the expression for \( e_j, 1 \leq j \leq 5 \), we can write the approximation in the components as

\[ \left\| u_1(t) - \frac{1}{2} \bar{u}^1(t) - e^{-2\tau} \left( \frac{\tau^2}{8} \tilde{w}^5 - \frac{\tau}{2} \tilde{w}^4 + \tilde{w}^3 \right) \right\|_{L_1(\mathbb{R}^+)} = O(\epsilon), \]
\[ \left\| u_2(t) - \frac{1}{2} \bar{u}^1(t) + e^{-2\tau} \left( \frac{\tau^2}{8} \bar{w}^5 - \frac{\tau}{2} \bar{w}^4 + \bar{w}^3 \right) + \frac{1}{2} e^{-2\tau} \tilde{w}^5 \right\|_{L_1(\mathbb{R}^+)} = O(\epsilon), \]
\[ \left\| u_3(t) - \frac{1}{2} \bar{u}^2(t) - e^{-2\tau} \left( -\frac{\tau}{2} \bar{w}^5 + \bar{w}^4 \right) \right\|_{L_1(\mathbb{R}^+)} = O(\epsilon), \]
\[ \left\| u_4(t) - \frac{1}{2} \bar{u}^2(t) + e^{-2\tau} \left( -\frac{\tau}{2} \bar{w}^5 + \bar{w}^4 \right) + \frac{1}{2} e^{-2\tau} \tilde{w}^5 \right\|_{L_1(\mathbb{R}^+)} = O(\epsilon), \]
\[ \left\| u_5(t) - e^{-2\tau} \tilde{w}^5 \right\|_{L_1(\mathbb{R}^+)} = O(\epsilon), \]

where \( \bar{u}^i, i = 1, 2, \) solve (59) and \( \bar{w}^j, j = 3, 4, 5, \) are defined by (60).

### 6. Numerical Results

For numerical simulations it is convenient to write system (1)–(3) in the integral equation form. As here it will not cause any misunderstanding, we shall drop the subscript \( \epsilon \) from the notation. Using the fundamental solution matrix \( L(a_0,a) \), as in (26), we find that, similarly to (27), the solution of (1)–(3) is given by
where the boundary condition \( u(t, 0) \) satisfies the Volterra integral equation

\[
u(t, 0) = \int_0^t B(a) L(0, a) u(t - a, 0) \, da + \int_0^\infty B(a + t) L(a, a + t) \hat{u}(a) \, da. \tag{62}\]

The numerical approximation to (61)–(62) is obtained in two steps: first, we solve Volterra integral equation (62) for \( u(t, 0) \) and then we recover \( u(t, a) \) by integrating linear ODEs along the characteristic lines using (61).

To solve (62) on \([0, T]\), we set \( F(t) = \int_0^\infty B(a + t) L(a, a + t) \hat{u}(a) \, da \), introduce a computational grid \( \{t_k\}_{k=1}^M \) and apply \( A(\alpha) \)-stable, 4-step, order 4 BDF formula, [13], to

\[
y(s, t) = F(t) + \int_0^s B(t - a) L(0, t - a) u(a, 0) \, da, \quad u(t, 0) = y(t, t). \]

This yields the algorithm

\[
\sum_{j=0}^4 a_j y(\cdot, t_{k-j}) = \tau_{k-1} B(\cdot - t_k) v(\cdot - t_k, t_k), \quad u(t_k, 0) = y(t_k, t_k),
\]

\[
\partial_s v(s, t_k) = -M(s) v(s, t_k) + \frac{1}{\epsilon} C(s) v(s, t_k), \quad v(0, t_k) = u(t_k, 0), \tag{63}\]

where \( \tau_k = t_k - t_{k-1} \) and \( a_j, 0 \leq j \leq 4 \) are coefficients of the BDF formula. The starting values \( y(\cdot, t_k), 0 \leq k \leq 3 \), are obtained with the 2-stage Radau IIA method, see [14]. Note that the algorithm requires one evaluation of \( F(t) \) per integration step. In our simulations this is done by means of the three-points, composite Gauss–Lobatto quadrature rule.

**Table 1.** The largest pointwise numerical error and the maximal stepsize vs. \( \epsilon \)

|        | \(10^{-1}\)     | \(10^{-2}\)     | \(10^{-3}\)     |
|--------|----------------|----------------|----------------|
| \(\tau\) | \(1.735 \times 10^{-2}\) | \(5.239 \times 10^{-3}\) | \(3.157 \times 10^{-3}\) |
| Error  | \(5.651 \times 10^{-5}\) | \(2.422 \times 10^{-4}\) | \(5.377 \times 10^{-4}\) |

**Table 2.** The largest \( L_1(\mathbb{R}^+, \mathbb{R}^n) \) errors vs. \( \epsilon \)

| \(\epsilon\) | \(10^{-1}\)     | \(10^{-2}\)     | \(10^{-3}\)     |
|--------------|----------------|----------------|----------------|
| \(\hat{E}\)  | \(6.321 \times 10^{-1}\) | \(6.321 \times 10^{-1}\) | \(6.321 \times 10^{-1}\) |
| \(\hat{E}\)  | \(3.482 \times 10^{-2}\) | \(5.677 \times 10^{-3}\) | \(2.434 \times 10^{-3}\) |
| \(\hat{E}\)  | \(3.601 \times 10^{-2}\) | \(5.016 \times 10^{-3}\) | \(2.230 \times 10^{-3}\) |
| \(\hat{E}\)  | \(1.926 \times 10^{-2}\) | \(5.016 \times 10^{-3}\) | \(2.230 \times 10^{-3}\) |
Figure 1. Components of the exact solution vs bulk approximation, $\varepsilon = 10^{-3}$

It can be shown that the above algorithm converges with order four to $u(t, 0)$ for sufficiently smooth $M, C$ and $B$ (i.e. the global error is $O(\max_k \tau_k^4)$) in any finite interval $[0, T]$; moreover, the convergence is uniform for all $\varepsilon > 0$. 
Figure 2. The bulk approximation error (left) and the effect of initial layer correction (right), $\varepsilon = 10^{-3}$

The described method is rather general and does not utilize the linear convolution structure of the Volterra integral operator. The side effect is that its computational cost grows quadratically with the number of grid points. That is a serious drawback. One obvious way to overcome the difficulty is to
use a spectral method combined with the Fourier–Laplace integral transform. However, the analysis of the resulting algorithm may require extra restrictions on the coefficients of the problem (1)–(3).

To illustrate the error analysis, we apply the numerical methods described above to (1)–(3) with the matrix $C$ given by (55) and specify
\( M = I \) and \( B = \text{diag}(1, 2, 1, 2, 1) \). Further, the initial conditions are taken to be \( \hat{u}(a) = (e^{-a}, e^{-2a}, e^{-a}, e^{-2a}, e^{-a}) \). The exact solution is given by (62), where \( L(0, a) = \{ l_{ij}(a) \}_{1 \leq i, j \leq 5} \) with

\[
\begin{align*}
l_{11}(a) &= l_{22}(a) = l_{33}(a) = l_{44}(a) = \frac{1}{2} e^{-a} (1 + e^{-2a}/\varepsilon), \\
l_{12}(a) &= l_{21}(a) = l_{34}(a) = l_{43}(a) = \frac{1}{2} e^{-a} (1 - e^{-2a}/\varepsilon), \\
l_{15}(a) &= l_{35}(a) = \frac{1}{4} e^{-a} (1 - e^{-2a}/\varepsilon (1 + 2a/\varepsilon)), \\
l_{25}(a) &= l_{45}(a) = \frac{1}{4} e^{-a} (1 - e^{-2a}/\varepsilon (1 - 2a/\varepsilon)), \\
l_{55}(a) &= e^{-(1+2/\varepsilon)a}
\end{align*}
\]

and \( l_{ij} = 0 \) otherwise. Hence, \( u(t, 0) = \frac{1}{\varepsilon \lambda_1 - \lambda_2} (u_j(t))_{1 \leq j \leq 5} \), where

\[
\begin{align*}
u_1(t) &= u_3(t) = (\lambda_1 - \lambda_2)e^{-2t/\varepsilon} - (\lambda_1 + \varepsilon (1 + \lambda_2)) e^{-\lambda_1 t/(2\varepsilon)} + (\lambda_2 + \varepsilon (1 + \lambda_1)) e^{-\lambda_2 t/(2\varepsilon)}, \\
u_2(t) &= u_4(t) = -2(\lambda_1 - \lambda_2)e^{-2t/\varepsilon} + (2\lambda_1 + \varepsilon (\lambda_1 - 6)) e^{-\lambda_1 t/(2\varepsilon)} - (2\lambda_2 + \varepsilon (\lambda_2 - 6)) e^{-\lambda_2 t/(2\varepsilon)}, \\
n_5(t) &= \varepsilon (\lambda_1 - \lambda_2)e^{-2t/\varepsilon}, \quad \lambda_{1,2} = 2 - \varepsilon \pm \sqrt{\varepsilon^2 + 4}.
\end{align*}
\]

Note that \( \hat{u}(0) = \int_0^\infty B \hat{u}(a) \, da \); hence the exact solution is continuous across the main characteristic line \( t = a \).

First, we solve the complete model (1)–(3) with coefficients described above, using numerical scheme (63). The results of calculations for \( \varepsilon = 10^{-1}, 10^{-2}, 10^{-3} \) are shown in Tables 1 and 2. The errors in Table 2 are given by

\[
\begin{align*}
\mathbf{E}_1 &= \max_t \| u - \hat{u} \|_{L_1}, \quad \mathbf{E}_2 = \max_t \| u - \hat{u} - \tilde{u} \|_{L_1}, \\
\mathbf{E}_3 &= \max_t \| u - \hat{u} - \tilde{u} \|_{L_1}, \quad \tilde{E}_4 = \max_t \| u - \hat{u} - \tilde{u} - \hat{u} \|_{L_1},
\end{align*}
\]

where the terms above were defined in Remark 4.1. As expected, the convergence rate improves with \( \varepsilon \). Note also that, in accordance with Theorem 4.1, we see a significant decrease in the error after inclusion of the initial layer correction, while the boundary and corner layer corrections only resulted in minor improvements. This seems to be inconsistent with Figs. 1–3, where the components of the numerical solution and various approximations for \( \varepsilon = 10^{-3} \) are plotted and which show a significant improvements of the approximation resulting from adding the boundary and corner layer corrections. This is, however, an apparent inconsistency. Indeed, the figures present pointwise errors and we can see that the errors on Fig. 2, which are removed by considering the boundary and corner layers in Fig. 3, are concentrated close to the boundary \( a = 0 \) and the corner \( a = 0, t = 0 \). Thus, they become insignificant after the integration with respect to \( a \), as seen in Table 2.

**Open Access.** This article is distributed under the terms of the Creative Commons Attribution License which permits any use, distribution, and reproduction in any medium, provided the original author(s) and the source are credited.
References

[1] Arino, O., Sánchez, E., Bravo de la Parra, R., Auger, P.: A singular perturbation in an age-structured population model. SIAM J. Appl. Math. 60(2), 408–436 (1999)

[2] Bailey, N.T.J.: The elements of stochastic processes. Wiley, New York (1964)

[3] Banasiak, J.: Asymptotic analysis of singularly perturbed dynamical systems. In: Abdulle, A., Banasiak, J., Damlamian, A. and Sango, M. (eds.), Multiscale Problems in Biomathematics, Physics and Mechanics: Modelling, Analysis and Numerics, GAKUTO Internat, pp. 221–255. Ser. Math. Sci. Appl. 31, Gakkotosho, Tokyo (2009)

[4] Banasiak, J., Bobrowski, A.: Interplay between degenerate convergence of semigroups and asymptotic analysis. J. Evol. Equ. 9(2), 293–314 (2009)

[5] Banasiak, J., Shindin, S.: Chapman–Enskog asymptotic procedure in structured population dynamics. Il Nuovo Cimento C 33(1), 31–38 (2010)

[6] Banasiak, J., Goswami, A., Shindin, S.: Aggregation in age and space structured population models: an asymptotic analysis approach. J. Evol. Equ. 11, 121–154 (2011)

[7] Banasiak, J., Namayanja, P.: Relative entropy and discrete Poincaré inequalities for reducible matrices. Appl. Math. Lett. 25(12), 2193–2197 (2012)

[8] Banasiak, J., Goswami, A.: Singularly perturbed models with reducible migration matrix. 1. Sova–Kurtz theorem and convergence to the aggregated model, submitted

[9] Bobrowski, A.: Convergence of one-parameter operator semigroups. In models of mathematical biology and elsewhere. Cambridge University Press, Cambridge, in preparation

[10] Caswell, H.: Matrix population models. Construction, analysis and interpretation. Sinauer Associates, Inc., Sunderland (1989)

[11] Gantmacher, F.R.: Applications of the theory of matrices. Interscience Publishers, New York (1959)

[12] Goswami, A.: Asymptotic analysis of singularly perturbed dynamical systems. PhD Thesis, University of KwaZulu-Natal, Durban (2011)

[13] Hairer, E., Norsett, S.P., Wanner, G.: Solving ordinary differential equations I. nonstiff problems. Springer-Verlag, Berlin (2008)

[14] Hairer, E., Wanner, G.: Solving ordinary differential equations. II. Stiff and differential-algebraic problems. Springer-Verlag, Berlin (2010)

[15] Inaba, H.: A semigroup approach to the strong ergodic theorem of the multi state stable population process. Math. Popul. Stud. 1(1), 49–77 (1988)

[16] Lisi, M., Totaro, S.: The Chapman–Enskog procedure for an age-structured population model: initial, boundary and corner layer corrections. Math. Biosci. 196(2), 153–186 (2005)

[17] Luenberger, D.G.: Introduction to dynamic systems. Theory, models and applications. Wiley, New York (1979)

[18] Meyer, C.D.: Matrix analysis and applied linear algebra. SIAM, Philadelphia (2000)

[19] Mika, J.R., Banasiak, J.: Singularly perturbed evolution equations with applications in kinetic theory. World Sci., Singapore (1995)
[20] Seneta, E.: Nonnegative matrices and Markov chains, 2nd edn. Springer-Verlag, New York (1981)
[21] Webb, G.F.: Theory of non-linear age dependent population dynamics. Marcel Dekker, New York (1985)

Jacek Banasiak and Sergey Shindin
School of Mathematics, Statistics and Computer Science
University of KwaZulu-Natal
Durban, South Africa
e-mail: banasiak@ukzn.ac.za

Sergey Shindin
e-mail: shindins@ukzn.ac.za

Jacek Banasiak
Institute of Mathematics
Technical University of Łódź
Łódź, Poland

Amartya Goswami
Department of Mathematical Sciences
University of Zululand
Richards Bay, South Africa
e-mail: goswami.amartya8@gmail.com

Received: January 26, 2013.
Accepted: June 4, 2013.