A finite delayed branching process with inhomogeneous regions
for infectious disease modeling

Andrew Hart∗    Servet Martínez†

Center for Mathematical Modeling, IRL 2807 CNRS-UCHILE, Facultad de Ciencias
Físicas y Matemáticas, Universidad de Chile, Santiago, Chile.

August 3, 2022

Abstract

We propose a model for the spread of an infectious agent throughout an inhomogeneous region that is a fusion between a multi-type branching process in which the types represent homogeneous sub-regions and a delayed branching process in which individuals are allotted lifetimes representing their convalescences. Infected individuals reproduce offspring during a finite time interval of random length bounded by $D > 1$ and within the construction of the process, individuals’ convalescence times and death/recovery outcomes form a random media where branching takes place. We show that the criterion for extinction is similar for the number of offspring, the infirm symptomatic and asymptomatic populations, and also for an associated multi-type branching process. With the imposition of the condition that the mean matrices at each delay offset share left and right Perron-Frobenius eigenvectors, we are able to give explicit analytic expressions for various quantities derived from the limit of the geometrically weighted mean evolution of the process.

Keywords: delayed branching process, multi-type branching process, Perron-Frobenius theory, renewal theory

2010 MSC: 60J80; 60J85; 92D30.

1 Introduction

Inspired by the 2019 SARS-CoV2 pandemic and agent-based simulation studies of the spread of SARS-CoV2 such as that reported in [9]), our aim here is to present a model for the epidemic spread of an infectious agent which possesses three key features, namely, spatial interaction, levels of infectiousness that vary according to the time elapsed since infection within a fixed finite time window and convalescence times that may be anywhere from none (indicating an asymptomatic individual) to far in the future (capturing phenomena such as long COVID which appeared during the SARS-CoV2 pandemic). The model is based on discrete-time branching processes (we shall write $\text{bp}$ for branching process in future), which find a natural application in describing disease spread, for instance, see [4]. We give a brief overview of the main characteristics of the model.

Varying Infectiousness. Individuals who contract an infectious disease are only contagious for a short time immediately following their infection, say, $D$ days. Thus, an individual born (infected) at time $s$ has the opportunity to reproduce offspring (infect others) at every time point in the window \{s+1, s+2, \ldots, s+D\}, and the number of offspring born (individuals infected) at time $s+d$ follows a law that depends on the time offset $0 < d \leq D$. The number of days that an infected person is contagious while not exhibiting symptoms is important for understanding the spread of a disease and for tailoring...
strategies of containment and/or mitigation to bring it under control. For many diseases, infection starts with a non-infectious incubation period during which the pathogen multiplies until a sufficient load is amassed for the infected individual to become contagious to others. The time-varying offspring law of the model allows the level of contagiousness to change and distinguish between symptomatic and asymptomatic states, take into account effects such as an initial incubation/latent phase and incorporate a final phase in which the degree of infectiousness falls as an individual recovers from the disease.

**Duration of illness.** Individuals may continue suffering the effects of the disease long after they have ceased to be infectious and the model captures this persistence by assigning each individual in the population a random lifetime (convalescence time) whose distribution is not bounded. This time may be zero, signifying that the individual is asymptomatic, or very large, corresponding to a complicated recovery entailing a prolonged convalescence. Symptomatic Individuals (having positive lifetimes) either die or recover according to a Bernoulli random variable at the end of their lifetimes. If they die before \( D \) days have lapsed, they cease to be infectious. Asymptomatic individuals (with zero lifetimes) do not die, but remain contagious for \( D \) days following infection.

**Spatial effects.** Within a country or city, there are inhomogeneities in the geographic and socioeconomic distribution of the population as well as in the daily patterns of movement that govern the frequency and location of contacts between individuals. The model captures such spatial disparity by including a multi-type component in which types represent regions and the numbers of offspring of all types produced by individuals of distinct types can have different distributions, that is, offspring distributions not only depend on the time elapsed since infection, but are allowed to depend on both the type of the parent (source) and the type of the offspring (infected).

Thus, the model is a fusion between a multi-type branching process and a branching process permitting delayed births (that is, having a finite reproductive period) on a random media. The random media encompasses the disease response of all individuals in the event they become infected. More precisely, it comprises the lifetimes and death/recovery case outcomes for each potentially infected individual. For the sake of brevity, we shall call the model a multi-type delayed \( \text{bp} \). The dependence of the offspring distributions on the offset \( d \leq D \) allows the model to be adjusted to emulate the infection profile of different infectious agents. Spatial interaction in the epidemic spread can be handled by using the mean matrices of infection to provide a graphical representation of transmission patterns or by relating the mean matrices to relevant data such as traffic flow or other population movement information.

The paper is organized as follows. We begin with a precis of the theory of discrete-time multi-type \( \text{bp} \)’s in Section 2. The delayed variant of the multi-type \( \text{bp} \) is introduced in Section 3. Delayed multi-type \( \text{bp} \)’s were described and studied in full generality in \([7], [8], [15], [5], [6]\). This body of work encompasses more general frameworks than that used here and mostly deal with the continuous-time setting.

As already mentioned, delayed multi-type \( \text{bp} \)’s incorporate two new elements: the individuals can produce new offspring during a bounded period of time according to a possibly time-dependent law and each individual will be ill for a random period of time. Our model comprises three closely related processes evolving in time: the offspring process which gives the number of offspring produced (incidence or number of new cases) at each time, the asymptomatic population size process, which models the number of contagious individuals who show no signs of illness, and the symptomatic population size process which tracks the number of individuals who are clearly suffering from the disease. All three processes either all become extinct with probability 1 together or with probability less than 1. Evolution equations are derived for the mean number of new cases (infections), the mean number of active asymptomatic cases and the mean number of active symptomatic cases as functions of time. It will be seen that the dynamics of all three equations are governed by the set of mean matrices at the delay times.

In Section 4 we show that the criterion for extinction of a delayed multi-type \( \text{bp} \) is the same as that for a non-delayed multi-type \( \text{bp} \). Our main result (Proposition 6) is then presented in Section 5 when the mean matrices at different delays share P-F eigenvectors, then the mean population size of our three processes (taking into account the lifetimes of individuals) weighted by an exponential of time and the Malthusian
parameter (which is the geometric rate of growth) converges to a multiple of the left P-F eigenvector. This is analogous to what happens in the classical multi-type bp. It holds in the supercritical case, but also in the critical case if the expected lifetime is finite, as well as in the subcritical case provided the exponential moment of the lifetime with parameter equal to the negative of the Malthusian parameter is finite.

In studying the asymptotic behavior of the mean population sizes, the main difficulty is to control a product matrix. A similar problem was encountered in [5] where very general conditions were imposed on supercritical delayed multi-type processes with countable types. The framework here is more restrictive—we assume that all the matrices share the same Perron-Frobenius (P-F) eigenvectors—but we are able to describe the limiting distribution similarly to the non-delayed multi-type bp, in particular, providing the limit of the mean evolution of types. We also interpret the quantities that are involved in the limit behavior. Note that in applications of the finite delayed process, the usual assumption would be to assume that the mean matrices governing the evolution are proportional to each other, in which case the limit can be easily identified. Proposition 6 shows that this limit result remains stable under the less stringent condition that the mean matrices share P-F eigenvectors.

Section 5.2 expands the mean evolution backwards in time to obtain classes of paths whose symbols are the delays in \( D \) governing the matrix products. We prove that the sequences of these paths have long runs of a symbol. The existence of long runs allows product of matrices that share P-F eigenvectors to be controlled. These results are at the core of the proof of Proposition 6 presented in Section 6.

In the final section, we examine some aspects of the model not previously considered. Section 7.1 discusses families of matrices sharing P-F eigenvectors in greater detail. It shows that the set of all such families is not trivially small and describes a straightforward way of constructing these families, while Proposition 12 shows that if a family of matrices share P-F eigenvectors is the most general (or weakest) condition we have found that obtains asymptotic behavior consistent with the simplest model in which the delay effects are products of a time-homogeneous region interaction and a time-dependent scale parameter. In Section 7.2, we compare our results to the Fibonacci-type branching process previously studied in [11] and [2] in which individuals live forever and produce offspring ad infinitum starting two time units after birth.

2 Multi-type branching processes

In a multi-type bp, each individual takes on a type \( i \) from a finite set of types \( I \) and the individuals produce offspring independently of all other individuals. Type \( i \) individuals produce \( \xi_{i,j} \) offspring of type \( j \) according to the law \( p^{i,j} \) on the non-negative integers. Each individual is born at some time \( s \) and produces all of its offspring at time \( s + 1 \). Individuals of type \( i \) born at the same time \( s \) are enumerated in some fixed way, so that any individual may be uniquely identified by a triple of the form \((i, s, l)\). The random variable \( \xi_{i,s,l}^{i,j} \) with distribution \( p^{i,j} \) indicates the number of type \( j \) offspring produced by the \( l \)-th individual of type \( i \) at time \( s \). All of these random variables are independent. To exclude uninteresting behavior, we assume that the process is non-singular. Singularity means that each individual reproduces exactly one offspring during its lifetime.

A multi-type bp is a process \( Z(s) = (Z_j(s) : j \in I) \) for \( s \in \mathbb{N}_0 = \{0, 1, 2, \ldots\} \), where the number of individuals \( Z_j(s) \) of type \( j \) born at time \( s \) is given by

\[
Z_j(s + 1) = \sum_{i \in I} \sum_{l=1}^{Z_i(s)} \xi_{i,s,l}^{i,j}.
\]

Since we will impose irreducibility conditions on the dynamics governing the propagation of types we will assume the simplest initial configuration which comprises a single individual of a fixed type \( i_0 \). Thus, \( P(Z(0) = e_{i_0}) = 1 \) where \( e_{i_0} \) is the vector having 1 as its \( i_0 \)-th coordinate and vanishes everywhere else in \( I \).
The mean number of offspring of type \( j \) generated by an individual of type \( i \) is

\[
M(i, j) = \mathbb{E}(\xi^{i,j}) = \sum_{n \geq 0} np^{i,j}(n), \ i, j \in I.
\]  

Then, the matrix of means is defined to be \( M = (M(i, j) : i, j \in I) \). We shall assume that \( M \) is irreducible, which suffices to guarantee that the process cannot be decomposed into multiple separate processes by type. The Perron-Frobenius theorem then asserts that \( M \) has a simple eigenvalue \( \rho > 0 \) which is the spectral radius of \( M \) and has (up to a constant multiple) unique, strictly positive left and right eigenvectors \( \nu = (\nu(i) : i \in I) \) and \( h = (h(i) : i \in I) \), that is,

\[
\nu'M = \rho\nu' \quad \text{and} \quad Mh = \rho h.
\]

See [15, Chapter 1]. Here the vectors are column vectors and ‘\( \cdot \)' signifies the transpose so that \( \nu' \) is a row vector. The vectors \( \nu \) and \( h \) are normalized such that \( \nu' h = \sum_{i \in I} \nu(i) h(i) = 1 \) and \( \nu' 1 = \sum_{i \in I} \nu(i) = 1 \).

The matrix \( \rho^{-1} M(i, j) h(j) / h(i) : i, j \in I \) is stochastic with stationary distribution \( (\nu(j) h(j) : j \in I) \) and convergence to the stationary distribution is geometrically fast. Thus for all norms \( \| \cdot \| \),

\[
\exists C < \infty, \delta \in (0, 1) \text{ such that } \forall t \geq 0 : \|\rho^{-1} M^t - h\nu'\| \leq C\delta^t
\]

and

\[
lim_{t \to \infty} \rho^{-1} M^t = h\nu' \text{ componentwise.}
\]

Various key properties of multi-type \( \mathbb{P} \)'s can be found in Chapter V of [3], in particular, the characterization of extinction. Let \( q_j = \mathbb{P}(\lim_{s \to \infty} Z_j(s) = 0) \) be the probability that the population of type \( j \) becomes extinct.

In [10] it was shown that

\[
\rho \leq 1 \implies q_j = 1 \text{ for all } j \in I,
\]

\[
\rho > 1 \implies q_j < 1 \text{ and } \mathbb{P}\left(\lim_{s \to \infty} Z_j(s) = \infty\right) = 1 - q_j > 0 \text{ for all } j \in I.
\]

Assume \( \sum_{i,j \in I} n \log n p^{i,j}(n) < \infty \). When \( \rho < 1 \) the processes \( (Z_j(s) : j \in I) \) are absorbed geometrically fast (see [12]). In [15] it was shown that in the supercritical case \( (\rho > 1) \) one has \( \lim_{s \to \infty} \rho^{-s} Z(s) = \nu W \) holds \( \mathbb{P} \)-a.s. for some non-negative random variable \( W \) satisfying \( \mathbb{P}(W > 0) > 0 \).

Let \( E_{i_0} \) denote the expected value of a multi-type \( \mathbb{P} \) which starts with a single individual of type \( i_0 \), but if there is no confusion we shall simply write \( E \). The expected value of the number of individuals at time \( s \) is \( E_{i_0}(Z(s)) = (E_{i_0}(Z_j(s)) : j \in I) \). Then, by conditioning relation (11) with respect to the \( \sigma \)-field of events defining individuals born up to time \( s - 1 \), we obtain \( E(Z_j(s)) = \sum_{i \in I} E(Z_i(s - 1)) M(i, j) \) for \( s \geq 1 \). In matrix form, this is \( E(Z(s))' = E(Z(s - 1))' M \), for \( s \geq 1 \) and iterating yields,

\[
E(Z(s))' = E(Z(0))' M^s, \text{ for } s \geq 0, \text{ with } E(Z(0)) = \epsilon_{i_0}.
\]

From (3), we have

\[
\rho < 1 \implies \lim_{s \to \infty} E(Z(s)) = 0, \quad \rho > 1 \implies \lim_{s \to \infty} E(Z_j(s)) = \infty \forall j \in I,
\]

\[
\rho = 1 \implies \lim_{s \to \infty} E(Z(s)) = E(Z(0))^h \nu' h.
\]

The probability-normalized left eigenvector \( \nu \) describes the stationary mean behavior of the types. More precisely, if \( E(Z(0)) = \kappa \nu \), where \( \kappa \) is the mean number of individuals at time 0, then \( E(Z(s))' = \kappa \rho^s \nu' \) so that \( \nu \) gives the mean limit behavior of the types. From (3) one gets \( \rho^{-t} E(Z_i(t)) = E(Z(0))^h \nu(j) + o(1) \) where \( o(1) \) denotes a function that vanishes at infinity. Hence,

\[
\lim_{t \to \infty} \left[ E(Z_j(t)) / \left( \sum_{i \in I} E(Z_i(t)) \right) \right] = \nu(j).
\]

Finally, we mention that a pointwise limit result on the type distribution of the population was presented in [14] in the form of the Theorem on the convergence of types which states that, almost surely on non-extinction for a supercritical multi-type \( \mathbb{P} \), one has \( \lim_{t \to \infty} Z_j(t) / \left( \sum_{i \in I} Z_i(t) \right) = \nu(j) \) for \( j \in I \).
3 Delayed multi-type branching processes

3.1 Definitions

Now we present the model which is a combination of delayed and multi-type \( \mathbf{bp} \)'s on random media. In this model, each individual is born at some time \( s \geq 0 \) and generates offspring independently of all other individuals in the process. The offspring can be of any type \( i \in I \) and each is born within a finite period of time \( D \subset \mathbb{N} = \{1, 2, \ldots\} \) following the birth of the parent. Thus, an individual born at some time \( s \) generates offspring at times in \( s + D \) and so the offspring generated at time \( s + d \) were born when the individual was age \( d \). The number of offspring of each type in \( I \) produced by the same individual at different ages in \( D \) are independent.

Let \( D = \max D < \infty \) be the maximum age at which an individual is able to produce offspring, regardless of type. One assumes that there exists more than one point in time at which an individual can produce offspring. Also, one can assume that g.c.d. \( D = 1 \), because in the contrary one can re-encode \( D \) using the initial elements of \( \mathbb{N}_0 \) to satisfy this condition. So \( D \) is of the form \( D = \{d_1, d_2, \ldots, d_\ell\} \), where \( 1 \leq d_1 < d_2 < \cdots < d_\ell = D \) and \( \ell > 1 \). Notice that a latency period of duration \( d_1 - 1 \) can be included which is a period of time following infection during which and individual is not yet contagious. In real-world applications, the set of times where individuals are contagious is typically of the form \( D = \{d_1, d_1 + 1, \ldots, d_1 + \ell - 1\} \) with \( d_1 \geq 1 \) and \( \ell \geq 2 \).

We wish to model the spread of infection as closely as possible. Toward this end, we define a random media that reproduces numbers of asymptomatic individuals together with statistics for the convalescence times of symptomatic individuals and their deaths during the infectious period. This introduces the natural constraint that individuals that die cease to be infectious, but also introduces dependence between offspring and individuals which are symptomatic. We will see later that this phenomenon can be handled with the usual tools when studying the mean behavior of the offspring process and the process which tracks the number of symptomatic individuals.

3.2 The random structure

We start by defining a space of nodes \( I \) which represents all the potential individuals involved in the process. This set has a tree structure which will be endowed with random elements in such a way as to guarantee that all individuals of the process behave independently with a common law for their convalescence times and probability of death, and with offspring laws only depending on age and type, in particular, the offspring laws do not depend on the time of birth.

A node \( b \in I \) is identified with a quartet of elements \( b = (a; i, t, l) \) where \( a \) identifies the parent node, \( i \in I \) is its type, \( t \) is its time of birth and \( l \) is an index that enumerates all nodes born to parent \( a \) of type \( i \) at time \( t \). A node born at time \( t = 0 \) has no parent, so it constitutes a root. We write \( a = \emptyset \) for such nodes. The full tree is grown by starting with a root and recursively enumerating all the child nodes of each existing node. For each node \( b = (a; i, t, l) \), we define all nodes \( c = (b; j, s, h) \) for \( j \in I \), \( s = t + d \) for \( d \in D \) and \( h \in \mathbb{N} \). Then, \( I \) denotes the set of all nodes and \( I(i) \) is the set of all type \( i \) nodes.

Next, we associate to each node \( b \in I \) independent random elements \( y(b) \) and \( (L, \varepsilon)(b) \) which are mutually independent for \( b \in I \). The random elements \( (L, \varepsilon)(b) \) are identically distributed for all \( b \in I \) and \( y(b) \) are identically distributed for the nodes \( b \in I \). For \( b = (a; i, t, l) \), \( y(b) = (y_{a,i,t,d}^{j}) : d \in D, j \in I \) is a vector comprised of \( |I| \times |D| \) random variables taking values in \( \mathbb{N}_0 \) called the potential offspring of \( b \). More precisely, \( y_{a,i,t,d}^{j} \) is the number of potential offspring of type \( j \) born to \( b \) at time \( t + d \) and its distribution only depends on \( d \) and \( (i, j) \). Let \( y_{a,i,t,d}^{j} \) be random variable with this distribution.

The random variable \( L(b) \) takes values in \( \mathbb{N}_0 \) and indicates the lifetime of \( b \) while the random variable \( \varepsilon(b) \) may depend on \( L(b) \), taking the value \( 1 \) or \( 0 \) according to whether or not \( b \) recovers or dies respectively after time \( L(b) \) has elapsed. When \( L(b) = 0 \) we set \( \varepsilon(b) = 1 \) so \( (L(b) = 0, \varepsilon(b) = 1) \). The reason for doing this is that, in the context of epidemic modelling, the lifetimes actually represent times of symptomatic infection and so any individual \( b \) for which \( L(b) = 0 \) is asymptomatic, it makes little sense
that such cases would die, particularly at the time of infection. Let \((\mathcal{L}, \varepsilon)\) denote a random element with the same law as \((\mathcal{L}, \varepsilon)(b)\). We assume
\[
\mathbb{P}(\mathcal{L} < \infty) = 1 \text{ and } \mathbb{P}(\mathcal{L} > 0) > 0,
\] the latter condition eliminating the possibility of the process having trivial dynamics. Also when modelling asymptomatic individuals we assume
\[
\mathbb{P}(\mathcal{L} = 0) > 0.
\]

Having equipped the set of nodes \(\mathcal{I}\) with a random structure, we now fix a random realization on \(\mathcal{I}^{l_0}\). Let \(b = (a; i, t, l)\) be a node. The offspring \(\xi_{a,t,l,d}^{i,j}\) generated by \(b\) are called potential because in order for \(b\) to produce offspring, one requires that it has not yet died. So, the total number of offspring of type \(j\) born at time \(t + d\) will be
\[
\xi_{a,t,l,d}^{i,j} = \zeta_{a,t,l,d}^{i,j} \left( 1 - 1 \left( \mathcal{L}(a; i, t, l) \leq d, \varepsilon(a; i, t, l) = 0 \right) \right).
\]
This depends on the pair \((\mathcal{L}, \varepsilon)\) specific to the individual. Variable \(\zeta_{a,t,l,d}^{i,j}\) incorporates the notion that no offspring are produced after the death of the parent. Let \(\xi(b) = \left( \xi_{a,t,l,d}^{i,j} : j \in \mathcal{I} \right)\). By definition the class of new random elements \((\xi(b) : b \in \mathcal{I})\) is independent. Since the distribution of \(\xi_{a,t,l,d}^{i,j}\) only depends on \(d\) and \((i, j)\), we define \(\xi_{a,t,l,d}^{i,j}\) to be a random variable having this distribution.

For each realization of the process, nodes (potential offspring) that are effectively generated are called individuals. Since the process generated by individuals do not depend on their parent except by birth time, individuals need only be identified by a triplet \((i, s, l)\) where \(i\) is the type, \(s\) the time of birth and \(l\) an index enumerating (in some fixed way) all the individuals of type \(i\) born at time \(s\). Henceforth we shall suppress the parent in the notation of the offspring counts and simply write \(\xi_{s,t,l,d}^{i,j}\) and \(\zeta_{s,t,l,d}^{i,j}\). Observe that if an individual \(b = (i, s, l)\) is generated by the process then it is ill in the interval of time \([s, s + L - 1]\) when \(\mathcal{L} > 0\). In contrast, if \(\mathcal{L} = 0\), then once infected the individual is asymptomatic during the time interval \([s, s + D]\) and is not counted as being ill (symptomatic).

### 3.3 The processes

Now we can define the offspring (or birth) process \(\mathcal{X}(s) = (\mathcal{X}_j(s) : j \in \mathcal{I})\) for \(s \in \mathbb{N}_0\): \(\mathcal{X}(s) = 0\) for \(s < 0\) and
\[
\mathcal{X}_j(s) = 1(j = i_0, s = 0) + \sum_{i \in \mathcal{I}} \sum_{d \in \mathcal{D}} \sum_{l = 1}^{\zeta_{s,t,l,d}^{i,j}} \mathcal{X}(s - d) \text{ for } s \geq 0.
\] \(\mathcal{X}_j(s)\) is the number of type \(j\) offspring born at time \(s\) for \(s > 0\). Since the process \(\mathcal{X} = (\mathcal{X}(s) : s \geq 0)\) only counts offspring, there is the implicit assumption that individuals live for a single unit of time, that is, at the time of birth.

Each individual \(b = (i, t, l)\) is endowed with a lifetime during which it is considered to be ill, written \(\mathcal{L}(i, t, l)\). All the lifetimes are independent copies of the variable \(\mathcal{L}\). Define \(\mathcal{U}(s) = (U_j(s) : j \in \mathcal{I})\) to be the number of individuals of each type ill at time \(s\). To derive an expression for \(\mathcal{U}(s)\), recall that when \(s \geq 1\), the set \(\{ (j, s, l) : l = 1, \ldots, \mathcal{X}_j(s) \}\) enumerates the type \(j\) offspring born at time \(s\). For \(s = 0\) and \(j = i_0\), \((i_0, 0, 1)\) denotes the initial individual. Since an individual with \(\mathcal{L}(i, t, l) = 0\) is never ill one has
\[
\mathcal{U}_j(s) = \sum_{c = 0}^{s - \mathcal{X}_j(s - c)} \sum_{l = 1}^{1} 1(\mathcal{L}(j, s - c, l) > c) \text{ for } s \geq 0
\] and \(\mathcal{U}(s) = 0\) for \(s < 0\). We shall call \(\mathcal{U} = (\mathcal{U}(s) : s \in \mathbb{N}_0)\) a delayed multi-type hp and \(\mathcal{U}_j(s)\) is the number of type \(j\) individuals ill at time \(s\).

In terms of the application that motivated this work, the processes \(\mathcal{X}\) and \(\mathcal{U}\) can be interpreted as follows. The birth of offspring equates to new infections while individuals that are alive correspond to clearly ill (that
is, symptomatic) individuals. Thus, $X$ gives the number of new cases (incidence under appropriate scaling) at each time point. Since the lifetime $L$ ascribed to each individual indicates how long that individual is sick, $U$ provides the total number of symptomatic cases. Individuals for whom $L = 0$ are not counted as ill by the model and since $\varepsilon = 1$ they are asymptomatic for $D$ time units and are able to infect others during that time. Note that the prevalence, in the epidemic sense, is larger than $U$ since $U$ does not count asymptomatic cases. The process that gives the number of asymptomatic cases present at each time point is $Y(s) = (Y_j(s) : j \in I)$ where

$$Y_j(s) = \sum_{c=0}^{D} \sum_{l=1}^{c} 1(L(j, s-c, l) = 0).$$

Note that the number of asymptomatic cases is slightly more than the number of asymptomatic cases capable of infecting others since newly born individuals do not reproduce until the following time unit. The prevalence is then given by the sum of $Y$ and $U$.

### 3.4 Criteria for extinction

In next result we will study extinction. It is interesting to consider the following related multi-type process $\tilde{X}$ which arranges births by generation rather than time. Let $b$ be an individual and $a$ be its parent. If $a = \emptyset$ we define $\tau(b) = 0$. Otherwise, if $a \neq \emptyset$ we recursively set $\tau(b) = \tau(a) + 1$. Thus, $\tau(b)$ connotes the depth of node $b$ in the random tree structure representation. Using $\tau$ to signify the flow of time, every individual is born precisely one time unit after its parent, which corresponds to the next generation, ignoring the instants at which births actually take place. For fixed $i \in I$ and $\tau \geq 0$, it is possible to use an index $k$ to enumerate all nodes $b = (i, t, l) \in I(i)$ of type $i$ for which $\tau(b) = \tau$, that is, born in generation $\tau$. Thus, in addition to each node being uniquely labeled by a triple $(i, t, l)$ in terms of birth instant, it also possesses a triple $(i, \tau, k)$ that uniquely identifies it from a generational standpoint and there is a bijective mapping between these two sets of labels. Consider an individual $(i, \tau, k)$. The number of offspring it produces can then be expressed in terms of the $s_{i, \tau, k}$'s introduced above as

$$s_{i, \tau, k} = \sum_{d \in D} \xi_{i, (i, \tau, k), l(i, \tau, k), d} = \sum_{d \in D} \chi_{i, (i, \tau, k), l(i, \tau, k), d}(1 - 1(L(i, t(i, \tau, k), l(i, \tau, k)) \leq d, \varepsilon(i, t(i, \tau, k), l(i, \tau, k)) = 0)).$$

Here, we have used the notation $(i, t(i, \tau, k), l(i, \tau, k))$ to map $(i, \tau, k)$ in the generational view of the random tree structure to the corresponding triple in the temporal view. Next we define the random media of the process $\tilde{X}$. It is useful to recall that $D = \{d_1, d_2, \ldots, d_\ell\}$ where $1 \leq d_1 < d_2 < \cdots < d_\ell = D$ and $\ell > 1$. Define the class of random variables $\chi(b)$ for $b = (i, t, l) \in I$ by:

$$\chi(b) = \begin{cases} \ell, & \text{if } L(b) > D \text{ or } \varepsilon(b) = 1; \\ q_i, & \text{if } d_q < L(b) \leq d_{q+1} \text{ and } \varepsilon(b) = 0. \end{cases}$$

So, the distribution of the offspring law of $s_{i, \tau, k}$ has a random distribution depending on the random media $\chi$, given by

$$s_{i, \tau, k} = \chi(i, t(i, \tau, k), l(i, \tau, k)) \sum_{q=1}^{\ell} \delta_{i, (i, \tau, k), l(i, \tau, k), d_q}.$$ 

Then the process $\tilde{X}(\tau) = (\tilde{X}_j(\tau) : j \in I)$ is defined by $\tilde{X}(\tau) = 0$ for $\tau < 0$, $\tilde{X}(0) = e_{i_0}$ and by recursion

$$\tilde{X}_j(\tau + 1) = \sum_{i \in I} \sum_{k=1}^{D} s_{i, \tau, k}, \quad \tau \geq 0. \quad (9)$$

Since the $s_{i, \tau, k}$'s are all independent, $\tilde{X}$ is a multi-type bp, see (11).
Proposition 1. Either all the processes $U$, $X$, $Y$ and $\tilde{X}$ become extinct almost surely or all of them have some positive probability of not dying out.

Proof. Solidarity of the extinction of the four processes will be established by examining four cases in which the extinction of one of the processes implies extinction of one of the others and a fifth case where the equivalence is directly established between two of the processes.

$X \Rightarrow U$. Based on (8) and the fact that lifetimes are almost surely finite, we will first show that if $X$ almost surely becomes extinct then $U$ also does. Assume that $X$ becomes extinct almost surely. This means that the extinction time

$$T^X = \inf \{ t \geq 0 : \sum_{j \in I} X_j(t + s) = 0 \text{ for all } s \geq 0 \}$$

is almost surely finite. Then, the total number of offspring born in a realization of the process $\sum_{j \in I} \sum_{t=0}^{T^X-1} X_j(t)$ is almost surely finite. Each individual $b = (j, t, l)$ generated during this process has an almost surely finite lifetime $L(j, t, l)$ (see (6)). Let $T^U$ be the extinction time of $U$. We have

$$\mathbb{P}(T^U = \infty, T^X < \infty) = \mathbb{P}(T^X < \infty, \exists j \in I, t < T^X, l \leq X_j(t) \text{ such that } L(j, t, l) = \infty).$$

Since $L(j, t, l)$ is finite we get $\mathbb{P}(T^U = \infty, T^X < \infty) = 0$. Then, as $\mathbb{P}(T^X < \infty) = 1$, one obtains that $U$ becomes extinct in a finite time almost surely.

$U \Rightarrow X$. Now we show the converse, that extinction of $U$ implies extinction of $X$. Assume $X$ does not become extinct, then

$$\mathbb{P}(T^U < \infty, \exists t_n \rightarrow \infty, \exists b_n = (i_n, t_n, 1)) > 0,$$

that is, with strictly positive probability there is an infinite sequence of times at which offspring are born after $T^U$. From (8) and the independence of lifetimes between different individuals, it follows that there is an infinite number of individuals born after $T^U$ with positive lifetimes,

$$\mathbb{P}(T^U < \infty, \exists t_n \rightarrow \infty : \exists b_n = (i_n, t_n, 1), L(b_n) > 0) > 0.$$

However, this is a contradiction. The last event must be empty since the existence of $b_n = (i_n, t_n, 1)$ with $L(b_n) > 0$ for arbitrarily large $t_n$ means that $U$ does not become extinct at time $T^U$. We conclude therefore that $X$ becomes extinct.

$X \Rightarrow Y$. Once again suppose that $X$ becomes extinct so that $T^X < \infty$ a.s. and the number of individuals born is almost surely finite. Observe that

$$Y_j(s) = \sum_{c=0}^{D} \sum_{l=1}^{X_j(s-c)} 1(L(j, s-c, l) = 0) \leq \sum_{c=0}^{D} X_j(s-c)$$

and so $T^Y \leq T^X + D < \infty$ a.s. and $Y$ becomes extinct.

$Y \Rightarrow X$. The argument to prove the converse is essentially the same as that used to prove that $U$ becomes extinct implies $X$ becomes extinct. Let $Y$ become extinct a.s. and suppose that $X$ does not become extinct. Thus,

$$\mathbb{P}(T^Y < \infty, \exists t_n \rightarrow \infty, \exists b_n = (i_n, t_n, 1)) > 0.$$

Thus, infinitely many individuals are born after time $T^Y$ with positive probability. Since lifetimes are identical and independent of all other variables, imposing the condition that $L(b_n) = 0$ merely thins the set of individuals born after time $T^Y$ and so

$$\mathbb{P}(T^Y < \infty, \exists t_n \rightarrow \infty : \exists b_n = (i_n, t_n, 1) \text{ and } L(b_n) > 0) > 0.$$

However, this is a contradiction since the set in this probability is empty. Therefore, $X$ must become extinct.
\(X \leftrightarrow \tilde{X}\). Suppose individual \(b\) is born at time \(t\). By definition
\[
[t/D] \leq \tau(b) \leq t,
\]
where \([x]\) denotes the smallest integer greater than or equal to \(x\). If \(X\) becomes extinct at time \(T_X\) then \(\tilde{X}\) becomes extinct at time \(T_X^\ast \leq T_X + 1 < \infty\). Conversely, if \(\tilde{X}\) becomes extinct at time \(T_X^\ast\), then \(X\) becomes extinct at time \(T_X \leq DT_X^\ast < \infty\). □

Remark 2. In the multi-type \(\mathcal{bp} \tilde{X}\) the random media is given by the field of independent Bernoulli random variables \(\chi(b)\) on the set of individuals \(b \in \mathcal{I}\). This is different from the random media studied in [17] where the media \(\chi\) is common to all individuals born at the same time.

3.5 Mean evolution

Now we average the offspring law of \(\xi^{i,j}_{t,d}\) over \((\mathcal{L}, \varepsilon)\). The resulting mean will only depend on \(i, j\) and \(d\). From now on, we will denote this mean law by \(p^{i,j}_d\) and, due to the independence between \(\xi(b)\) and \((\mathcal{L}, \varepsilon)(b)\), it is given by
\[
p^{i,j}_d(n) = \mathbb{E}_{(\mathcal{L}, \varepsilon)} \left( \mathbb{P}(\xi^{i,j}_d = n) \right)
= \begin{cases} 
\mathbb{P}(\xi^{i,j}_d = n)(1 - \mathbb{P}(\mathcal{L} \leq d, \varepsilon = 0)), & \text{if } n > 0, \\
\mathbb{P}(\xi^{i,j}_d = 0)(1 - \mathbb{P}(\mathcal{L} \leq d, \varepsilon = 0)) + \mathbb{P}(\mathcal{L} \leq d, \varepsilon = 0), & \text{if } n = 0.
\end{cases}
\]

Naturally, \(\sum_{n \geq 0} p^{i,j}_d(n) = 1\) and non-singularity holds.

The mean number of offspring of type \(j\) produced by an individual of type \(i\) and age \(d\) is given by
\[
M_d(i, j) = \mathbb{E}(\xi^{i,j}_d) = \sum_{n \geq 0} np^{i,j}_d(n), \text{ for } i, j \in I, d \in \mathcal{D}.
\]

This gives rise to the matrix of the mean number of offspring of each type born after an offset of \(d\) time units: \(M_d = (M_d(i, j) : i, j \in I)\).

We assume \(M_d\) is irreducible for all \(d \in \mathcal{D}\). By convention, set \(M_d = \mathbf{0}\) for any \(d \notin \mathcal{D}\), in which case the corresponding offspring distributions are of the form \(p^{i,j}_d(n) = 1(n = 0)\). Moreover, the left and right eigenvectors \(\nu_d\) and \(h_d\) of each mean matrix \(M_d\) are assumed to be normalized such that \(\nu'_d h_d = 1\) and \(\nu'_d \mathbf{1} = 1\) for all \(d \in \mathcal{D}\). Let \(\rho_d\) denote the P-F eigenvalue (spectral radius) of \(M_d\), so \(M_d h_d = \rho_d h_d\) and \(\nu'_d M_d = \rho_d \nu'_d\).

We begin by computing the mean \(\mathbb{E}_{i_0}(X(s))\). Let \(\mathcal{F}(s)\) be the \(\sigma\)-field of events up to time \(s\) generated by the random variables defining individuals born up to time \(s\). Now we condition relation (7) with respect to \(\mathcal{F}(s - 1)\), take expectations and use \(\mathbb{E}(\xi^{i,j}_{s-d,t,d}) = M_d(i, j)\) for \(s - d \geq 0, l \in \mathbb{N}\), and \(\mathbb{E}(X(s - d)) = 0\) when \(s - d < 0\), to obtain
\[
\mathbb{E}_{i_0}(X(s)) = \mathbf{1}(j = i_0, s = 0) + \sum_{i \in I} \sum_{d \in \mathcal{D}} \mathbb{E}_{i_0}(X_i(s - d))\mathbb{E}(\xi^{i,j}_{s-d,t,d})
= \mathbf{1}(j = i_0, s = 0) + \sum_{i \in I} \sum_{d \in \mathcal{D}} \mathbb{E}_{i_0}(X_i(s - d))M_d(i, j).
\]

The sequence of means satisfy \(\mathbb{E}(X(s)) = 0\) for \(s < 0\) and we obtain the following evolution equation,
\[
\mathbb{E}_{i_0}(X(s))' = \epsilon_{i_0}' \mathbf{1}(s = 0) + \sum_{d \in \mathcal{D}} \mathbb{E}_{i_0}(X(s - d))'M_d, \text{ for } s \geq 0. \quad (10)
\]
To evaluate the expected number of individuals of each type ill at time $s$, $E_{i_0}(U_j(s))$, we can use relation (8) together with $M_d = 0$ for $d \not\in \mathcal{D}$ to get

$$E_{i_0}(U_j(s)) = \sum_{c=0}^{s} E_{i_0} \left( \sum_{l=1}^{j_c} 1(\mathcal{L}(j, s - c, l) > c) \right) = \sum_{c=0}^{s} E_{i_0}(X_j(s - c))P(\mathcal{L} > c).$$

Then, substituting $E(X_j(s))$ into this gives

$$E_{i_0}(U_j(s)) = \sum_{c=0}^{s} \left[ 1(j = i_0, s - c = 0) + \sum_{i \in \mathcal{I}} \sum_{d \in \mathcal{D}} E_{i_0}(X_i(s - d - c))M_d(i, j) \right] P(\mathcal{L} > c)$$

$$= \sum_{c=0}^{s} 1(j = i_0, s = c)P(\mathcal{L} > c) + \sum_{i \in \mathcal{I}} \sum_{d \in \mathcal{D}} E_{i_0}(X_i(s - d - c))M_d(i, j)P(\mathcal{L} > c)$$

$$= 1(j = i_0)P(\mathcal{L} > s) + \sum_{i \in \mathcal{I}} \sum_{d \in \mathcal{D}} E_{i_0}(U_i(s - d))M_d(i, j).$$

We therefore have

$$E_{i_0}(U(s))' = e_{i_0}' P(\mathcal{L} > s) + \sum_{d \in \mathcal{D}} E_{i_0}(U(s - d))'M_d,$$  \hspace{1cm} s \geq 0. \hspace{1cm} (11)$$

Similarly, the expected number of asymptomatic cases is straightforward to compute and is

$$E_{i_0}(Y_j(s)) = \sum_{c=0}^{D} E_{i_0} \left( \sum_{l=1}^{j_c} 1(\mathcal{L}(j, s - c, l) = 0) \right) = P(\mathcal{L} = 0) \sum_{c=0}^{D} E_{i_0}(X_j(s - c))$$

so that

$$E_{i_0}(Y(s))' = P(\mathcal{L} = 0) \sum_{c=0}^{D} E_{i_0}(X(s - c))'.$$  \hspace{1cm} (12)$$

Once again making a substitution for $E_{i_0}(X(s))'$ gives

$$E_{i_0}(Y(s))' = P(\mathcal{L} = 0) \sum_{c=0}^{D} \left( e_{i_0}' 1(s - c = 0) + \sum_{d \in \mathcal{D}} E_{i_0}(X(s - c - d))'M_d \right)$$

$$= e_{i_0}' 1(0 \leq s \leq D)P(\mathcal{L} = 0) + \sum_{d \in \mathcal{D}} P(\mathcal{L} = 0) \sum_{c=0}^{D} E_{i_0}(X(s - c - d))M_d$$

$$= e_{i_0}' 1(0 \leq s \leq D)P(\mathcal{L} = 0) + \sum_{d \in \mathcal{D}} E_{i_0}(Y(s - d))'M_d \text{ for } s \geq 0.$$

### 4 An associated finite delayed branching process

The mean evolution of the offspring process $X$ coincides with the mean evolution of the following finite delayed branching process $\hat{X}$ whose definition includes no concept of lifetime. In $\hat{X}$, each individual $b = (i, t, l)$ generates $\hat{X}_{b,i,t,l}^j$ offspring of type $j$ at time $t + l$ according to law $p_{i,t}^{j,l}$ and the descendants are enumerated by $l \in \mathbb{N}$. This occurs independently among all individuals. The mean number of offspring generated by individuals in $\hat{X}$ at offset $d$ is the same as that for $X$ because it is

$$M_d^{i,j} = \sum_{n \geq 0} np_{d,i}^{j,n}.$$
Then, one has
\[ \mathbb{E}(\hat{X}(s)) = \mathbb{E}(X(s)), \ s \geq 0. \]
Therefore, criteria for determining mean extinction or mean explosion of the process \( X \) will be the same as for \( \hat{X} \).

Below, we will use \( \rho_A \) to denote the spectral radius of matrix \( A \). If \( A \) is a non-negative irreducible matrix, then the P-F eigenvalue is simple and it is characterized as the unique eigenvalue having strictly positive left and right eigenvectors. We make use of two facts about irreducible non-negative matrices. First, if \( A \) is irreducible and non-negative, then the P-F eigenvalue is simple and it is characterized as the unique eigenvalue having strictly positive left and right eigenvectors. We make use of two facts about irreducible non-negative matrices. First, if \( \hat{A} \) and \( \hat{B} \) are two irreducible non-negative matrices with \( \hat{A} \leq \hat{B} \) componentwise and such that \( \hat{A}(i, j) < \hat{B}(i, j) \) when \( \hat{A}(i, j) > 0 \), then \( \rho_A < \rho_B \). Second, if \( A \) is a non-negative matrix having a strictly positive left eigenvector \( \nu \) with strictly positive eigenvalue \( \alpha \), then \( \rho_A = \alpha \). In fact the stochastic matrix defined by \( P(i, j) = \alpha^{-1}\nu(j)A(j, i)/\nu(i) \) is irreducible, so it has a stationary probability measure \( \pi \). The strictly positive vector \( h(j) = \pi(j)/\nu(j) \) defines a right eigenvector of \( A \) with eigenvalue \( \alpha \). Thus \( \rho_A = \alpha \).

**Lemma 3.** There exists a unique value \( \hat{\rho} > 0 \) such that

The P-F eigenvalue of \( \sum_{d \in \mathcal{D}} \hat{\rho}^{-d} M_d \) is 1. \hfill (13)

Furthermore,
\[
\left[ \lim_{s \to \infty} \mathbb{E}(X(s)) = 0 \iff \hat{\rho} < 1 \right], \quad \left[ \lim_{s \to \infty} \mathbb{E}(X_j(s)) = \infty \forall j \in I \iff \hat{\rho} > 1 \right]
\]
\[
\lim_{s \to \infty} \mathbb{E}(X(s)) = C \text{ a finite, strictly positive vector } \iff \hat{\rho} = 1.
\]

**Proof.** Set \( [D] = \{1, \ldots, D\} \). The process \( \hat{X} \) may be viewed as a multi-type \( \mathbf{b} \) by encoding the time since birth in the type. Consider the new set of types \( \hat{I} = [D] \times I \) and define the process
\[
\hat{Z}(s) = (\hat{Z}_{d,j}(s) : (d, j) \in \hat{I}), s \geq 0, \text{ with } \hat{Z}_{d,j}(s) = \hat{X}_j(s + D - d).
\] \hfill (14)

For \( 1 < d \leq D \) one has
\[
\hat{Z}_{d,j}(s + 1) = \hat{Z}_{d-1,j}(s)
\]
while for \( d = 1 \) one obtains from (7) that
\[
\hat{Z}_{1,j}(s + 1) = \hat{X}_j(s + D) = \sum_{i \in I} \sum_{e \in \mathcal{D}} \sum_{l=1}^{\hat{X}_e(s+D-e)} \hat{Z}_{e,i}(s) M_{\xi_{s+D-e,l,e}} = \sum_{i \in I} \sum_{e \in \mathcal{D}} \sum_{l=1}^{\hat{Z}_{e,i}(s)} \hat{Z}_{e,i}(s+D-e,l,e).
\]

From (1) and (2), the mean matrix is given by
\[
\begin{cases}
\hat{M}((e, i), (d, j)) = 1(i = j)1(e = d - 1); & \text{if } e \in \mathcal{D}; \\
\hat{M}((e, i), (1, j)) = \left\{ \begin{array}{ll}
0 & \text{if } e \notin \mathcal{D} \\
\mathbb{E}(\xi_{s+D-e,l,e}) M_e(i, j) & \text{if } e \in \mathcal{D}.
\end{array} \right.
\end{cases}
\] \hfill (15)

The matrix \( \hat{M} \) is irreducible because paths starting from \( (e, i) \) can reach any other state \( (f, k) \). Moreover, \( \hat{M} \) is also mixing because g.c.d. \( \mathcal{D} = 1 \). Let \( \hat{\nu} = (\hat{\nu}(e, i) : (e, i) \in \hat{I}) \) be a left eigenvector of \( \hat{M} \) corresponding to its P-F eigenvalue \( \hat{\rho} \). For \( d > 1 \) using the first part of (15) yields
\[
\hat{\rho} \hat{\nu}(d, j) = (\hat{\nu} \hat{M})(d, j) = \sum_{(e, i) \in \hat{I}} \hat{\nu}(e, i) \hat{M}((e, i), (d, j)) = \hat{\nu}(d - 1, j),
\]
and so
\[
\hat{\nu}(d, j) = \rho^{-(d-1)} \hat{\nu}(1, j), \ d \in [D], j \in I.
\]
Using this last relation together with the second part of (15) then obtains
\[
\rho \hat{\nu}(1, j) = (\hat{\nu}^\top \hat{M})(1, j) = \sum_{(e, i) \in I} \hat{\nu}(e, i) \hat{M}((e, i), (1, j)) \\
= \sum_{e \in D} \sum_{i \in I} \hat{\nu}(e, i) M_e(i, j) = \sum_{e \in D} \sum_{i \in I} \rho^{-e-1} \hat{\nu}(1, i) M_e(i, j) \\
= \sum_{i \in I} \hat{\nu}(1, i) \left( \sum_{e \in D} \rho^{-e-1} M_e(i, j) \right).
\]

Therefore,
\[
\nu = (\nu(j) = \hat{\nu}(1, j) : j \in I)
\]
satisfies \( \nu^\top M_\rho = \nu' \), where
\[
M_\rho = \sum_{e \in D} \rho^{-e} M_e.
\]

In other words, \( \nu \) is a strictly positive left eigenvector of the matrix \( M_\rho \) whose eigenvalue is 1. It therefore must have a right eigenvector corresponding to the eigenvalue 1 and so 1 is the P-F eigenvalue of \( M_\rho \).

Next, define \( \hat{\rho} \) to be the unique strictly positive value satisfying (13), that is, \( \hat{\rho} \) is such that the P-F eigenvalue of \( M_\rho \) is unity. Its uniqueness is a result of the matrix \( M_\rho \) strictly decreasing componentwise in its positive components as \( \rho \) increases. It follows that \( \hat{\rho} \) is also the P-F eigenvalue of \( \hat{M} \).

Now, since \( E(\mathcal{X}(s)) = E(\hat{\mathcal{X}}(s)) = E(\hat{\mathcal{Z}}_{D, \nu}(s)) \), one has \( \lim_{s \to \infty} E(\mathcal{X}(s)) = 0 \) if and only if \( \lim_{s \to \infty} E(\hat{\mathcal{Z}}(s)) = 0 \). According to (9) this occurs if and only if \( \hat{\rho} < 1 \), which provides the condition for extinction. Characterizing when the limit is \( \infty \) is completely analogous to the case for extinction. Now when \( \hat{\rho} = 1 \) one gets \( \lim_{s \to \infty} E(\mathcal{Z}(s)) = E(\hat{\mathcal{Z}}(0)) h \hat{\nu} \) where \( h \) is the right Perron-Frobenius eigenvector of \( \hat{M} \) and so \( \lim_{s \to \infty} E(\hat{\mathcal{Z}}(s)) \) is a strictly positive vector. By definition (13) this implies that \( \lim_{s \to \infty} E(\mathcal{X}(s)) \) is also a strictly positive vector. \( \square \)

The Malthusian parameter is
\[
\theta = \log \hat{\rho},
\]
where \( \hat{\rho} \) is the unique solution to (13). Observe that \( \text{sgn}(\hat{\rho} - 1) = \text{sgn}(\theta) \) since the log function is strictly monotonic. Therefore, the subcritical, critical or supercritical nature of the process is determined by the Malthusian parameter \( \theta \) being negative, zero or positive respectively.

**Remark 4.** The definition of \( \hat{\rho} \) in (13) allows us to establish the following equivalences between conditions on \( \hat{\rho} \) and conditions on the P-F eigenvalue \( \overline{\nu} \) of \( \sum_{d \in D} M_d \):
\[
\hat{\rho} < 1 \Leftrightarrow \overline{\nu} < 1, \quad \hat{\rho} = 1 \Leftrightarrow \overline{\nu} = 1, \quad \hat{\rho} > 1 \Leftrightarrow \overline{\nu} > 1,
\]
or equivalently,
\[
\theta < 0 \Leftrightarrow \overline{\nu} < 1, \quad \theta = 0 \Leftrightarrow \overline{\nu} = 1, \quad \theta > 0 \Leftrightarrow \overline{\nu} > 1.
\]
The upshot of this is it is possible to determine if a multi-type delayed \( bp \) becomes extinct or explodes without knowledge of \( \hat{\rho} \) or \( \theta \). It suffices to examine the spectral radius \( \overline{\nu} \) of \( \sum_{d \in D} M_d \). This sum of matrices corresponds to the mean matrix of the multi-type \( bp \) process defined in (9) used to establish the extinction criterion in Proposition 4.

**Remark 5.** Analogous to the multi-type \( bp \) without delay, there exists a distribution of the mean evolution of types for the process \( \mathcal{X} \) which is stationary. Recall that \( M_\rho = \sum_{d \in D} \rho^{-d} M_d \) has spectral radius 1 and let \( \nu \) be the corresponding left eigenvalue normalized to sum to unity. Set \( E(\mathcal{X}(s))' = K \nu' \rho^s \) for \( s = 0, 1, \ldots, D-1 \), where \( K \) is the total initial population size. Then, (10) yields
\[
E_\nu(\mathcal{X}(s))' = \sum_{d \in D} K \nu' M_d \rho^{-d} = K \rho^s \nu' \sum_{d \in D} M_d \rho^{-d} = K \nu' \rho^s.
\]
for all \( s \geq D \), so that

\[
\forall s \geq 0, \quad \frac{\mathbb{E}(\mathcal{X}(s))'}{\mathbb{E}(\mathcal{X}(s))'1} = \nu'.
\]

5 Limit results for matrices sharing P-F eigenvectors

Before presenting the main result, we describe a condition we impose on the family of matrices \( (M_d : d \in \mathcal{D}) \) which makes it possible to handle products of matrices \( \prod_{i=1}^{L} M_d_i \) that appear when equations such as (10), (11) and (12) are recursively expanded.

The family of mean matrices \( (M_d : d \in \mathcal{D}) \) is said to share P-F eigenvectors if \( h_d = h \) and \( \nu_d = \nu \) for \( d \in \mathcal{D} \). We assume that \( h \) and \( \nu \) are scaled so that \( \nu'1 = 1 \) and \( \nu'h = 1 \). When the \( M_d \)'s share P-F eigenvectors, \( h \) satisfies \( M_d h = \rho_d h \) for all \( d \in \mathcal{D} \) and one obtains an explicit value

\[
\bar{\rho} = \sum_{d \in \mathcal{D}} \rho_d
\]

for the spectral radius of \( \sum_{d \in \mathcal{D}} M_d \). Moreover \( \nu \) and \( h \) are left and right eigenvectors of \( \sum_{d \in \mathcal{D}} \rho^{-d} M_d \) with corresponding eigenvalue \( \sum_{d \in \mathcal{D}} \rho^{-d} \rho_d \) and so we find that

\[
\sum_{d \in \mathcal{D}} \hat{\rho}^{-d} \rho_d = 1 \text{ or equivalently } \sum_{d \in \mathcal{D}} e^{-\theta d} \rho_d = 1,
\]

where \( \hat{\rho} \) was defined in (13) and \( \theta = \log \hat{\rho} \) is the Malthusian parameter. The main result considers the long-term behavior of the evolution equations (11) and (12) when the mean matrices share P-F eigenvectors. The key notion is the Malthusian parameter which is positive, zero or negative according as \( \sum_{d \in \mathcal{D}} \rho_d \) is greater than, equal to or less than unity, respectively. As was mentioned in the previous section, the subcritical, critical or supercritical nature of the process is determined by the sign of the Malthusian parameter, which is negative, zero or positive, respectively.

Taking advantage of (16), one can define a probability vector \( \vec{\beta} = (\beta_d : d \in \mathcal{D}) \) by

\[
\beta_d = \rho_d e^{-\theta d}, \quad d \in \mathcal{D}.
\]

The mean of \( \vec{\beta} \) is \( \mu(\vec{\beta}) = \sum_{d \in \mathcal{D}} d \beta_d \).

**Proposition 6.** Assume that the matrices \( (M_d : d \in \mathcal{D}) \) share P-F eigenvectors and let \( \theta \in \mathbb{R} \) be the unique solution to (17). In the subcritical case \( (\theta < 0) \), assume that \( \mathbb{E}(e^{-\theta L}) < \infty \), and in the critical case \( (\theta = 0) \), assume that \( \mathbb{E}(L) < \infty \). Then the long-term behavior of the number of symptomatic individuals is

\[
\lim_{s \to \infty} e^{-\theta s} \mathbb{E}(\mathcal{U}(s))' = (\mu(\vec{\beta}))^{-1} \mathbb{E}(\mathcal{(U(0))'} \left( \sum_{c \geq 0} \mathbb{P}(L > c) e^{-\theta c} \right) \nu'.
\]

while the limiting number of asymptomatic individuals is given by

\[
\lim_{s \to \infty} e^{-\theta s} \mathbb{E}(\mathcal{Y}(s))' = (\mu(\vec{\beta}))^{-1} (D + 1) \mathbb{P}(L = 0) \mathbb{E}(\mathcal{X(0))'} \nu'.
\]

The asymptotic behavior of the expected number of offspring is described by

\[
\lim_{s \to \infty} e^{-\theta s} \mathbb{E}(\mathcal{X}(s))' = (\mu(\vec{\beta}))^{-1} \mathbb{E}(\mathcal{X(0))'} \nu',
\]

Moreover, \( \nu' \) is the limit of the mean evolution of types for the processes \( \mathcal{X}, \mathcal{Y} \) and \( \mathcal{U} \), that is,

\[
\lim_{s \to \infty} \frac{\mathbb{E}(\mathcal{X}(s))'}{\mathbb{E}(\mathcal{X}(s))'1} = \lim_{s \to \infty} \frac{\mathbb{E}(\mathcal{Y}(s))'}{\mathbb{E}(\mathcal{Y}(s))'1} = \lim_{s \to \infty} \frac{\mathbb{E}(\mathcal{U}(s))'}{\mathbb{E}(\mathcal{U}(s))'1} = \nu'.
\]
Finally, the proportion of the expected actively reproducing (infectious) population of type $j$ that is of age $d$ at time $s$, which is given by

$$E_d^j(s) = \frac{\mathbb{E}(X_j(s-d))\mathbb{P}(L > d)}{\sum_{d' \in D} \mathbb{E}(X_j(s-d'))\mathbb{P}(L > d')}$$

satisfies

$$\lim_{s \to \infty} E_d^j(s) = \frac{\mathbb{P}(L > d)e^{-\theta d}}{\sum_{d' \in D} \mathbb{P}(L > d')e^{-\theta d'}} \text{ for } d \in D, \ j \in I. \quad (21)$$

**Remark 7.** For the result of Proposition 6 pertaining to process $\mathcal{Y}$ to hold, it is essential that $\mathbb{P}(L = 0) > 0$ to avoid trivial evolution with $\mathcal{Y}(s) = 0$ for all $s \geq 1$. In contrast, the asymptotic results concerning the number of new cases $\mathcal{X}$ and the number of active symptomatic cases $\mathcal{U}$ are valid with or without $\mathbb{P}(L = 0) > 0$.

In the following sections, we discuss the main elements involved in Proposition 6 and furnish its proof.

### 5.1 Expansion over the set of paths

Let $s \geq 1$. We can iterate equations (10), (11) and (12) backwards through time from $s$ to $0$. This is done by taking a path with elements in $D$. For $r \leq s$, the set of paths of length $r$ from $s$ to $0$ with elements in $D$ is defined by

$$\Gamma(s, r) = \{(d_1, \ldots, d_r) \in D^r : \sum_{l=1}^r d_i = s\}. \quad (22)$$

The elements in a path $(d_1, \ldots, d_r)$ represent step sizes in $D$ so that a path is represented as a sequence of $r$ steps that span the range from $0$ to $s$, so $d_r$ connects $s-d_r$ to $s$ and so on. Note that if $\Gamma(s, r) \neq \emptyset$ then $r \geq \lceil s/D \rceil$. From (10) and (11) we find

$$\mathbb{E}(\mathcal{X}(s))' = \mathbb{1}(s = 0)\mathbb{E}(\mathcal{X}(0))' + \mathbb{E}(\mathcal{X}(0))' \left(\sum_{1 \leq r \leq s} \sum_{(d_1, \ldots, d_r) \in \Gamma(s, r)} \prod_{l=1}^r M_{d_l}\right) \quad (23)$$

and

$$\mathbb{E}(\mathcal{U}(s))' = \mathbb{E}(\mathcal{U}(0))'\mathbb{P}(L \geq s) + \mathbb{E}(\mathcal{U}(0))' \sum_{c=0}^{s-1} \mathbb{P}(L > c) \left(\sum_{1 \leq r \leq s-c} \sum_{(d_1, \ldots, d_r) \in \Gamma(s-c, r)} \prod_{l=1}^r M_{d_l}\right).$$

These can be written in the form

$$\mathbb{E}(\mathcal{X}(s))' = \mathbb{1}(s = 0)\mathbb{E}(\mathcal{X}(0))' + \mathbb{E}(\mathcal{X}(0))' \left(\sum_{\lceil s/D \rceil \leq r \leq s} \Xi(s, r)\right) \quad \text{and}$$

$$\mathbb{E}(\mathcal{U}(s))' = \mathbb{E}(\mathcal{U}(0))'\mathbb{P}(L \geq s) + \mathbb{E}(\mathcal{U}(0))' \sum_{c=0}^{s-1} \mathbb{P}(L > c) \left(\sum_{\lceil (s-c)/D \rceil \leq r \leq s-c} \Xi(s-c, r)\right) \quad (24)$$

where $\Xi(s, r) = \sum_{(d_1, \ldots, d_r) \in \Gamma(s, r)} \left(\prod_{l=1}^r M_{d_l}\right).$

We will group the paths in $\Gamma(s, r)$ according to the tuples $\vec{k} = (k_d : d \in D)$ where $k_d = |\{l : d_l = d\}|$ is the number of steps of size $d$ in the path.

We need to introduce some notation. For vectors $\vec{k} = (k_d : d \in D) \in \mathbb{N}_0^D$ with non-negative coordinates, we define

$$|\vec{k}| = \sum_{d \in D} k_d, \quad m(\vec{k}) = \sum_{d \in D} d k_d.$$
We consider the following classes of tuples of paths,
\[ \Lambda(s) = \{ \vec{k} \in \mathbb{N}^D : \mathbf{m}(\vec{k}) = s \}, \quad \Lambda(s, r) = \{ \vec{k} \in \Lambda(s) : |\vec{k}| = r \}, \]
the first is the set of all paths from \( s \) to 0, and the second one is the set of these paths having length \( r \). Note that \( \Lambda(s, r) \neq \emptyset \) only when \( \lfloor s/D \rfloor \leq r \leq s \). To each \( \vec{k} \in \Lambda(s, r) \) one associates the set of sequences
\[ S(\vec{k}) = \{ (x_1, \ldots, x_r) \in D^r : \sum_{i=1}^r \mathbf{1}(x_i = d) = k_d, \forall d \in D \}. \tag{25} \]
This is the set of all the different permutations of one fixed sequence \((y_1, \ldots, y_r) \in S(\vec{k})\). So,
\[ |S(\vec{k})| = \frac{r!}{\prod_{d \in D} k_d}, \text{ where } r = |\vec{k}|. \tag{26} \]
Then, the expression \( \Xi(s, r) \) in (24) satisfies
\[ \Xi(s - c, r) = \sum_{(d_1, \ldots, d_r) \in \Gamma(s - c, r)} \left( \prod_{l=1}^r M_{d_l} \right) \tag{27} \]
\[ = \sum_{\vec{k} \in \Lambda(s - c, r)} \prod_{d \in D} \rho_d^{k_d} \sum_{(d_1, \ldots, d_r) \in S(\vec{k})} \prod_{l=1}^r \rho_{d_l}^{-1} M_{d_l}. \]
The products of subsets of the matrices \((M_d : d \in D)\) appearing in (23), (24) and (27) are a hindrance to the analytic treatment of the equations. In order to proceed, we shall impose the afore-mentioned condition that the mean matrices share P-F eigenvectors.

5.2 Matrices sharing P-F eigenvectors and runs

First we will study runs in sequences of symbols in \( D \) that will provide control over the product of matrices that appears in (24) when the matrices \((M_d : d \in D)\) share the same eigenvectors, that is, when \( h_d = h \) and \( \nu_d = \nu \). We will also assume that the additional normalizing condition \( \nu' h = 1 \) holds.

A first consequence of the fact that the matrices \((M_d : d \in D)\) share P-F eigenvectors is,
\[ \prod_{d \in D} h_d \nu_d = h \nu'. \]
This follows straightforwardly from \( h_d = h, \nu_d = \nu \) and \( \nu' h_d = \nu' h = 1 \) for all \( d, d' \in D \).

Let \( \| \cdot \| \) be the \( \infty \)-norm on vectors as well as the associated spectral norm on matrices, that is, \( \|A\| = \max_{i \in I} \sum_{j \in I} |A(i, j)| \). In particular, all stochastic matrices \( P \) have \( \|P\| = 1 \).

**Lemma 8.** Assume the matrices \((M_d : d \in D)\) share the right P-F eigenvector \( h \). Then,
\[ \forall r \geq 1, (d_1, \ldots, d_r) \in D^r : \| \prod_{l=1}^r \rho_{d_l}^{-1} M_{d_l} \| \leq \mathbf{w}(h) \text{ where } \mathbf{w}(h) = \max_{i, j \in I} (h_i / h_j). \tag{28} \]

**Proof.** From the assumptions we get that \( h \) is a right eigenvector of \( A = \prod_{l=1}^r \rho_{d_l}^{-1} M_{d_l} \) corresponding to eigenvalue 1. Then, \( P = (P(i, j : i, j \in I) \text{ defined by } P(i, j) = A(i, j)h(j)/h(i), \) is a stochastic matrix so \( \|P\| = 1 \) and
\[ \|A\| = \max_{i \in I} \sum_{j \in I} |P(i, j)h(i)/h(j)| \leq \max_{i, j \in I} (h(i)/h(j)). \]
So, (28) holds. \( \square \)
Similarly, one can show that if \((M_d : d \in \mathcal{D})\) share the left P-F eigenvector \(\nu\), then, \(\forall r \geq 1, (d_1, \ldots, d_r) \in \mathcal{D}^r, \|\prod_{l=1}^{r} \rho_d^{-1} M_{d_l} \|_1 \leq w(\nu)\), where \(\|A\|_1 = \|A\|\) is the \(L^1\)-norm.

Below we introduce the notion of runs in a sequence and we show that they allow us to handle the product of matrices appearing in \( (27) \) by controlling the shape of most part of the elements in the set of tuples \(S(\hat{k})\) defined in \( (26) \).

**Definition 9.** Let \(\kappa > 1\). We say that a sequence \((d_l : l = 1, \ldots, s)\) has a \(\kappa\)-run if there exists some \(l_0 \in \{1, \ldots, s - \kappa + 1\}\) such that \(d_l = d_{l_0}\) for \(l = l_0, \ldots, l_0 + \kappa - 1\).

**Lemma 10.** Assume \((M_d : d \in \mathcal{D})\) share right and left P-F eigenvectors \(h\) and \(\nu\) respectively. Then, for every \(\epsilon \in (0, 1)\) there exists \(\kappa(\epsilon)\) such that for all \(r \geq \kappa(\epsilon)\), the following property holds: if the sequence \((d_l : l = 1, \ldots, r)\) has a \(\kappa(\epsilon)\)-run, then

\[
\|\prod_{l=1}^{r} \rho_d^{-1} M_{d_l} - h\nu\| \leq \epsilon. \tag{29}
\]

**Proof.** Since \(\lim_{k \to \infty} \rho_d^{-k} M_d^k = h\nu\) componentwise and the matrix \(h\nu\) is strictly positive, for every \(\epsilon' \in (0, 1)\) there exists \(k(\epsilon')\) such that

\[
\forall d \in \mathcal{D}, \forall k \geq k(\epsilon') : \quad (1 - \epsilon')h\nu \leq \rho_d^{-k} M_d^k - h\nu \leq (1 + \epsilon')h\nu \quad \text{componentwise}. \tag{30}
\]

Further, if for \((d_l : l = 1, \ldots, r)\) there exists some \(l_0 \in \{1, \ldots, r - k(\epsilon') + 1\}\) such that \(d_l = d_{l_0}\) for \(l = l_0, \ldots, l_0 + k(\epsilon') - 1\), then

\[
\rho_{d_{l_0}^{-1}} M_{d_{l_0}^{-1}} (1 + \epsilon')h\nu = (1 + \epsilon')h\nu \quad \text{for } l_0 > 1,
\]

\[
(1 + \epsilon')h\nu (\rho_{d_{l_0+k(\epsilon')}}^{-1} M_{d_{l_0+k(\epsilon')}} = (1 + \epsilon')h\nu \quad \text{for } l_0 + k(\epsilon') \leq r.
\]

Therefore, an inductive argument shows that

\[
(1 - \epsilon')h\nu \leq \prod_{l=1}^{r} \rho_d^{-1} M_{d_l} \leq (1 + \epsilon')h\nu,
\]

componentwise. Hence, \(\|\prod_{l=1}^{r} \rho_d^{-1} M_{d_l} - h\nu\| \leq \epsilon'\|h\nu\|\). We take \(\kappa(\epsilon) = k(\epsilon')\) with \(\epsilon' = \epsilon/\|h\nu\|\). \(\square\)

### 5.3 Frequency of runs

Now we will study the frequency of runs in some sets of sequences that allow the product of matrices appearing in \( (27) \) to be controlled. This is done by showing that the set of tuples \(S(\hat{k})\) defined in \( (26) \) possesses long runs.

We will present some properties of finite classes of sequences whose symbols belong to \(\mathcal{D}\). Let \((x_1, \ldots, x_{r(x)})\) denote a finite sequence whose elements are in \(\mathcal{D}\).

Let \(\kappa > 1\). For a class of finite sequences \(\Gamma \subseteq \bigcup_{r \in \mathbb{N}} \mathcal{D}^r\) one defines the class of the sequences in \(\Gamma\) that avoid \(\kappa\)-runs as:

\[
\Gamma^{-\kappa} = \{(x_1, \ldots, x_{r(x)}) \in \Gamma : \exists m \leq r(x) - (\kappa - 1) \text{ such that } x_m = \ldots = x_{m+\kappa-1}\}.
\]

Note that all sequences in \(\Gamma\) with lengths strictly less than \(\kappa\) necessarily belong to \(\Gamma^{-\kappa}\). It is also useful to introduce the class of sequences that avoid \(\kappa\)-runs of a particular symbol \(d\):

\[
\Gamma^{-\kappa:d} = \{(x_1, \ldots, x_{r(x)}) \in \Gamma : \exists m \leq r(x) - (\kappa - 1) \text{ such that } x_m = \ldots = x_{m+\kappa-1} = d\} \text{ for } d \in \mathcal{D}.
\]
Of course, $\Gamma^{-\kappa} = \bigcap_{d \in D} \Gamma^{-\kappa;d}$. Now, let $\Gamma_1$ and $\Gamma_2$ be two classes of sequences contained in $\bigcup_{r \in \mathbb{N}} D^r$. First, since $(\Gamma_1 \times \Gamma_2)^{-\kappa} \subseteq \Gamma_1^{-\kappa} \times \Gamma_2^{-\kappa}$, one obtains

$$
\frac{|(\Gamma_1 \times \Gamma_2)^{-\kappa}|}{|\Gamma_1 \times \Gamma_2|} \leq \frac{|\Gamma_1^{-\kappa}|}{|\Gamma_1|} \cdot \frac{|\Gamma_2^{-\kappa}|}{|\Gamma_2|} \quad \text{and} \quad \forall d \in D, \quad \frac{|(\Gamma_1 \times \Gamma_2)^{-\kappa;d}|}{|\Gamma_1 \times \Gamma_2|} \leq \frac{|\Gamma_1^{-\kappa;d}|}{|\Gamma_1|} \cdot \frac{|\Gamma_2^{-\kappa;d}|}{|\Gamma_2|}.
$$

On the other hand, when the classes $\Gamma_1$ and $\Gamma_2$ are disjoint one has $(\Gamma_1 \cup \Gamma_2)^{-\kappa} = \Gamma_1^{-\kappa} \cup \Gamma_2^{-\kappa}$ and $(\Gamma_1 \cup \Gamma_2)^{-\kappa;d} = \Gamma_1^{-\kappa;d} \cup \Gamma_2^{-\kappa;d}$. Then,

$$
\frac{|(\Gamma_1 \cup \Gamma_2)^{-\kappa}|}{|\Gamma_1 \cup \Gamma_2|} \leq \max \left( \frac{|\Gamma_1^{-\kappa}|}{|\Gamma_1|}, \frac{|\Gamma_2^{-\kappa}|}{|\Gamma_2|} \right) \quad \text{and} \quad \forall d \in D, \quad \frac{|(\Gamma_1 \cup \Gamma_2)^{-\kappa;d}|}{|\Gamma_1 \cup \Gamma_2|} \leq \max \left( \frac{|\Gamma_1^{-\kappa;d}|}{|\Gamma_1|}, \frac{|\Gamma_2^{-\kappa;d}|}{|\Gamma_2|} \right).
$$

We have associated the set of sequences $S(\vec{k})$ defined in (25) to each $\vec{k} \in \Lambda(s)$. Now, we can associate with the class of sequences $S(\vec{k})$ the classes $S(\vec{k})^{-\kappa}$ and $S(\vec{k})^{-\kappa;d}$ for $d \in D$. Then, $|S(\vec{k})^{-\kappa}| \leq |S(\vec{k})^{-\kappa;d}|$.

**Lemma 11.** Let $\kappa > 1$ be fixed. Then, for all $\epsilon > 0$ there exists $s(\epsilon)$ such that for all $s \geq s(\epsilon)$, the inequality $|S(\vec{k})^{-\kappa}|/|S(\vec{k})| \leq \epsilon$ holds for every tuple $\vec{k} \in \Lambda(s)$.

**Proof.** Note that the structure of the classes $S(\vec{k})^{-\kappa}$ and $S(\vec{k})^{-\kappa;d}$ only depend on the unordered multiset \{\{k_d : d \in D\}\}. In fact when \{\{k_d : d \in D\}\} = \{\lambda_d : d \in D\}$, one can define a permutation \pi in $D$, such that $k_d = \lambda_{\pi(d)}$. This permutation induces a bijection between $S(\vec{k})$ and $S(\vec{\lambda})$ while its restriction maps $S(\vec{k})^{-\kappa}$ into $S(\vec{\lambda})^{-\kappa}$ and $S(\vec{k})^{-\kappa;d}$ into $S(\vec{\lambda})^{-\kappa;d}$. A $\vec{k} = (k_d : d \in D)$ is called uniform if the $(k_d)$ are as equal as they can be. More precisely, the following is required. If $r$ is a multiple of $|D|$, then $k_d = r/|D|$ for all $d \in D$. If $r$ is not a multiple of $|D|$, then for a set $D_0 \subset D$ one has: $k_d = \lfloor r/|D|\rfloor$ for $d \in D_0$ and $k_d = \lceil r/|D|\rceil$, the largest integer smaller than or equal to $r/|D|$, for $d \in D \setminus D_0$. Naturally, $D_0$ must satisfy $|D_0| = r - \lfloor r/|D|\rfloor \cdot |D|$.

Let $\Lambda^U$ denote the set of uniform tuples. For a tuple $\vec{k} \in \Lambda(s,r)$, there always exists a uniform tuple $\vec{\lambda} \in \Lambda^U$ that satisfies $|\vec{\lambda}| = |\vec{k}|$. Any such uniform tuple will be denoted by $\vec{\lambda} = U(\vec{k})$.

Take $\vec{k} \in \Lambda(s,r)$ and $d_0 \in D$ to be such that $k_d \leq k_{d_0}$ for all $d \in D$. Note that for all $\vec{\lambda} = U(\vec{k})$ one has $\lambda_{d_0} \leq k_{d_0}$. Also, it can be proved that

$$
\forall \vec{\lambda} = U(\vec{k}) : \quad |S(\vec{k})^{-\kappa;d_0}| \leq |S(\vec{\lambda})^{-\kappa;d_0}|.
$$

To see this, one uses the fact that $\lambda_{d_0} \leq k_{d_0}$ to construct a bijection from $S(\vec{k})$ into $S(\vec{\lambda})$ such that every sequence in $S(\vec{\lambda})$ that contains a $\kappa$-run of symbol $d_0$ is mapped to a sequence in $S(\vec{k})$ that also contains a $\kappa$-run of symbol $d_0$. Consequently, if $\kappa$-runs of symbol $d_0$ are avoided in $S(\vec{k})$, then they are also avoided in $S(\vec{\lambda})$ and (33) follows.

Let

$$
\delta_0 = \max\{|S(\vec{k})^{-\kappa;d}|/|S(\vec{k})| : \vec{k} \in \Lambda^U, |\vec{k}| \in \{\kappa D, \kappa D + 1\}, d \in D\}
$$

be the maximal frequency of uniform sequences of length $\kappa D$ and $\kappa D + 1$ that avoid $\kappa$-runs of symbol $d \in D$. Note that in the case $|\vec{k}| = \kappa D$, the ratios $|S(\vec{k})^{-\kappa;d}|/|S(\vec{k})|$ are all the same for every $d \in D$. Next, $\delta_0 \in (0,1)$ and one fixes $t_0$ such that $\delta_0^U \leq \epsilon$.

Take $\vec{k} \in \Lambda(s,r)$. Let $1 \leq m < m + v \leq r$ and $u > 0$. We define

$$
S(\vec{k}; u; m, m + v) = \{(y_1, \ldots, y_v) : \exists (x_1, \ldots, x_r) \in S(\vec{k}) \text{ such that } (x_m, \ldots, x_{m+v-1}) = (y_1, \ldots, y_v), \sum_{i=m}^{m+v-1} x_i = u\}.
$$
Assume \((y_1, \ldots, y_v) \in S(\bar{k}; u; m, m + v)\). Then, every permutation \((y_{\pi(1)}, \ldots, y_{\pi(v)})\) of \((y_1, \ldots, y_v)\) also belongs to \(S(\bar{k}; u; m, m + v)\). So \(S(\widetilde{\lambda}) \subseteq S(\bar{k}; u; m, m + v)\) where \(\widetilde{\lambda} \in \Lambda(u, v)\) is such that \(\lambda_d = \sum_{i=1}^{v} 1(y_i = d)\) for \(d \in D\). In this case we write \(\widetilde{\lambda} \in \Lambda(\bar{k}; u; m, m + v)\) and we have \(S(\bar{k}; u; m, m + v) = \bigcup_{\tilde{\lambda} \in \Lambda(\bar{k}; u; m, m + v)} S(\tilde{\lambda})\), where this last union is disjoint.

Now take a tuple \(\vec{k} \in \Lambda(s)\) where \(s \geq t_0 \kappa D^2\). Then we must necessarily have \(r = |\vec{k}| \geq t_0 \kappa D\). Suppose that \(r = t \kappa D + m\) where \(t \geq t_0\) and \(0 \leq m < \kappa D\). Firstly, we shall assume that \(m = 0\). We partition the class of sequences \(S(\vec{k})\) as follows:

\[
S(\vec{k}) = \bigcup_{(s_1, \ldots, s_t): \sum_{i=1}^t s_i = s} \prod_{j=1}^t S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D),
\]

where the union is disjoint and we assume the classes in the products are non-empty. From relations \((31)\) and \((32)\) one gets,

\[
|S(\vec{k})^{-\kappa}|/|S(\vec{k})| \leq \max_{(s_1, \ldots, s_t): \sum_{i=1}^t s_i = s} \prod_{j=1}^t |S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)^{-\kappa}|/|S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)|.
\]

Now, we know that \(S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D) = \bigcup_{\tilde{\lambda} \in \Lambda(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)} S(\tilde{\lambda})\) is a disjoint union. So,

\[
\frac{|S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)^{-\kappa}|}{|S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)|} \leq \max_{\tilde{\lambda} \in \Lambda(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)} \left\{ \frac{|S(\tilde{\lambda})^{-\kappa}|}{|S(\tilde{\lambda})|} \right\}.
\]

Now, from \((33)\) one has that \(|S(\tilde{\lambda})^{-\kappa}|/|S(\tilde{\lambda})| \leq |S(\widetilde{\lambda})^{-\kappa}|/|S(\widetilde{\lambda})|\) where \(\widetilde{\lambda}\) is a uniform tuple of the same length as \(\tilde{\lambda}\). Since \(|\tilde{\lambda}| = \kappa D\) from \((34)\), we get \(|S(\tilde{\lambda})^{-\kappa}|/|S(\tilde{\lambda})| \leq \delta_0\). Since \(t \geq t_0\) one finds that

\[
\prod_{j=1}^t |S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)^{-\kappa}|/|S(\vec{k}; s_j; (j-1)\kappa D + 1, j\kappa D)| \leq \delta_t^k \leq e.
\]

When \(m > 0\), instead of the \(t\) blocks \([j-1\kappa D + 1, j\kappa D]\), \(j = 1, \ldots, t\) of length \(\kappa D\), one takes \(t - m\) blocks of length \(\kappa D\) together with \(m\) of length \(\kappa D + 1\) and performs a similar construction to that carried out in the case \(m = 0\). This gives the result. \(\square\)

## 6 Proof of Proposition 6

We now have everything we need to prove Proposition 6. First, since we are considering the mean evolution of \(\mathcal{U} = (\mathcal{U}(s) : s \geq 0)\) we will relate it to a 1-type delayed \(\mathbb{P} \xi = (\xi(s) : s \geq 0)\) with offspring means \((\rho_d : d \in D)\) starting from \(\xi(0) = 1\). We assume each individual is ill for \(\mathcal{L}\) units of time. The mean \(a(s) = \mathbb{E}(\xi(s))\) of the total number of individuals born at time \(s\) satisfies a similar evolution equation to \((11)\), which is

\[
a(s) = \mathbb{P}(\mathcal{L} > s) + \sum_{d \in D} a(s - d)\rho_d, \quad s \geq 0.
\]

The 1-type process shares the same Malthusian parameter \(\theta\) as \(\mathcal{U}\) and thus the two processes are either subcritical, critical or supercritical together. Recall the probability vector \(\vec{\beta} = (\beta_d = \rho_de^{-\theta d} : d \in D)\) and consider the sequence \((c(s) = a(s)e^{-\theta s} : s \geq 0)\). This satisfies the renewal equation

\[
c(s) = e^{-\theta s}\mathbb{P}(\mathcal{L} > s) + \sum_{k \geq 0} c(s - k)\beta_k, \quad s \geq 0.
\]
The solution to this equation has the form

\[ c(s) = \sum_{c=0}^{s} e^{-\theta(s-c)} \mathbb{P}(L > s - c) u(c), \quad s \geq 0, \]

where \( \sum_{t=1}^{\infty} u(t) \) is the renewal function associated with \( \bar{\beta} \) and \( u(t) \) is a discrete analogue to the renewal density. On one hand, if \( \theta = 0 \), then \( \sum_{s \geq 0} \mathbb{P}(L > s) < \mathbb{E}(L) < \infty \) by assumption. On the other hand when \( \theta \neq 0 \), one has

\[ \sum_{s \geq 0} \mathbb{P}(L > s) e^{-\theta s} = \frac{\mathbb{E}(e^{-\theta L}) - 1}{e^{-\theta} - 1}. \]

In the supercritical case this is finite because \( \theta > 0 \) and in the subcritical case this is also finite since \( \mathbb{E}(e^{-\theta L}) \) is assumed to be finite. Thus, \( \sum_{s \geq 0} \mathbb{P}(L > s) e^{-\theta s} < \infty \) and this makes \( \mathbb{P}(L > s) e^{-\theta s} \) directly Riemann integrable. From the renewal theorem, see Proposition 4.7 in Chapter V of [1], we then get

\[ \lim_{s \to \infty} e^{-\theta s} \mathbb{E}(\zeta(s)) = \frac{1}{\mu(\bar{\beta})} \sum_{t \geq 0} \mathbb{P}(L > t) e^{-\theta t}. \] (35)

Recall that \( \mu(\bar{\beta}) = \sum_{d \in \mathcal{D}} d \beta_d \) is the mean value of the distribution \( \bar{\beta} \). Lemma 1 of [8] showed this for the non-arithmetic case in a general framework. Also see Proposition 1.1 in [15].

Expression (24) also shows that the mean of the associated 1-type delayed bp starting from \( \zeta(0) = 1 \) is

\[ \mathbb{E}(\zeta(s)) = \mathbb{P}(L > s) + \sum_{c=0}^{s-1} \mathbb{P}(L > c) \sum_{1 \leq r \leq s-c} \sum_{k \in \Lambda(s-c,r)} r! \prod_{d \in \mathcal{D}} \rho_d^{k_d}. \]

Then, by using (20) one obtains

\[
e^{-\theta s} \mathbb{E}(\zeta(s)) = \mathbb{P}(L > s) e^{-\theta s} + e^{-\theta s} \sum_{c=0}^{s-1} \mathbb{P}(L > c) \left( \sum_{1 \leq r \leq s-c} \sum_{k \in \Lambda(s-c,r)} |S(k)| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) = \sum_{c=0}^{s} \mathbb{P}(L > c) e^{-\theta c} u(s-c)\]

where \( u(0) = 1 \) and for \( t > 0 \),

\[ u(t) = \left( \sum_{1 \leq r \leq t} \sum_{k \in \Lambda(t,r)} |S(k)| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) e^{-\theta t} = \sum_{1 \leq r \leq t} \sum_{k \in \Lambda(t,r)} |S(k)| \prod_{d \in \mathcal{D}} \beta_d^{k_d}, \]

which is the renewal density for \( (\beta_d : d \in \mathcal{D}) \). From (35) we have

\[
\lim_{s \to \infty} e^{-\theta s} \mathbb{E}(\zeta(s)) = \lim_{s \to \infty} e^{-\theta s} \sum_{c=0}^{s} \mathbb{P}(L > c) \left( \sum_{1 \leq r \leq s-c} \sum_{k \in \Lambda(s-c,r)} |S(k)| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) = \mu(\bar{\beta})^{-1} \sum_{t \geq 0} \mathbb{P}(L > t) e^{-\theta t}. \] (36)

On the other hand, notice that from (27) one has

\[ \Xi(s - c, r) = \sum_{k \in \Lambda(s-c,r)} \prod_{d \in \mathcal{D}} \rho_d^{k_d} \sum_{(d_1, \ldots, d_r) \in S(k)} \prod_{l=1}^{r} \rho_d^{-1} M_{d_l}. \]
Assume for the moment that \( \mathcal{L} \) is bounded by \( L \), so \( \mathbb{P}(\mathcal{L} \geq t) = 0 \) for \( t > L \). From (30) it holds for all \( k \geq \kappa(\varepsilon) \) that \( \| \rho_d^{-k} M_d^{k - \text{th}} - h' \| \leq \varepsilon \) for all \( d \in \mathcal{D} \). On the other hand, from Lemma 11 there exists \( s(\varepsilon) \) such that for all \( s \geq s(\varepsilon) + L \) one has that every tuple \( \tilde{k} \in \Lambda(s - c, r) \) with \( c \leq L \) satisfies \( |S(\tilde{k})|^{-\kappa(\varepsilon)} / |S(\tilde{k})| \leq \varepsilon \). Since for every sequence \( (d_1, \ldots, d_r) \in S(\tilde{k}) \setminus S(\tilde{k})^{-\kappa(\varepsilon)} \) there exists a \( \kappa(\varepsilon) \)-run, relation (29) in Lemma 10 guarantees that

\[
\| \prod_{l=1}^{r} \rho_d^{-1} M_{d_l} - h' \| \leq \varepsilon.
\]

Hence, for all \( s \geq s(\varepsilon) + L \),

\[
\| \Xi(s - c, r) - \left( \sum_{\tilde{k} \in \Lambda(s - c, r)} |S(\tilde{k})| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) h' \| \leq \left( \sum_{\tilde{k} \in \Lambda(s - c, r)} |S(\tilde{k})| \setminus S(\tilde{k})^{-\kappa(\varepsilon)} \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) \varepsilon
\]

\[
+ \sum_{\tilde{k} \in \Lambda(s - c, r)} \prod_{d \in \mathcal{D}} \rho_d^{k_d} \sum_{(d_1, \ldots, d_r) \in S(\tilde{k})^{-\kappa(\varepsilon)}} \| \prod_{l=1}^{r} \rho_d^{-1} M_{d_l} \|.
\]

From (28) one has \( \prod_{l=1}^{r} \rho_d^{-1} M_{d_l} \| \leq \omega(h) \). On the other hand \( |S(\tilde{k})|^{-\kappa(\varepsilon)} \leq \varepsilon |S(\tilde{k})| \). Setting \( C = 1 + \omega(h) \), we thus have

\[
\| \Xi(s - b, r) - \left( \sum_{\tilde{k} \in \Lambda(s - b, r)} |S(\tilde{k})| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) h' \| \leq C \varepsilon \left( \sum_{\tilde{k} \in \Lambda(s - b, r)} |S(\tilde{k})| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right).
\]

Therefore,

\[
\sum_{c=0}^{L} \mathbb{P}(\mathcal{L} > c) \sum_{(s - c)/D \leq r \leq s - c} \| \Xi(s - c, r) - \left( \sum_{\tilde{k} \in \Lambda(s - c, r)} |S(\tilde{k})| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) h' \|
\]

\[
\leq C \varepsilon \left( \sum_{c=0}^{L} \mathbb{P}(\mathcal{L} > c) \sum_{(s - c)/D \leq r \leq s - c} \sum_{\tilde{k} \in \Lambda(s - c, r)} |S(\tilde{k})| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right).
\]

So, from (30) we get

\[
\lim_{s \to \infty} e^{-\theta s} \sum_{c=0}^{L} \mathbb{P}(\mathcal{L} > c) \sum_{(s - c)/D \leq r \leq s - c} \| \Xi(s - c, r) - \left( \sum_{\tilde{k} \in \Lambda(s - c, r)} |S(\tilde{k})| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) h' \|
\]

\[
\leq C \varepsilon \mu(\beta)^{-1} \sum_{c=0}^{L} \mathbb{P}(\mathcal{L} > c) e^{-\theta c}.
\]

Letting \( \varepsilon \to 0 \) and using (30) again shows that

\[
\lim_{s \to \infty} e^{-\theta s} E(\mathcal{U}(s)) = E(\mathcal{U}(0))' h' e_j \mu(\beta)^{-1} \left( \sum_{c=0}^{L} \mathbb{P}(\mathcal{L} > c) e^{-\theta c} \right).
\]

So, we have proved (17) in the case of bounded lifetimes \( \mathcal{L} \leq L \).

Next, we shall extend this to the general case. Take \( \mathcal{L}^n \) to be an increasing sequence of bounded lifetimes that converge to \( \mathcal{L} \). Then, for all \( c \geq 0 \), \( \mathbb{P}(\mathcal{L}^n > c) \) increases to \( \mathbb{P}(\mathcal{L} > c) \). Let us add a superscript \( n \) on all quantities in which the lifetime is \( \mathcal{L}^n \). We have

\[
E(\zeta(s)) - E(\zeta^n(s)) = \left( \mathbb{P}(\mathcal{L} > s) - \mathbb{P}(\mathcal{L}^n > s) \right)
\]

\[
+ \sum_{c=0}^{s-1} \left( \mathbb{P}(\mathcal{L} > c) - \mathbb{P}(\mathcal{L}^n > c) \right) \left( \sum_{(s - c)/D \leq r \leq s - c} \sum_{\tilde{k} \in \Lambda(s - c, r)} |S(\tilde{k})| \prod_{d \in \mathcal{D}} \rho_d^{k_d} \right) \left( \sum_{c=0}^{L} \mathbb{P}(\mathcal{L} > c) e^{-\theta c} \right).
\]
Observe that the Malthusian parameter \( \theta \) does not depend on \( L \). By using (36) yet again and taking advantage of the fact that \( \mu(\bar{\beta})^{-1} \sum_{t \geq 0} \mathbb{P}(L^n > t)e^{-\theta t} \geq \mu(\bar{\beta})^{-1} \sum_{t \geq 0} \mathbb{P}(L > t)e^{-\theta t} \) as \( n \to \infty \), we obtain
\[
\lim_{n \to \infty} \lim_{s \to \infty} e^{-\theta s} (\mathbb{E}(\zeta(s)) - \mathbb{E}(\zeta^n(s))) = 0.
\] (39)

On the other hand
\[
\mathbb{E}(U(s))' - \mathbb{E}(U^n(s))' = (\mathbb{P}(L > s) - \mathbb{P}(L^n > s)) + \sum_{c=0}^{s-1} (\mathbb{P}(L > c) - \mathbb{P}(L^n > c)) \left( \sum_{|c(D)| \leq r \leq s-c} \Xi(s-c, r) \right),
\]
which is non-negative componentwise. Consequently, we have
\[
e^{-\theta s} (\mathbb{E}(U(s))' - \mathbb{E}(U^n(s))') = e^{-\theta s} (\mathbb{P}(L > s) - \mathbb{P}(L^n > s)) \mathbb{E}(U(0))' + \sum_{c=0}^{s-1} (\mathbb{P}(L > c) - \mathbb{P}(L^n > c)) \sum_{|c(D)| \leq r \leq s-c} \sum_{k \in \Lambda(s-c, r)} \rho^k_d \prod_{d \in D} \rho^k_d.
\]
Then,
\[
e^{-\theta s} ||\mathbb{E}(U(s))' - \mathbb{E}(U^n(s))'|| \leq e^{-\theta s} (\mathbb{P}(L > s) - \mathbb{P}(L^n > s)) ||\mathbb{E}(U(0))'|| + h(\bar{\beta})e^{-\theta s} \left( \sum_{c=0}^{s-1} (\mathbb{P}(L > c) - \mathbb{P}(L^n > c)) \sum_{|c(D)| \leq r \leq s-c} |S(k)| \prod_{d \in D} \rho^k_d \right).
\]
From (38) and (39), it follows that
\[
\lim_{n \to \infty} \lim_{s \to \infty} e^{-\theta s} ||\mathbb{E}(U(s))' - \mathbb{E}(U^n(s))'|| = 0
\]
and hence
\[
\lim_{s \to \infty} e^{-\theta s} \mathbb{E}(U(s))' = \lim_{n \to \infty} \lim_{s \to \infty} e^{-\theta s} \mathbb{E}(U^n(s))'.
\]
At this point, an application of (37) completes the proof of (17).

Next, to prove the asymptotic behavior of the mean offspring process, observe that the offspring counts can be viewed as \( \xi_{l,t,i} = \xi_{l,t,0} (1 - 1(\bar{Z}(i, t, l) \leq d, \tau(i, t, l) = 0)) = \xi_{l,t,0} \), where \( \xi_{l,t,0} \) and \( \tau(i, t, l) = 1 \) a.s. In other words, \( \mathcal{X}_j(s) \) is identical to an offspring process in which all individuals have lifetimes and always recover and the original random media is incorporated into the potential offspring counts. Thus, it satisfies both (7) and (8). Hence (19) can be established by applying (17) directly to \( \mathcal{X}_j(s) \) and noting that \( \mathbb{P}(\bar{Z} > 0) = 1 \).

Now, considering the mean number of asymptotic cases, we have
\[
\lim_{s \to \infty} e^{-\theta s} \mathbb{E}(\mathcal{X}(s))' = \lim_{s \to \infty} e^{-\theta s} \mathbb{P}(\mathcal{L} = 0) \sum_{c=0}^{D} \mathbb{E}(\mathcal{X}(s))' = \mathbb{P}(\mathcal{L} = 0) \sum_{c=0}^{D} \lim_{s \to \infty} e^{-\theta s} \mathbb{E}(\mathcal{X}(s))'
\]
\[
= \mathbb{P}(\mathcal{L} = 0) \sum_{c=0}^{D} \mu(\bar{\beta})^{-1} \mathbb{E}(\mathcal{X}(0))' \mathcal{h} \nu'
\]
\[
= \mu(\bar{\beta})^{-1}(D + 1)\mathbb{P}(\mathcal{L} = 0) \mathbb{E}(\mathcal{X}(0))' \mathcal{h} \nu'.
\]

21
Focusing on the mean evolution of types, relation (20) follows straightforwardly from (17), (19) and (18) since \( \nu' \mathbf{1} = 1 \). As the calculation of the limits is the same for all three processes, we only present the case of \( \mathcal{U} \).

\[
\lim_{s \to \infty} \frac{\mathbb{E}(\mathcal{U}(s))'}{\mathbb{E}(\mathcal{U}(s))} = \frac{\mu(\tilde{\beta})^{-1}(\mathbb{E}(\mathcal{U}(0))'/h) \left( \sum_{i \geq 0} \mathbb{P}(L > t)e^{-\theta t} \right) \nu'}{\mu(\tilde{\beta})^{-1}(\mathbb{E}(\mathcal{U}(0))'/h) \left( \sum_{i \geq 0} \mathbb{P}(L > t)e^{-\theta t} \right) \nu' \mathbf{1}} = \nu'.
\]

Finally, one can write \( \mathcal{E}_d^i(s) \) as

\[
\mathcal{E}_d^i(s) = \frac{e^{-\theta(s-d)}\mathbb{E}(\chi_j(s-d))\mathbb{P}(L > d)e^{-\theta d}}{\sum_{b' \in \mathcal{D}} e^{-\theta(s-d')}\mathbb{E}(\chi_j(s-b'))\mathbb{P}(L > b')e^{-\theta d'}}.
\] (40)

As \( e^{-\theta(s-d)}\mathbb{E}(\chi_j(s-d)) \) tends to the same limit for all \( d \in \mathcal{D} \) and \( j \in I \), taking limits as \( s \to \infty \) in (40) yields (21). □

7 Final comments and results

We conclude by discussing two subjects related to the model:

- Some properties of classes of matrices sharing the same P-F eigenvectors; and
- The relationship of the delayed process with Fibonacci sequences as well as a Fibonacci-type process.

7.1 Matrices sharing P-F eigenvectors

Let \( M \) be an irreducible non-negative matrix which has maximal eigenvalue \( \rho > 0 \) and corresponding left and right P-F eigenvectors \( \nu \) and \( h \) respectively. The associated stochastic matrix \( P_M \) defined by \( P_M(i, j) = \rho^{-1}M(i, j)h(j)/h(i) \) has the componentwise product \( \nu h = (\nu(i)h(i) : i \in I) \) as its unique stationary distribution. Conversely, let \( P \) be an irreducible stochastic matrix with stationary distribution \( \pi \). Fix \( \rho > 0 \), a vector \( h > 0 \) and define

\[
M(i, j) = \rho P(i, j)h(i)/h(j).
\] (41)

Then, \( M \) is an irreducible non-negative matrix. It is straightforward to see that \( Mh = \rho h \) and so \( \rho \) is the P-F eigenvalue corresponding to the right eigenvector \( h \). It is easily checked that \( \nu = \pi/h = (\pi(i)/h(i) : i \in I) \) is the left P-F eigenvector for \( M \). By definition, \( P \) is the stochastic matrix associated with \( M \). Therefore, the class of irreducible non-negative matrices with fixed P-F eigenvectors \( \nu \) and \( h \) can be described by combining arbitrary values \( \rho > 0 \) and stochastic matrices with stationary distribution \( \pi = (\pi(i) = h(i)\nu(i) : i \in I) \). Thus, the class of families of matrices sharing P-F eigenvectors is non-trivial and there is a straightforward mechanism for generating such families on demand. Incidentally, an alternative construction to (11) (based on the left eigenvector instead of the right eigenvector) can be carried out by taking \( M(i, j) = \rho(\nu(j)/\nu(i))P(j, i) \). This is a time-reversed construction because \( P^* \) given by \( P^*(i, j) = (\pi(j)/\pi(i))P(j, i) \) is the transition matrix of the time reverse of the Markov chain defined by \( P \).

Next, we examine a commutativity relation between matrices \( (M : d \in \mathcal{D}) \). This property requires that the order of the types of the individuals in chains of disease transmission should not matter. In the following result, we show that if the mean matrices commute, then they share P-F eigenvectors.

**Proposition 12.** Assume that the family of matrices \( (M_d : d \in \mathcal{D}) \) is commutative, that is \( M_d M_{d'} = M_{d'} M_d \) for \( d, d' \in \mathcal{D} \). Then, the matrices share the same right and left P-F eigenvectors.
Proof. Let \( d, d' \in D \). The commutativity property and relation (1) implies

\[
\gamma_d \gamma_{d'} = \lim_{t \to \infty} \gamma_d^{-t} M_d^t \gamma_{d'} = \lim_{t \to \infty} \gamma_{d'}^{-t} M_{d'}^t \gamma_d = \gamma_{d'} \gamma_d.
\]

So, \( \delta \gamma_d \gamma_{d'} = c \delta \gamma_d \gamma_{d'} \) where \( \delta = \gamma_d \gamma_{d'} \) and \( \epsilon = \gamma_{d'} \). Then, \( \delta \gamma_d(i) \gamma_{d'}(j) = c \gamma_d(i) \gamma_{d'}(j) \) for all \( i, j \in I \). Summing over \( j \) leads to \( \gamma_d = (\delta / \epsilon) \gamma_{d'} \) and hence \( \delta = \gamma_{d'} \gamma_d = \gamma_d \delta / \epsilon = \delta / \epsilon \). Similarly, one can show that \( \epsilon = \epsilon / \delta \). From this we obtain \( \delta = 1 = \epsilon \) and hence we can conclude that \( \gamma_d = \gamma_{d'} \). It also follows that \( \nu_d = \nu_{d'}(j) \) for all \( j \in I \), that is, \( \nu_d = \nu_{d'} \). \( \square \)

In the context of epidemiological modeling the family of mean matrices \( (M_d : d \in D) \) encompasses two different aspects of the system. Each matrix describes the effect of geographical/regional interaction: component \( (i, j) \) indicates the mean number of infections caused in type \( j \) individuals by type \( i \) individuals who are already infected. The series of matrices itself encapsulates temporal effects such as variation in contagiousness after contracting an infectious disease. The simplest way of modeling this is to treat geographical and temporal effects as orthogonal, in which case \( M_d = a_d M \), where the matrix \( M \) expresses geographical interaction effects while the \( a_d \)'s provide time-dependent changes in the overall level of contagiousness. Due to their very simple form, the \( M_d \)'s all have the same left and right P-F eigenvectors, but their P-F eigenvalues vary according to the \( a_d \)'s. Here, as the mean matrices are constructed in a very restricted way, they share P-F eigenvectors and they also commute.

However, Proposition 14.3 shows that families of mean matrices that are commutative also share P-F eigenvectors. Thus, rescalings of a single matrix, commutativity and sharing P-F eigenvectors are increasingly more relaxed conditions on the structure of the mean matrices. The requirement that the family of mean matrices share P-F eigenvectors is the most general condition we have found that allows the product of mean matrices to be controlled as the product of a single matrix.

### 7.2 Relationship to Fibonacci sequences

Consider the special 1-type supercritical case in which \( \rho_d = 1 \) for \( d \in D \). From now on assume \( D = |D| = \{1, \ldots, D\} \) for a \( D > 1 \).

Recall that the number of paths of length \( r \) from \( s \) to \( 0 \) with elements in \( D \) is \(|\Gamma(s, r)|\), see (22). From (23) \( \mathbb{E}(X(s)) = \mathbb{E}(X(0)) \gamma(s) \), where \( \gamma(s) = \sum_{|s/D| \leq r \leq s} |\Gamma(s, r)| \) is the total number of paths from \( s \) to \( 0 \). It is straightforward to check that \( \gamma(s) \) satisfies the following properties:

\[
\gamma(s) = \sum_{d=1}^{D} \gamma(s - d) \quad \text{with initial conditions } \gamma(0) = 1 \text{ and } \gamma(s) = 0 \text{ for } s < 0.
\]

Therefore, \( \gamma(s) : s \geq 1 \) is the \( D \)-Fibonacci sequence. It is increasing in \( s \) and may be expressed analytically in the form \( \gamma(s) = \sum_{d=1}^{D} c_d \phi_d^s \) with \( (c_d : d \in D) \) real constants and \( (\phi_d : d \in D) \) are the roots of \( g(y) = 0 \), where \( g(y) = y^D - \sum_{d=0}^{D-1} g_d y^d \). Since \( g(1) \neq 0 \) and \( g(y) = (-y^{D+1} + 2y^D - 1) / (1 - y) \) for \( y \neq 1 \), it follows that

\[
g(y) = 0 \Rightarrow [y^{D+1} - 2y^D + 1 = 0, y \neq 1].
\]

By Lemma 3.6 of 13, \( g(s) \) has a unique real root in the interval \((1, 2)\), say \( \Phi_D \), which tends to 2 as \( D \to \infty \). All the other roots have moduli strictly smaller than 1. The root \( \Phi_D \) corresponds to the \( D \)-Fibonacci constant. The 2-Fibonacci constant is the golden ratio \( \Phi_2 = (1 + \sqrt{5}) / 2 \) and Theorem 3.9 of 13 provides a series representation of \( \Phi_D \). The \( D \)-Fibonacci constant dominates \( \gamma(s) \) as \( s \) becomes large, which gives rise to the following limit behavior (see Corollary 3.7 in 13):

\[
\Phi_D = \lim_{s \to \infty} \frac{\gamma(s + 1)}{\gamma(s)} \quad \text{and} \quad \lim_{s \to \infty} \frac{\gamma(s)}{\Phi_D} = C' \text{ a constant.}
\]

In the current set-up, the relation (13) defining the Malthusian parameter becomes \( \sum_{d=1}^{D} e^{-d \theta} = 1 \), so \( \theta > 0 \). Let \( \eta = e^{-\theta} \). It is evident that \( \eta^{-1} = \Phi_D \), because the \( D \)-Fibonacci constant is the unique real root larger
than 1. We have
\[
\mu(\beta)^{-1} = \sum_{d=1}^{s} d \eta^d = \frac{\eta}{(1-\eta)^2} ((D+1)(1-\eta^D) - D + D\eta^{D+1}) = \frac{D+1 - 2D\eta}{1-\eta},
\]
where relation \((12)\) was used to get \(1 - 2\eta + \eta^{D+1} = 0\), which is equivalent to \(\eta(1-\eta^D) = 1 - \eta\) for \(\eta \in (0,1)\).

Finally, we obtain \(\mu(\beta)^{-1} = (\Phi_D - 1)/((D + 1)\Phi_D - 2D)\).

Now, we compare the 1-type delayed \(bp\) with a reproductive period of 2 time units to the process introduced by Heyde in [11] called a Fibonacci or lagged Bienaymé-Galton-Watson) \(bp\) in the singular case where the offspring distribution is degenerate and produces exactly one individual. The lagged Fibonacci process, denoted by \((W(s): s \geq 0)\) starts with initial conditions \(W(0)\) and \(W(1)\). For \(s \geq 2\), it is defined by
\[
W(s) - W(s-1) = \sum_{l=1}^{W(s-2)} \xi_{s,l},
\]
where the random variables \((\xi_{s,l} : l \geq 1, s \geq 2)\) are i.i.d. with common distribution \(p = (p(n) : n \geq N_0)\) having mean \(m = \sum_{n \geq 1} n p(n)\). We take the initial conditions to be \(W(0) = 1\) and \(W(1) \sim p\).

On the other hand, a delayed 1-type \(bp\) \(\hat{X}\) with \(D = \{1,2\}\) and offspring laws \(p_1 = p_2 = p\) where each individual is alive for one unit of time satisfies the relation \((7)\) which becomes
\[
\hat{X}(s) = \sum_{l=1}^{\hat{X}(s-1)} \xi_{s-1,l;1} + \sum_{l=1}^{\hat{X}(s-2)} \xi_{s-2,l;2}.
\]

Then, when \(p(1) = 1\) or equivalently the variables are deterministic with \(\xi \equiv 1\), then both processes \(W\) and \(\hat{X}\) coincide and satisfy the 2-Fibonacci evolution \(\Psi(s) = \Psi(s-1) + \Psi(s-2)\) with \(\Psi(0) = 1 = \Psi(1)\). A similar comparison can be made between the means of the two processes when the mean of the offspring distribution is \(m = 1\). In this case, \(E(W(s)) = E(\hat{X}(s)) = \Psi(s)\) for \(s \geq 0\).

However, for other cases the processes \(W\) and \(\hat{X}\) behave differently from each other. In fact for \(s > 0\) one has the following comparison between mean values: if \(m > 1\) then \(E(\hat{X}(s)) > E(W(s))\) and if \(m < 1\) then \(E(\hat{X}(s)) < E(W(s))\).

Acknowledgments

This work was supported by the Center for Mathematical Modeling ANID Basal Projects ACE210010 and FB210005.

References

[1] S. Asmussen. Applied Probability and Queues, volume 51 of Stochastic Modelling and Applied Probability. Springer-Verlag, New York, second edition edition, 2010.

[2] Soren Asmussen. On the role of a certain eigenvalue in estimating the growth rate of a branching process. Austral. J. Statist., 24(2):151–159, 1982.

[3] Krishna B. Athreya and Peter N. Ney. Branching Processes. Springer-Verlag, New York, 1972.

[4] N. Becker. Estimation for discrete time branching processes with application to epidemics. Biometrics, 33(3):515–522, 1977.

[5] J.D. Biggins, H. Cohn, and O. Nerman. Multi-type branching in varying environment. Stoch. Process. Applic., 83(2):357–400, 1999.
[6] J.D. Biggins and A.E. Kyprianou. Measure change in multi-type branching. *Adv. Appl. Probab.*, 36(2):544–581, 2004.

[7] R.A. Doney. A limit theorem for a class of supercritical branching processes. *J. Appl. Probab.*, 9(4):707–724, 1972.

[8] R.A. Doney. On single- and multi-type general age-dependent branching processes. *J. Appl. Probab.*, 13(2):239–246, 1976.

[9] N.M. Ferguson, D. Laydon, G. Nedjati-Gilani, N. Imai, K. Ainslie, M. Baguelin, S. Bhattacharya, A. Boonyasiri, Z. Cucunubá, G. Cuomo-Dannenburg, and A. Dighe. Report 9: Impact of non-pharmaceutical interventions (npis) to reduce covid-19 mortality and healthcare demand. Imperial College COVID-19 Response Team, London, 16 March 2020. 
https://www.imperial.ac.uk/media/imperial-college/medicine/sph/ide/gida-fellowships/Imperial-College-COVID19-NPI-modelling-16-03-2020.pdf

[10] T. Harris. *The Theory of Branching Processes*. Springer-Verlag, Berlin, 1963.

[11] C.C. Heyde. On Fibonacci (or lagged Bienaymé-Galton-Watson) branching processes. *J. Appl. Probab.*, 18(3):583–591, 1981.

[12] A. Joffe and F. Spitzer. On multitype markov branching processes with \( \rho \leq 1 \). *J. Math. Anal. Applic.*, 19(3):409–430, 1967.

[13] H. Kesten and B. Stigum. A limit theorem for multidimensional galton-watson processes. *Ann. Math. Statist.*, 37:1211–1223, 1966.

[14] T.G. Kurtz, R. Lyons, R. Pemantle, and Y. Peres. A conceptual proof of the Kesten-Stigum theorem for multi-type branching processes. In K. Athreya and P. Jagers, editors, *Classical and Modern Branching Processes*, pages 181–185, New York, 1997. Springer.

[15] O. Nerman. On the convergence of supercritical general (C-M-J) branching processes. *Z. Wahrscheinlichkeitstheorie verw. Gebiete*, 57:365–395, 1981.

[16] E. Seneta. *Nonnegative Matrices and Markov Chains*. Springer Series in Statistics. Springer-Verlag, New York, 2nd edition, 1981.

[17] David Tanny. On multitype branching processes in a random environment. *Adv. Appl. Probab.*, 13(3):464–497, 1981.

[18] D.A. Wolfram. Solving generalized fibonacci recurrences. *Fibonacci Quart.*, 36(2):129–145, 1998.