Probab. Theory Relat. Fields (2012) 153:727–757
DOI 10.1007/s00440-011-0359-2

A law of large numbers approximation for Markov population processes with countably many types

A. D. Barbour · M. J. Luczak

Received: 2 September 2010 / Revised: 24 March 2011 / Published online: 7 April 2011

Abstract When modelling metapopulation dynamics, the influence of a single patch on the metapopulation depends on the number of individuals in the patch. Since the population size has no natural upper limit, this leads to systems in which there are countably infinitely many possible types of individual. Analogous considerations apply in the transmission of parasitic diseases. In this paper, we prove a law of large numbers for quite general systems of this kind, together with a rather sharp bound on the rate of convergence in an appropriately chosen weighted ℓ_1 norm.

Keywords Epidemic models · Metapopulation processes · Countably many types · Quantitative law of large numbers · Markov population processes

Mathematics Subject Classification (2000) 92D30 · 60J27 · 60B12

1 Introduction

There are many biological systems that consist of entities that differ in their influence according to the number of active elements associated with them, and can be divided into types accordingly. In parasitic diseases [2,7,11–13], the infectivity of a host

A.D. Barbour was supported in part by Schweizerischer Nationalfonds Projekt Nr. 20–107935/1.
M.J. Luczak was supported in part by a STICERD grant.

A. D. Barbour (✉)
Angewandte Mathematik, Universität Zürich,
Winterthurerstrasse 190, 8057 Zürich, Switzerland
e-mail: a.d.barbour@math.uzh.ch

M. J. Luczak
Department of Mathematics, London School of Economics,
Houghton St., London WC2A 2AE, UK
e-mail: m.j.luczak@lse.ac.uk

depends on the number of parasites that it carries; in metapopulations, the migration pressure exerted by a patch is related to the number of its inhabitants [1]; the behaviour of a cell may depend on the number of copies of a particular gene that it contains ([6], Chapter 7); and so on. In none of these examples is there a natural upper limit to the number of associated elements, so that the natural setting for a mathematical model is one in which there are countably infinitely many possible types of individual. In addition, transition rates typically increase with the number of associated elements in the system—for instance, each parasite has an individual death rate, so that the overall death rate of parasites grows at least as fast as the number of parasites—and this leads to processes with unbounded transition rates. This paper is concerned with approximations to density-dependent Markov models of this kind, when the typical population size N becomes large.

In density-dependent Markov population processes with only finitely many types of individual, a law of large numbers approximation, in the form of a system of ordinary differential equations, was established by Kurtz [8], together with a diffusion approximation [9]. In the infinite dimensional case, the law of large numbers was proved for some specific models [1,2,13], see also [10], using individually tailored methods. A more general result was then given by Eibeck and Wagner [5]. In Barbour and Luczak [3], the law of large numbers was strengthened by the addition of an error bound in ℓ_1 that is close to optimal order in N. Their argument makes use of an intermediate approximation involving an independent particles process, for which the law of large numbers is relatively easy to analyse. This process is then shown to be sufficiently close to the interacting process of actual interest, by means of a coupling argument. However, the generality of the results obtained is limited by the simple structure of the intermediate process, and the model of Arrigoni [1], for instance, lies outside their scope.

In this paper, we develop an entirely different approach, which circumvents the need for an intermediate approximation, enabling a much wider class of models to be addressed. The setting is that of families of Markov population processes $X_N := (X_N(t), t \geq 0)$, $N \geq 1$, taking values in the countable space $\mathcal{X}_+ := \{X \in \mathbb{Z}_+^{\mathbb{Z}_+};\ \sum_{m\geq 0} X^m < \infty\}$. Each component represents the number of individuals of a particular type, and there are countably many types possible; however, at any given time, there are only finitely many individuals in the system. The process evolves as a Markov process with state-dependent transitions

$$X \to X + J \quad \text{at rate} \quad N\alpha_J(N^{-1}X), \quad X \in \mathcal{X}_+,\ J \in \mathcal{J}, \tag{1.1}$$

where each jump is of bounded influence, in the sense that

$$\mathcal{J} \subset \{X \in \mathbb{Z}^{\mathbb{Z}_+};\ \sum_{m\geq 0} |X^m| \leq J_* < \infty\}, \quad \text{for some fixed} \quad J_* < \infty, \tag{1.2}$$

so that the number of individuals affected is uniformly bounded. Density dependence is reflected in the fact that the arguments of the functions α_J are counts normalised by the 'typical size' N. Writing $\mathcal{R} := R_+^{\mathbb{Z}_+}$, the functions $\alpha_J\colon\ \mathcal{R} \to R_+$ are assumed to satisfy

$$\sum_{J \in \mathcal{J}} \alpha_J(\xi) < \infty, \quad \xi \in \mathcal{R}_0, \tag{1.3}$$

where $\mathcal{R}_0 := \{\xi \in \mathcal{R} \colon \xi_i = 0$ for all but finitely many $i\}$; this assumption implies that the processes X_N are pure jump processes, at least for some non-zero length of time. To prevent the paths leaving $\mathcal{X}_+$, we also assume that $J_l \geq -1$ for each l, and that $\alpha_J(\xi) = 0$ if $\xi^l = 0$ for any $J \in \mathcal{J}$ such that $J^l = -1$. Some remarks on the consequences of allowing transitions J with $J^l \leq -2$ for some l are made at the end of Sect. 4.

The law of large numbers is then formally expressed in terms of the system of *deterministic equations*

$$\frac{d\xi}{dt} = \sum_{J \in \mathcal{J}} J\alpha_J(\xi) =: F_0(\xi), \tag{1.4}$$

to be understood componentwise for those $\xi \in \mathcal{R}$ such that

$$\sum_{J \in \mathcal{J}} |J^l|\alpha_J(\xi) < \infty, \quad \text{for all } l \geq 0,$$

thus by assumption including $\mathcal{R}_0$. Here, the quantity F_0 represents the infinitesimal average drift of the components of the random process. However, in this generality, it is not even immediately clear that equations (1.4) have a solution.

In order to make progress, it is assumed that the unbounded components in the transition rates can be assimilated into a linear part, in the sense that F_0 can be written in the form

$$F_0(\xi) = A\xi + F(\xi), \tag{1.5}$$

again to be understood componentwise, where A is a constant $\mathbb{Z}_+ \times \mathbb{Z}_+$ matrix. These equations are then treated as a perturbed linear system ([14], Chapter 6). Under suitable assumptions on A, there exists a measure μ on $\mathbb{Z}_+$, defining a weighted ℓ_1 norm $\|\cdot\|_\mu$ on $\mathcal{R}$, and a strongly $\|\cdot\|_\mu$–continuous semigroup $\{R(t), t \geq 0\}$ of transition matrices having pointwise derivative $R'(0) = A$. If F is locally $\|\cdot\|_\mu$–Lipschitz and $\|x(0)\|_\mu < \infty$, this suggests using the solution x of the integral equation

$$x(t) = R(t)x(0) + \int_0^t R(t-s)F(x(s))\,ds \tag{1.6}$$

as an approximation to $x_N := N^{-1}X_N$, instead of solving the deterministic equations (1.4) directly. We go on to show that the solution X_N of the stochastic system can be expressed using a formula similar to (1.6), which has an additional stochastic component in the perturbation:

$$x_N(t) = R(t)x_N(0) + \int_0^t R(t-s)F(x_N(s))\,ds + \widetilde{m}_N(t), \tag{1.7}$$

where

$$\widetilde{m}_N(t) := \int_0^t R(t-s)\,dm_N(s), \tag{1.8}$$

and m_N is the local martingale given by

$$m_N(t) := x_N(t) - x_N(0) - \int_0^t F_0(x_N(s))\,ds. \tag{1.9}$$

The quantity m_N can be expected to be small, at least componentwise, under reasonable conditions.

To obtain tight control over $\widetilde{m}_N$ in all components simultaneously, sufficient to ensure that $\sup_{0\le s\le t} \|\widetilde{m}_N(s)\|_\mu$ is small, we derive Chernoff-like bounds on the deviations of the most significant components, with the help of a family of exponential martingales. The remaining components are treated using some general *a priori* bounds on the behaviour of the stochastic system. This allows us to take the difference between the stochastic and deterministic equations (1.7) and (1.6), after which a Gronwall argument can be carried through, leading to the desired approximation.

The main result, Theorem 4.7, guarantees an approximation error of order $O(N^{-1/2}\sqrt{\log N})$ in the weighted ℓ_1 metric $\|\cdot\|_\mu$, except on an event of probability of order $O(N^{-1}\log N)$. More precisely, for each $T > 0$, there exist constants $K_T^{(1)}, K_T^{(2)}, K_T^{(3)}$ such that, for N large enough, if

$$\|N^{-1}X_N(0) - x(0)\|_\mu \le K_T^{(1)}\sqrt{\frac{\log N}{N}},$$

then

$$\mathbf{P}\left(\sup_{0\le t\le T} \|N^{-1}X_N(t) - x(t)\|_\mu > K_T^{(2)}\sqrt{\frac{\log N}{N}}\right) \le K_T^{(3)}\frac{\log N}{N}. \tag{1.10}$$

The error bound is sharper, by a factor of $\log N$, than that given in Barbour and Luczak [3], and the theorem is applicable to a much wider class of models. However, the method of proof involves moment arguments, which require somewhat stronger assumptions on the initial state of the system, and, in models such as that of Barbour and Kafetzaki [2], on the choice of infection distributions allowed. The conditions under which the theorem holds can be divided into three categories: growth conditions on the transition rates, so that the *a priori* bounds, which have the character of moment bounds, can be established; conditions on the matrix A, sufficient to limit the growth of the semigroup R, and (together with the properties of F) to determine the weights defining the metric in which the approximation is to be carried out; and conditions on the initial state of the system. The *a priori* bounds are derived in Sect. 2, the semigroup analysis is conducted in Sect. 3, and the approximation proper is carried out in Sect. 4. The paper concludes in Sect. 5 with some examples.

The form (1.8) of the stochastic component $\widetilde{m}_N(t)$ in (1.7) is very similar to that of a key element in the analysis of stochastic partial differential equations (see, e.g., [4, Sect. 6.6]). The SPDE arguments used for its control are, however, typically conducted in a Hilbert space context. Our setting is quite different in nature, and it does not seem clear how to translate the SPDE methods into our context.

2 A priori bounds

We begin by imposing further conditions on the transition rates of the process X_N, sufficient to constrain its paths to bounded subsets of $\mathcal{X}_+$ during finite time intervals, and in particular to ensure that only finitely many jumps can occur in finite time. The conditions that follow have the flavour of moment conditions on the jump distributions. Since the index $j \in \mathbb{Z}_+$ is symbolic in nature, we start by fixing an $\nu \in \mathcal{R}$, such that $\nu(j)$ reflects in some sense the 'size' of j, with most indices being 'large':

$$\nu(j) \geq 1 \quad \text{for all } j \geq 0 \quad \text{and} \quad \lim_{j\to\infty} \nu(j) = \infty. \tag{2.1}$$

We then define the analogues of higher empirical moments using the quantities $\nu_r \in \mathcal{R}$, defined by $\nu_r(j) := \nu(j)^r$, $r \geq 0$, setting

$$S_r(x) := \sum_{j\geq 0} \nu_r(j) x^j = x^T \nu_r, \quad x \in \mathcal{R}_0, \tag{2.2}$$

where, for $x \in \mathcal{R}_0$ and $y \in \mathcal{R}$, $x^T y := \sum_{l\geq 0} x_l y_l$. In particular, for $X \in \mathcal{X}_+$, $S_0(X) = \|X\|_1$. Note that, because of (2.1), for any $r \geq 1$,

$$\#\{X \in \mathcal{X}_+ \colon\ S_r(X) \leq K\} < \infty \quad \text{for all} \quad K > 0. \tag{2.3}$$

To formulate the conditions that limit the growth of the empirical moments of $X_N(t)$ with t, we also define

$$U_r(x) := \sum_{J\in\mathcal{J}} \alpha_J(x) J^T \nu_r; \quad V_r(x) := \sum_{J\in\mathcal{J}} \alpha_J(x) (J^T \nu_r)^2, \quad x \in \mathcal{R}. \tag{2.4}$$

The assumptions that we shall need are then as follows.

Assumption 2.1 There exists a ν satisfying (2.1) and $r_{\max}^{(1)}, r_{\max}^{(2)} \geq 1$ such that, for all $X \in \mathcal{X}_+$,

$$\sum_{J \in \mathcal{J}} \alpha_J(N^{-1}X)|J^T \nu_r| < \infty, \quad 0 \leq r \leq r_{\max}^{(1)}, \tag{2.5}$$

the case $r = 0$ following from (1.2) and (1.3); furthermore, for some non-negative constants k_{rl}, the inequalities

$$\begin{aligned} U_0(x) &\leq k_{01} S_0(x) + k_{04}, \\ U_1(x) &\leq k_{11} S_1(x) + k_{14}, \\ U_r(x) &\leq \{k_{r1} + k_{r2} S_0(x)\} S_r(x) + k_{r4}, \quad 2 \leq r \leq r_{\max}^{(1)}; \end{aligned} \tag{2.6}$$

and

$$\begin{aligned} V_0(x) &\leq k_{03} S_1(x) + k_{05}, \\ V_r(x) &\leq k_{r3} S_{p(r)}(x) + k_{r5}, \qquad 1 \leq r \leq r_{\max}^{(2)}, \end{aligned} \tag{2.7}$$

are satisfied, where $1 \leq p(r) \leq r_{\max}^{(1)}$ for $1 \leq r \leq r_{\max}^{(2)}$.

The quantities $r_{\max}^{(1)}$ and $r_{\max}^{(2)}$ usually need to be reasonably large, if Assumption 4.2 below is to be satisfied.

Now, for X_N as in Sect. 1, we let $t_n^{X_N}$ denote the time of its nth jump, with $t_0^{X_N} = 0$, and set $t_\infty^{X_N} := \lim_{n\to\infty} t_n^{X_N}$, possibly infinite. For $0 \leq t < t_\infty^{X_N}$, we define

$$S_r^{(N)}(t) := S_r(X_N(t)); \quad U_r^{(N)}(t) := U_r(x_N(t)); \quad V_r^{(N)}(t) := V_r(x_N(t)), \tag{2.8}$$

once again with $x_N(t) := N^{-1} X_N(t)$, and also

$$\tau_r^{(N)}(C) := \inf\left\{t < t_\infty^{X_N} : \; S_r^{(N)}(t) \geq NC\right\}, \quad r \geq 0, \tag{2.9}$$

where the infimum of the empty set is taken to be ∞. Our first result shows that $t_\infty^{X_N} = \infty$ a.s., and limits the expectations of $S_0^{(N)}(t)$ and $S_1^{(N)}(t)$ for any fixed t.

In what follows, we shall write $\mathcal{F}_s^{(N)} = \sigma(X_N(u), 0 \leq u \leq s)$, so that $(\mathcal{F}_s^{(N)} : s \geq 0)$ is the natural filtration of the process X_N.

Lemma 2.2 *Under Assumptions 2.1, $t_\infty^{X_N} = \infty$ a.s. Furthermore, for any $t \geq 0$,*

$$\begin{aligned} \mathbf{E}\left\{S_0^{(N)}(t)\right\} &\leq \left(S_0^{(N)}(0) + N k_{04} t\right) e^{k_{01} t}; \\ \mathbf{E}\left\{S_1^{(N)}(t)\right\} &\leq \left(S_1^{(N)}(0) + N k_{14} t\right) e^{k_{11} t}. \end{aligned}$$

Proof Introducing the formal generator $\mathbb{A}_N$ associated with (1.1),

$$\mathbb{A}_N f(X) := N \sum_{J \in \mathcal{J}} \alpha_J(N^{-1}X)\{f(X+J) - f(X)\}, \quad X \in \mathcal{X}_+, \tag{2.10}$$

we note that $NU_l(x) = \mathbb{A}_N S_l(Nx)$. Hence, if we define $M_l^{(N)}$ by

$$M_l^{(N)}(t) := S_l^{(N)}(t) - S_l^{(N)}(0) - N \int_0^t U_l^{(N)}(u)\,du, \quad t \geq 0, \tag{2.11}$$

for $0 \leq l \leq r_{\max}^{(1)}$, it is immediate from (2.3), (2.5) and (2.6) that the process $(M_l^{(N)}(t \wedge \tau_1^{(N)}(C)), t \geq 0)$ is a zero mean $\mathcal{F}^{(N)}$-martingale for each $C > 0$.

In particular, considering $M_1^{(N)}(t \wedge \tau_1^{(N)}(C))$, it follows in view of (2.6) that

$$\begin{aligned}\mathbf{E}\{S_1^{(N)}(t \wedge \tau_1^{(N)}(C))\} &\leq S_1^{(N)}(0) + \mathbf{E}\left\{\int_0^{t \wedge \tau_1^{(N)}(C)} \{k_{11} S_1^{(N)}(u) + Nk_{14}\}\,du\right\} \\ &\leq S_1^{(N)}(0) + \int_0^t (k_{11}\mathbf{E}\{S_1^{(N)}(u \wedge \tau_1^{(N)}(C))\} + Nk_{14})\,du.\end{aligned}$$

Using Gronwall's inequality, we deduce that

$$\mathbf{E}\left\{S_1^{(N)}(t \wedge \tau_1^{(N)}(C))\right\} \leq \left(S_1^{(N)}(0) + Nk_{14}t\right)e^{k_{11}t}, \tag{2.12}$$

uniformly in $C > 0$, and hence that

$$\mathbf{P}\left[\sup_{0 \leq s \leq t} S_1(X_N(s)) \geq NC\right] \leq C^{-1}(S_1(x_N(0)) + k_{14}t)e^{k_{11}t} \tag{2.13}$$

also. Hence $\sup_{0 \leq s \leq t} S_1(X_N(s)) < \infty$ a.s. for any t, $\lim_{C \to \infty} \tau_1^{(N)}(C) = \infty$ a.s., and, from (2.3) and (1.3), it thus follows that $t_\infty^{X_N} = \infty$ a.s. The bound on $\mathbf{E}\{S_1^{(N)}(t)\}$ is now immediate, and that on $\mathbf{E}\{S_0^{(N)}(t)\}$ follows by applying the same Gronwall argument to $M_0^{(N)}(t \wedge \tau_1^{(N)}(C))$. □

The next lemma shows that, if any $T > 0$ is fixed and C is chosen large enough, then, with high probability, $N^{-1}S_0^{(N)}(t) \leq C$ holds for all $0 \leq t \leq T$.

Lemma 2.3 *Assume that Assumptions 2.1 are satisfied, and that $S_0^{(N)}(0) \le NC_0$ and $S_1^{(N)}(0) \le NC_1$. Then, for any $C \ge 2(C_0 + k_{04}T)e^{k_{01}T}$, we have*

$$\mathbf{P}\left[\{\tau_0^{(N)}(C) \le T\}\right] \le (C_1 \vee 1)K_{00}/(NC^2),$$

where K_{00} depends on T and the parameters of the model.

Proof It is immediate from (2.11) and (2.6) that

$$\begin{aligned} S_0^{(N)}(t) &= S_0^{(N)}(0) + N\int_0^t U_0^{(N)}(u)\,du + M_0^{(N)}(t) \\ &\le S_0^{(N)}(0) + \int_0^t \left(k_{01}S_0^{(N)}(u) + Nk_{04}\right)du + \sup_{0\le u\le t} M_0^{(N)}(u). \end{aligned} \tag{2.14}$$

Hence, from Gronwall's inequality, if $S_0^{(N)}(0) \le NC_0$, then

$$S_0^{(N)}(t) \le \left\{N(C_0 + k_{04}T) + \sup_{0\le u\le t} M_0^{(N)}(u)\right\} e^{k_{01}t}. \tag{2.15}$$

Now, considering the quadratic variation of $M_0^{(N)}$, we have

$$\mathbf{E}\left\{\left\{M_0^{(N)}\left(t\wedge\tau_1^{(N)}(C')\right)\right\}^2 - N\int_0^{t\wedge\tau_1^{(N)}(C')} V_0^{(N)}(u)\,du\right\} = 0 \tag{2.16}$$

for any $C' > 0$, from which it follows, much as above, that

$$\begin{aligned} \mathbf{E}\left(\left\{M_0^{(N)}\left(t\wedge\tau_1^{(N)}(C')\right)\right\}^2\right) &\le \mathbf{E}\left\{N\int_0^t V_0^{(N)}\left(u\wedge\tau_1^{(N)}(C')\right)du\right\} \\ &\le \int_0^t \left\{k_{03}\mathbf{E}S_1^{(N)}\left(u\wedge\tau_1^{(N)}(C')\right) + Nk_{05}\right\}du. \end{aligned}$$

Using (2.12), we thus find that

$$\mathbf{E}\left(\left\{M_0^{(N)}\left(t\wedge\tau_1^{(N)}(C')\right)\right\}^2\right) \le \frac{k_{03}}{k_{11}} N\left(C_1 + k_{14}T\right)\left(e^{k_{11}t} - 1\right) + Nk_{05}t, \tag{2.17}$$

uniformly for all C'. Doob's maximal inequality applied to $M_0^{(N)}(t \wedge \tau_1^{(N)}(C'))$ now allows us to deduce that, for any $C', a > 0$,

$$\mathbf{P}\left[\sup_{0\le u\le T} M_0^{(N)}\left(u \wedge \tau_1^{(N)}(C')\right) > aN\right]$$
$$\le \frac{1}{Na^2}\left\{\frac{k_{03}}{k_{11}}\,(C_1 + k_{14}T)\left\{e^{k_{11}T} - 1\right\} + k_{05}T\right\} =: \frac{C_1K_{01} + K_{02}}{Na^2},$$

say, so that, letting $C' \to \infty$,

$$\mathbf{P}\left[\sup_{0\le u\le T} M_0^{(N)}(u) > aN\right] \le \frac{C_1K_{01} + K_{02}}{Na^2}$$

also. Taking $a = \frac{1}{2}Ce^{-k_{01}T}$ and putting the result into (2.15), the lemma follows. $\square$

In the next theorem, we control the 'higher ν-moments' $S_r^{(N)}(t)$ of $X_N(t)$.

Theorem 2.4 *Assume that Assumptions 2.1 are satisfied, and that* $S_1^{(N)}(0) \le NC_1$ *and* $S_{p(1)}^{(N)}(0) \le NC_1'$. *Then, for* $2 \le r \le r_{\max}^{(1)}$ *and for any* $C > 0$, *we have*

$$\mathbf{E}\left\{S_r^{(N)}\left(t \wedge \tau_0^{(N)}(C)\right)\right\} \le \left(S_r^{(N)}(0) + Nk_{r4}t\right)e^{(k_{r1}+Ck_{r2})t}, \quad 0 \le t \le T. \tag{2.18}$$

Furthermore, if for $1 \le r \le r_{\max}^{(2)}$, $S_r^{(N)}(0) \le NC_r$ *and* $S_{p(r)}^{(N)}(0) \le NC_r'$, *then, for any* $\gamma \ge 1$,

$$\mathbf{P}\left[\sup_{0\le t\le T} S_r^{(N)}\left(t \wedge \tau_0^{(N)}(C)\right) \ge N\gamma C''_{rT}\right] \le K_{r0}\gamma^{-2}N^{-1}, \tag{2.19}$$

where

$$C''_{rT} := \left(C_r + k_{r4}T + \sqrt{(C_r' \vee 1)}\right)e^{(k_{r1}+Ck_{r2})T}$$

and K_{r0} *depends on* C, T *and the parameters of the model.*

Proof Recalling (2.11), use the argument leading to (2.12) with the martingales $M_r^{(N)}(t \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C))$, for any $C' > 0$, to deduce that

$$\mathbf{E}S_r^{(N)}\left(t \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C)\right)$$
$$\le S_r^{(N)}(0) + \int_0^t \left(\{k_{r1} + Ck_{r2}\}\,\mathbf{E}\left\{S_r^{(N)}\left(u \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C)\right)\right\} + Nk_{r4}\right)du,$$

for $1 \le r \le r_{\max}^{(1)}$, since $N^{-1} S_0^{(N)}(u) \le C$ when $u \le \tau_0^{(N)}(C)$: define $k_{12} = 0$. Gronwall's inequality now implies that

$$\mathbf{E} S_r^{(N)} \left(t \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C) \right) \le \left(S_r^{(N)}(0) + N k_{r4} t \right) e^{(k_{r1} + C k_{r2}) t}, \tag{2.20}$$

for $1 \le r \le r_{\max}^{(1)}$, and (2.18) follows by Fatou's lemma, on letting $C' \to \infty$.

Now, also from (2.11) and (2.6), we have, for $t \ge 0$ and each $r \le r_{\max}^{(1)}$,

$$\begin{aligned} & S_r^{(N)} \left(t \wedge \tau_0^{(N)}(C) \right) \\ & = S_r^{(N)}(0) + N \int_0^{t \wedge \tau_0^{(N)}(C)} U_r^{(N)}(u)\, du + M_r^{(N)} \left(t \wedge \tau_0^{(N)}(C) \right) \\ & \le S_r^{(N)}(0) + \int_0^t \left(\{k_{r1} + C k_{r2}\} S_r^{(N)} \left(u \wedge \tau_0^{(N)}(C) \right) + N k_{r4} \right) du \\ & \quad + \sup_{0 \le u \le t} M_r^{(N)} \left(u \wedge \tau_0^{(N)}(C) \right). \end{aligned}$$

Hence, from Gronwall's inequality, for all $t \ge 0$ and $r \le r_{\max}^{(1)}$,

$$S_r^{(N)} \left(t \wedge \tau_0^{(N)}(C) \right) \le \left\{ N(C_r + k_{r4} t) + \sup_{0 \le u \le t} M_r^{(N)} \left(u \wedge \tau_0^{(N)}(C) \right) \right\} e^{(k_{r1} + C k_{r2}) t}. \tag{2.21}$$

Now, as in (2.16), we have

$$\mathbf{E} \left\{ \left\{ M_r^{(N)} \left(t \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C) \right) \right\}^2 - N \int_0^{t \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C)} V_r^{(N)}(u)\, du \right\} = 0, \tag{2.22}$$

from which it follows, using (2.7), that, for $1 \le r \le r_{\max}^{(2)}$,

$$\begin{aligned} & \mathbf{E} \left(\left\{ M_r^{(N)} \left(t \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C) \right) \right\}^2 \right) \\ & \le \mathbf{E} \left\{ N \int_0^{t \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C))} V_r^{(N)}(u)\, du \right\} \end{aligned}$$

$$\leq \int_0^t \left\{ k_{r3} \mathbf{E} S^{(N)}_{p(r)} \left(u \wedge \tau_1^{(N)}(C') \wedge \tau_0^{(N)}(C) \right) + N k_{r5} \right\} du$$
$$\leq \frac{N(C'_r + k_{p(r),4}T) k_{r3}}{k_{p(r),1} + C k_{p(r),2}} (e^{(k_{p(r),1} + C k_{p(r),2} t)} - 1) + N k_{r5} T,$$

this last by (2.20), since $p(r) \leq r^{(1)}_{\max}$ for $1 \leq r \leq r^{(2)}_{\max}$. Using Doob's inequality, it follows that, for any $a > 0$,

$$\mathbf{P}\left[\sup_{0 \leq u \leq T} M^{(N)}_r \left(u \wedge \tau_0^{(N)}(C) \right) > aN \right]$$
$$\leq \frac{1}{Na^2} \left\{ \frac{k_{r3}(C'_r + k_{p(r),4}T)}{k_{p(r),1} + C k_{p(r),2}} (e^{(k_{p(r),1} + C k_{p(r),2} T)} - 1) + k_{r5} T \right\}$$
$$=: \frac{C'_r K_{r1} + K_{r2}}{Na^2}.$$

Taking $a = \gamma \sqrt{(C'_r \vee 1)}$ and putting the result into (2.21) gives (2.19), with $K_{r0} = (C'_r K_{r1} + K_{r2})/(C'_r \vee 1)$. □

Note also that $\sup_{0 \leq t \leq T} S^{(N)}_r(t) < \infty$ a.s. for all $0 \leq r \leq r^{(2)}_{\max}$, in view of Lemma 2.3 and Theorem 2.4.

In what follows, we shall particularly need to control quantities of the form $\sum_{J \in \mathcal{J}} \alpha_J(x_N(s)) d(J, \zeta)$, where $x_N := N^{-1} X_N$ and

$$d(J, \zeta) := \sum_{j \geq 0} |J^j| \zeta(j), \tag{2.23}$$

for $\zeta \in \mathcal{R}$ chosen such that $\zeta(j) \geq 1$ grows fast enough with j [see (4.12)]. Defining

$$\tau^{(N)}(a, \zeta) := \inf \left\{ s : \sum_{J \in \mathcal{J}} \alpha_J(x_N(s)) d(J, \zeta) \geq a \right\}, \tag{2.24}$$

infinite if there is no such s, we show in the following corollary that, under suitable assumptions, $\tau^{(N)}(a, \zeta)$ is rarely less than T.

Corollary 2.5 *Assume that Assumptions 2.1 hold, and that ζ is such that*

$$\sum_{J \in \mathcal{J}} \alpha_J(N^{-1}X) d(J, \zeta) \leq \{k_1 N^{-1} S_r(X) + k_2\}^b \tag{2.25}$$

for some $1 \leq r := r(\zeta) \leq r^{(2)}_{\max}$ and some $b = b(\zeta) \geq 1$. For this value of r, assume that $S^{(N)}_r(0) \leq NC_r$ and $S^{(N)}_{p(r)}(0) \leq NC'_r$ for some constants C_r and C'_r. Assume *further that* $S_0^{(N)}(0) \le NC_0$, $S_1^{(N)}(0) \le NC_1$ *for some constants* C_0, C_1*, and define* $C := 2(C_0 + k_{04}T)e^{k_{01}T}$. *Then*

$$\mathbf{P}[\tau^{(N)}(a,\zeta) \le T] \le N^{-1}\{K_{r0}\gamma_a^{-2} + K_{00}(C_1 \vee 1)C^{-2}\},$$

for any $a \ge \{k_2 + k_1 C''_{rT}\}^b$*, where* $\gamma_a := (a^{1/b} - k_2)/\{k_1 C''_{rT}\}$*,* K_{r0} *and* C''_{rT} *are as in Theorem 2.4, and* K_{00} *is as in Lemma 2.3.*

Proof In view of (2.25), it is enough to bound the probability

$$\mathbf{P}[\sup_{0\le t\le T} S_r^{(N)}(t) \ge N(a^{1/b} - k_2)/k_1].$$

However, Lemma 2.3 and Theorem 2.4 together bound this probability by

$$N^{-1}\left\{K_{r0}\gamma_a^{-2} + K_{00}(C_1 \vee 1)C^{-2}\right\},$$

where γ_a is as defined above, as long as $a^{1/b} - k_2 \ge k_1 C''_{rT}$. □

If (2.25) is satisfied, $\sum_{J\in\mathcal{J}} \alpha_J(x_N(s))d(J,\zeta)$ is a.s. bounded on $0 \le s \le T$, because $S_r^{(N)}(s)$ is. The corollary shows that the sum is then bounded by $\{k_2 + k_1 C''_{r,T}\}^b$, except on an event of probability of order $O(N^{-1})$. Usually, one can choose $b = 1$.

3 Semigroup properties

We make the following initial assumptions about the matrix A: first, that

$$A_{ij} \ge 0 \quad \text{for all } i \ne j \ge 0; \quad \sum_{j\ne i} A_{ji} < \infty \quad \text{for all } i \ge 0, \tag{3.1}$$

and then that, for some $\mu \in R_+^{\mathbb{Z}_+}$ such that $\mu(m) \ge 1$ for each $m \ge 0$, and for some $w \ge 0$,

$$A^T\mu \le w\mu. \tag{3.2}$$

We then use μ to define the μ-norm

$$\|\xi\|_\mu := \sum_{m\ge 0} \mu(m)|\xi^m| \quad \text{on} \quad \mathcal{R}_\mu := \{\xi \in \mathcal{R}\colon\ \|\xi\|_\mu < \infty\}. \tag{3.3}$$

Note that there may be many possible choices for μ. In what follows, it is important that F be a Lipschitz operator with respect to the μ-norm, and this has to be borne in mind when choosing μ.

Setting

$$Q_{ij} := A^T_{ij}\mu(j)/\mu(i) - w\delta_{ij}, \tag{3.4}$$

where δ is the Kronecker delta, we note that $Q_{ij} \geq 0$ for $i \neq j$, and that

$$0 \leq \sum_{j \neq i} Q_{ij} = \sum_{j \neq i} A^T_{ij} \mu(j)/\mu(i) \leq w - A_{ii} = -Q_{ii},$$

using (3.2) for the inequality, so that $Q_{ii} \leq 0$. Hence Q can be augmented to a conservative Q–matrix, in the sense of Markov jump processes, by adding a coffin state ∂, and setting $Q_{i\partial} := -\sum_{j \geq 0} Q_{ij} \geq 0$. Let $P(\cdot)$ denote the semigroup of Markov transition matrices corresponding to the minimal process associated with Q; then, in particular,

$$Q = P'(0) \quad \text{and} \quad P'(t) = QP(t) \quad \text{for all } t \geq 0 \tag{3.5}$$

([15], Theorem 3). Set

$$R^T_{ij}(t) := e^{wt} \mu(i) P_{ij}(t)/\mu(j). \tag{3.6}$$

Theorem 3.1 *Let A satisfy Assumptions (3.1) and (3.2). Then, with the above definitions, R is a strongly continuous semigroup on $\mathcal{R}_\mu$, and*

$$\sum_{i \geq 0} \mu(i) R_{ij}(t) \leq \mu(j) e^{wt} \quad \textit{for all } j \textit{ and } t. \tag{3.7}$$

Furthermore, the sums $\sum_{j \geq 0} R_{ij}(t) A_{jk} = (R(t)A)_{ik}$ are well defined for all i, k, and

$$A = R'(0) \quad \textit{and} \quad R'(t) = R(t)A \quad \textit{for all } t \geq 0. \tag{3.8}$$

Proof We note first that, for $x \in \mathcal{R}_\mu$,

$$\begin{aligned} \|R(t)x\|_\mu &\leq \sum_{i \geq 0} \mu(i) \sum_{j \geq 0} R_{ij}(t)|x_j| = e^{wt} \sum_{i \geq 0} \sum_{j \geq 0} \mu(j) P_{ji}(t)|x_j| \\ &\leq e^{wt} \sum_{j \geq 0} \mu(j)|x_j| = e^{wt} \|x\|_\mu, \end{aligned} \tag{3.9}$$

since $P(t)$ is substochastic on $\mathbb{Z}_+$; hence $R\colon \mathcal{R}_\mu \to \mathcal{R}_\mu$. To show strong continuity, we take $x \in \mathcal{R}_\mu$, and consider

$$\begin{aligned} \|R(t)x - x\|_\mu &= \sum_{i \geq 0} \mu(i) \left| \sum_{j \geq 0} R_{ij}(t) x_j - x_i \right| = \sum_{i \geq 0} \left| e^{wt} \sum_{j \geq 0} \mu(j) P_{ji}(t) x_j - \mu(i) x_i \right| \\ &\leq (e^{wt} - 1) \sum_{i \geq 0} \sum_{j \geq 0} \mu(j) P_{ji}(t) x_j + \sum_{i \geq 0} \sum_{j \neq i} \mu(j) P_{ji}(t) x_j + \sum_{i \geq 0} \mu(i) x_i (1 - P_{ii}(t)) \\ &\leq (e^{wt} - 1) \sum_{j \geq 0} \mu(j) x_j + 2 \sum_{i \geq 0} \mu(i) x_i (1 - P_{ii}(t)), \end{aligned}$$

from which it follows that $\lim_{t\to 0} \|R(t)x - x\|_\mu = 0$, by dominated convergence, since $\lim_{t\to 0} P_{ii}(t) = 1$ for each $i \ge 0$.

The inequality (3.7) follows from the definition of R and the fact that P is substochastic on $\mathbb{Z}_+$. Then

$$\left(A^T R^T(t)\right)_{ij} = \sum_{k\neq i} Q_{ik} \frac{\mu(i)}{\mu(k)} e^{wt} \frac{\mu(k)}{\mu(j)} P_{kj}(t) + (Q_{ii} + w) e^{wt} \frac{\mu(i)}{\mu(j)} P_{ij}(t)$$
$$= \frac{\mu(i)}{\mu(j)} \left[(QP(t))_{ij} + wP_{ij}(t)\right] e^{wt},$$

with $(QP(t))_{ij} = \sum_{k\ge 0} Q_{ik} P_{kj}(t)$ well defined because $P(t)$ is sub-stochastic and Q is conservative. Using (3.5), this gives

$$(A^T R^T(t))_{ij} = \frac{\mu(i)}{\mu(j)} \frac{d}{dt}[P_{ij}(t) e^{wt}] = \frac{d}{dt} R^T_{ij}(t),$$

and this establishes (3.8). □

4 Main approximation

Let X_N, $N \ge 1$, be a sequence of pure jump Markov processes as in Sect. 1, with A and F defined as in (1.4) and (1.5), and suppose that $F\colon \mathcal{R}_\mu \to \mathcal{R}_\mu$, with $\mathcal{R}_\mu$ as defined in (3.3), for some μ such that Assumption (3.2) holds. Suppose also that F is locally Lipschitz in the μ-norm: for any $z > 0$,

$$\sup_{x\neq y\colon\ \|x\|_\mu, \|y\|_\mu \le z} \|F(x) - F(y)\|_\mu / \|x - y\|_\mu \le K(\mu, F; z) < \infty. \tag{4.1}$$

Then, for $x(0) \in \mathcal{R}_\mu$ and R as in (3.6), the integral equation

$$x(t) = R(t)x(0) + \int_0^t R(t-s) F(x(s))\, ds \tag{4.2}$$

has a unique continuous solution x in $\mathcal{R}_\mu$ on some non-empty time interval $[0, t_{\max})$, such that, if $t_{\max} < \infty$, then $\|x(t)\|_\mu \to \infty$ as $t \to t_{\max}$ ([14], Theorem 1.4, Chapter 6). Thus, if A were the generator of R, the function x would be a *mild solution* of the deterministic equations (1.4). We now wish to show that the process $x_N := N^{-1} X_N$ is close to x. To do so, we need a corresponding representation for X_N.

To find such a representation, let $W(t)$, $t \ge 0$, be a pure jump path on $\mathcal{X}_+$ that has only finitely many jumps up to time T. Then we can write

$$W(t) = W(0) + \sum_{j\colon \sigma_j \le t} \Delta W(\sigma_j), \quad 0 \le t \le T, \tag{4.3}$$

where $\Delta W(s) := W(s) - W(s-)$ and σ_j, $j \geq 1$, denote the times when W has its jumps. Now let A satisfy (3.1) and (3.2), and let $R(\cdot)$ be the associated semigroup, as defined in (3.6). Define the path $W^*(t)$, $0 \leq t \leq T$, from the equation

$$W^*(t) := R(t)W(0) + \sum_{j\,:\,\sigma_j \leq t} R(t-\sigma_j)\Delta_j - \int_0^t R(t-s)AW(s)\,ds, \tag{4.4}$$

where $\Delta_j := \Delta W(\sigma_j)$. Note that the latter integral makes sense, because each of the sums $\sum_{j\geq 0} R_{ij}(t)A_{jk}$ is well defined, from Theorem 3.1, and because only finitely many of the coordinates of W are non-zero.

Lemma 4.1 $W^*(t) = W(t)$ *for all* $0 \leq t \leq T$.

Proof Fix any t, and suppose that $W^*(s) = W(s)$ for all $s \leq t$. This is clearly the case for $t = 0$. Let $\sigma(t) > t$ denote the time of the first jump of W after t. Then, for any $0 < h < \sigma(t) - t$, using the semigroup property for R and (4.4),

$$\begin{aligned} &W^*(t+h) - W^*(t) \\ &\quad = (R(h) - I)R(t)W(0) + \sum_{j\,:\,\sigma_j \leq t} (R(h) - I)R(t-\sigma_j)\Delta_j \\ &\qquad - \int_0^t (R(h) - I)R(t-s)AW(s)\,ds - \int_t^{t+h} R(t+h-s)AW(t)\,ds, \end{aligned} \tag{4.5}$$

where, in the last integral, we use the fact that there are no jumps of W between t and $t+h$. Thus we have

$$\begin{aligned} &W^*(t+h) - W^*(t) \\ &\quad = (R(h) - I)\left\{R(t)W(0) + \sum_{j\,:\,\sigma_j \leq t} R(t-\sigma_j)\Delta_j - \int_0^t R(t-s)AW(s)\,ds\right\} \\ &\qquad - \int_t^{t+h} R(t+h-s)AW(t)\,ds \\ &\quad = (R(h) - I)W(t) - \int_t^{t+h} R(t+h-s)AW(t)\,ds. \end{aligned} \tag{4.6}$$

But now, for $x \in \mathcal{X}_+$,

$$\int_t^{t+h} R(t+h-s)Ax\,ds = (R(h) - I)x,$$

from (3.8), so that $W^*(t+h) = W^*(t)$ for all $t+h < \sigma(t)$, implying that $W^*(s) = W(s)$ for all $s < \sigma(t)$. On the other hand, from (4.4), we have $W^*(\sigma(t)) - W^*(\sigma(t)-) = \Delta W(\sigma(t))$, so that $W^*(s) = W(s)$ for all $s \leq \sigma(t)$. Thus we can prove equality over the interval $[0, \sigma_1]$, and then successively over the intervals $[\sigma_j, \sigma_{j+1}]$, until $[0, T]$ is covered. □

Now suppose that W arises as a realization of X_N. Then X_N has transition rates such that

$$M_N(t) := \sum_{j\,:\,\sigma_j \leq t} \Delta X_N(\sigma_j) - \int_0^t A X_N(s)\,ds - \int_0^t N F(x_N(s))\,ds \tag{4.7}$$

is a zero mean local martingale. In view of Lemma 4.1, we can use (4.4) to write

$$X_N(t) = R(t)X_N(0) + \widetilde{M}_N(t) + N \int_0^t R(t-s)F(x_N(s))\,ds, \tag{4.8}$$

where

$$\begin{aligned} \widetilde{M}_N(t) := &\sum_{j\,:\,\sigma_j \leq t} R(t-\sigma_j)\Delta X_N(\sigma_j) \\ &- \int_0^t R(t-s)AX_N(s)\,ds - \int_0^t R(t-s)NF(x_N(s))\,ds. \end{aligned} \tag{4.9}$$

Thus, comparing (4.8) and (4.2), we expect x_N and x to be close, for $0 \leq t \leq T < t_{\max}$, provided that we can show that $\sup_{t \leq T} \|\widetilde{m}_N(t)\|_\mu$ is small, where $\widetilde{m}_N(t) := N^{-1}\widetilde{M}_N(t)$. Indeed, if $x_N(0)$ and $x(0)$ are close, then

$$\begin{aligned} &\|x_N(t) - x(t)\|_\mu \\ &\quad \leq \|R(t)(x_N(0) - x(0))\|_\mu \\ &\qquad + \int_0^t \|R(t-s)[F(x_N(s)) - F(x(s))]\|_\mu\,ds + \|\widetilde{m}_N(t)\|_\mu \\ &\quad \leq e^{wt}\|x_N(0) - x(0)\|_\mu \\ &\qquad + \int_0^t e^{w(t-s)}K(\mu, F; 2\Xi_T)\|x_N(s) - x(s)\|_\mu\,ds + \|\widetilde{m}_N(t)\|_\mu, \end{aligned} \tag{4.10}$$

by (3.9), with the stage apparently set for Gronwall's inequality, assuming that $\|x_N(0) - x(0)\|_\mu$ and $\sup_{0 \leq t \leq T} \|\widetilde{m}_N(t)\|_\mu$ are small enough that then $\|x_N(t)\|_\mu \leq 2\Xi_T$ for $0 \leq t \leq T$, where $\Xi_T := \sup_{0 \leq t \leq T} \|x(t)\|_\mu$.

Bounding $\sup_{0 \le t \le T} \|\widetilde{m}_N(t)\|_\mu$ is, however, not so easy. Since $\widetilde{M}_N$ is not itself a martingale, we cannot directly apply martingale inequalities to control its fluctuations. However, since

$$\widetilde{M}_N(t) = \int_0^t R(t-s)\, dM_N(s), \tag{4.11}$$

we can hope to use control over the local martingale M_N instead. For this and the subsequent argument, we introduce some further assumptions.

Assumption 4.2 1. There exists $r = r_\mu \le r_{\max}^{(2)}$ such that $\sup_{j \ge 0}\{\mu(j)/\nu_r(j)\} < \infty$.
2. There exists $\zeta \in \mathcal{R}$, with $\zeta(j) \ge 1$ for all j, such that (2.25) is satisfied for some $b = b(\zeta) \ge 1$ and $r = r(\zeta)$ such that $1 \le r(\zeta) \le r_{\max}^{(2)}$, and such that

$$Z := \sum_{k \ge 0} \frac{\mu(k)(|A_{kk}| + 1)}{\sqrt{\zeta(k)}} < \infty. \tag{4.12}$$

The requirement that ζ satisfies (4.12) as well as satisfying (2.25) for some $r \le r_{\max}^{(2)}$ implies in practice that it must be possible to take $r_{\max}^{(1)}$ and $r_{\max}^{(2)}$ to be quite large in Assumption 2.1 (see the examples in Sect. 5).

Note that part 1 of Assumption 4.2 implies that $\lim_{j\to\infty}\{\mu(j)/\nu_r(j)\} = 0$ for some $r = \tilde{r}_\mu \le r_\mu + 1$. We define

$$\rho(\zeta, \mu) := \max\{r(\zeta), p(r(\zeta)), \tilde{r}_\mu\}, \tag{4.13}$$

where $p(\cdot)$ is as in Assumptions 2.1. We can now prove the following lemma, which enables us to control the paths of $\widetilde{M}_N$ by using fluctuation bounds for the martingale M_N.

Lemma 4.3 *Under Assumption 4.2,*

$$\widetilde{M}_N(t) = M_N(t) + \int_0^t R(t-s) A M_N(s)\, ds.$$

Proof From (3.8), we have

$$R(t-s) = I + \int_0^{t-s} R(v) A\, dv.$$

Substituting this into (4.11), we obtain

$$\begin{aligned}\widetilde{M}_N(t) &= \int_0^t R(t-s)\,dM_N(s)\\ &= M_N(t) + \int_0^t \left\{ \int_0^t R(v)A\mathbf{1}_{[0,t-s]}(v)\,dv \right\} dM_N(s)\\ &= M_N(t) + \int_0^t \left\{ \int_0^t R(v)A\mathbf{1}_{[0,t-s]}(v)\,dv \right\} dX_N(s)\\ &\quad - \int_0^t \left\{ \int_0^t R(v)A\mathbf{1}_{[0,t-s]}(v)\,dv \right\} F_0(x_N(s))\,ds.\end{aligned}$$

It remains to change the order of integration in the double integrals, for which we use Fubini's theorem.

In the first, the outer integral is almost surely a finite sum, and at each jump time $t_l^{X_N}$ we have $dX_N(t_l^{X_N}) \in \mathcal{J}$. Hence it is enough that, for each i, m and t, $\sum_{j\ge 0} R_{ij}(t)A_{jm}$ is absolutely summable, which follows from Theorem 3.1. Thus we have

$$\int_0^t \left\{ \int_0^t R(v)A\mathbf{1}_{[0,t-s]}(v)\,dv \right\} dX_N(s) = \int_0^t R(v)A\{X_N(t-v) - X_N(0)\}\,dv. \tag{4.14}$$

For the second, the kth component of $R(v)AF_0(x_N(s))$ is just

$$\sum_{j\ge 0} R_{kj}(v) \sum_{l\ge 0} A_{jl} \sum_{J\in\mathcal{J}} J^l \alpha_J(x_N(s)). \tag{4.15}$$

Now, from (3.7), we have $0 \le R_{kj}(v) \le \mu(j)e^{wv}/\mu(k)$, and

$$\sum_{j\ge 0} \mu(j)|A_{jl}| \le \mu(l)(2|A_{ll}| + w), \tag{4.16}$$

because $A^T\mu \le w\mu$. Hence, putting absolute values in the summands in (4.15) yields at most

$$\frac{e^{wv}}{\mu(k)} \sum_{J\in\mathcal{J}} \alpha_J(x_N(s)) \sum_{l\ge 0} |J^l|\mu(l)(2|A_{ll}| + w).$$

Now, in view of (4.12) and since $\zeta(j) \ge 1$ for all j, there is a constant $K < \infty$ such that $\mu(l)(2|A_{ll}| + w) \le K\zeta(l)$. Furthermore, ζ satisfies (2.25), so that, by Corollary 2.5,

$\sum_{J\in\mathcal{J}} \alpha_J(x_N(s)) \sum_{l\geq 0} |J^l|\zeta(l)$ is a.s. uniformly bounded in $0 \leq s \leq T$. Hence we can apply Fubini's theorem, obtaining

$$\int_0^t \left\{ \int_0^t R(v)A\mathbf{1}_{[0,t-s]}(v)\,dv \right\} F_0(x_N(s))\,ds = \int_0^t R(v)A \left\{ \int_0^{t-v} F_0(x_N(s))\,ds \right\} dv,$$

and combining this with (4.14) proves the lemma. □

We now introduce the exponential martingales that we use to bound the fluctuations of M_N. For $\theta \in R^{\mathbb{Z}_+}$ bounded and $x \in \mathcal{R}_\mu$,

$$Z_{N,\theta}(t) := e^{\theta^T x_N(t)} \exp\left\{ -\int_0^t g_{N\theta}(x_N(s-))\,ds \right\}, \quad t \geq 0,$$

is a non-negative finite variation local martingale, where

$$g_{N\theta}(\xi) := \sum_{J\in\mathcal{J}} N\alpha_J(\xi)\left(e^{N^{-1}\theta^T J} - 1\right).$$

For $t \geq 0$, we have

$$\begin{aligned}\log Z_{N,\theta}(t) &= \theta^T x_N(t) - \int_0^t g_{N\theta}(x_N(s-))\,ds \\ &= \theta^T m_N(t) - \int_0^t \varphi_{N,\theta}(x_N(s-), s)\,ds, \end{aligned} \tag{4.17}$$

where

$$\varphi_{N,\theta}(\xi) := \sum_{J\in\mathcal{J}} N\alpha_J(\xi)\left(e^{N^{-1}\theta^T J} - 1 - N^{-1}\theta^T J\right), \tag{4.18}$$

and $m_N(t) := N^{-1} M_N(t)$. Note also that we can write

$$\varphi_{N,\theta}(\xi) = N \int_0^1 (1-r) D^2 v_N(\xi, r\theta)[\theta, \theta]\,dr, \tag{4.19}$$

where

$$v_N(\xi, \theta') := \sum_{J\in\mathcal{J}} \alpha_J(\xi) e^{N^{-1}(\theta')^T J},$$

and $D^2 v_N$ denotes the matrix of second derivatives with respect to the second argument:

$$D^2 v_N(\xi, \theta')[\zeta_1, \zeta_2] := N^{-2} \sum_{J \in \mathcal{J}} \alpha_J(\xi) e^{N^{-1}(\theta')^T J} \zeta_1^T J J^T \zeta_2 \tag{4.20}$$

for any $\zeta_1, \zeta_2 \in \mathcal{R}_\mu$.

Now choose any $B := (B_k,\ k \geq 0) \in \mathcal{R}$, and define $\tilde{\tau}_k^{(N)}(B)$ by

$$\tilde{\tau}_k^{(N)}(B) := \inf \left\{ t \geq 0\colon \sum_{J: J_k \neq 0} \alpha_J(x_N(t-)) > B_k \right\}.$$

Our exponential bound is as follows.

Lemma 4.4 *For any $k \geq 0$,*

$$\mathbf{P}\left[\sup_{0 \leq t \leq T \wedge \tilde{\tau}_k^{(N)}(B)} |m_N^k(t)| \geq \delta \right] \leq 2 \exp(-\delta^2 N / 2 B_k K_* T),$$

for all $0 < \delta \leq B_k K_* T$, where $K_* := J_*^2 e^{J_*}$, and J_* is as in (1.2).

Proof Take $\theta = e^{(k)}\beta$, for β to be chosen later. We shall argue by stopping the local martingale $Z_{N,\theta}$ at time $\sigma^{(N)}(k, \delta)$, where

$$\sigma^{(N)}(k, \delta) := T \wedge \tilde{\tau}_k^{(N)}(B) \wedge \inf\{t\colon\ m_N^k(t) \geq \delta\}.$$

Note that $e^{N^{-1}\theta^T J} \leq e^{J_*}$, so long as $|\beta| \leq N$, so that

$$D^2 v_N(\xi, r\theta)[\theta, \theta] \leq N^{-2} \left(\sum_{J: J_k \neq 0} \alpha_J(\xi) \right) \beta^2 K_*.$$

Thus, from (4.19), we have

$$\varphi_{N,\theta}(x_N(u-)) \leq \frac{1}{2} N^{-1} B_k \beta^2 K_*, \qquad u \leq \tilde{\tau}_k^{(N)}(B),$$

and hence, on the event that $\sigma^{(N)}(k, \delta) = \inf\{t\colon\ m_N^k(t) \geq \delta\} \leq (T \wedge \tilde{\tau}_k^{(N)}(B))$, we have

$$Z_{N,\theta}(\sigma(k, \delta)) \ \geq\ \exp\left\{ \beta\delta - \frac{1}{2} N^{-1} B_k \beta^2 K_* T \right\}.$$

But since $Z_{N,\theta}(0) = 1$, it now follows from the optional stopping theorem and Fatou's lemma that

$$\begin{aligned} 1 &\geq \mathbf{E}\{Z_{N,\theta}(\sigma^{(N)}(k,\delta))\} \\ &\geq \mathbf{P}\left[\sup_{0\leq t\leq T\wedge\tilde{\tau}_k^{(N)}(B)} m_N^k(t) \geq \delta\right] \exp\left\{\beta\delta - \frac{1}{2}N^{-1}B_k\beta^2K_*T\right\}. \end{aligned}$$

We can choose $\beta = \delta N/B_kK_*T$, as long as $\delta/B_kK_*T \leq 1$, obtaining

$$\mathbf{P}\left(\sup_{0\leq t\leq T\wedge\tilde{\tau}_k^{(N)}(B)} m_N^k(t) \geq \delta\right) \leq \exp(-\delta^2N/2B_kK_*T).$$

Repeating with

$$\tilde{\sigma}^{(N)}(k,\delta) := T \wedge \tilde{\tau}_k^{(N)}(B) \wedge \inf\{t\colon\ -m_N^k(t) \geq \delta\},$$

and choosing $\beta = \delta N/B_kK_*T$, gives the lemma. □

The preceding lemma gives a bound for each individual component of M_N. We need first to translate this into a statement for all components simultaneously. For ζ as in Assumption 4.2, we start by writing

$$Z_*^{(1)} := \max_{k\geq 1} k^{-1}\#\{m\colon\ \zeta(m) \leq k\};\quad Z_*^{(2)} := \sup_{k\geq 0}\frac{\mu(k)(|A_{kk}|+1)}{\sqrt{\zeta(k)}}. \tag{4.21}$$

$Z_*^{(2)}$ is clearly finite, because of Assumption 4.2, and the same is true for $Z_*^{(1)}$ also, since Z of Assumption 4.2 is at least $\#\{m\colon\ \zeta(m) \leq k\}/\sqrt{k}$, for each k. Then, using the definition (2.24) of $\tau^{(N)}(a,\zeta)$, note that, for every k,

$$\sum_{J\colon\ J^k\neq 0}\alpha_J(x_N(t))h(k) \leq \sum_{J\colon\ J^k\neq 0}\frac{\alpha_J(x_N(t))h(k)d(J,\zeta)}{|J^k|\zeta(k)} \leq \frac{ah(k)}{\zeta(k)}, \tag{4.22}$$

for any $t < \tau^{(N)}(a,\zeta)$ and any $h \in \mathcal{R}$, and that, for any $\mathcal{K} \subseteq \mathbb{Z}_+$,

$$\begin{aligned} \sum_{k\in\mathcal{K}}\sum_{J\colon\ J^k\neq 0}\alpha_J(x_N(t))h(k) &\leq \sum_{k\in\mathcal{K}}\sum_{J\colon\ J^k\neq 0}\frac{\alpha_J(x_N(t))h(k)d(J,\zeta)}{|J^k|\zeta(k)} \\ &\leq \frac{a}{\min_{k\in\mathcal{K}}(\zeta(k)/h(k))}. \end{aligned} \tag{4.23}$$

From (4.22) with $h(k) = 1$ for all k, if we choose $B := (a/\zeta(k), k \geq 0)$, then $\tau^{(N)}(a,\zeta) \leq \tilde{\tau}_k^{(N)}(B)$ for all k. For this choice of B, we can take

$$\delta_k^2 := \delta_k^2(a) := \frac{4aK_*T\log N}{N\zeta(k)} = \frac{4B_kK_*T\log N}{N} \tag{4.24}$$

in Lemma 4.4 for $k \in \kappa_N(a)$, where

$$\kappa_N(a) := \left\{k\colon\ \zeta(k) \le \frac{1}{4} a K_* T N / \log N\right\} = \{k\colon\ B_k \ge 4 \log N / K_* T N\}\,, \quad (4.25)$$

since then $\delta_k(a) \le B_k K_* T$. Note that then, from (4.12),

$$\sum_{k \in \kappa_N(a)} \mu(k)\delta_k(a) \le 2Z\sqrt{a K_* T N^{-1} \log N}, \quad (4.26)$$

with Z as defined in Assumption 4.2, and that

$$|\kappa_N(a)| \le \frac{1}{4} a Z_*^{(1)} K_* T N / \log N. \quad (4.27)$$

Lemma 4.5 *If Assumptions 4.2 are satisfied, taking $\delta_k(a)$ and $\kappa_N(a)$ as defined in (4.24) and (4.25), and for any $\eta \in \mathcal{R}$, we have*

1. $$\mathbf{P}\left[\bigcup_{k \in \kappa_N(a)} \left\{\sup_{0 \le t \le T \wedge \tau^{(N)}(a,\zeta)} |m_N(t)| \ge \delta_k(a)\right\}\right] \le \frac{a Z_*^{(1)} K_* T}{2N \log N};$$
2. $$\mathbf{P}\left[\sum_{k \notin \kappa_N(a)} X_N^k(t) = 0 \textit{ for all } 0 \le t \le T \wedge \tau^{(N)}(a,\zeta)\right] \ge 1 - \frac{4 \log N}{K_* N};$$
3. $$\sup_{0 \le t \le T \wedge \tau^{(N)}(a,\zeta)} \left\{\sum_{k \notin \kappa_N(a)} \eta(k) |F^k(x_N(t))|\right\} \le \frac{a J_*}{\min_{k \notin \kappa_N(a)} (\zeta(k)/\eta(k))}.$$

Proof For part 1, use Lemma 4.4 together with (4.24) and (4.27) to give the bound. For part 2, the total rate of jumps into coordinates with indices $k \notin \kappa_N(a)$ is

$$\sum_{k \notin \kappa_N(a)} \sum_{J\colon\ J^k \ne 0} \alpha_J(x_N(t)) \le \frac{a}{\min_{k \notin \kappa_N(a)} \zeta(k)},$$

if $t \le \tau^{(N)}(a,\zeta)$, using (4.23) with $\mathcal{K} = (\kappa_N(a))^c$, which, combined with (4.25), proves the claim. For the final part, if $t \le \tau^{(N)}(a,\zeta)$,

$$\sum_{k \notin \kappa_N(a)} \eta(k) |F^k(x_N(t))| \le \sum_{k \notin \kappa_N(a)} \eta(k) \sum_{J\,:\,J^k \ne 0} \alpha_J(x_N(t)) J_*,$$

and the inequality follows once more from (4.23). $\square$

Let $B_N^{(1)}(a)$ and $B_N^{(2)}(a)$ denote the events

$$\begin{aligned} B_N^{(1)}(a) &:= \left\{ \sum_{k \notin \kappa_N(a)} X_N^k(t) = 0 \text{ for all } 0 \le t \le T \wedge \tau^{(N)}(a,\zeta) \right\}; \\ B_N^{(2)}(a) &:= \left(\bigcap_{k \in \kappa_N(a)} \left\{ \sup_{0 \le t \le T \wedge \tau^{(N)}(a,\zeta)} |m_N(t)| \le \delta_k(a) \right\} \right), \end{aligned} \tag{4.28}$$

and set $B_N(a) := B_N^{(1)}(a) \cap B_N^{(2)}(a)$. Then, by Lemma 4.5, we deduce that

$$\mathbf{P}[B_N(a)^c] \le \frac{aZ_*^{(1)} K_* T}{2N \log N} + \frac{4 \log N}{K_* N}, \tag{4.29}$$

of order $O(N^{-1} \log N)$ for each fixed a. Thus we have all the components of M_N simultaneously controlled, except on a set of small probability. We now translate this into the desired assertion about the fluctuations of $\widetilde{m}_N$.

Lemma 4.6 *If Assumptions 4.2 are satisfied, then, on the event $B_N(a)$,*

$$\sup_{0 \le t \le T \wedge \tau^{(N)}(a,\zeta)} \|\widetilde{m}_N(t)\|_\mu \le \sqrt{a}\, K_{4.6} \sqrt{\frac{\log N}{N}},$$

where the constant $K_{4.6}$ depends on T and the parameters of the process.

Proof From Lemma 4.3, it follows that

$$\begin{aligned} &\sup_{0 \le t \le T \wedge \tau^{(N)}(a,\zeta)} \|\widetilde{m}_N(t)\|_\mu \\ &\quad \le \sup_{0 \le t \le T \wedge \tau^{(N)}(a,\zeta)} \|m_N(t)\|_\mu + \sup_{0 \le t \le T \wedge \tau^{(N)}(a,\zeta)} \int_0^t \|R(t-s)Am_N(s)\|_\mu \, ds. \end{aligned} \tag{4.30}$$

For the first term, on $B_N(a)$ and for $0 \le t \le T \wedge \tau^{(N)}(a,\zeta)$, we have

$$\|m_N(t)\|_\mu \le \sum_{k \in \kappa_N(a)} \mu(k)\delta_k(a) + \int_0^t \sum_{k \notin \kappa_N(a)} \mu(k) |F^k(x_N(u))| \, du.$$

The first sum is bounded using (4.26) by $2Z\sqrt{aK_*T}\, N^{-1/2}\sqrt{\log N}$, the second, from Lemma 4.5 and (4.25), by

$$\frac{TaJ_*}{\min_{k \notin \kappa_N(a)} (\zeta(k)/\mu(k))} \le Z_*^{(2)} 2J_* \sqrt{\frac{Ta}{K_*}} \sqrt{\frac{\log N}{N}}.$$

For the second term in (4.30), from (3.7) and (4.16), we note that

$$\begin{aligned}\|R(t-s)Am_N(s)\|_\mu &\le \sum_{k\ge 0}\mu(k)\sum_{l\ge 0}R_{kl}(t-s)\sum_{r\ge 0}|A_{lr}||m_N^r(s)| \\ &\le e^{w(t-s)}\sum_{l\ge 0}\mu(l)\sum_{r\ge 0}|A_{lr}||m_N^r(s)| \\ &\le e^{w(t-s)}\sum_{r\ge 0}\mu(r)\{2|A_{rr}|+w\}|m_N^r(s)|.\end{aligned}$$

On $B_N(a)$ and for $0\le s\le T\wedge\tau^{(N)}(a,\zeta)$, from (4.12), the sum for $r\in\kappa_N(a)$ is bounded using

$$\begin{aligned}&\sum_{r\in\kappa_N(a)}\mu(r)\{2|A_{rr}|+w\}|m_N^r(s)| \\ &\quad\le \sum_{r\in\kappa_N(a)}\mu(r)\{2|A_{rr}|+w\}\delta_r(a) \\ &\quad\le \sum_{r\in\kappa_N(a)}\mu(r)\{2|A_{rr}|+w\}\sqrt{\frac{4aK_*T\log N}{N\zeta(r)}} \\ &\quad\le (2\vee w)Z\sqrt{4aK_*T}\sqrt{\frac{\log N}{N}}.\end{aligned}$$

The remaining sum is then bounded by Lemma 4.5, on the set $B_N(a)$ and for $0\le s\le T\wedge\tau^{(N)}(a,\zeta)$, giving at most

$$\begin{aligned}&\sum_{r\notin\kappa_N(a)}\mu(r)\{2|A_{rr}|+w\}|m_N^r(s)| \\ &\quad\le \sum_{r\notin\kappa_N(a)}\mu(r)\{2|A_{rr}|+w\}\int_0^s|F^r(x_N(t))|\,dt \\ &\quad\le \frac{(2\vee w)saJ_*}{\min_{k\notin\kappa_N(a)}(\zeta(k)/\mu(k)\{|A_{kk}|+1\})} \\ &\quad\le (2\vee w)Z_*^{(2)}2J_*\sqrt{\frac{Ta}{K_*}}\sqrt{\frac{\log N}{N}}.\end{aligned}$$

Integrating, it follows that

$$\begin{aligned}&\sup_{0\le t\le T\wedge\tau^{(N)}(a,\zeta)}\int_0^t\|R(t-s)Am_N(s)\|_\mu\,ds \\ &\quad\le (2T\vee 1)e^{wT}\left\{\sqrt{4aK_*T}Z+Z_*^{(2)}J2J_*\sqrt{\frac{Ta}{K_*}}\right\}\sqrt{\frac{\log N}{N}},\end{aligned}$$

and the lemma follows. □

This has now established the control on $\sup_{0\le t\le T}\|\widetilde{m}_N(t)\|_\mu$ that we need, in order to translate (4.10) into a proof of the main theorem.

Theorem 4.7 *Suppose that* (1.2)*,* (1.3)*,* (3.1)*,* (3.2) *and* (4.1) *are all satisfied, and that Assumptions* 2.1 *and* 4.2 *hold. Recalling the definition* (4.13) *of* $\rho(\zeta,\mu)$*, for* ζ *as given in Assumption* 4.2*, suppose that* $S^{(N)}_{\rho(\zeta,\mu)}(0)\le NC_*$ *for some* $C_*<\infty$*.*

Let x *denote the solution to* (4.2) *with initial condition* $x(0)$ *satisfying* $S_{\rho(\zeta,\mu)}(x(0))<\infty$*. Then* $t_{\max}=\infty$*.*

Fix any T*, and define* $\Xi_T := \sup_{0\le t\le T}\|x(t)\|_\mu$*. If* $\|x_N(0)-x(0)\|_\mu \le \frac{1}{2}\Xi_T e^{-(w+k_*)T}$*, where* $k_* := e^{wT}K(\mu,F;2\Xi_T)$*, then there exist constants* c_1, c_2 *depending on* C_*, T *and the parameters of the process, such that for all* N *large enough*

$$\mathbf{P}\left(\sup_{0\le t\le T}\|x_N(t)-x(t)\|_\mu > \left(e^{wT}\|x_N(0)-x(0)\|_\mu + c_1\sqrt{\frac{\log N}{N}}\right)e^{k_*T}\right) \le \frac{c_2\log N}{N}. \tag{4.31}$$

Proof As $S^{(N)}_{\rho(\zeta,\mu)}(0)\le NC_*$, it follows also that $S^{(N)}_r(0)\le NC_*$ for all $0\le r\le \rho(\zeta,\mu)$. Fix any $T<t_{\max}$, take $C := 2(C_*+k_{04}T)e^{k_{01}T}$, and observe that, for $r\le\rho(\zeta,\mu)\wedge r^{(2)}_{\max}$, and such that $p(r)\le\rho(\zeta,\mu)$, we can take

$$C''_{rT} \le \widetilde{C}_{rT} := \{2(C_*\vee 1)+k_{r4}T\}e^{(k_{r1}+Ck_{r2})T}, \tag{4.32}$$

in Theorem 2.4, since we can take C_* to bound C_r and C'_r. In particular, $r=r(\zeta)$ as defined in Assumption 4.2 satisfies both the conditions on r for (4.32) to hold. Then, taking $a := \{k_2+k_1\widetilde{C}_{r(\zeta)T}\}^{b(\zeta)}$ in Corollary 2.5, it follows that for some constant $c_3>0$, on the event $B_N(a)$,

$$\mathbf{P}[\tau^{(N)}(a,\zeta)\le T]\le c_3N^{-1}.$$

Then, from (4.29), for some constant c_4, $\mathbf{P}[B_N(a)^c]\le c_4N^{-1}\log N$. Here, the constants c_3, c_4 depend on C_*, T and the parameters of the process.

We now use Lemma 4.6 to bound the martingale term in (4.10). It follows that, on the event $B_N(a)\cap\{\tau^{(N)}(a,\zeta)>T\}$ and on the event that $\|x_N(s)-x(s)\|_\mu\le\Xi_T$ for all $0\le s\le t$,

$$\|x_N(t)-x(t)\|_\mu \le \left(e^{wT}\|x_N(0)-x(0)\|_\mu + \sqrt{a}\,K_{4.6}\sqrt{\frac{\log N}{N}}\right) + k_*\int_0^t \|x_N(s)-x(s)\|_\mu\,ds,$$

where $k_* := e^{wT}K(\mu, F; 2\Xi_T)$. Then from Gronwall's inequality, on the event $B_N(a) \cap \{\tau^{(N)}(a, \zeta) > T\}$,

$$\|x_N(t) - x(t)\|_\mu \leq \left(e^{wT}\|x_N(0) - x(0)\|_\mu + \sqrt{a}\, K_{4.6}\sqrt{\frac{\log N}{N}} \right) e^{k^* t}, \qquad (4.33)$$

for all $0 \leq t \leq T$, provided that

$$\left(e^{wT}\|x_N(0) - x(0)\|_\mu + \sqrt{a}\, K_{4.6}\sqrt{\frac{\log N}{N}} \right) \leq \Xi_T e^{-k^* T}.$$

This is true for all N sufficiently large, if $\|x_N(0) - x(0)\|_\mu \leq \frac{1}{2}\Xi_T e^{-(w+k^*)T}$, which we have assumed. We have thus proved (4.31), since, as shown above, $\mathbf{P}(B_N(a)^c \cup \{\tau^{(N)}(a, \zeta) > T\}^c) = O(N^{-1}\log N)$.

We now use this to show that in fact $t_{\max} = \infty$. For $x(0)$ as above, we can take $x_N^j(0) := N^{-1}\lfloor Nx^j(0)\rfloor \leq x^j(0)$, so that $S^{(N)}_{\rho(\zeta,\mu)}(0) \leq NC_*$ for $C_* := S_{\rho(\zeta,\mu)}(x(0)) < \infty$. Then, by (4.13), $\lim_{j\to\infty}\{\mu(j)/\nu_{\rho(\zeta,\mu)}(j)\} = 0$, so it follows easily using bounded convergence that $\|x_N(0) - x(0)\|_\mu \to 0$ as $N \to \infty$. Hence, for any $T < t_{\max}$, it follows from (4.31) that $\|x_N(t) - x(t)\|_\mu \to_D 0$ as $N \to \infty$, for $t \leq T$, with uniform bounds over the interval, where '$\to_D$' denotes convergence in distribution. Also, by Assumption 4.2, there is a constant c_5 such that $\|x_N(t)\|_\mu \leq c_5 N^{-1} S^{(N)}_{r_\mu}(t)$ for each t, where $r_\mu \leq r^{(2)}_{\max}$ and $r_\mu \leq \rho(\zeta, \mu)$. Hence, using Lemma 2.3 and Theorem 2.4, $\sup_{0\leq t\leq 2T}\|x_N(t)\|_\mu$ remains bounded in probability as $N \to \infty$. Hence it is impossible that $\|x(t)\|_\mu \to \infty$ as $T \to t_{\max} < \infty$, implying that in fact $t_{\max} = \infty$ for such $x(0)$. □

Remark The dependence on the initial conditions is considerably complicated by the way the constant C appears in the exponent, for instance in the expression for $\widetilde{C}_{rT}$ in the proof of Theorem 4.7. However, if k_{r2} in Assumptions 2.1 can be chosen to be zero, as for instance in the examples below, the dependence simplifies correspondingly.

There are biologically plausible models in which the restriction to $J^l \geq -1$ is irksome. In populations in which members of a given type l can fight one another, a natural possibility is to have a transition $J = -2e^{(l)}$ at a rate proportional to $X^l(X^l - 1)$, which translates to $\alpha_J = \alpha_J^{(N)} = \gamma x^l(x^l - N^{-1})$, a function depending on N. Replacing this with $\alpha_J = \gamma(x^l)^2$ removes the N-dependence, but yields a process that can jump to negative values of X^l. For this reason, it is useful to be able to allow the transition rates α_J to depend on N.

Since the arguments in this paper are not limiting arguments for $N \to \infty$, it does not require many changes to derive the corresponding results. Quantities such as A, F, $U_r(x)$ and $V_r(x)$ now depend on N; however, Theorem 4.7 continues to hold with constants c_1 and c_2 that do not depend on N, provided that μ, w, ν, the k_{lm} from Assumption 2.1 and ζ from Assumption 4.2 can be chosen to be independent of N, and that the quantities $Z_*^{(l)}$ from (4.21) can be bounded uniformly in N. On the other hand, the solution $x = x^{(N)}$ of (4.2) that acts as approximation to x_N in

Theorem 4.7 now itself depends on N, through $R = R^{(N)}$ and $F = F^{(N)}$. If A (and hence R) can be taken to be independent of N, and $\lim_{N\to\infty} \|F^{(N)} - F\|_\mu = 0$ for some fixed μ–Lipschitz function F, a Gronwall argument can be used to derive a bound for the difference between $x^{(N)}$ and the (fixed) solution x to equation (4.2) with N-independent R and F. If A has to depend on N, the situation is more delicate.

5 Examples

We begin with some general remarks, to show that the assumptions are satisfied in many practical contexts. We then discuss two particular examples, those of Kretzschmar [7] and of Arrigoni [1], that fitted poorly or not at all into the general setting of Barbour and Luczak [3], though the other systems referred to in Sect. 1 could also be treated similarly. In both of our chosen examples, the index j represents a number of individuals—parasites in a host in the first, animals in a patch in the second—and we shall for now use the former terminology for the preliminary, general discussion.

Transitions that can typically be envisaged are: births of a few parasites, which may occur either in the same host, or in another, if infection is being represented; births and immigration of hosts, with or without parasites; migration of parasites between hosts; deaths of parasites; deaths of hosts; and treatment of hosts, leading to the deaths of many of the host's parasites. For births of parasites, there is a transition $X \to X + J$, where J takes the form

$$J_l = 1; \quad J_m = -1; \quad J_j = 0, \quad j \neq l,\ m, \tag{5.1}$$

indicating that one m-host has become an l-host. For births of parasites within a host, a transition rate of the form $b_{l-m} m X_m$ could be envisaged, with $l > m$, the interpretation being that there are X_m hosts with parasite burden m, each of which gives birth to s offspring at rate b_s, for some small values of s. For infection of an m-host, a possible transition rate would be of the form

$$X_m \sum_{j\geq 0} N^{-1} X_j \lambda p_{j,l-m},$$

since an m-host comes into contact with j-hosts at a rate proportional to their density in the host population, and p_{jr} represents the probability of a j-host transferring r parasites to the infected host during the contact. The probability distributions $p_{j\cdot}$ can be expected to be stochastically increasing in j. Deaths of parasites also give rise to transitions of the form (5.1), but now with $l < m$, the simplest form of rate being just dmX_m for $l = m - 1$, though $d = d_m$ could also be chosen to increase with parasite burden. Treatment of a host would lead to values of l much smaller than m, and a rate of the form κX_m for the transition with $l = 0$ would represent fully successful treatment of randomly chosen individuals. Births and deaths of hosts and immigration all lead to transitions of the form

$$J_l = \pm 1; \quad J_j = 0, \quad j \neq l. \tag{5.2}$$

For deaths, $J_l = -1$, and a typical rate would be $d'X_l$. For births, $J_l = 1$, and a possible rate would be $\sum_{j\geq 0} X_j b'_{jl}$ (with $l = 0$ only, if new-born individuals are free of parasites). For immigration, constant rates λ_l could be supposed. Finally, for migration of individual parasites between hosts, transitions are of the form

$$J_l = J_m = -1; \quad J_{l+1} = 1; \quad J_{m-1} = 1; \quad J_j = 0, \quad j \neq l, m, l+1, m-1, \tag{5.3}$$

a possible rate being $\gamma m X_m N^{-1} X_l$.

For all the above transitions, we can take $J_* = 2$ in (1.2), and (1.3) is satisfied in biologically sensible models. (3.1) and (3.2) depend on the way in which the matrix A can be defined, which is more model specific; in practice, (3.1) is very simple to check. The choice of μ in (3.2) is influenced by the need to have (4.1) satisfied. For Assumptions 2.1, a possible choice of ν is to take $\nu(j) = (j+1)$ for each $j \geq 0$, with $S_1(X)$ then representing the number of hosts plus the number of parasites. Satisfying (2.5) is then easy for transitions only involving the movement of a single parasite, but in general requires assumptions as to the existence of the rth moments of the distributions of the numbers of parasites introduced at birth, immigration and infection events. For (2.6), in which transitions involving a net reduction in the total number of parasites and hosts can be disregarded, the parasite birth events are those in which the rates typically have a factor mX_m for transitions with $J_m = -1$, with m in principle unbounded. However, at such events, an m-individual changes to an $m+s$ individual, with the number s of offspring of the parasite being typically small, so that the value of $J^T \nu_r$ associated with this rate has magnitude m^{r-1}; the product $mX_m\, m^{r-1}$, when summed over m, then yields a contribution of magnitude $S_r(X)$, which is allowable in (2.6). Similar considerations show that the terms $N^{-1}S_0(X)S_r(X)$ accommodate the migration rates suggested above. Finally, in order to have Assumptions 4.2 satisfied, it is in practice necessary that Assumptions 2.1 are satisfied for large values of r, thereby imposing restrictions on the distributions of the numbers of parasites introduced at birth, immigration and infection events, as above.

5.1 Kretzschmar's model

Kretzschmar [7] introduced a model of a parasitic infection, in which the transitions from state X are as follows:

$$\begin{array}{lll}
J = e^{(i-1)} - e^{(i)} & \text{at rate} & Ni\mu x^i, \quad i \geq 1; \\
J = -e^{(i)} & \text{at rate} & N(\kappa + i\alpha)x^i, \quad i \geq 0; \\
J = e^{(0)} & \text{at rate} & N\beta \sum_{i\geq 0} x^i \theta^i; \\
J = e^{(i+1)} - e^{(i)} & \text{at rate} & N\lambda x^i \varphi(x), \quad i \geq 0,
\end{array}$$

where $x := N^{-1}X$, $\varphi(x) := \|x\|_{11}\{c + \|x\|_1\}^{-1}$ with $c > 0$, and $\|x\|_{11} := \sum_{j\geq 1} j|x|^j$; here, $0 \leq \theta \leq 1$, and θ^i denotes its ith power (our θ corresponds to the constant ξ in [7]). Both (1.2) and (1.3) are obviously satisfied. For Assumptions

(3.1), (3.2) and (4.1), we note that equation corresponding to (1.5) has

$$\begin{aligned} A_{ii} &= -\{\kappa + i(\alpha+\mu)\}; \quad A^T_{i,i-1} = i\mu \quad \text{and} \quad A^T_{i0} = \beta\theta^i, \quad i \geq 2; \\ A_{11} &= -\{\kappa + \alpha + \mu\}; \quad A^T_{10} = \mu + \beta\theta; \\ A_{00} &= -\kappa + \beta, \quad i \geq 1, \end{aligned}$$

with all other elements of the matrix equal to zero, and

$$F^i(x) = \lambda(x^{i-1} - x^i)\varphi(x), \quad i \geq 1; \quad F^0(x) = -\lambda x^0 \varphi(x).$$

Hence Assumption (3.1) is immediate, and Assumption (3.2) holds for $\mu(j) = (j+1)^s$, for any $s \geq 0$, with $w = (\beta - \kappa)_+$. For the choice $\mu(j) = j + 1$, F maps elements of $\mathcal{R}_\mu$ to $\mathcal{R}_\mu$, and is also locally Lipschitz in the μ-norm, with $K(\mu, F; \Xi) = c^{-2}\lambda\Xi(2c + \Xi)$.

For Assumptions 2.1, choose $\nu = \mu$; then (2.5) is a finite sum for each $r \geq 0$. Turning to (2.6), it is immediate that $U_0(x) \leq \beta S_0(x)$. Then, for $r \geq 1$,

$$\begin{aligned} \sum_{i\geq 0} \lambda\varphi(N^{-1}X)X^i\{(i+2)^r - (i+1)^r\} &\leq \lambda \frac{S_1(X)}{S_0(X)} \sum_{i\geq 0} rX^i(i+2)^{r-1} \\ &\leq r2^{r-1}\lambda S_r(X), \end{aligned}$$

since, by Jensen's inequality, $S_1(X)S_{r-1}(X) \leq S_0(X)S_r(X)$. Hence we can take $k_{r2} = k_{r4} = 0$ and $k_{r1} = \beta + r2^{r-1}\lambda$ in (2.6), for any $r \geq 1$, so that $r^{(1)}_{\max} = \infty$. Finally, for (2.7),

$$V_0(x) \leq (\kappa + \beta)S_0(x) + \alpha S_1(x),$$

so that $k_{03} = \kappa + \beta + \alpha$ and $k_{05} = 0$, and

$$V_r(x) \leq r^2(\kappa S_{2r}(x) + \alpha S_{2r+1}(x) + \mu S_{2r-1}(x) + 2^{2(r-1)}\lambda S_{2r-1}(x)) + \beta S_0(x),$$

so that we can take $p(r) = 2r + 1$, $k_{r3} = \beta + r^2\{\kappa + \alpha + \mu + 2^{2(r-1)}\lambda\}$, and $k_{r5} = 0$ for any $r \geq 1$, and so $r^{(2)}_{\max} = \infty$. In Assumptions 4.2, we can clearly take $r_\mu = 1$ and $\zeta(k) = (k+1)^7$, giving $r(\zeta) = 8$, $b(\zeta) = 1$ and $\rho(\zeta, \mu) = 17$.

5.2 Arrigoni's model

In the metapopulation model of Arrigoni [1], the transitions from state X are as follows:

$$\begin{array}{ll} J = e^{(i-1)} - e^{(i)} & \text{at rate } Nix^i(d_i + \gamma(1-\rho)), \quad i \geq 2; \\ J = e^{(0)} - e^{(1)} & \text{at rate } Nx^1(d_1 + \gamma(1-\rho) + \kappa); \\ J = e^{(i+1)} - e^{(i)} & \text{at rate } Nib_i x^i, \quad i \geq 1; \\ J = e^{(0)} - e^{(i)} & \text{at rate } Nx^i\kappa, \quad i \geq 2; \\ J = e^{(k+1)} - e^{(k)} + e^{(i-1)} - e^{(i)} & \text{at rate } Nix^i x^k \rho\gamma, \quad k \geq 0, \quad i \geq 1; \end{array}$$

as before, $x := N^{-1}X$. Here, the total number $N = \sum_{j\geq 0} X_j = S_0(X)$ of patches remains constant throughout, and the number of animals in any one patch changes by at most one at each transition; in the final (migration) transition, however, the numbers in two patches change simultaneously. In the above transitions, γ, ρ, κ are non-negative, and (d_i), (b_i) are sequences of non-negative numbers.

Once again, both (1.2) and (1.3) are obviously satisfied. The equation corresponding to (1.4) can now be expressed by taking

$$A_{ii} = -\{\kappa + i(b_i + d_i + \gamma)\}; \quad A^T_{i,i-1} = i(d_i + \gamma); \quad A^T_{i,i+1} = ib_i, \quad i \geq 1;$$
$$A_{00} = -\kappa,$$

with all other elements of A equal to zero, and

$$F^i(x) = \rho\gamma\|x\|_{11}(x^{i-1} - x^i), \quad i \geq 1; \qquad F^0(x) = -\rho\gamma x^0\|x\|_{11} + \kappa,$$

where we have used the fact that $N^{-1}\sum_{j\geq 0} X_j = 1$. Hence Assumption (3.1) is again immediate, and Assumption (3.2) holds for $\mu(j) = 1$ with $w = 0$, for $\mu(j) = j + 1$ with $w = \max_i(b_i - d_i - \gamma - \kappa)_+$ (assuming (b_i) and (d_i) to be such that this is finite), or indeed for $\mu(j) = (j+1)^s$ with any $s \geq 2$, with appropriate choice of w. With the choice $\mu(j) = j + 1$, F again maps elements of $\mathcal{R}_\mu$ to $\mathcal{R}_\mu$, and is also locally Lipschitz in the μ-norm, with $K(\mu, F; \Xi) = 3\rho\gamma\,\Xi$.

To check Assumptions 2.1, take $\nu = \mu$; once again, (2.5) is a finite sum for each r. Then, for (2.6), it is immediate that $U_0(x) = 0$. For any $r \geq 1$, using arguments from the previous example,

$$\begin{aligned} U_r(x) &\leq r2^{r-1}\left\{\sum_{i\geq 1} ib_i x^i(i+1)^{r-1} + \sum_{i\geq 1}\sum_{k\geq 0} i\rho\gamma x^i x^k(k+1)^{r-1}\right\} \\ &\leq r2^{r-1}\{\max_i b_i\, S_r(x) + \rho\gamma S_1(x)S_{r-1}(x)\} \\ &\leq r2^{r-1}\{\max_i b_i\, S_r(x) + \rho\gamma S_0(x)S_r(x)\}, \end{aligned}$$

so that, since $S_0(x) = 1$, we can take $k_{r1} = r2^{r-1}(\max_i b_i + \rho\gamma)$ and $k_{r2} = k_{r4} = 0$ in (2.6), and $r^{(1)}_{\max} = \infty$. Finally, for (2.7), $V_0(x) = 0$ and, for $r \geq 1$,

$$\begin{aligned} V_r(x) \leq r^2 \Big\{ & 2^{2(r-1)}\max_i b_i\, S_{2r-1}(x) + \max_i(i^{-1}d_i)S_{2r}(x) + \gamma(1-\rho)S_{2r-1}(x) \\ & + \rho\gamma(2^{2(r-1)}S_1(x)S_{2r-2}(x) + S_0(x)S_{2r-1}(x))\Big\} + \kappa S_{2r}(x), \end{aligned}$$

so that we can take $p(r) = 2r$, and (assuming $i^{-1}d_i$ to be finite)

$$k_{r3} = \kappa + r^2\{2^{2(r-1)}(\max_i b_i + \rho\gamma) + \max_i(i^{-1}d_i) + \gamma\},$$

and $k_{r5} = 0$ for any $r \geq 1$, and $r^{(2)}_{\max} = \infty$. In Assumptions 4.2, we can again take $r_\mu = 1$ and $\zeta(k) = (k+1)^7$, giving $r(\zeta) = 8$, $b(\zeta) = 1$ and $\rho(\zeta, \mu) = 16$.

Acknowledgments We wish to thank a referee for recommendations that have substantially streamlined our arguments. ADB wishes to thank both the Institute for Mathematical Sciences of the National University of Singapore and the Mittag–Leffler Institute for providing a welcoming environment while part of this work was accomplished. MJL thanks the University of Zürich for their hospitality on a number of visits.

References

1. Arrigoni, F.: Deterministic approximation of a stochastic metapopulation model. Adv. Appl. Prob. **35**, 691–720 (2003)
2. Barbour, A.D., Kafetzaki, M.: A host–parasite model yielding heterogeneous parasite loads. J. Math. Biol. **31**, 157–176 (1993)
3. Barbour, A.D., Luczak, M.J.: Laws of large numbers for epidemic models with countably many types. Ann. Appl. Probab. **18**, 2208–2238 (2008)
4. Chow, P.-L.: Stochastic Partial Differential Equations. Chapman and Hall, Boca Raton (2007)
5. Eibeck, A., Wagner, W.: Stochastic interacting particle systems and non-linear kinetic equations. Ann. Appl. Probab. **13**, 845–889 (2003)
6. Kimmel, M., Axelrod, D.E.: Branching Processes in Biology. Springer, Berlin (2002)
7. Kretzschmar, M.: Comparison of an infinite dimensional model for parasitic diseases with a related 2-dimensional system. J. Math. Anal. Appl. **176**, 235–260 (1993)
8. Kurtz, T.G.: Solutions of ordinary differential equations as limits of pure jump Markov processes. J. Appl. Probab. **7**, 49–58 (1970)
9. Kurtz, T.G.: Limit theorems for sequences of jump Markov processes approximating ordinary differential processes. J. Appl. Probab. **8**, 344–356 (1971)
10. Léonard, C.: Some epidemic systems are long range interacting particle systems. In: Gabriel, J.-P., Lefèvre, C., Picard, P. (eds.) Stochastic Processes in Epidemic Theory, Lecture Notes in Biomathematics, vol. 86, pp. 170–183. Springer, New York (1990)
11. Luchsinger, C.J.: Mathematical Models of a Parasitic Disease. Ph.D. thesis, University of Zürich (1999)
12. Luchsinger, C.J.: Stochastic models of a parasitic infection, exhibiting three basic reproduction ratios. J. Math. Biol. **42**, 532–554 (2001)
13. Luchsinger, C.J.: Approximating the long term behaviour of a model for parasitic infection. J. Math. Biol. **42**, 555–581 (2001)
14. Pazy, A.: Semigroups of Linear Operators and Applications to Partial Differential Equations. Springer, Berlin (1983)
15. Reuter, G.E.H.: Denumerable Markov processes and the associated contraction semigroups on l. Acta Math. **97**, 1–46 (1957)