Research article

How fast is the linear chain trick? A rigorous analysis in the context of behavioral epidemiology

Alessia Andò, Dimitri Breda* **and Giulia Gava**

Computational Dynamics Laboratory, Department of Mathematics, Computer Science and Physics University of Udine, via delle scienze 206, 33100 Udine, Italy

* **Correspondence:** Email: dimitri.breda@uniud.it; Tel: +390432558756.

Abstract: A prototype SIR model with vaccination at birth is analyzed in terms of the stability of its endemic equilibrium. The information available on the disease influences the parents' decision on whether vaccinate or not. This information is modeled with a delay according to the Erlang distribution. The latter includes the degenerate case of fading memory as well as the limiting case of concentrated memory. The linear chain trick is the essential tool used to investigate the general case. Besides its novel analysis and that of the concentrated case, it is showed that through the linear chain trick a distributed delay approaches a discrete delay at a linear rate. A rigorous proof is given in terms of the eigenvalues of the associated linearized problems and extension to general models is also provided. The work is completed with several computations and relevant experimental results.

Keywords: linear chain trick; stability analysis; distributed delay; SIR models with vaccination; behavioral epidemiology

1. Introduction

It is widely known that delay equations originate dynamical systems of infinite dimension [1]. Their treatment in view of stability and bifurcation analyses thus poses several challenges. Then it is not a surprise that the issue still attracts the attention of many researchers after several decades of efforts. A major direction to attack the problem is undoubtedly represented by the reduction to finite dimension, either exactly or through approximation.

The recent work [2] represents an elegant and up-to-date reference for the first option, also useful as a starting point for the relevant branch of literature. In particular, it shows up as a modern account of the possibility of reducing a delay equation to an equivalent finite-dimensional system of Ordinary Differential Equations (ODEs). This has been well known since long for the case of delays modeled according to the Erlang distribution. In particular, the corresponding methodology originated from the

pioneering work [3] in queueing theory and nowadays, in the context of delays, it undergoes the name of *linear chain trick*. The term has been probably proposed by MacDonald [4, 5], even though the technique dates back to the late sixties [6–8], at least as far as fields like epidemiology, population dynamics and evolutionary systems in general are concerned.

When an exact reduction is not available, the study of relevant approximations through ODEs becomes central. In this sense, Breda, et al. [9] proposes a general numerical approach to address a broad class of functional problems. Nevertheless, in the early eighties the subject was already alive. Busenberg and Travis [10] and Cooke and Grossman [11] represent important contributions from those years, with particular reference to biological models, where indeed delays emerge naturally. Therein, the linear chain trick is investigated in view of analyzing the stability of steady states (in the first) or as a way to gain insight into the pros and cons of using discrete or distributed delays in modeling (in the second).

Essentially, and on the one hand, the linear chain trick is a smart way to construct a system of ODEs equivalent to a problem with a delay modeled according to the Erlang distribution*. The number of ODEs corresponds to the (integer) shape parameter of the latter. As this number grows, the variance of the distribution decreases and the delay values tend to concentrate around their mean value. If we let this number grow unbounded, in the limit we obtain the Dirac's delta distribution, corresponding to a problem with a single discrete delay. Thus, on the other hand, the linear chain trick can be interpreted as a mean to approach a fixed delay through a sequence of distributed delays. Under this point of view, it is natural to ask how the relevant stability properties behave, a question already central in [10, 11], but still contemporary, see, e.g., [12] or [13] concerning the computation of the basic reproduction number.

Eventually, despite the evident interest that the literature devoted to the argument, there has been no rigorous study on the rate at which the linear chain trick makes the sequence of distributed delays converge to the limit of concentrated delay, together with the associated stability features. It is anyway worth recalling from page 169 in [5] that a general degree should be kept in mind when one applies the linear chain trick, this degree being indeed the number of ODEs used to "decompose" the delay. Let us point out that some effort in this direction is reported in [14, 15] on investigating experimentally the effect of increasing the degree of approximation. As a general outcome, it is observed that on the way towards a fixed delay the chances of instability grow, mainly due to sustained oscillations.

In the works just mentioned the delay represents the infectious period of childhood viral diseases. The author addresses the study of more realistic distributions of the latter (viz. Erlang with larger shape parameter) by using the linear chain trick, therein named *method of stages* after the original work [3]. The biological field of infectious diseases and epidemics is a major area of investigation. Recently it gained even more attention, also outside the scientific community, due to the impact of vaccination strategies in modern societies as opposed to the emerging of often clamoring campaigns of vaccine hesitancy. It is not surprising then that the discipline of *behavioral epidemiology* arose lively, with the broad target of studying the effect of the "individual behavior as a key determinant of infection trajectories" [16]. Besides the two cornerstones [17, 18] and many other contributions (see, e.g., [19–26]), d'Onofrio, et al. [27] is a thorough representative of this new branch of mathematical biology. In it, the authors study the effect of the so called *rational exemption*, the phenomenon by

---

*Note that the linear chain trick works as the Erlang distribution is in fact a sum of independent exponential distributions.

which parents relate the decision to whether vaccinate or not their children to the current as well as past information available on the disease. They propose to model the information variable through the Erlang distribution, even though in the end only the degenerate exponential case of fading memory is elaborated. The main focus is on the stability of the endemic equilibrium and the appearance of stable oscillations through Hopf bifurcations is demonstrated. We mention also [28], which is seemingly the first work in the field about distributed delays, and [29] on Hopf bifurcations in structured populations.

In view of the above and taking inspiration from [27], the current work chases two objectives.

On the one side, we aim at extending the analysis of stability of the endemic equilibrium of the SIR model studied in [27], investigating both the general Erlang case with varying shape parameter, as well as the limiting case of concentrated memory. The latter alone already represents a novel contribution in that the relevant analysis makes use of modern numerical methods for the stability analysis of delay systems, e.g., [30–32]. As such, it is our hope that this first attempt will open the door to further extensions towards more realistic models including even different (type and number of) delays.

On the other side, as the linear chain trick is used to deal with the general Erlang distribution, we perform a rigorous study of how the latter approaches the concentrated case by increasing its shape parameter. In particular, we analyze the behavior of the rightmost eigenvalues of the relevant variational problems, eventually proving that those of the Erlang case converge to those of the concentrated case at a *linear* rate in terms of the number of ODEs used to reduce the distributed delay. To the best of our knowledge, this question has never been addressed in the literature.

In addition to the contributions described above, our results confirm that when the delay distribution presents a lower variance, systems are more prone to instability. This instability shows up as a series of stability switches through couples of destabilizing/stabilizing Hopf bifurcations. In particular, as a clear outcome, we obtain that the number of these switches increases from a single one in the case of exponential distribution to possibly infinitely-many in the case of concentrated memory. We also give a rigorous argument justifying the latter result based on [33].

Finally, we show that the proof of linear convergence mentioned above can be extended to models described by a general system of ODEs in which there is a dependence on an "information" variable modeled through a distributed delay. Therefore, the applicability of the proposed analysis goes far beyond the SIR model used here as a prototype.

In the sequel we present in section 2 the SIR model just mentioned, analyzing the case of fading memory in section 2.1, the case of concentrated memory in section 2.2 and the general Erlang distribution in section 2.3. In section 2.4 we collect the outcome of the experimental analysis in terms of stability charts. The convergence of the linear chain trick in terms of eigenvalues is proved in section 3 for the prototype SIR model, and then it is extended in section 4. Some concluding remarks close the work in section 5.

2. Prototype model and kernels

Following [27], let us consider the SIR model for a non-fatal pediatric disease in a constant homogeneously mixing population with vaccination at birth

$$\begin{cases} S'(t) = \mu[1 - p(M(t))] - \mu S(t) - \beta S(t)I(t) \\ I'(t) = \beta S(t)I(t) - (\mu + \nu)I(t) \\ R'(t) = \nu I(t) - \mu R(t) \\ V'(t) = \mu p(M(t)) - \mu V(t), \end{cases}$$

where μ represents both the constant fertility and mortality rates, β the constant contact rate, ν the constant recovery rate and $p : [0, +\infty) \to \mathbb{R}$ the vaccination coverage function. The latter gives the fraction of vaccinated newborns, which depends on the information about the disease available to their parents. This information is described by the map $M : [0, +\infty) \to \mathbb{R}$ given by

$$M(t) := \int_0^{+\infty} g(S(t - \theta), I(t - \theta))K(\theta) \, d\theta, \tag{2.1}$$

which considers the past values of the disease prevalence $g : [0, +\infty) \times [0, +\infty) \to \mathbb{R}$ as weighted through a memory kernel $K : [0, +\infty) \to \mathbb{R}$, non-negative and normalized as

$$\int_0^{+\infty} K(\theta) d\theta = 1. \tag{2.2}$$

The coverage function p is assumed to be piecewise smooth and non-decreasing in M, so that the decision to vaccinate is directly correlated to the disease prevalence. The latter is a smooth function of both susceptible and infected individuals, assumed to be increasing in the second. Moreover, $g(S, 0) = 0$ independently of the values of S: In absence of infection there is indeed no information. Note that S, I, R and V represent fractions with respect to the total (constant) population.

The equations for R and V are linear inhomogeneous ODEs, so that their solution is straightforward through the variation of constants formula once I and M are given. Therefore, in view of investigating the relevant dynamics, it is enough to consider the reduced model

$$\begin{cases} S'(t) = \mu[1 - p(M(t))] - \mu S(t) - \beta S(t)I(t) \\ I'(t) = \beta S(t)I(t) - (\mu + \nu)I(t) \\ M(t) = \int_0^{+\infty} g(S(t - \theta), I(t - \theta))K(\theta) \, d\theta, \end{cases} \tag{2.3}$$

where we directly add (2.1) to the equations for S and I.

It is not difficult to see that system (2.3) presents a unique endemic equilibrium $(\overline{S}, \overline{I}, \overline{M})$ with $\overline{S} = (\mu + \nu)/\beta$, $\overline{I}$ the solution of

$$\mu[1 - p(g(\overline{S}, I))] = (\mu + \beta I)\overline{S} \tag{2.4}$$

and $\overline{M} = g(\overline{S}, \overline{I})$. Note that $\overline{I}$ is unique since in Eq (2.4) the left-hand side is decreasing in I while the right-hand side is increasing. Moreover, and more important, the equilibrium is independent of the choice of the memory kernel K, at least as long as Eq (2.2) holds. The same cannot be said for its local asymptotic stability and, indeed, the focus of this work is at investigating how the choice of the memory kernel influences the latter.

To this aim, we first analyze two prototype choices, i.e., the case of *fading* memory

$$K(\theta) = ae^{-a\theta}, \tag{2.5}$$

where $a > 0$ plays the role of decaying rate, and the case of *concentrated* memory

$$K(\theta) = \delta(\theta - \tau), \tag{2.6}$$

where δ is the Dirac's delta at $\tau > 0$. We call *delay* the latter, observing that it represents a fixed time instant in the past where all the information about the disease is assumed to be concentrated.

The two cases above can be seen as the extrema of a spectrum of choices, in which the information is distributed on $[0, +\infty)$ around a mean value with a certain variance. In this rispect, we take into consideration also the kernel

$$K(\theta) = \mathrm{Erl}_{n,a}(\theta), \tag{2.7}$$

where

$$\mathrm{Erl}_{n,a}(\theta) := \frac{a^n \theta^{n-1}}{(n-1)!} e^{-a\theta}, \tag{2.8}$$

is the Erlang distribution, i.e., a Gamma distribution with integer shape parameter $n \in \mathbb{N} \setminus \{0\}$ and rate parameter $a > 0$. In particular, $\mathrm{Erl}_{n,a}$ has mean value n/a and variance n/a^2, so that by fixing

$$a = \frac{n}{\tau}, \tag{2.9}$$

we end up with a mean value τ and variance τ^2/n. As anticipated, by letting $n = 1$ we recover exactly the case of fading memory Eq (2.5), while by letting $n \to \infty$ we obtain the case of concentrated memory Eq (2.6), see Figure 1.

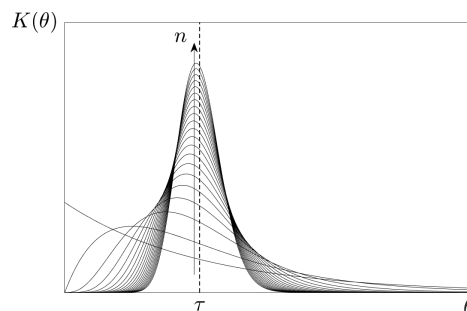**Figure 1.** Erlang distribution (2.8) according to Eq (2.9) for fixed τ and increasing n (in this plot $\tau = 10$ and $n = 1, 2, 4, 6, \ldots, 40$).

As far as the coverage function is concerned, in the sequel we initially fix

$$p(M) := p_0 + p_1(M) \tag{2.10}$$

with the specific choice

$$p_1(M) = \min\{cM, 1 - p_0\}. \tag{2.11}$$

In (2.10) the constant $p_0 \in (0, 1)$ represents the fixed fraction of population that vaccinate their children independently of the available information M, while the function $p_1 : [0, +\infty) \to \mathbb{R}$ portrays the concept of *rational exemption* (vaccinate less/more with more/less immune individuals) page 304, note 1 in [27]. With (2.11) p is piecewise smooth with slope $c > 0$ in the first part and it reaches the saturation value $1 - p_0$ for larger values of M. Other choices are considered in section 2.4.

Similarly, as for the disease prevalence, we perform most of the analysis by choosing

$$g(S, I) = kI \tag{2.12}$$

for k some positive constant[†].

In the following sections, keeping the choices above, we investigate the local asymptotic stability of the endemic equilibrium $(\overline{S}, \overline{I}, \overline{M})$ for varying τ (and possibly c) with either Eq (2.5) in section 2.1, Eq (2.6) in section 2.2 and Eq (2.7) in section 2.3.

Remark 2.1. *Condition* (2.9) *is implicitly assumed in the sequel, even when not explicitly used.*

2.1. Fading memory

Let us consider the choice (2.5) and rewrite the third of system (2.3) as

$$M(t) = ae^{-at} \int_{-\infty}^{t} g(S(\sigma), I(\sigma))e^{a\sigma} \, d\sigma.$$

By differentiating the latter system (2.3) reduces to the system of 3 ODEs

$$\begin{cases} S'(t) = \mu[1 - p(M(t))] - \mu S(t) - \beta S(t)I(t) \\ I'(t) = \beta S(t)I(t) - (\mu + \nu)I(t) \\ M'(t) = ag(S(t), I(t)) - aM(t). \end{cases}$$

The local asymptotic stability of the endemic equilibrium $(\overline{S}, \overline{I}, \overline{M})$ is analyzed in [27] by resorting to the Routh-Hurwitz criterion as applied to the matrix associated to the relevant linearization, i.e.,

$$A_1 := \begin{pmatrix} -\mu - \beta\overline{I} & -\beta\overline{S} & -\mu p'(\overline{M}) \\ \beta\overline{I} & \beta\overline{S} - (\mu + \nu) & 0 \\ ag_S(\overline{S}, \overline{I}) & ag_I(\overline{S}, \overline{I}) & -a \end{pmatrix} \in \mathbb{R}^{3\times3}, \tag{2.13}$$

where g_S and g_I denote the partial derivatives of g with respect to S and I, respectively. Here, instead, we compute the eigenvalues of matrix (2.13) in view of generalizing this approach in the forthcoming sections: A rightmost eigenvalue with negative real part guarantees asymptotic stability, otherwise any eigenvalue in the right-half of $\mathbb{C}$ determines instability.

Figure 2 reproduces in this sense the same results of Figure 1 in [27] for the choices (2.10), (2.11) and (2.12), and for the values of the concerned parameters as in Table 1 (which depict measles in a low-mortality population [27]). Let us highlight that for sufficiently large c at most one instability interval occurs in the domain of τ (recall Eq (2.9)). This is originated by a couple of Hopf bifurcations, the first destabilizing and the other stabilizing, Figure 3. In the latter, the eigenvalues are computed by standard routines for matrix eigenvalues (specifically, `eig` in Matlab) and the bifurcation values for the delay τ are found by standard nonlinear solvers (specifically, `fzero` in Matlab).

[†]This choice is motivated by the aim of reproducing the results of [27]. Note, however, that the general setting discussed in section 4 includes also the choice $g(S, I) = kSI$, as well as more general ones

Table 1. Values of the parameters used in the simulations.

Parameter	Value
fertility/mortality rates [1/days]	$\mu = 1/(75 \cdot 365)$
contact rate [1/days]	$\beta = 1.4289$
recovery rate [1/days]	$\nu = 1/7$
coverage baseline	$p_0 = 0.75$
coverage slope	c free bifurcation parameter
prevalence constant	$k = 1$
delay [days]	τ free bifurcation parameter
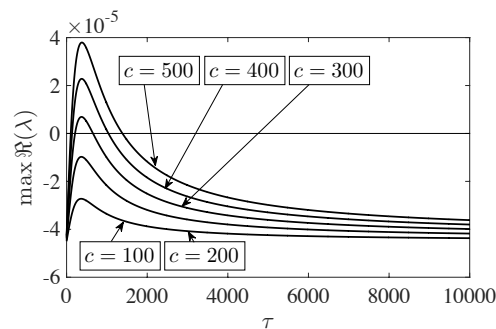
Figure 2. Real part of the rightmost eigenvalue(s) of matrix (2.13) for varying τ and several values of c.
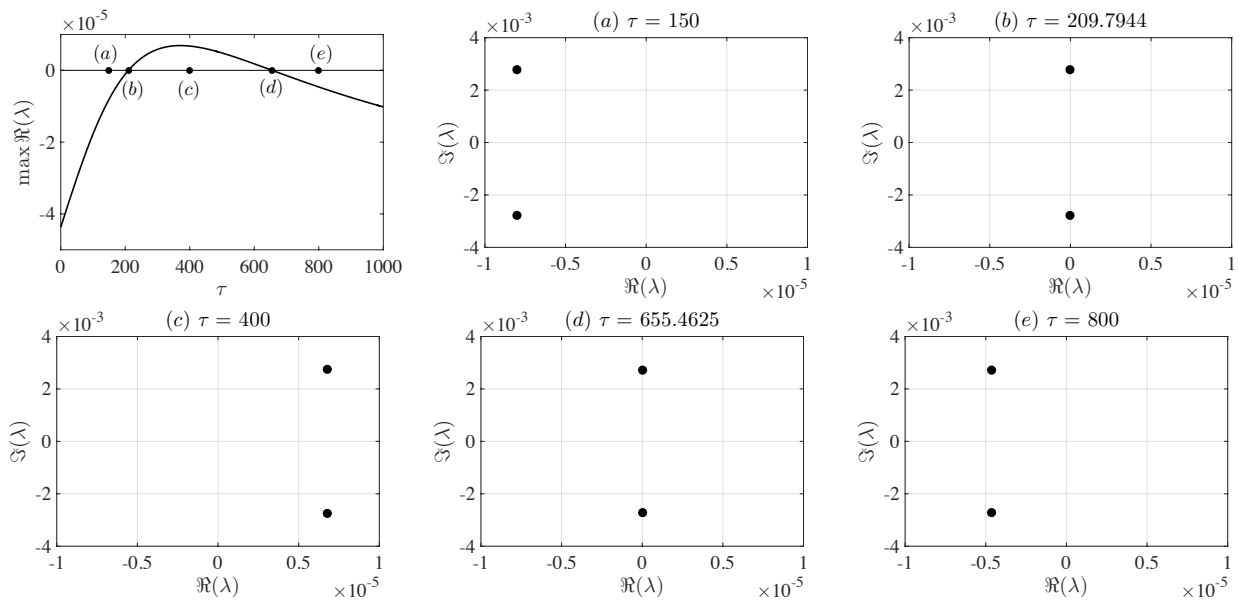
Figure 3. Real part of the rightmost eigenvalue(s) of matrix (2.13) for varying τ and $c = 300$ (top-left panel); eigenvalues corresponding to the selected values of τ in the top-left panel (remaining panels).

2.2. Concentrated memory

Let us consider now the choice Eq (2.6), by which model (2.3) reduces to

$$\begin{cases} S'(t) = \mu[1 - p(M(t))] - \mu S(t) - \beta S(t)I(t) \\ I'(t) = \beta S(t)I(t) - (\mu + \nu)I(t) \\ M(t) = g(S(t - \tau), I(t - \tau)). \end{cases} \qquad (2.14)$$

System (2.14) is made of 2 ODEs coupled to an algebraic equation with delay. This makes the corresponding dynamical system infinite-dimensional, with state space a set of functions defined on the delay interval $[-\tau, 0]$. For this and the following related issues we refer to [34]. Alternatively, one can substitute the third equation for M into the first one, obtaining a more conventional DDE for S coupled to the ODE for I, for which we refer to [32] as far as the following is concerned. Anyway we feel free to use one formulation or the other depending on the convenience from time to time.

The endemic equilibrium $(\overline{S}, \overline{I}, \overline{M})$ of system (2.14) remains unchanged, although the analysis of its local asymptotic stability is far less trivial due to the infinite dimension. The principle of linearized stability still holding, the problem reduces to determine the spectrum of suitable linear operators associated to the linearized system. These are either the semigroup of solution operators or its infinitesimal generator. Of course, numerical approximation is needed, as they are operators on Banach spaces of functions. In this regard, we rely on the pseudospectral methods in either [35] or [32] depending on the chosen formulation, both furnishing accurate approximations of the rightmost eigenvalue, which is the stability determining one. The results are illustrated in Figure 4, which is the analogous of Figure 2 but for Eq (2.6) replacing Eq (2.5). Now infinitely-many intervals of instability can appear, always due to Hopf bifurcations, Figure 5 (the analogous of Figure 3). A rigorous argument behind this claim is given next.
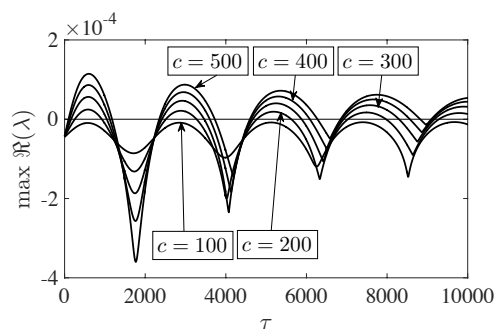

Figure 4. Real part of the rightmost eigenvalue(s) relevant to system (2.14) for varying τ and several values of c.

Let us resume from [33] the concepts of *kernel* and *offspring* sets, introduced therein for linear autonomous systems with delay. Let us stress that this *kernel* has nothing to deal with the memory kernel: To avoid confusion we use italic fonts when dealing with the former. We anticipate from Eq (3.8) in section 3 that the characteristic equation associated to the linearized version of system (2.14) contains the complex exponential $e^{-\lambda \tau}$, a well-known fact indeed for systems with a discrete delay τ. Nontrivial conjugated roots on the imaginary axis, say $\lambda = \pm \beta i$, make the exponential periodic in $\beta \tau$.
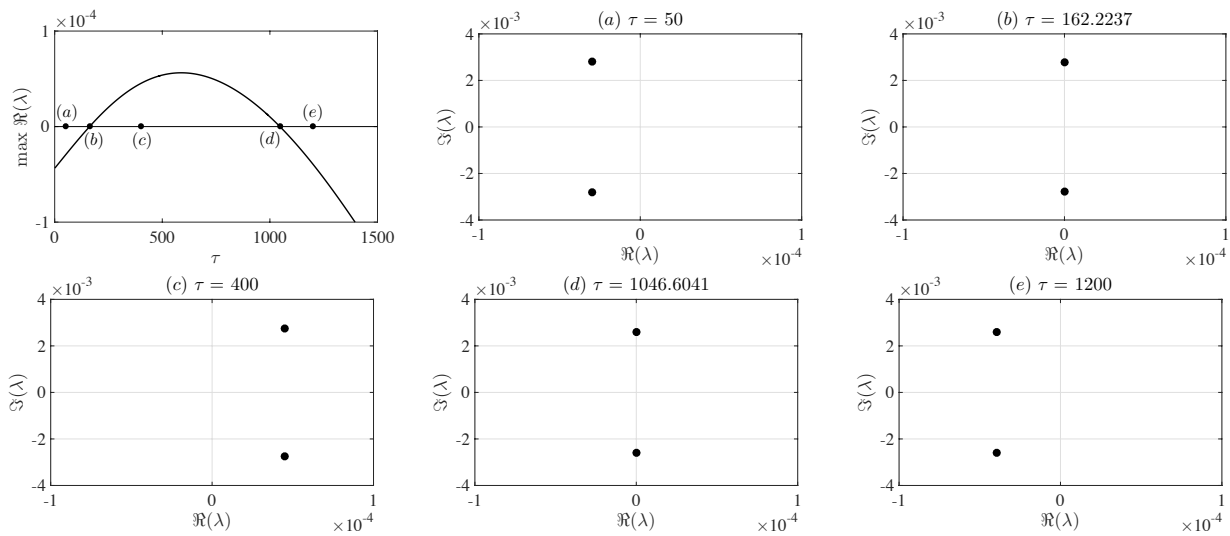
Figure 5. Real part of the rightmost eigenvalue(s) relevant to system (2.14) for varying τ and $c = 300$ (top-left panel); eigenvalues corresponding to the selected values of τ in the top-left panel (remaining panels).

Given β, let us define the *kernel* set

$$\mathcal{K}(\beta) := \{\tau > 0 \; : \; \beta\tau \in (0, 2\pi)\}.$$

This in turn generates the *offspring* set

$$O(\beta) := \{\tau + 2\pi j/\beta \; : \; \tau \in \mathcal{K}(\beta), \; j = 1, 2, \ldots\} \,.$$

Therefore any value in the *kernel* set gives rise to a periodic sequence of delay values in the relevant *offspring* set. Note however that one has to compute only the *kernel* set, since the *offspring* set follows automatically. In the specific case (2.14), for any given value of c over a certain threshold, there are just two values of β giving rise to Hopf bifurcations, one destabilizing and one stabilizing. Correspondingly we can find the relevant *kernel* sets, each of which consists of only a single delay value. As a convention, in the sequel we use the index 1 for the delay value τ_1 originating the destabilizing Hopf bifurcation and the index 2 for the delay value τ_2 originating the stabilizing Hopf bifurcation. These two values of τ correspond to the first two crossings of the horizontal axis in Figure 4. All the other crossings are obtained as *offspring* sets, and are indeed periodically distributed along the delay axis with period $2\pi/\beta$. Let us observe anyway that the instability intervals created by the two sequences of τ values may merge or disappear at some points, given that the period $2\pi/\beta$ changes with β. Eventually, we remark that this theoretical argument is validated experimentally with rather high accuracy, see Table 2 and Figure 6. In Table 2 *exact* means computed as $\tau + 2\pi j/\beta$ once β and τ are computed for the *kernel* set.

Table 2. Kernel and offspring delay values for system (2.14) with $c = 300$.

delay	τ_1	
kernel	162.2237119784580	
offspring	exact	numerical
$j = 1$	2 415.766179177166	2 415.766179177342
$j = 2$	4 669.308646375874	4 669.308646376010
$j = 3$	6 922.851113574581	6 922.851113574772
delay	τ_2	
kernel	1 046.604140032942	
offspring	exact	numerical
$j = 1$	3 462.713668448635	3 462.713668448839
$j = 2$	5 878.823196864328	5 878.823196864712
$j = 3$	8 294.932725280021	8 294.932725280545
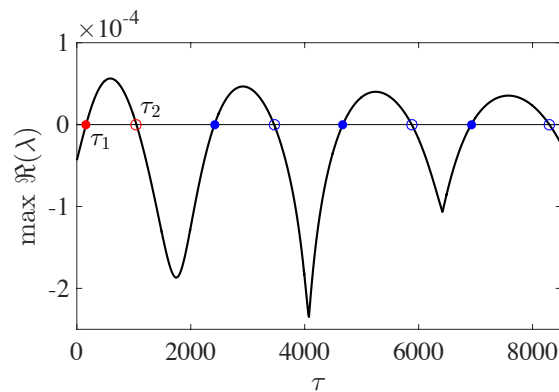
Figure 6. Kernel (red • and ○) and offspring (blue • and ○) delay values for system (2.14) with $c = 300$.

2.3. Erlang distribution and the linear chain trick

Let us consider now the choice Eq (2.7) and define the auxiliary variables

$$Z_i(t) := \int_0^{+\infty} g(S(t - \theta), I(t - \theta)) \mathrm{Erl}_{i,a}(\theta) d\theta, \quad i = 1, \ldots, n.$$

Clearly $Z_n(t) = M(t)$, while differentiation allows to transform system (2.3) into the system of $n + 2$ ODEs

$$\begin{cases} S'(t) = \mu[1 - p(Z_n(t))] - \mu S(t) - \beta S(t)I(t) \\ I'(t) = \beta S(t)I(t) - (\mu + v)I(t) \\ Z_1'(t) = ag(S(t), I(t)) - aZ_1(t) \\ Z_i'(t) = aZ_{i-1}(t) - aZ_i(t), \quad i = 2, \ldots, n. \end{cases} \tag{2.15}$$

This procedure is known as *linear chain trick*, as already anticipated and discussed in the Introduction.

The endemic equilibrium is now $(\overline{S}, \overline{I}, \overline{M}, \ldots, \overline{M})$: It differs just in the presence of the auxiliary components which, being all equal to the information compartment, do not bring any novelty in. By proceeding as in section 2.1, we compute the eigenvalues of the matrix of the linearized system, i.e.,

$$
A_n := \begin{pmatrix}
-\mu - \beta\overline{I} & -\beta\overline{S} & 0 & 0 & \cdots & -\mu p'(\overline{M}) \\
\beta\overline{I} & \beta\overline{S} - \mu - \nu & 0 & 0 & \cdots & 0 \\
ag_S(\overline{S}, \overline{I}) & ag_I(\overline{S}, \overline{I}) & -a & 0 & \cdots & 0 \\
0 & 0 & a & -a & \cdots & 0 \\
\vdots & \vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & 0 & a & -a
\end{pmatrix}
\tag{2.16}
$$

in $\mathbb{R}^{(n+2)\times(n+2)}$. Figure 7 shows the analogous results of Figures 2 and 4. It emerges clear that the curve representing the real part of the rightmost eigenvalue(s) of matrix (2.16) gets close to that relevant to the case of concentrated memory as n increases: Indeed, more and more instability intervals due to Hopf bifurcations appear in the direction of increasing τ as long as n grows. We omit to include the analogous of Figures 3 and 5 since qualitatively unchanged. We instead show in Figure 8 the values of τ for varying n corresponding respectively to the first (destabilizing) and the second (stabilizing) Hopf bifurcations. Some of these values are provided in Table 3 for reference.
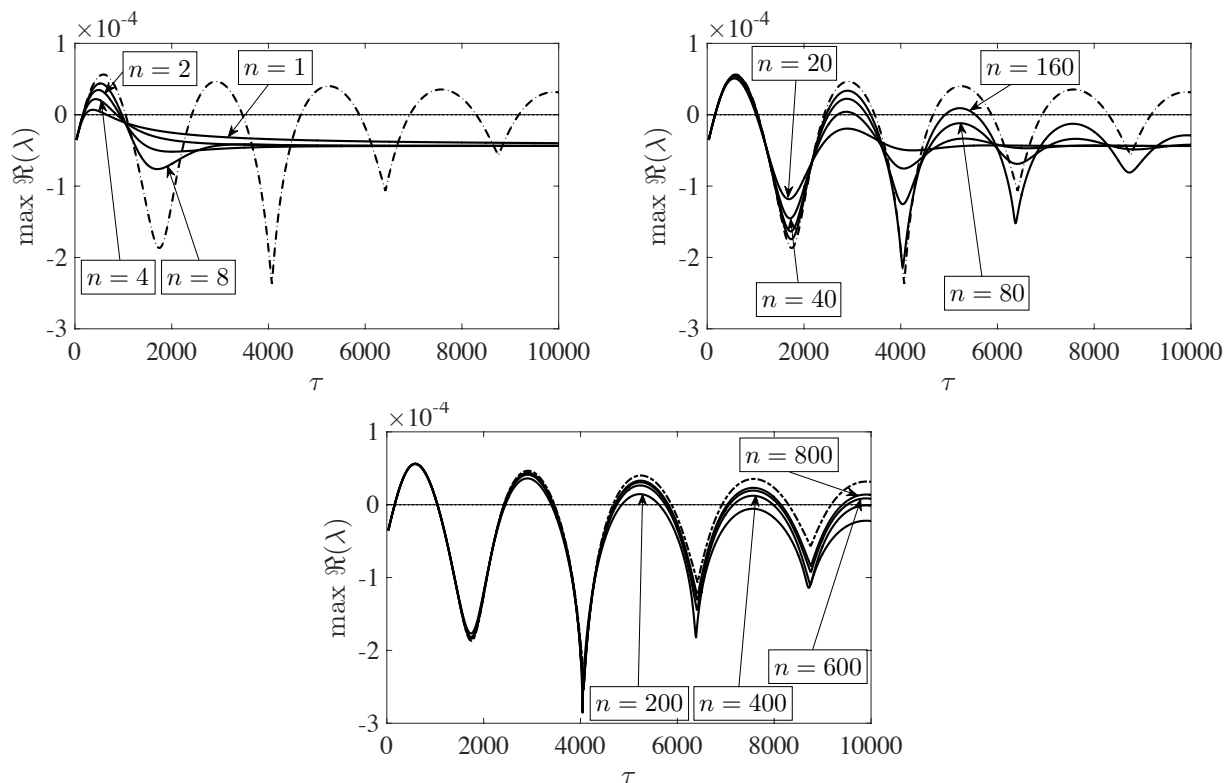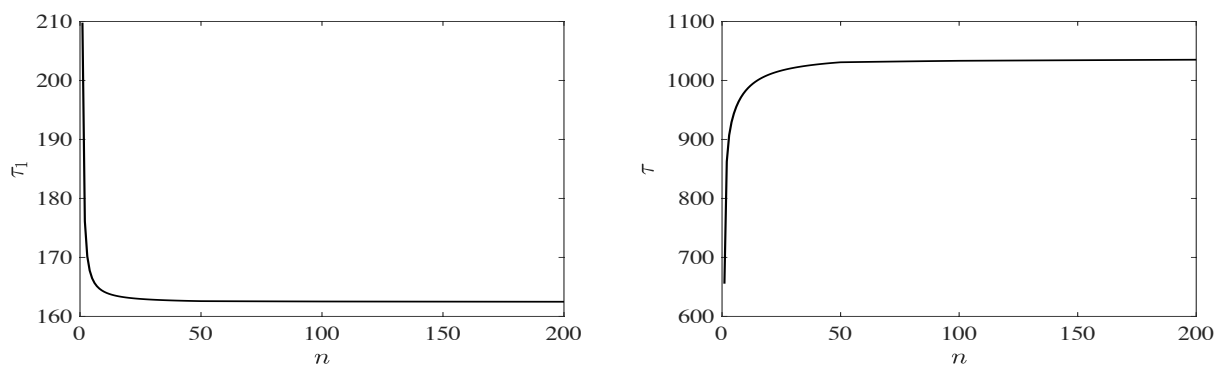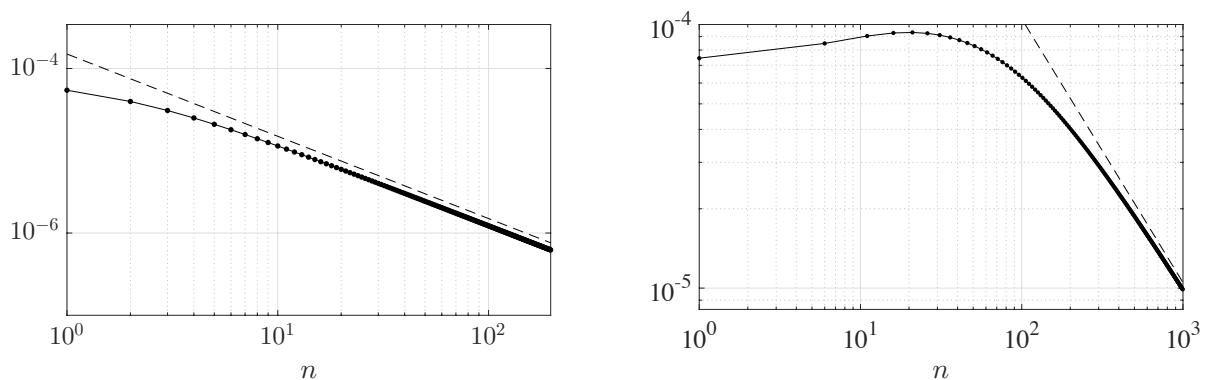

Figure 7. Real part of the rightmost eigenvalue(s) of matrix (2.16) for $c = 300$, varying τ and increasing n (solid lines), compared to the concentrated case (dashed-dotted line).

Table 3. Delay values for system (2.15) for varying n and $c = 300$.

n	τ_1	τ_2
1	209.7943611962976	655.462477398103
5	166.4696842243129	944.028737310052
20	163.1531571152444	1 010.230663791640
50	162.5856775785959	1 030.759177525466
100	162.4030841436188	1 038.443502844732
200	162.3129976629166	1 042.462005449525
400	162.2682549880071	1 044.517323668264
∞	162.2237119789684	1 046.604140033032

Figure 8. Delay graphs for system (2.15) for varying n and $c = 300$.

Although we leave a detailed investigation of this behavior to section 3, let us anticipate that, experimentally, we find that the difference between the real part of the rightmost eigenvalue(s) of cases (2.6) and (2.7) for fixed values of τ is $O(1/n)$, Figure 9. In section 3, as a major contribution of the present work, we rigorously show that this is indeed the case.

Figure 9. Difference between the real part of the rightmost eigenvalues relevant to (2.14) and that of (2.15) for increasing n, compared to a line of slope -1 (dashed): $c = 300$, $\tau = 600$ (left) and $\tau = 2800$ (right).
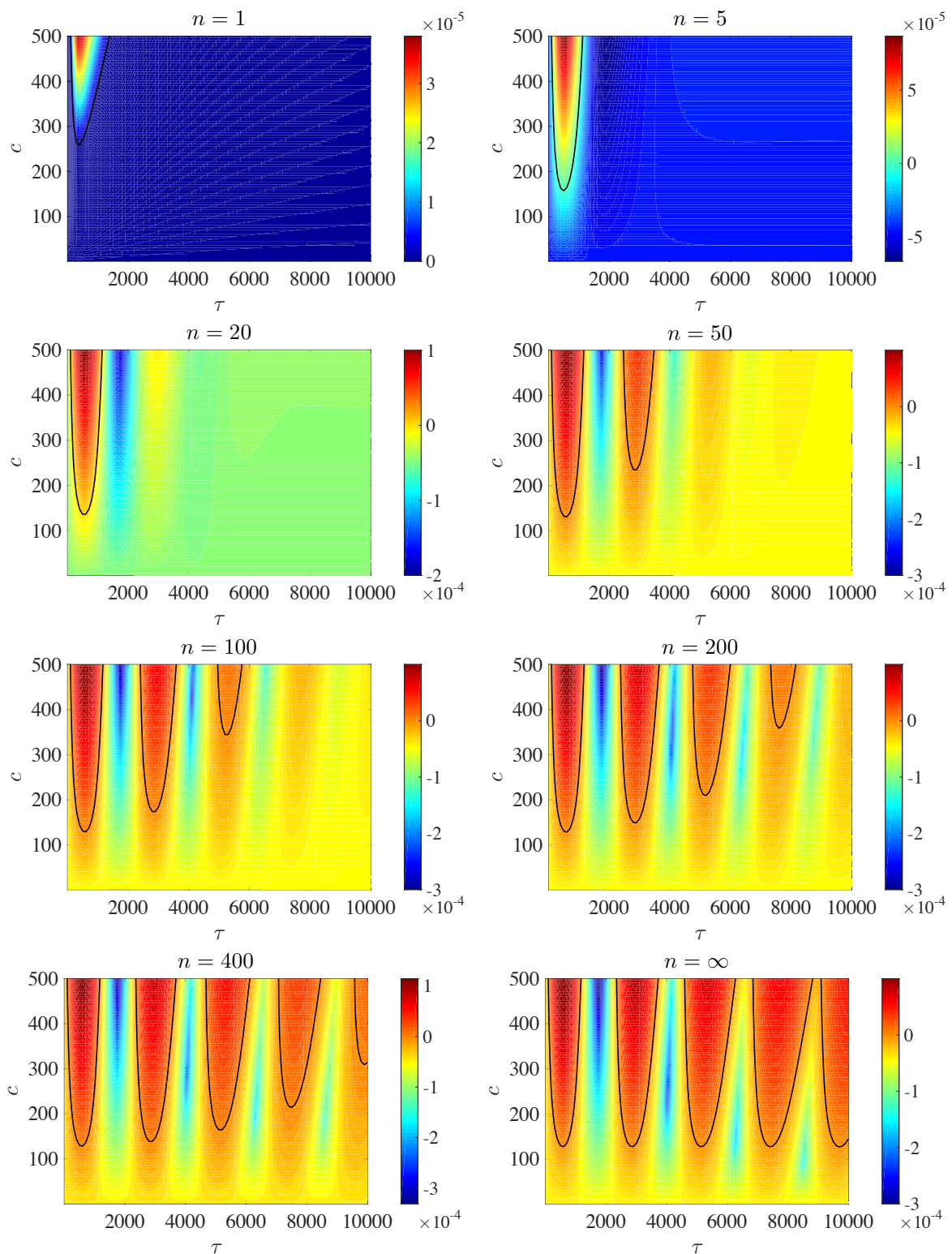
Figure 10. Real part of the rightmost eigenvalue(s) in the (τ, c)-plane for varying $n = 1$, i.e., Eq (2.5), $n = 5, 20, 50, 100, 200, 400$ in Eq (2.7) and $n = \infty$, i.e., Eq (2.6), with relevant stability boundaries (solid thick).

2.4. Stability charts and other coverage functions

To give a panorama of the situation described at the end of the preceding section, we illustrate in Figure 10 a sketch of the variation of the stability properties of the endemic equilibrium in the (τ, c)-plane as n increases from 1 to ∞, i.e., passing from the case of fading memory Eq (2.5) ($n = 1$, top-left panel) to that of concentrated memory Eq (2.6) ($n = \infty$, bottom-right-panel) through the cases of Erlang distributions Eq (2.7) (intermediate panels). The obtained stability boundaries (through `contour` in Matlab) denote the presence of so-called *stability lobes*. In particular, there is only one such lobe for $n = 1$, while more and more lobes appear as n grows, leading to a possibly infinite series of lobes for $n = \infty$. Below these lobes the endemic equilibrium is asymptotically stable, above it becomes unstable in favor of a limit cycle. The latter is stable, at least in proximity of the boundary of the lobe (recall that eigenvalues provide local results). Of course, the stability boundaries represent loci of Hopf bifurcations.

Below we report the results of the same analysis as carried out for different choices of the coverage function p and of the disease prevalence g. As for the former and with reference to Eq (2.10), we use the functions of the examples in [27, section 6], i.e., besides Eq (2.11), we consider both the Michaelis-Menten type coverage

$$p_1(M) = \frac{c_1 M}{d_1 M + 1} \tag{2.17}$$

and the Holling-type 2 coverage

$$p_1(M) = \frac{c_2 M^2}{d_2 M^2 + 1}, \tag{2.18}$$

where c_1, c_2, d_1 and d_2 are positive constants. In all cases, p_1 is (piecewise) smooth and increasing in M, $p_1(0) = 0$ and $0 \leq p_1(M) \leq 0.99 - p_0$ ($= c_i/d_i$, $i = 1, 2$, see [27, section 6]), with an S-shaped profile portraying a slow increase at low levels of M followed by a rapid increase towards saturation. The latter takes into account for the fraction of population that is never reached by vaccination [27, section 2]. Examples of these profiles are given in Figure 11 for $c_1 = c_2 = c$, $d_1 = d_2 = d$ and $c/d = 1 - p_0$ with $c = 300$. As for g, instead, besides the choice (2.12), we consider also
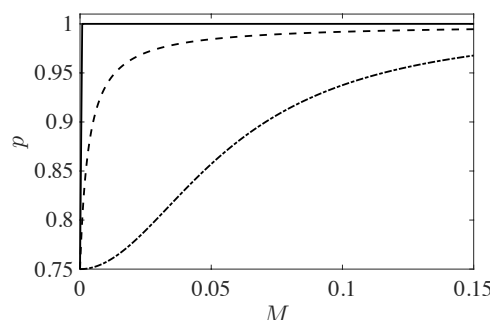
$$g(S, I) = kSI. \tag{2.19}$$

Figure 11. Examples of coverage functions p in (2.10) with S-shaped profiles: (2.11) (solid), (2.17) (dashed) and (2.18) (dashed-dotted).

Figure 12 illustrates the stability charts and boundaries for the choice (2.17) and both cases (2.12) and (2.19) in the same guise of Figure 10, just limiting the choice of values of n to the set

{1, 50, 200, ∞}, given that the results are qualitatively unchanged. Figure 13 is the same but for Eq (2.18): Again, the quality of the results is unchanged. The values of the parameters are those described above for Figure 11, except for c which is now varying.
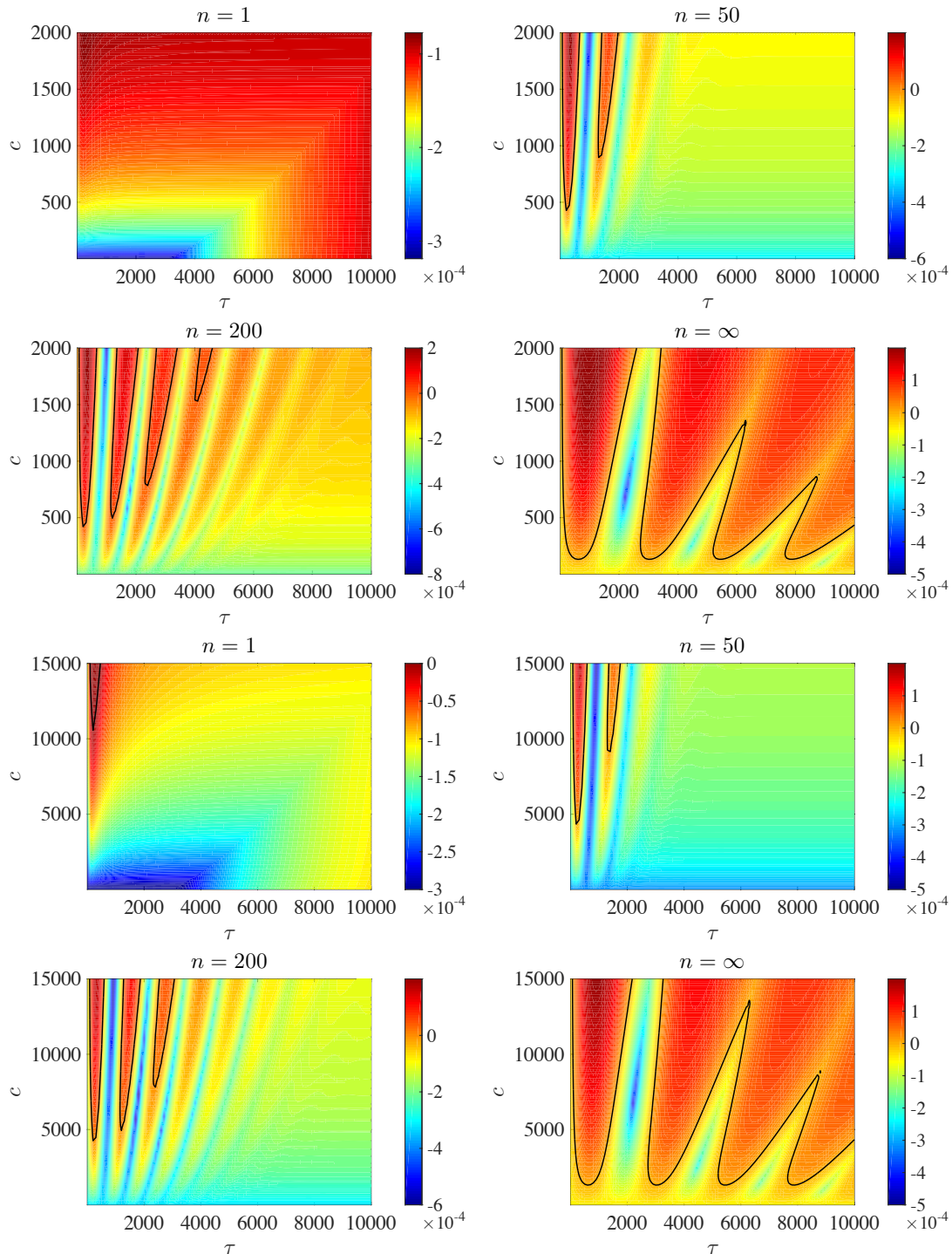

Figure 12. Real part of the rightmost eigenvalue(s) in the (τ, c)-plane for varying $n \in \{1, 50, 200, \infty\}$ with relevant stability boundaries (solid thick) for (2.17) and (2.12) (first 4 panels in the top) and (2.19) (last 4 panels in the bottom).
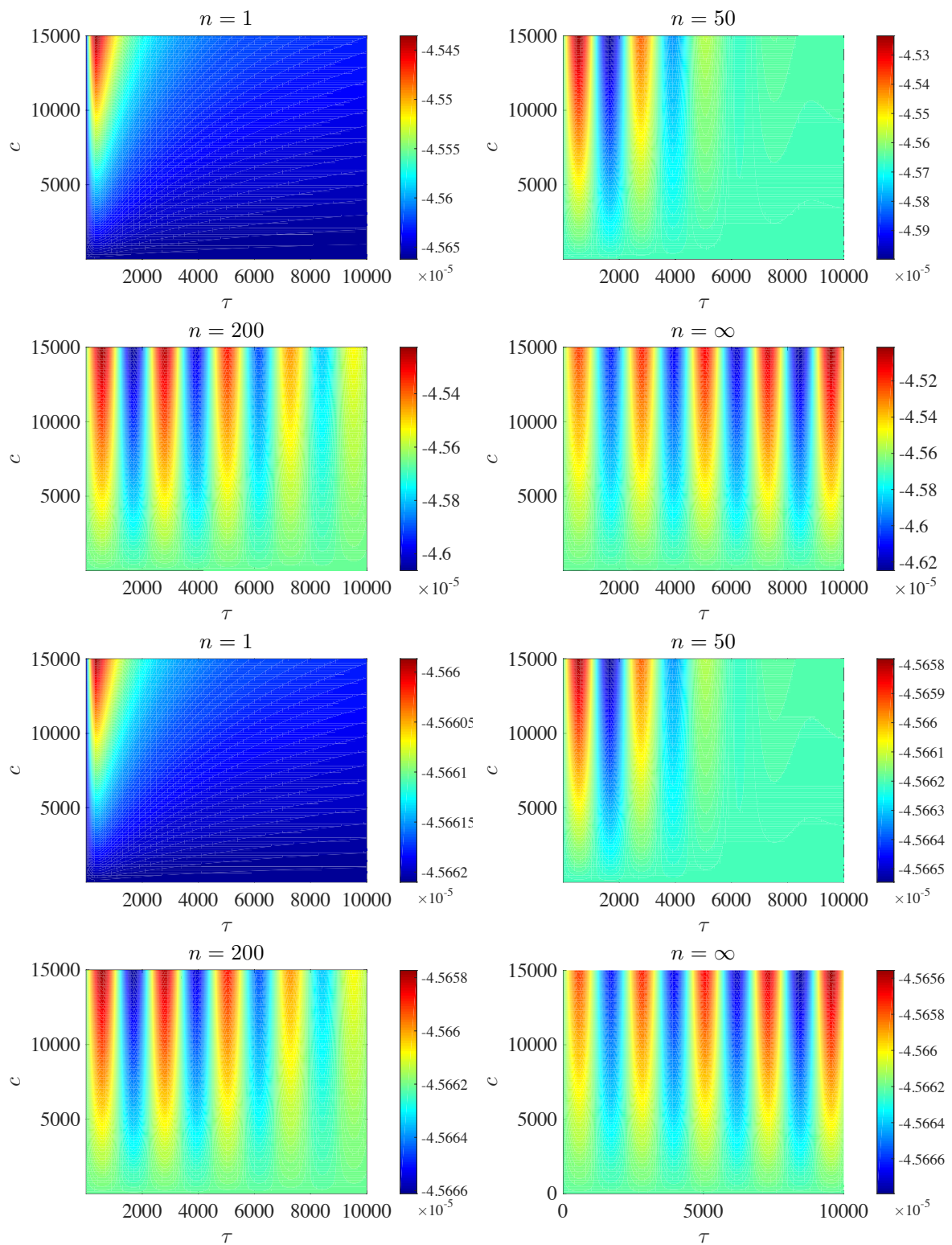
Figure 13. Real part of the rightmost eigenvalue(s) in the (τ, c)-plane for varying $n \in \{1, 50, 200, \infty\}$ for (2.18) and (2.12) (first 4 panels in the top) and (2.19) (last 4 panels in the bottom).

3. From fading to concentrated memory

In section 2.3 we have seen that by using Eq (2.7) and increasing n starting from 1, more and more instability intervals due to Hopf bifurcations appear, possibly becoming infinitely-many in the limit as $n \to \infty$. Here we study the behavior of the relevant rightmost eigenvalues, showing that they converge to those of Eq (2.6) linearly with respect to n, i.e., their difference decreases as $O(1/n)$. We thus provide the theoretical confirmation to the experimental results illustrated in Figure 9. Then, in section 4, we show that these results easily extend to models far more general than system (2.3).

The following analysis of convergence is based on the characteristic equations relevant to systems (2.14) and (2.15). The corresponding characteristic functions are compared by applying Rouché's Theorem on zeros of holomorphic functions (see, e.g., [36]). By following the arguments used in [31], we eventually show that the difference between the eigenvalues of the intermediate case and those of the concentrated case is $O(1/n)$, thus proving also their coincidence in the limit as $n \to \infty$.

Let us start from the matrix matrix (2.16) at the right-hand side of the linearization of system (2.15) around the endemic equilibrium. The associated characteristic function is

$$h_n(\lambda) := \det(\lambda I_{n+2} - A_n).$$

By applying the Laplace rule along the first row we get

$$\begin{aligned} h_n(\lambda) = &(\lambda + \mu + \beta\overline{I})(\lambda - \beta\overline{S} + \mu + \nu)(\lambda + a)^n + \beta^2\overline{S}\,\overline{I}(\lambda + a)^n \\ &+(-1)^n\mu p'(\overline{M})[(-\beta\overline{I})\left(-ag_I(\overline{S},\overline{I})\right)(-a)^{n-1} \\ &+(-\lambda + \beta\overline{S} - \mu - \nu)(-ag_S(\overline{S},\overline{I})(-a)^{n-1}]. \end{aligned}$$

Equivalently,

$$h_n(\lambda) = f_1(\lambda)(\lambda + a)^n + f_2(\lambda)a^n, \tag{3.1}$$

where

$$f_1(\lambda) := (\lambda + \mu + \beta\overline{I})(\lambda - \beta\overline{S} + \mu + \nu) + \beta^2\overline{S}\,\overline{I} \tag{3.2}$$

and

$$f_2(\lambda) := \mu p'(\overline{M})[\beta\overline{I}g_I(\overline{S},\overline{I}) + (\lambda - \beta\overline{S} + \mu + \nu)g_S(\overline{S},\overline{I})] \tag{3.3}$$

are analytic functions. Eventually, let us write

$$h_n(\lambda) = (\lambda + a)^n \overline{h}_n(\lambda) \tag{3.4}$$

for

$$\overline{h}_n(\lambda) := f_1(\lambda) + f_2(\lambda)\left(1 + \frac{\lambda}{a}\right)^{-n}. \tag{3.5}$$

Now let us look for the characteristic function of the model with concentrated memory, i.e., (2.14). The system linearized at the nontrivial equilibrium $(\overline{S}, \overline{I}, \overline{M})$ reads

$$\begin{cases} S'(t) = (-\mu - \beta\overline{I})S(t) - \beta\overline{S}I(t) - \mu p'(\overline{M})M(t) \\ I'(t) = \beta\overline{I}S(t) + [\beta\overline{S} - (\mu + \nu)]I(t) \\ M(t) = g_S(\overline{S}, \overline{I})S(t - \tau) + g_I(\overline{S}, \overline{I})I(t - \tau). \end{cases}$$

Let $X := C([-\tau, 0]; \mathbb{R}^3)$ be the state space of the corresponding dynamical system. With reference to [34], the associated infinitesimal generator is the linear unbounded operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subseteq X \to X$ defined as $\mathcal{A}\varphi = \varphi'$ with domain

$$\mathcal{D}(\mathcal{A}) = \{\varphi \in X : \varphi' \in X \text{ and } \varphi_1'(0) = L_1\varphi, \ \varphi_2'(0) = L_2\varphi, \ \varphi_3(0) = L_3\varphi\}$$

for

$$\begin{aligned}
L_1\varphi &:= [-\mu - \beta\bar{I}]\varphi_1(0) - \beta\bar{S}\varphi_2(0) - \mu p'(\overline{M})\varphi_3(0), \\
L_2\varphi &:= \beta\bar{I}\varphi_1(0) + [\beta\bar{S} - (\mu + \nu)]\varphi_2(0), \\
L_3\varphi &:= g_S(\bar{S}, \bar{I})\varphi_1(-\tau) + g_I(\bar{S}, \bar{I})\varphi_2(-\tau)\}.
\end{aligned} \tag{3.6}$$

Since $\lambda \in \sigma(\mathcal{A})$ if and only if there exists $\varphi \in \mathcal{D}(\mathcal{A}) \setminus \{0\}$ such that $\mathcal{A}\varphi = \lambda\varphi$, it follows that φ satisfies

$$\begin{cases}
\varphi'(\theta) = \lambda\varphi(\theta), & \theta \in [-\tau, 0], \\
\varphi_1'(0) = L_1\varphi \\
\varphi_2'(0) = L_2\varphi \\
\varphi_3(0) = L_3\varphi.
\end{cases}$$

The unique solution has the form $\varphi(\theta) = e^{\lambda\theta}\varphi(0)$, so that

$$\begin{aligned}
L_1(e^{\lambda\cdot}\varphi(0)) &= L_1\varphi = \varphi_1'(0) = \lambda\varphi_1(0) \\
L_2(e^{\lambda\cdot}\varphi(0)) &= L_2\varphi = \varphi_2'(0) = \lambda\varphi_2(0) \\
L_3(e^{\lambda\cdot}\varphi(0)) &= L_3\varphi = \varphi_3(0).
\end{aligned} \tag{3.7}$$

Therefore, the characteristic function turns out to be

$$h_\infty(\lambda) := \det\left[\begin{pmatrix} \lambda & 0 & 0 \\ 0 & \lambda & 0 \\ 0 & 0 & 1 \end{pmatrix} - A_\infty(\lambda)\right]$$

for

$$A_\infty(\lambda) := \begin{pmatrix} -\mu - \beta\bar{I} & -\beta\bar{S} & -\mu p'(\overline{M}) \\ \beta\bar{I} & \beta\bar{S} - (\mu + \nu) & 0 \\ g_S(\bar{S}, \bar{I})e^{-\lambda\tau} & g_I(\bar{S}, \bar{I})e^{-\lambda\tau} & 0 \end{pmatrix}$$

coming from Eqs (3.6) and (3.7). Then, by applying Laplace rule along the first row and by using Eqs (3.2) and (3.3), we get

$$h_\infty(\lambda) = f_1(\lambda) + f_2(\lambda)e^{-\lambda\tau}. \tag{3.8}$$

In order to study how h_n in Eq (3.1) approaches h_∞ in Eq (3.8) as $n \to \infty$ we need to evaluate their difference for varying n. However, thanks to the following lemma, we compare $\bar{h}_n$ in Eq (3.5) with h_∞.

Lemma 3.1. *If*

$$a \neq \mu + \nu - \beta\bar{S} + \beta\bar{I}\frac{g_I(\bar{S}, \bar{I})}{g_S(\bar{S}, \bar{I})} \tag{3.9}$$

then $h_n(\lambda) = 0 \iff \bar{h}_n(\lambda) = 0$.

Proof. It is enough to observe that condition (3.9) guarantees $f_2(-a) \neq 0$, so that Eq (3.1) ensures that $\lambda = -a$ cannot be a characteristic root of h_n. The thesis follows from Eq (3.4). □

Note that condition (3.9) is satisfied generically when Eq (2.9) holds, given that n is integer.

The following lemmas provide all the ingredients necessary to prove the final convergence result. We denote $\overline{B}(\lambda, \rho)$ the closed ball in $\mathbb{C}$ of center λ and radius ρ.

Lemma 3.2 (see Lemma 3.5 in [31]). *Let λ^* be a zero of h_∞ with algebraic multiplicity m. Then there exists $\rho_1 = \rho_1(\lambda^*) > 0$ and $C_1 = C_1(\lambda^*) > 0$ such that for all $\lambda \in \overline{B}(\lambda^*, \rho_1) \setminus \{\lambda^*\}$*

$$|h_\infty(\lambda)| > C_1 |\lambda - \lambda^*|^m.$$

Proof. By considering the Taylor series of $h_\infty(\lambda)$ around λ^* and by taking into account the multiplicity m we obtain $h_\infty(\lambda) = h_\infty^{(m)}(\lambda^*)(\lambda - \lambda^*)^m/m! + O\left(|\lambda - \lambda^*|^{m+1}\right)$ with $h_\infty^{(m)}(\lambda^*) \neq 0$. Then $\lim_{\lambda \to \lambda^*} |h_\infty(\lambda)|/|\lambda - \lambda^*|^m = |h_\infty^{(m)}(\lambda^*)|m!$. Let us set $C_1' := |h_\infty^{(m)}(\lambda^*)|/m!$. Since $C_1' > 0$, there exists $\rho_1 > 0$ and $C_1 > 0$, both depending on λ^*, such that, for all $\lambda \in \overline{B}(\lambda^*, \rho_1) \setminus \{\lambda^*\}$, $|h_\infty(\lambda)|/|\lambda - \lambda^*|^m > C_1$. □

Lemma 3.3. *Let $\lambda^* \in \mathbb{C}$ and $\rho_0 > 0$. Then there exists $C_0 = C_0(\lambda^*, \rho_0) > 0$ such that for all $\lambda \in \overline{B}(\lambda^*, \rho_0)$ and all $n \in \mathbb{N} \setminus \{0\}$*

$$\left| \left(1 + \frac{\lambda\tau}{n}\right)^n - e^{\lambda\tau} \right| \leq \frac{C_0}{n}.$$

Proof. Let us fix $z \in \mathbb{C}$ and consider the initial value problem

$$\begin{cases} y'(t) = zy(t), & t \in [0, 1], \\ y(0) = 1, \end{cases} \tag{3.10}$$

whose unique solution is $y(t) = e^{zt}$. The proof is based on analyzing the error committed by solving (3.10) with the explicit Euler method with step $h = 1/n$, $n \in \mathbb{N} \setminus \{0\}$. Let y_i be the relevant approximation of $y(t_i)$, $t_i = ih$, $i = 1, \ldots, n$, and let $y_0 = y(0)$. The method reads $y_{i+1} = (1 + hz)y_i$, $i = 0, \ldots, n-1$, with $y_0 = 1$. The local truncation error is $\varepsilon_{i+1} := (1 + hz)y(t_i) - y(t_{i+1})$, $i = 0, \ldots, n-1$, and the second-order Taylor expansion of $y(t_{i+1})$ with Lagrange remainder gives

$$\begin{aligned} \varepsilon_{i+1} &= (1 + hz)y(t_i) - y(t_i) - y'(t_i)h - y''(\xi_i)\frac{h^2}{2} \\ &= (1 + hz)y(t_i) - y(t_i) - zy(t_i)h - y''(\xi_i)\frac{h^2}{2} \\ &= -y''(\xi_i)\frac{h^2}{2} \end{aligned}$$

for some $\xi_i \in (t_i, t_{i+1})$. The global truncation error at $t_n = 1$ follows as

$$\begin{aligned} e_n &:= y_n - y(t_n) \\ &= (1 + hz)y_{n-1} - (1 + hz)y(t_{n-1}) + \varepsilon_n \\ &= (1 + hz)e_{n-1} + \varepsilon_n \\ &= (1 + hz)[(1 + hz)e_{n-2} + \varepsilon_{n-1}] + \varepsilon_n \\ &\ \ \vdots \\ &= (1 + hz)^n e_0 + (1 + hz)^{n-1}\varepsilon_1 + \cdots + (1 + hz)\varepsilon_{n-1} + \varepsilon_n. \end{aligned}$$

Since $e_0 = y(0) - y_0 = 0$, we get $e_n = \sum_{i=0}^{n-1}(1 + hz)^i \varepsilon_{n-i}$. Then the triangle inequality gives $|e_n| \leq \max_{i=1,...,n} |\varepsilon_i| \sum_{i=0}^{n-1}(1 + h|z|)^i$. The first factor at the right-hand side above is bounded as

$$\max_{i=1,...,n} |\varepsilon_i| \leq \frac{h^2}{2}\|y''\|_\infty = \frac{h^2}{2}|z^2| \max_{t \in [0,1]} |e^{zt}| = \frac{h^2}{2}|z|^2 \max\{1, e^{\Re z}\}.$$

As for the second factor we get

$$\sum_{i=0}^{n-1}(1 + h|z|)^i \leq \sum_{i=0}^{n-1} e^{ih|z|} = \frac{1}{h}\sum_{i=0}^{n-1} he^{ih|z|} \leq \frac{1}{h}\int_0^1 e^{|z|t}dt = \frac{e^{|z|} - 1}{h|z|}$$

thanks to the (left) rectangular quadrature formula for the function $e^{|z|t}$ in $t \in [0, 1]$ with quadrature points t_i, $i = 0, \ldots, n - 1$. Therefore $|e_n| \leq C_0'h$ holds for $C_0' = C_0'(z) := \frac{1}{2}(e^{|z|} - 1)|z| \max\{1, e^{\Re z}\}$. We conclude that $|(1 + hz)^n - e^z| = |e_n| \leq C_0'/n$ and the thesis follows by choosing $z = \lambda\tau$ and $C_0 = C_0(\lambda^*, \rho_0) := \max_{\lambda \in \overline{B}(\lambda^*, \rho_0)} C_0'(\lambda\tau)$. □

Lemma 3.4. *Let $\lambda^* \in \mathbb{C}$ and $\rho_0 > 0$. Then there exists $C_2 = C_2(\lambda^*, \rho_0) > 0$ such that for all $\lambda \in \overline{B}(\lambda^*, \rho_0)$ and all $n \in \mathbb{N} \setminus \{0\}$*

$$\left|\overline{h}_n(\lambda) - h_\infty(\lambda)\right| \leq \frac{C_2}{n}.$$

Proof. Thanks to Lemma 3.3 there exists $C_0 = C_0(\lambda^*, \rho_0) > 0$ such that $(1 + \lambda\tau/n)^n = e^{\lambda\tau} + \gamma(n)$ with $\gamma(n)$ such that $|\gamma(n)| \leq C_0/n$. Then $[(1 + \lambda\tau/n)^n]^{-1} = [e^{\lambda\tau} + \gamma(n)]^{-1} = e^{-\lambda\tau} - e^{-\lambda\tau}\gamma(n) + O(|\gamma(n)|^2)$ follows by a second-order Taylor expansion. Thus, for all $\lambda \in \overline{B}(\lambda^*, \rho_0)$, $|(1 + \lambda\tau/n)^{-n} - e^{-\lambda\tau}| \leq C_2'/n$, where $C_2' := C_0 \cdot \max_{\lambda \in \overline{B}(\lambda^*, \rho_0)} |e^{-\lambda\tau}|$. Now, by Eqs (3.5) and (3.8), we have $|\overline{h}_n(\lambda) - h_\infty(\lambda)| \leq \max_{\lambda \in \overline{B}(\lambda^*, \rho_0)} |f_2(\lambda)| \cdot |(1 + \lambda\tau/n)^{-n} - e^{-\lambda\tau}|$ and the thesis follows by choosing $C_2 = C_2(\lambda^*, \rho_0) := C_2' \cdot \max_{\lambda \in \overline{B}(\lambda^*, \rho_0)} |f_2(\lambda)|$. □

Eventually, we state and prove the convergence of the eigenvalues of A_n to those of $\mathcal{A}$ in terms of zeros of h_n and h_∞, respectively.

Theorem 3.5 (see Theorem 3.6 in [31]). *Let λ^* be a zero of h_∞ with algebraic multiplicity m and $r > 0$ be such that λ^* is the only zero of h_∞ in $\overline{B}(\lambda^*, r)$. Then there exists $n^* \in \mathbb{N} \setminus \{0\}$ sufficiently large such that h_{n^*} has exactly m zeros $\lambda_{n^*,1}, \ldots, \lambda_{n^*,m}$ (counted with multiplicities) in $\overline{B}(\lambda^*, r)$ and*

$$\max_{i=1,...,m} |\lambda^* - \lambda_{n^*,i}| \leq \rho(n^*),$$

with $\rho(n^*) = O((1/n^*)^{1/m})$.

Proof. Thanks to Lemma 3.2 there exist $\rho_1 = \rho_1(\lambda^*)$ and $C_1 = C_1(\lambda^*, \rho_1) > 0$ such that, for all $\lambda \in \overline{B}(\lambda^*, \rho_1) \setminus \{\lambda^*\}$, $|h_\infty(\lambda)| > C_1|\lambda - \lambda^*|^m$. We can assume $\rho_1 < r$ without loss of generality. Let us define $\rho(n) := [C_2/(C_1n)]^{1/m}$ for C_2 as in Lemma 3.4 with $\rho_0 = \rho_1$. Since $\rho(n) \to 0$ as $n \to \infty$, there exists n^* sufficiently large such that $\rho(n^*) < \rho_1$. Then

$$|h_\infty(\lambda)| > C_1|\lambda - \lambda^*|^m = C_1\rho(n^*)^m = \frac{C_2}{n^*} \geq \left|\overline{h}_n(\lambda) - h_\infty(\lambda)\right|.$$

follows by taking $\lambda \in \overline{B}(\lambda^*, \rho_1)$ such that $|\lambda - \lambda^*| = \rho(n^*)$. Since both $\overline{h}_n$ and h_∞ are holomorphic in $\overline{B}(\lambda^*, \rho_1)$, Rouché's Theorem (see, e.g., [36]) ensures that they have the same number of zeros in $\overline{B}(\lambda^*, \rho(n^*))$ counted with multiplicities. The thesis follows by Lemma 3.1. □

Remark 3.6. *The linear chain trick can be interpreted as a way to approximate a constant delay τ through a sequence of time steps of constant size $h := \tau/n = 1/a$. If one thinks about the finite difference scheme corresponding to explicit Euler with step h, the bidiagonal matrix of entries $1/h = a$ and $-1/h = -a$ arises naturally, compare with matrix (2.16) or matrix (4.2) below. This translates then into the Taylor's truncation of the exponential function, on which the given convergence proof is indeed based.*

4. A general framework

It is not difficult to realize that the results of the preceding section are independent of the specific model at hands (in our case (2.3)), and thus they can be extended to the much more general modeling setting

$$
\begin{cases}
x'(t) = f(x(t), M(t)) \\
M(t) = \int_0^{+\infty} g(x(t-\theta))K(\theta)\,\mathrm{d}\theta
\end{cases}
\tag{4.1}
$$

as we illustrate next. Above, for d some positive integer, $f : \mathbb{R}^{d+1} \to \mathbb{R}^d$ and $g : \mathbb{R}^d \to \mathbb{R}$ are smooth. Moreover, we assume the existence of a nontrivial equilibrium $(\overline{x}, \overline{M}) \in \mathbb{R}^{d+1}$, i.e., $f(\overline{x}, \overline{M}) = 0$ for $\overline{M} = g(\overline{x})$. Note that, beyond population models, this setting includes also economic models with time delay, see, e.g., [37, 38].

Following section 2.3, for K in Eq (2.7) we apply the linear chain trick to obtain the system of $d+n$ ODEs

$$
\begin{cases}
x'(t) = f(x(t), M(t)) \\
Z_1'(t) = ag(x(t)) - aZ_1(t) \\
Z_i'(t) = aZ_{i-1}(t) - aZ_i(t), \quad i = 2, \ldots, n.
\end{cases}
$$

Linearizing around the nontrivial equilibrium $(\overline{x}, \overline{M})$ gives a linear system with matrix

$$
A_n := \begin{pmatrix}
D_x f(\overline{x}, \overline{M}) & 0 & 0 & \cdots & D_M f(\overline{x}, \overline{M}) \\
aDg(\overline{x}) & -a & 0 & \cdots & 0 \\
0 & a & -a & \cdots & 0 \\
\vdots & \vdots & \ddots & \ddots & \vdots \\
0 & 0 & 0 & a & -a
\end{pmatrix} \in \mathbb{R}^{(d+n)\times(d+n)},
\tag{4.2}
$$

where the entries have the appropriate dimension. In order to recover the characteristic function it is convenient to partition the characteristic matrix as

$$
\lambda I_{d+n} - A_n = \begin{pmatrix} A & B \\ C & D \end{pmatrix}
$$

with blocks

$$A := \lambda I_d - D_x f(\bar{x}, \overline{M}) \in \mathbb{R}^{d \times d}, \qquad B := \begin{pmatrix} 0 \cdots 0 & -D_M f(\bar{x}, \overline{M}) \end{pmatrix} \in \mathbb{R}^{d \times n},$$

$$C := \begin{pmatrix} -aDg(\bar{x}) \\ 0 \\ \vdots \\ 0 \end{pmatrix} \in \mathbb{R}^{n \times d}, \qquad D := \begin{pmatrix} \lambda + a & 0 & \cdots & 0 \\ -a & \lambda + a & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ 0 & 0 & -a & \lambda + a \end{pmatrix} \in \mathbb{R}^{n \times n}.$$

We can assume the block D to be nonsingular, i.e., $\lambda \neq -a$, in parallel with Lemma 3.2. Therefore its inverse reads

$$D^{-1} = \begin{pmatrix} \dfrac{1}{\lambda + a} & 0 & \cdots & 0 \\ \dfrac{a}{(\lambda + a)^2} & \dfrac{1}{\lambda + a} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \dfrac{a^{n-1}}{(\lambda + a)^n} & \cdots & \dfrac{a}{(\lambda + a)^2} & \dfrac{1}{\lambda + a} \end{pmatrix}$$

and we can compute the Schur complement

$$S(D) := A - BD^{-1}C$$

$$= \lambda I_d - D_x f(\bar{x}, \overline{M})$$

$$- \begin{pmatrix} 0 \cdots 0 & -D_M f(\bar{x}, \overline{M}) \end{pmatrix} \begin{pmatrix} \dfrac{1}{\lambda + a} & 0 & \cdots & 0 \\ \dfrac{a}{(\lambda + a)^2} & \dfrac{1}{\lambda + a} & \cdots & 0 \\ \vdots & \ddots & \ddots & \vdots \\ \dfrac{a^{n-1}}{(\lambda + a)^n} & \cdots & \dfrac{a}{(\lambda + a)^2} & \dfrac{1}{\lambda + a} \end{pmatrix} \begin{pmatrix} -aDg(\bar{x}) \\ 0 \\ \vdots \\ 0 \end{pmatrix}$$

$$= \lambda I_d - D_x f(\bar{x}, \overline{M}) - D_M f(\bar{x}, \overline{M}) Dg(\bar{x}) \left(1 + \dfrac{\lambda}{a}\right)^{-n}.$$

The Schur determinant identity gives then the characteristic function as

$$h_n(\lambda) = \det(\lambda I_{d+n} - A_n) = \det(D) \det(S(D)) = (\lambda + a)^n \bar{h}_n(\lambda),$$

where

$$\bar{h}_n(\lambda) := \det\left(\lambda I_d - D_x f(\bar{x}, \overline{M}) - D_M f(\bar{x}, \overline{M}) Dg(\bar{x}) \left(1 + \dfrac{\lambda}{a}\right)^{-n}\right).$$

As for K in Eq (2.6), i.e., in the concentrated case, following section 3 we arrive at the linear delay system

$$\begin{cases} x'(t) = D_x f(\bar{x}, \overline{M}) x(t) + D_M f(\bar{x}, \overline{M}) M(t) \\ M(t) = Dg(\bar{x}) x(t - \tau), \end{cases}$$

whose characteristic matrix reads

$$\lambda I_{d+1} - A_\infty(\lambda) := \begin{pmatrix} \lambda I_d - D_x f(\bar{x}, \overline{M}) & -D_M f(\bar{x}, \overline{M}) \\ -Dg(\bar{x})e^{-\lambda\tau} & 1 \end{pmatrix}.$$

By using the Schur complement of 1 we obtain the characteristic function as

$$h_\infty(\lambda) := \det(\lambda I_{d+1} - A_\infty(\lambda))$$
$$= \det\left(\lambda I_d - D_x f(\bar{x}, \overline{M}) - D_M f(\bar{x}, \overline{M})Dg(\bar{x})e^{-\lambda\tau}\right).$$

Now we can apply all the results of section 3, ending up with the same Theorem 3.5 through Lemmas 3.2, 3.3 and 3.4. The latter, in particular, is proved along the same lines just by noting that the determinant is a continuously differentiable function of its argument, so that there exists a constant $C = C(\lambda^*, \rho_0)$ such that $|\bar{h}_n(\lambda) - h_\infty(\lambda)| \leq C|(1 + \lambda\tau/n)^{-n} - e^{-\lambda\tau}|$.

5. Conclusions

In this paper we extended the stability analysis of [27] concerning an SIR model with vaccination at birth, in which the decision to vaccinate depends on the information on the disease. Indeed, we passed from the case of fading memory treated therein to that of concentrated memory through the Erlang distribution with varying shape parameter. The Erlang case is dealt with by using the linear chain trick. We obtained stability charts where it emerges clearly how the number of instability intervals along the delay axis increases to possibly infinitely-many, concluding that delay distributions with lower variance augment the chances of sustained oscillations.

In parallel, we gave a rigorous proof of the linear rate (in terms of the shape parameter) at which the Erlang distribution approaches the Dirac's delta. This is done by analyzing the convergence of the eigenvalues of the variational problems obtained by linearizing the original models around their endemic equilibrium. Moreover, this analysis is showed to be more general than that of the case of the concerned SIR framework.

The above outcomes would not have been obtained without resorting to recent techniques for treating equations with delays, e.g., [30–32]. We hope that in the future this could serve as a basis to analyze other models involving more or different delays. This would represent indeed an opportunity in view of the many possible directions to extend the modeling of the information term. In particular, it is plan of the authors and colleagues to tackle more realistic memory kernels. Examples can be multimodal distributions, or distributions with unconstrained variance (recall Eq (2.9)), as well as laws of power type or Gaussian-derived kernels. Of course, also going beyond the stability of the endemic equilibrium should be considered, and in this regard we remark that efficient techniques are already available to analyze the stability of periodic behaviors of systems with delays [39, 40], as well as a more complete bifurcation analysis [9].

Acknowledgments

this work. AA and DB are members of INdAM Research group GNCS. AA is supported by the PhD program in Computer Science, Mathematics and Physics, University of Udine; DB is supported by the INdAM GNCS project "Approssimazione numerica di problemi di evoluzione: aspetti deterministici e stocastici" (2018) and by the project PSD_2015_2017_DIMA_PRID_2017_ZANOLIN "SIDIA – SIstemi DInamici e Applicazioni" (UNIUD).

Conflict of interest

All authors declare no conflicts of interest in this paper.

References

1. F. Curtain, H. Zwart, *An Introduction to Infinite-Dimensional Linear Systems Theory*, Texts in Applied Mathematics **21**, Springer-Verlag, New York, 1995.

2. O. Diekmann, M. Gyllenberg, J. A. J. Metz, Finite dimensional state representation of linear and nonlinear delay systems, *J. Dynam. Differ. Equations*, **30** (2018), 1439–1467.

3. A. K. Erlang, Solution of some problems in the theory of probabilities of significance in automatic telephone exchanges, *Post Office Elec. Eng.*, (1917), 189–197.

4. N. MacDonald, *Time Lags in Biological Models*, Lecture Notes in Biomathematics **27**, Springer Verlag, Berlin, 1978.

5. N. MacDonald, *Biological Delay Systems: Linear Stability Theory*, Cambridge Studies in Mathematical Biology **8**, Cambridge Univeristy Press, Cambridge, 1989.

6. D. Fargue, Reductibilitè des systèmes héréditaires à des systèmes dynamics, *C.R. Acad. Sci. Paris Sér. A-B*, **277** (1973), 471–473.

7. D. Fargue, Reductibilitè des systèmes héréditaires, *Int. J. Nonlin. Mech.*, **9** (1974), 331–338.

8. T. Vogel, *Théorie Des Systèmes Evolutifs*, Gautier Villars, Paris, 1965.

9. D. Breda, O. Diekmann, M. Gyllenberg, F. Scarabel, R. Vermiglio, Pseudospectral discretization of nonlinear delay equations: New prospects for numerical bifurcation analysis, *SIAM J. Appl. Dyn. Sys.*, **15** (2016), 1–23.

10. S. Busenberg, C. Travis, On the use of reducible-functional differential equations in biological models, *J. Math. Anal. Appl.*, **89** (1982), 46–66.

11. K. L. Cooke, Z. Grossman, Discrete delay, distributed delays and stability switches, *J. Math. Anal. Appl.*, **86** (1982), 592–627.

12. E. Beretta, D. Breda, Discrete or distributed delay? Effects on stability of population growth, *Math. Biosci. Eng.*, **13** (2016), 19–41.

13. C. Barril, A. Calsina, J. Ripoll, A practical approach to R_0 in continuous-time ecological models, *Math. Meth. Appl. Sci.*, **41** (2018), 8432–8445.

14. A. Lloyd, Destabilization of epidemic models with the inclusion of realistic distributions of infectious periods, *Proc. Roy. Soc. Lond. B*, **268** (2001), 985–993.

15. A. Lloyd, Realistic distributions of infectious periods in epidemic models: Changing patterns of persistence and dynamics, *Theor. Popul. Biol.*, **60** (2001), 59–71.

16. C. Bauch, A. d'Onofrio, P. Manfredi, Behavioral epidemiology of infectious diseases: An overview, in *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases* (eds. P. Manfredi and A. d'Onofrio), Springer-Verlag, New York, (2013), 1–19.

17. P. Manfredi, A. d'Onofrio, *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases*, Springer-Verlag, New York, 2013.

18. Z. Wang, T. C. Bauch, S. Bhattacharyya, A. d'Onofrio, P. Manfredi, M. Percg, et al., Statistical physics of vaccination, *Phys. Rep.*, **664** (2016), 1–113.

19. B. Buonomo, G. Carbone, A. d'Onofrio, Effect of seasonality on the dynamics of an imitation-based vaccination model with public health intervention, *Math. Biosci.*, **15** (2018), 299–321.

20. B. Buonomo, A. d'Onofrio, D. Lacitignola, Global stability of an sir epidemic model with information dependent vaccination, *Math. Biosci.*, **216** (2008), 9–16.

21. B. Buonomo, A. d'Onofrio, D. Lacitignola, The geometric approach to global stability in behavioral epidemiology, in *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases* (eds. P. Manfredi and A. d'Onofrio), Springer-Verlag, New York, (2013), 289–308.

22. A. d'Onofrio, P. Manfredi, Information-related changes in contact patterns may trigger oscillations in the endemic prevalence of infectious diseases, *J. Theoret. Biol.*, **256** (2009), 473–478.

23. A. d'Onofrio, P. Manfredi, Vaccine demand driven by vaccine side effects: Dynamic implications for sir diseases, *J. Theoret. Biol.*, **264** (2010), 237–252.

24. A. d'Onofrio, P. Manfredi, P. Poletti, The impact of vaccine side effects on the natural history of immunization programmes: an imitation-game approach, *J. Theoret. Biol.*, **273** (2011), 63–71.

25. A. d'Onofrio, P. Manfredi, E. Salinelli, Vaccinating behaviour and the dynamics of vaccine preventabe infections, in *Modeling the Interplay Between Human Behavior and the Spread of Infectious Diseases* (eds. P. Manfredi and A. d'Onofrio), Springer-Verlag, New York, (2013), 267–287.

26. S. Funk, M. Salathé, V. A. Jansen, Modelling the influence of human behaviour on the spread of infectious diseases: A review, *J. R. Soc. Interface*, **7** (2010), 1247–1256.

27. A. d'Onofrio, P. Manfredi, E. Salinelli, Vaccinating behaviour, information, and the dynamics of SIR vaccine preventable diseases, *Theor. Popul. Biol.*, **71** (2007), 301–317.

28. A. d'Onofrio, Mixed pulse vaccination strategy in epidemic model with realistically distributed infectious and latent times, *Appl. Math. Comput*, **151** (2004), 181–187.

29. A. Calsina, J. Ripoll, Hopf bifurcation in a structured population model for the sexual phase of monogonont rotifers, *J. Math. Biol.*, **45** (2002), 22–36.

30. D. Breda, O. Diekmann, S. Maset, R. Vermiglio, A numerical approach for investigating the stability of equilibria for structured population models, *J. Biol. Dyn.*, **7** (2013), 4–20.

31. D. Breda, S. Maset, R. Vermiglio, Pseudospectral differencing methods for characteristic roots of delay differential equations, *SIAM J. Sci. Comput.*, **27** (2005), 482–495.

32. D. Breda, S. Maset, R. Vermiglio, *Stability of Linear Delay Differential Equations—A Numerical Approach with MATLAB*, Springer, New York, 2015.

33. N. Olgac, R. Sipahi, Kernel and offspring concepts for the stability robustness of multiple time delayed systems (MTDS), *J. Dyn. Syst. T. ASME*, **129** (2006), 245–251.

34. O. Diekmann, P. Getto, M. Gyllenberg, Stability and bifurcation analysis of Volterra functional equations in the light of suns and stars, *SIAM J. Math. Anal.*, **39** (2008), 1023–1069.

35. D. Breda, P. Getto, J. Sánchez Sanz, R. Vermiglio, Computing the eigenvalues of realistic Daphnia models by pseudospectral methods, *SIAM J. Sci. Comput.*, **37** (2015), 2607–2629.

36. H. A. Priestley, *Introduction to Complex Analysis*, Oxford University Press, New York, 1990.

37. L. Fanti, P. Manfredi, The Solow's model with endogenous population: A neoclassical growth cycle model, *J. Econ. Dev.*, **28** (2003), 103–115.

38. P. Manfredi, L. Fanti, Cycles in dynamic economic modelling, *Econ. Model.*, **21** (2004), 573–594.

39. D. Breda, D. Liessi, Approximation of eigenvalues of evolution operators for linear renewal equations, *SIAM J. Numer. Anal.*, **56** (2018), 1456–1481.

40. D. Breda, S. Maset, R. Vermiglio, Approximation of eigenvalues of evolution operators for linear retarded functional differential equations, *SIAM J. Numer. Anal.*, **50** (2012), 1456–1483.