ON SIR-MODELS WITH MARKOV-MODULATED EVENTS: LENGTH OF AN OUTBREAK, TOTAL SIZE OF THE EPIDEMIC AND NUMBER OF SECONDARY INFECTIONS

E. Almaraz
Department of Statistics and Operations Research
School of Mathematical Sciences
Complutense University of Madrid, 28040 Madrid, Spain

A. Gómez-Corral*
Instituto de Ciencias Matemáticas CSIC-UAM-UC3M-UCM
Calle Nicolás Cabrera 13-15, Campus de Cantoblanco UAM
28049 Madrid, Spain

(Communicated by Gail Wolkowicz)

Abstract. A Markov-modulated framework is presented to incorporate correlated inter-event times into the stochastic susceptible-infectious-recovered (SIR) epidemic model for a closed finite community. The resulting process allows us to deal with non-exponential distributional assumptions on the contact process between the compartment of infectives and the compartment of susceptible individuals, and the recovery process of infected individuals, but keeping the dimensionality of the underlying Markov chain model tractable. The variability between SIR-models with distinct level of correlation is discussed in terms of extinction times, the final size of the epidemic, and the basic reproduction number, which is defined here as a random variable rather than an expected value.

1. Introduction. The SIR-model is a well-studied epidemic model that, together with its generalizations, has been widely applied to infectious diseases such as measles, chickenpox, or mumps, among other situations where infection confers (typically lifelong) immunity; a good discussion about SIR-models analyzed from deterministic and stochastic perspectives can be found in the lucid texts by Allen [1, 2], Andersson and Britton [8], Bailey [16] and Keeling and Rohani [36]. The SIR-model is first analyzed in depth by Kermack and McKendrick [37] in 1927 in order to study the evolution of a disease in a closed community of finite size where, at time $t$, individuals are classified into three categories: $S(t)$ susceptibles, $I(t)$ infectives, and $R(t)$ removed individuals. The general description in [37] is related to an homogeneously mixed population (i.e., any infective can infect any susceptible with equal probability), where the infection and recovery rates of a given infective depend on the total time that this individual has been infected for. The analytical
difficulties of addressing this time-dependent assumption lead Kermack and McKendrick [37] to the consideration of a number of special cases. Specifically, Special Case B in [37] amounts to the particular assumption that infection and recovery rates are constant, becoming the origin of the basic SIR-model; see e.g. [37, Eq. (29)]. Therefore, in the basic SIR-model the rate of new infections is proportional to the current numbers of susceptibles and infectives in the population; as time passes, some infectives recover and go into the removed category and, at the same time, infectives contact susceptibles so that \( S(t) \) decreases and \( I(t) \) increases.

In recent years, there has been a rapidly growing interest in SIR-models and their applications; a comprehensive review of the existing literature and some results can be found in the survey papers by Britton [21], and Isham [34, 35]. In the stochastic setting, three important quantities of the SIR-model include the final size distribution, the expected duration of an epidemic, and the basic reproduction number \( R_0 \), which are analyzed by using a variety of tools in applied probability; see, for example, the articles by Allen [3], Allen and Burgin [4], Artalejo et al. [9], and Artalejo and López-Herrero [12], among others. Gani and Purdue [27] outline a matrix-geometric method for determining the total size distribution in a recursive manner, and El Maroufy et al. [25] explore how this method can be used to study SIR-models with a generalized mechanism of infection. A recursive algorithm used to compute the distribution of the epidemic size, under the assumption of a large finite initial population of susceptibles, is described by Gordillo et al. [30], who also derive a diffusion limit in the near-critical case. Unlike Neuts and Li [40] who derive algorithms to compute the time to extinction of the infectious disease, Grasman [31] deals with a diffusion approximation for continuous states. More concretely, Grasman [31] presents a method to approximate the long-term stochastic dynamics of an epidemic modeled by state variables denoting the various classes of individuals within the population such as in the SIR-model and a variant with spatial structure, including epidemics in populations distributed among different locations with migration of individuals between these locations. The concepts of local infectious clump and local susceptibility set are used by Ball and Neal [18] to develop a unified approach to the threshold behavior of SIR epidemics with two levels of mixing. The objective of Artalejo and López-Herrero [12] is to investigate quasi-stationarity and the ratio of expectations as two conceptually different approaches for understanding the dynamics of the SIR-model and its variant with demography before the extinction of an epidemic. New descriptors, including the time to reach specific numbers of infectives and susceptible individuals [15, Section 3.1], and the time to reach a critical number of infections [15, Section 3.2] are also investigated by using efficient numerical tools; for a related work, see the paper [5] where the interest is in the time to reach a critical number of infections in the SIR-model with infective and susceptible immigrants. Results based on the maximum entropy (ME) formalism are first derived analytically in [13] for the SIR-model, and they are then applied to outbreaks of extended spectrum beta lactamase organisms in intensive care units of hospitals. Recently, Clancy [23], and Gómez-Corral and López-García [29] show how SIR epidemics can be naturally modeled as piecewise deterministic Markov processes when individuals’ infectious periods are not necessarily exponentially distributed; note that the epidemic model in [23, 29] is previously analyzed using special constructions by Ball [17], and Picard and Lefèvre [41]. Non-exponential tolerance to infection in SIR- and SEIR-models is analyzed by Streftaris and Gibson [45], who explore the use of a non-exponential model for individuals’ thresholds
to infection (see Sellke [44]) as an alternative to the traditionally employed Poisson streams for occurrence of infectious contacts.

With the exception of the ME principle [13], SIR-models with general infectious period distribution [23, 29] and non-exponential tolerance to infection [45], the key for methods used in the above papers lies in the fact that the events are generated by Poisson streams and so new infections and recoveries, being exponentially distributed, possess the Markov property but do not incorporate correlations. Our major achievement in this article is the presentation of an SIR-model with correlated inter-event times. Specifically, we define a Markov-modulated process that allows us to reflect non-exponential distributional assumptions on the contact process between the compartment of infectives and the compartment of susceptible individuals, and the recovery process affecting the compartment of infectives, but keeping the dimensionality of the underlying Markov chain tractable. This means that inter-event times in our Markov-modulated SIR-model are scheduled at a compartmental level, rather than according to the individual level, which is the case of the extensions studied by Clancy [23], Gómez-Corral and López-García [29], and Streftaris and Gibson [45]. Thus, instead of an extension of the SIR-model, the resulting Markov-modulated model should be considered as an approximating model.

Up till now, relatively little work has been reported on epidemics exhibiting correlation in the process generating events. The use of a state-dependent (SD) Markov-modulated mechanism – including the Markovian arrival process (MAP, [10, Section 2]), and block-structured state events (BSDE, [10, Section 3.5]) –, as a point process characterized by a small number of matrices, provides us with a highly tractable benchmark for the study of SIR-models with versatile qualities. From a conceptual perspective, by using a Markov-modulated framework one deals with a weakened version of the lack-of-memory property allowing a Markovian dependence on one of a finite number of phases and, consequently, incorporating correlated and non-exponentially generated events. The BSDE approach (Artalejo and Gómez-Corral [10]) and related SD formulations can be used to extend stochastic systems that use Markov chains to model a biological population. For example, a SD-SIS model is first formulated in [11] as an augmented process, and then analyzed according to the time until the extinction, and the coefficient of correlation between successive events; as a related work, the article [28] deals with non-exponential distributional assumptions on a host birth, a parasite death, and parasitism in bivariate host-parasitoid models. Amador and Artalejo [6] use the BSDE formalism to study the number of cases of infection in a computer network, with an application to the propagation of the CodeRed II virus.

In the following sections we consider stochastic SIR-models of increasing complexity. In Section 2, we present the basic SIR-model as a bivariate stochastic model with the susceptible and infective compartments as state variables. In Section 3, we first construct SIR-models with correlated inter-event times by applying a Markov-modulated approach that provides a methodological tool to model SD Markovian transitions operating in the presence of phases. Next, we focus on a concrete specification of the underlying phase process such that relevant properties of the basic SIR-model are appropriately satisfied by the resulting Markov-modulated model. We analyze the ultimate extinction of the epidemic, the length of an outbreak and the final size of the epidemic, as well as two random variants of the basic reproduction number $\mathcal{R}_0$. Numerical results are presented in Section 4 to illustrate the effects of the correlation structure on the dynamics of Markov-modulated versions.
of two matrices, respectively, and the exponential of a square matrix $W$ by $\exp$ order $a$ of order $a$ by $A$ represented by lowercase letters, like by bold letters; in particular, vectors are assumed to be column vectors and are those Markovian arrival streams used in Section 4 are specified in Appendix B.

Finally, some conclusions are given in Section 5. For ease of reference, Appendix A contains the proof of certain properties of the Markov-modulated SIR-model; and those Markovian arrival streams used in Section 4 are specified in Appendix B.

**Glossary of notation** Throughout the paper, vectors and matrices are denoted by bold letters; in particular, vectors are assumed to be column vectors and are represented by lowercase letters, like $a$, and matrices are denoted in uppercase, like $A$, with the transpose of $A$ denoted by $A^T$. The identity matrix of order $a$ is given by $1_a$, and $0_{a \times b}$ is the null matrix of dimension $a \times b$. The null and unit vectors of order $a$ are denoted by $0_a$ and $e_a$, respectively, and $e_a(b)$ is a column vector of order $a$ such that all entries equal 0, except for the $b$th one which is equal to 1.

The symbols $\otimes$ and $\oplus$ stand for the Kronecker product and the Kronecker sum of two matrices, respectively, and the exponential of a square matrix $W$ is defined by \[ \exp\{W\} = \sum_{x=0}^{\infty}(x!)^{-1}W^x. \]

The Kronecker delta $\delta_{a,b}$ is defined by 1 if $a = b$, and 0 if $a \neq b$.

### 2. The basic SIR-model

In the SIR-model, individuals develop an immunity to the disease. At time $t$, the population consists of $I(t) = i$ infectives, $S(t) = s$ susceptibles, and $R(t) = r$ removed individuals, and the only possible events (Table 1) correspond to contacts between an infective and a susceptible, and the removal of an infective. The rates $\lambda_{i,s}$ and $\mu_i$ can be specified in infinitely many ways. For example, the choice $\lambda_{i,s} = \beta is$ and $\mu_i = \eta i$ yields the general stochastic epidemic (Bailey [16, Chapter 6], Gani and Purdue [27]), and Models 1 and 2 analyzed by Neuts and Li [40] are related to the respective infection rates $\lambda_{i,s} = \beta i^a s$ and $\lambda_{i,s} = \beta i \min\{s, \epsilon n\}$, where the parameter $a \in (0, 1)$ quantifies the degree of interaction between infectives and susceptibles, and the value $\epsilon$ specifies the fraction of susceptible population that is exposed to each infective; the transmission of myxomatosis among rabbits is analyzed by Saunders [43] by selecting the infection rates $\lambda_{i,s} = (i + s)^{-1/2} is$.

In the case of a closed population that is homogeneously mixed, two state variables suffice and, under the assumption of initial numbers of $m > 0$ infectives and $n > 0$ susceptibles, the SIR-model is formulated as a time-homogeneous continuous-time Markov chain (CTMC) $X = \{X(t) = (I(t), S(t)) : t \geq 0\}$ defined on a finite state space $S(m, n) = \{(i, s) \in \mathbb{N}_0 \times \mathbb{N}_0 : 0 \leq i \leq m + n, 0 \leq s \leq \min\{n, m + n - i\}\}$. The process $X$, which is from now on termed basic, is uniquely specified by the transition rates $\lambda_{i,s}$ and $\mu_i$ in such a way that states in $C_0(n) = \{(0, 0), (0, 1), ..., (0, n)\}$ are absorbing, and the class $C(m, n) = \{(i, s) \in \mathbb{N}_0 \times \mathbb{N}_0 : 1 \leq i \leq m + n, 0 \leq s \leq \min\{n, m + n - i\}\}$ consists of transient states. States in $C_0(n)$ are associated with the ultimate extinction of the epidemic, and the class $C(m, n)$ can be expressed in terms of levels as $\bigcup_{s=0}^n I(s)$, where the $s$th level corresponds to states with $s$ susceptibles present, that is, $I(s) = \{(i, s) \in \mathbb{N}_0 \times \mathbb{N}_0 : 1 \leq i \leq m + n - s\}$, for $0 \leq s \leq n$.

| Transitions       | Events         | Rates  |
|-------------------|----------------|--------|
| $i \rightarrow i + 1$, $s \rightarrow s - 1$, $r \rightarrow r$, for $i, s \in \mathbb{N}$ | A new infection | $\lambda_{i,s}$ |
| $i \rightarrow i - 1$, $s \rightarrow s$, $r \rightarrow r + 1$, for $i \in \mathbb{N}, s \in \mathbb{N}_0$ | A removal       | $\mu_i$ |

Table 1. Stochastic transitions, events and rates in the basic SIR-model
For later use, we recall the structured form of the infinitesimal generator of $X$, which is seen to be the keystone to derive efficient algorithms for the distributions of the total size and the maximum size of the epidemic (Neuts and Li [40, Sections 3-4]), and the joint distribution of the maximum and the time instant of its occurrence (Neuts and Li [40, Section 5]). Specifically, states in $S(m, n)$ are assumed to be labeled so that transient states precede absorbing states, in such a way that

$$l(n) \prec l(n - 1) \prec ... \prec l(1) \prec l(0) \prec C_0(n),$$

where states $(i, s) \in l(s)$, with $s \in \{0, 1, ..., n\}$, are ordered as

$$(m + n - s, s) \prec (m + n - s - 1, s) \prec ... \prec (2, s) \prec (1, s)$$

and absorbing states are ordered as

$$(0, n) \prec (0, n - 1) \prec ... \prec (0, 1) \prec (0, 0).$$

This labeling of states results in the infinitesimal generator $Q$ of $X$

$$Q = \left( \begin{array}{cccc} Q_{C(m, n), C(m, n)} & Q_{C(m, n), C_0(n)} \\ 0_{(n+1) \times J(m, n)} & 0_{(n+1) \times (n+1)} \end{array} \right),$$

(1)

where $J(m, n) = 2^{-1}(n + 1)(2m + n)$ is the cardinality of the class $C(m, n)$ of transient states, and

$$Q_{C(m, n), C(m, n)} = \begin{pmatrix} A(n) & B(n) \\ A(n - 1) & B(n - 1) \\ ... & ... \\ A(1) & B(1) \\ A(0) \end{pmatrix},$$

(2)

$$Q_{C(m, n), C_0(n)} = \begin{pmatrix} t(n) \\ t(n - 1) \\ ... \\ t(1) \end{pmatrix}.$$  

(3)

In Eqs. (2)-(3), the sub-matrices $A(s)$ and $B(s)$ record transition rates related to jumps of $X$ from states of the $s$th level to states of $l(s)$ and $l(s - 1)$, respectively, and diagonal elements of $A(s)$ are given by $-q_{i,s} = -(\lambda_{i,s} + \mu_i)$, for $1 \leq i \leq m + n - s$; similarly, the column vector $t(s)$ contains transition rates related to jumps of $X$ from states of the $s$th level to the absorbing state $(0, s)$. They are given by $t(s) = \mu_1 e_{m+n-s}(m + n - s)$, and

$$A(s) = \begin{pmatrix} -q_{m+n-s,s} & \mu_{m+n-s} \\ -q_{m+n-s-1,s} & \mu_{m+n-s-1} \\ ... & ... \\ -q_{2,s} & \mu_2 \\ -q_{1,s} & \mu_1 \end{pmatrix},$$

(4)

$$B(s) = \begin{pmatrix} \lambda_{m+n-s,s} \\ \lambda_{m+n-s-1,s} \\ ... \end{pmatrix}.$$  

(5)
3. Markov-modulated events in the SIR epidemic model. The dynamics in the SIR-model (Table 1) can be explained in terms of scheduled events. Let us assume that the process $\mathcal{X}$ enters state $(i, s) \in \mathcal{C}(m, n)$ at time $t$. The next transition $(i, s) \rightarrow (i', s')$ is triggered by two scheduled events $E_{(i,s)}(1,-1)$ and $E_{(i,s)}(-1,0)$, where the former amounts to the infection of a susceptible and the latter is related to the removal of an infective. The occurrence instants of $E_{(i,s)}(1,-1)$ and $E_{(i,s)}(-1,0)$ are independent and exponentially distributed with mean values $\lambda_{i,s}^{-1}$ and $\mu_i^{-1}$, respectively, and they are also independent of the history of $\mathcal{X}$ up to time $t$. The basic process $\mathcal{X}$ spends in $(i, s)$ an exponentially distributed period of time, with expected length $(\lambda_{i,s}+\mu_i)^{-1}$, and then moves to state $(i', s') = (i+z_1, s+z_2)$ with probability $(\lambda_{i,s}+\mu_i)^{-1}\lambda_{i,s}$ if $(z_1, z_2) = (1,-1)$ (i.e., a new infection), and $(\lambda_{i,s}+\mu_i)^{-1}\mu_i$ if $(z_1, z_2) = (-1,0)$ (i.e., the removal of an infective). Therefore, the basic state $(i, s) \in \mathcal{C}(m, n)$ is updated in the light of the observed value $(z_1, z_2)$ of a bivariate vector $(Z_1, Z_2)$ that, as a function of $(i, s)$, records scheduled events taking place when the sojourn time that $\mathcal{X}$ spends in $(i, s)$ expires. More concretely, we may express the resulting basic state $(i', s')$ as $(i', s') = f((i, s), (Z_1, Z_2))$, with $f((i, s), (z_1, z_2)) = (i+z_1, s+z_2)$ for pairs $(z_1, z_2) \in \{(1,-1), (-1,0)\}$. Distributed assumptions for $(Z_1, Z_2)$ are inherently linked to the superposition of two independent Poisson streams with arrival rates $\lambda_{i,s}$ and $\mu_i$.

In what follows, we replace the basic process $\mathcal{X}$ by an augmented version $(\mathcal{X}', \mathcal{Y})$ allowing a Markovian dependence on a finite number of phases. To construct the augmented process $(\mathcal{X}', \mathcal{Y})$, every basic state $\mathcal{X} \in \mathcal{C}(m, n)$ is first replaced by a set of augmented states $(i, s, y, y_2)$ with $1 \leq y_k \leq L_k$ for predetermined values $L_k \in \mathbb{N}$ and $k \in \{1, 2\}$. Poisson assumptions for $(Z_1, Z_2)$ are then generalized to incorporate correlated inter-event times into the SIR-model; specifically, we construct the superposition from two independent MAPs of orders $L_1$ and $L_2$ (Artalejo et al. [11]) instead of two Poisson processes, and we introduce a third pair $(z_1, z_2) = (0, 0)$ that reflects a transition between phases, but it does not imply any transition in the basic state $(i, s)$.

We briefly recall that an MAP (He [33, Section 2.2]) is related to a bivariate CTMC $(N, K) = \{(N(t), K(t)) : t \geq 0\}$, which is defined on a denumerable state space $\mathcal{S}_{(N, K)} = \{(n, k) : n \in \mathbb{N}_0, k \in \{1, \ldots, L\}\}$ and has infinitesimal generator

$$Q_{(N, K)} = \begin{pmatrix} E_0 & E_1 \\ E_0 & E_1 \\ \vdots & \vdots \end{pmatrix}.$$ 

The matrices $E_0$ and $E_1$ are square matrices of order $L_E$ with $E_1 \geq 0_{L_E \times L_E}$, the entries of $E_0$ satisfy $(E_0)_{k,k} < 0$ and $(E_0)_{k,k'} \geq 0$ if $k' \neq k$, for $k, k' \in \{1, \ldots, L\}$, and $(E_0 + E_1)_{L_E} = 0_{L_E}$; for technical reasons, it is assumed that the resulting infinitesimal generator $E_0 + E_1$ is irreducible. In the CTMC $(N, K)$, the counting process $N = \{N(t) : t \geq 0\}$ describes the number of births (or occurrences of a predetermined event) during the interval $(0, t]$, provided that the CTMC starts from state $(0, k)$ at time 0, and is modulated by the phase process $K = \{K(t) : t \geq 0\}$, whose phase $k$ dictates the instantaneous rates for births\footnote{The selection $L_E = 1$ and $(E_0, E_1) = (-\sigma, \sigma)$ yields the counting process $N = \{N(t) : t \geq 0\}$ of a Poisson process with rate $\sigma > 0$, whose main simplifying feature is the independence and exponential equidistribution of successive inter-birth intervals.}. By partitioning the
state space $\mathcal{S}(\mathcal{X}, \mathcal{K})$ into subsets $l(\mathcal{X}, \mathcal{K})(n) = \{(n, k) : k \in \{1, \ldots, L_E\}\}$, for integers $n \in \mathbb{N}_0$, the process $(\mathcal{X}, \mathcal{K})$ allows for single births so that a transition is possible in one step from each subset $l(\mathcal{X}, \mathcal{K})(n)$ to states of the subset $l(\mathcal{X}, \mathcal{K})(n+1)$, and it may allow phase changes at birth epochs to occur without any restriction; more particularly, the birth rate may depend on the phases, but not on the current subset. A generalization of the MAP that accommodates processes with different types of births is the marked Markovian arrival process (MMAP, He and Neuts [32], He [33, Section 2.5]). The main idea is to assign different interpretations to the birth rates, which can be done by considering a set $Z$ of indices, a family \{$E_z : z \in Z$\} of non-negative matrices and a matrix $E_0$ with non-negative off-diagonal elements and negative diagonal elements verifying that $E_0 + \sum_{z \in Z} E_z$ is an irreducible infinitesimal generator of order $L_Z$. The MMAP with characteristic matrices \{$E_0, E_z : z \in Z$\} of order $L_Z$ is then related to a multidimensional CTMC $\{(N_1(t), \ldots, N_{L_Z}(t), K(t)) : t \geq 0\}$ where $N_z = \{N_z(t) : t \geq 0\}$ is the marginal counting process of type-$z$ births, for $z \in Z$, and $K = \{K(t) : t \geq 0\}$ is the underlying phase process. This means that, for $z \in Z$, the matrix $E_z$ consists of infinitesimal rates for the increase of one unit in $N_z(t)$; needless to say, MAPs are obtained as special cases of MMAPs with a single type of births.

The above construction leads to a regular, time-homogeneous CTMC $(\mathcal{X}, \mathcal{Y})$, where the phase process $\mathcal{Y} = \{Y(t) = (Y_1(t), Y_2(t)) : t \geq 0\}$ can be thought of as a SD MMAP with two categories of marked events: a type-(1, 2) event with either $(z_1, z_2) = (1, -1)$ if the infection of a susceptible individual is observed, or $(z_1, z_2) = (-1, 0)$ if the removal of an infective is registered. Similarly to $\mathcal{X}$, the state space $\mathcal{S}^*$ of $(\mathcal{X}, \mathcal{Y})$ can be partitioned into the set $C_0(n)$ of absorbing states, and a class $C^*(m, n)$ of transient states, where $C^*(m, n) = \cup_{s=0}^n C^*(s)$ and the $s$th augmented level consists of states $(i, s, y_1, y_2)$ with $1 \leq i \leq m+n-s$ and $1 \leq y_k \leq L_k$ for $k \in \{1, 2\}$.

In our approach, we use two scaled MAPs that, for a basic state $(i, s) \in C(m, n)$, are specified by characteristic matrices $(C_{0.(i,s)}^*, C_{1.(i,s)}^*)$ for the occurrence of an infection, and $(D_{0.(i,s)}^*, D_{1.(i,s)}^*)$ for the removal of an infective, with

\[
C_{k,.(i,s)}^* = \hat{\lambda}^{-1}(1)\lambda_{i,s} C_k, \\
D_{k,.(i,s)}^* = \hat{\lambda}^{-1}(2)\mu_i D_k,
\]

for $k \in \{0, 1\}$, where $(C_0, C_1)$ and $(D_0, D_1)$ are the characteristic matrices of two independent MAPs of orders $L_1$ and $L_2$, respectively, and fundamental arrival rates $\hat{\lambda}(1)$ and $\hat{\lambda}(2)$. This means that the infinitesimal generator $Q^*$ of $(\mathcal{X}, \mathcal{Y})$ has the structured form of $Q$ in (1) with $A(s)$, $B(s)$, and $t(s)$ in (2)-(3) replaced by suitably defined sub-matrices $A^*(s)$ and $B^*(s)$, and vectors $t^*(s)$. They are specified as follows:

(i) The sub-matrix $A^*(0)$ has the structured form in (4) with $-q_{m+n-s',s}$ and $\mu_{m+n-s}$ replaced by $(C_{0,(m+n-s',0)}^* + C_{1,(m+n-s',0)}^*) \oplus D_{0,(m+n-s',0)}^*$ for $0 \leq s' \leq m + n - 1$, and $I_{L_1} \otimes D_{1,(m+n-s',0)}^*$ for $0 \leq s' \leq m + n - 2$, respectively.

(ii) For $1 \leq s \leq n$, the sub-matrices $A^*(s)$ are derived from (4) by replacing the elements $-q_{m+n-s-s',s}$ and $\mu_{m+n-s-s'}$ by $C_{0,(m+n-s-s',s)}^* + D_{0,(m+n-s-s',s)}^*$ for $0 \leq s' \leq m + n - 1$, and $I_{L_1} \otimes D_{1,(m+n-s-s',s)}^*$ for $0 \leq s' \leq m + n - 2$, respectively.
(iii) For \(1 \leq s \leq n\), the sub-matrices \(B^*(s)\) are obtained from (5) by replacing the rates \(\lambda_{m+n-s-s',s}\) by the matrices \(C^*_1,(m+n-s-s',s) \otimes I_{L_2}\) for \(0 \leq s' \leq m + n - s - 1\).
(iv) For \(0 \leq s \leq n\), the vector \(t^*(s)\) is given by

\[
t^*(s) = \left( 0_{(m+n-s-1)L_1L_2}, e_{L_1} \otimes (D^*_1,(1,s)e_{L_2}) \right).
\]

It is worth noting that the basic process \(\mathcal{X}\) and its augmented counterpart \((\mathcal{X}, \mathcal{Y})\) possess identical structural properties, but only \((\mathcal{X}, \mathcal{Y})\) exhibits correlation between events. In particular, for every basic state \((i, s) \in \mathcal{C}(m, n)\), the fundamental arrival rate of type-\((z_1, z_2)\) marks is given by

\[
\bar{\lambda}_{i,s}(z_1, z_2) = \begin{cases} 
\lambda_{i,s}, & \text{if } (z_1, z_2) = (1, -1), \\
\mu_i, & \text{if } (z_1, z_2) = (-1, 0),
\end{cases}
\]

and the fundamental arrival rate \(\bar{\lambda}_{i,s}\) of the SD MMAP is given by

\[
\bar{\lambda}_{i,s} = q_{i,s},
\]

regardless of the initial MAPs with matrices \((C_0, C_1)\) and \((D_0, D_1)\).

Moreover, the correlation structures of the initial MAPs and their scaled variants are identical; that is, if \(\rho_k = \rho(X_1, X_k)\) denotes the correlation coefficient between the first and the \(k\)th inter-arrival times \(X_1\) and \(X_k\) in the MAP with characteristic matrices \((C_0, C_1)\), and \(\rho_{k,(i,s)}\) is its counterpart for the scaled MAP with matrices \((C^*_0,(i,s), C^*_1,(i,s))\), then it can be verified that

\[
\rho_k = \rho_{k,(i,s)},
\]

regardless of the basic state \((i, s) \in \mathcal{C}(m, n)\). A similar result holds for the MAPs with characteristic matrices \((D_0, D_1)\), and \((D^*_0,(i,s), D^*_1,(i,s))\). We address mathematically inclined readers to Appendix A for a proof of Eqs. (6)-(8).

3.1. **Length of an outbreak and total size of the epidemic.** An outbreak begins when a single individual chosen appropriately from the population becomes infected. The disease spreads from this infective to susceptible individuals at a certain rate, in such a way that new infectives attempt to infect other susceptibles and then recover in accordance to the model description in Section 2. The outbreak ends when no infectives remain.

The distributional assumptions in the basic SIR-model imply that the process \(\mathcal{X}\) will reach states \((0, s')\) with \(0 \leq s' \leq n\), starting from any numbers of \(m > 0\) infectives and \(n > 0\) susceptibles, and the epidemic will always die out. This means that the epidemic extinction is always certain and the expected times till extinction are all finite, regardless of the initial basic state. Let \(T_{(i,s)}\) be the time until the epidemic dies out, provided that the population currently contains \(i\) infectives and \(s\) susceptibles, for \((i, s) \in \mathcal{C}(m, n)\). Then, the cumulative distribution function of the length \(T_{(i,s)}\) of the residual outbreak\(^2\) has the form

\[
P(T_{(i,s)} \leq t) = 1 - \bar{e}_{J(m,n)}(i,s) \exp\{Q_{\mathcal{C}(m,n), \mathcal{C}(m,n)} t\} e_{J(m,n)}, \quad t \geq 0,
\]

\(^2\)Since \(T_{(i,s)}\) corresponds to the time till absorption of \(\mathcal{X}\) into the set \(\mathcal{C}_0(n)\), the random length \(T_{(i,s)}\) of a residual outbreak can be seen as a continuous phase-type (PH) random variable of order \(J(m, n)\) with representation \((\bar{e}_{J(m,n)}(i,s), Q_{\mathcal{C}(m,n), \mathcal{C}(m,n)});\) see e.g. Latouche and Ramaswami [38, Section 2.3], and Neuts [39, Chapter 2].
where $\tilde{e}_{J(m,n)}(i,s)$ denotes the initial distribution on the class $C(m,n)$ of transient states, and the joint distribution of $(T_{(i,s)}, S(T_{(i,s)}))$ can be expressed as

$$P(T_{(i,s)} \leq t, S(T_{(i,s)}) = s') = \mu_1 e_{J(m,n)}(i,s) \left( I_{J(m,n)} - \exp \left\{ Q_{C(m,n), C(m,n)}' \right\} \right) \times \left( -Q_{C(m,n), C(m,n)}^{-1} \right) e_{J(m,n)}(J'(m, n; s')), \quad (10)$$

for integers $0 \leq s' \leq s$, where $J'(m, n; s') = 2^{-1}(2m + n - s')(n - s' + 1)$. Note that, since $(i, s) \in C(m,n)$ is the currently visited state, we may specify $\tilde{e}_{J(m,n)}(i,s)$ as a row vector of order $J(m,n)$ such that all entries equal zero, with the exception of the entry for the initial state $(i, s)$ which is equal to one.

The proof of (9)-(10) is based on the structured form of $Q$ given in (1). More concretely, the transition function $P(t)$ of $X$ with elements $P((I(t), S(t)) = (i', s'))(I(0), S(0) = (i, s))$, for states $(i, s), (i', s') \in S(m, n)$, is given by $P(t) = \exp\{Qt\}$, and it is easy to verify that

$$P(t) = \begin{pmatrix} P_{C(m,n), C(m,n)}(t) & P_{C(m,n), C_0(n)}(t) \\ 0_{(n+1) \times J(m,n)} & I_{n+1} \end{pmatrix},$$

with sub-matrices $P_{C(m,n), C(m,n)}(t) = \exp\{Q_{C(m,n), C(m,n)}t\}$ and $P_{C(m,n), C_0(n)}(t) = (I_{J(m,n)} - \exp\{Q_{C(m,n), C(m,n)}t\})^{-1}Q_{C(m,n), C_0(n)}$. Then, the term $e_{J(m,n)}(i,s) \exp\{Q_{C(m,n), C_0(n)}t\} e_{J(m,n)}$ on the right side of (9) amounts to the conditional probability of being in a non-absorbing state at time $t$, provided that $(i, s)$ is the current state of $X$, and $P(T_{(i,s)} \leq t, S(T_{(i,s)}) = s')$ corresponds to the entry of $P_{C(m,n), C_0(n)}(t)$ associated with the initial transient state $(i, s)$ and the absorbing state $(0, s')$ visited by the process $X$ at time $t$; note that this entry is given by $e_{J(m,n)}(i,s)P_{C(m,n), C_0(n)}(t)e_{n+1}(n+1 - s')$, and (10) is then proved from the identity $Q_{C(m,n), C_0(n)}e_{n+1}(n+1 - s') = \mu_1 e_{J(m,n)}(J'(m, n; s'))$. It is readily seen that, in Eq. (10), the matrix $Q_{C(m,n), C_0(n)}$ is non-singular since the absorption into states of $C_0(n)$ occurs with probability one. Then, by using the jargon of the labeling of states yielding Eqs. (1)-(5), the entry of $-Q_{C(m,n), C(m,n)}^{-1}$ located at the $(i, s)$th row and the $(i', s')$th column is the expected total time spent in state $(i', s')$ – which is linked to a population consisting of $i'$ infectives and $s'$ susceptible individuals – during an outbreak, given that the outbreak starts with $i$ infectives and $s$ susceptibles.

In the case of Markov-modulated events, the length $T_{(i,s,y_1,y_2)}$ of the resulting residual outbreak, provided that $(i, s, y_1, y_2) \in C^*(m,n)$ is the current state of $(X, Y)$, follows a continuous PH law of order $J^*(m,n) = L_1L_2J(m,n)$ and a representation in terms of the vector $e_{J^*(m,n)}(i,s,y_1,y_2)$ of initial probabilities and the sub-matrix $Q_{C^*(m,n), C^*(m,n)}$ recording transition rates related to jumps of $(X, Y)$ from states in the class $C^*(m,n)$ of transient states to states in $C^*(m,n)$. Note that $Q_{C^*(m,n), C^*(m,n)}$ has the structured form of $Q_{C(m,n), C(m,n)}$ in (2) with $A(s)$ and $B(s)$ replaced by the sub-matrices $A^*(s)$ and $B^*(s)$ defined for the augmented process in Section 3. Then, it is a matter of routine to derive expressions for $P(T_{(i,s,y_1,y_2)} \leq t)$ and $P(T_{(i,s,y_1,y_2)} \leq t, S(t) = s')$ from Eqs. (9)-(10) by replacing the matrices $I_{J(m,n)}$ and $Q_{C(m,n), C(m,n)}$ by $I_{J^*(m,n)}$ and $Q_{C^*(m,n), C^*(m,n)}$, respectively, and the vectors $e_{J(m,n)}(i,s), e_{J(m,n)}$ and $e_{J(m,n)}(J'(m,n,s'))$ by $e_{J^*(m,n)}(i,s,y_1,y_2), e_{J^*(m,n)}$ and $e_{J^*(m,n)}(J'(m,n,s'))$, respectively, with $J''(m,n,s') = L_1L_2J'(m,n,s')$.

The final size of the epidemic in the basic SIR-model can be readily determined from the mass function $P(S(T_{(i,s)}) = s')$ of the number of survivors, for integers $0 \leq s' \leq s$ and states $(i, s) \in C(m,n)$. To evaluate the probability $P(S(T_{(i,s)}) = s')$,
we may first observe that the number of jumps of the basic process $X$ till absorption into $C_0(n)$ follows a discrete PH law (Latouche and Ramaswami [38, Section 2.5]) of order $J(m,n)$, and representation in terms of the vector $e_{J(m,n)}(i,s)$ of initial probabilities and the sub-matrix

$$P_{C(m,n),C(m,n)} = \begin{pmatrix} \overline{A}(n) & \overline{B}(n) \\ \overline{A}(n-1) & \overline{B}(n-1) \\ \vdots & \vdots \\ \overline{A}(1) & \overline{B}(1) \\ \overline{A}(0) & \overline{B}(0) \end{pmatrix},$$

which contains one-step transition probabilities among transient states in the underlying jump embedded chain; specifically, non-null entries in the $k$th row of $\overline{A}(s)$ and $\overline{B}(s)$ are obtained from the sub-matrices $A(s)$ and $B(s)$ by dividing their infinitesimal counterparts in (4) and (5) by the sojourn rate $\eta_{m+n-s+1-k}$, for $1 \leq k \leq m+n-s$ and $0 \leq s \leq n$, and diagonal elements of $\overline{A}(s)$ are all null. Then, we may notice that the event $\{S(T(i,s)) = s - j\}$ amounts to the fact that the number of jumps till absorption into $C_0(n)$ equals $i + 2j$ (i.e., $j$ new infections and $i + j$ recoveries), for integers $0 \leq j \leq s$ and states $(i,s) \in C(m,n)$. Hence, we may express

$$P(S(T(i,s)) = s - j) = e_{J(m,n)}(i,s)P_{C(m,n),C(m,n)}^{i+2j-1}P_{C_0(n),C_0(n)}, \quad (11)$$

for integers $0 \leq j \leq s$, where the column vector $p_{C_0(n)}$ consists of one-step transition probabilities from states in $C(m,n)$ to absorbing states, and it takes the form

$$p_{C_0(n)} = \begin{pmatrix} \tilde{t}(n) \\ \tilde{t}(n-1) \\ \vdots \\ \tilde{t}(0) \end{pmatrix},$$

with $\tilde{t}(k) = (\lambda_{1,k} + \mu_1)^{-1}\mu_1 e_{m+n-k}(m+n-k)$, for $0 \leq k \leq n$. From (11), it is seen that

$$E[S(T(i,s))] = s + 1 - e_{J(m,n)}(i,s) \left( I_{J(m,n)} - P_{C(m,n),C(m,n)}^2 \right)^{-1} e_{J(m,n)}, \quad (12)$$

for states $(i,s) \in C(m,n)$. More concretely, the expected number $E[S(T(i,s))]$ of surviving susceptibles can be evaluated from (11) as

$$E[S(T(i,s))] = (s + 1) \sum_{j=0}^{s} e_{J(m,n)}(i,s)P_{C(m,n),C(m,n)}^{i+2j-1}p_{C_0(n)} - \sum_{j=1}^{s+1} j e_{J(m,n)}(i,s)P_{C(m,n),C(m,n)}^{2(j-1)}q_{C_0(n)},$$

where $q_{C_0(n)} = \left( P_{C(m,n),C(m,n)}^{-1}P_{C_0(n),C_0(n)} \right)$. Eq. (12) is then derived by noting that the first term on the right side equals $s + 1$ since $\sum_{j=0}^{s} e_{J(m,n)}(i,s)P_{C(m,n),C(m,n)}^{i+2j-1}P_{C_0(n)} = 1$ by (11), and the second term $\sum_{j=1}^{s+1} j e_{J(m,n)}(i,s)P_{C(m,n),C(m,n)}^{2(j-1)}q_{C_0(n)}$ corresponds to the first moment of a discrete PH random variable (see [38, Section 2.5]) of order $J(m,n)$ with representation $(e_{J(m,n)}(i,s), P_{C(m,n),C(m,n)}^2)$. 
Similarly to the above arguments, the mass function of the number \( S(T^*_1(i,s,y_1,y_2)) \) of survivors in the Markov-modulated SIR-model and its expectation can be expressed from Eqs. (11)-(12) by using \( \mathbf{e}_j^*(m,n) \) instead of the vectors \( \mathbf{e}_j(n,m) \) and \( \mathbf{e}_j(m,n) \), respectively, and by replacing the matrices \( \mathbf{I}_{j(m,n)} \) and \( \mathbf{P}_{C(m,n),C(m,n)} \), and the vector \( \mathbf{p}_{C_0(n)} \) by \( \mathbf{L}_j^*(m,n) \), \( \mathbf{P}_{C^*(m,n),C^*(m,n)} \) and \( \mathbf{p}_{C_0(n)}^* \), where the matrix \( \mathbf{P}_{C^*(m,n),C^*(m,n)} \) and the column vector \( \mathbf{p}_{C_0(n)}^* \) are suitably defined from the jump embedded chain of the augmented process \((\mathcal{X},\mathcal{Y})\).

Eqs. (11)-(12) are easy to implement and one may expect them to be numerically well behaved in the case of small numbers \((m,n)\) of initial infectives and susceptibles. Large values of \((m,n)\) imply large cardinalities of the class \( \mathcal{C}(m,n) \) (respectively, \( \mathcal{C}^*(m,n) \)) of transient states in the basic process \( \mathcal{X} \) (respectively, its augmented counterpart \((\mathcal{X},\mathcal{Y})\)), which often turns the computation of powers of \( \mathbf{P}_{C(m,n),C(m,n)} \) (respectively, \( \mathbf{P}_{C^*(m,n),C^*(m,n)} \)) in (11)-(12) into instable. As an alternative to (11)-(12), Black and Ross [20] recently derive, for a variety of epidemic models including the basic SIR-model, an iterative method for computing the final size distribution, which is mainly based on the replacement of the basic SIR-model with generalized infection rates, establish two fast and efficient recursive methods to evaluate the final size of the epidemic, and show how those methods are numerically stable for large numbers \((m,n)\) of initial infectives and susceptibles. However, the extension of the approaches in [20, 26] to the augmented process \((\mathcal{X},\mathcal{Y})\) for Markov-modulated SIR-models is not immediate.

### 3.2. The exact reproduction number \( R_{exact,0} \)

In the SIR-model defined by the rates \( \lambda_{i,s} = \beta i s \) and \( \mu_i = \eta i \) with \( \beta = (m+n)^{-1}\beta' \), the basic reproductive ratio \( R_0 = \eta^{-1}\beta' \) separates the growing and shrinking behaviors and, consequently, it marks the epidemic threshold between regimes in which the disease either increases or dies out in the long run. Despite its simplicity and robustness, it is known that, due to the occurrence of repeated contacts taking place between a focal infective and other individuals already infected before the time \( t \) and a co-lexicographic labeling of states for the resulting process. Similarly, El Maroufy et al. [26] give an explicit formula of the final outcome probabilities in the SIR-model with generalized infection rates, establish two fast and efficient recursive methods to evaluate the final size of the epidemic, and show how those methods are numerically stable for large numbers \((m,n)\) of initial infectives and susceptibles. However, the extension of the approaches in [20, 26] to the augmented process \((\mathcal{X},\mathcal{Y})\) for Markov-modulated SIR-models is not immediate.

In this section, we focus on an exact variant of \( R_0 \) first introduced by Ross [42] in the setting of SIS and SIR epidemics for a homogeneously mixed population of finite size and then presented by Artalejo and López-Herrero [14] in a more general setting, which is defined as a random variable instead of an expected value. To be concrete, based on [14, Section 3], we define \( R_{exact,0} \) as the exact number of secondary cases generated by a focal infective during its infectious period. An important property of \( R_{exact,0} \) is that, unlike \( R_0 \) which is related to the time of invasion (i.e., for a community with \((I(0), S(0)) = (1, n), R(0) = 0 \) and \( n \geq 1 \)), the exact version \( R_{exact,0} \) can be appropriately defined at every time instant. Nevertheless, we from now on set the initial state \((1,n)\), with \( n \geq 1 \), for comparison purposes.

In the Markov-modulated SIR-model, the basic state \((1,n)\), with \( n \geq 1 \), is replaced by the set of augmented states \((1,n,y_1,y_2)\) with phases \( 1 \leq y_k \leq L_k \) and \( k \in \mathbb{N} \).
and the conditional probability mass function $x_{(1,n,y_1,y_2)} = x_{(1,n,y_1,y_2)}(j) : 0 \leq j \leq n$ of $R_{exact,0}$, provided that $(1,n,y_1,y_2) \in C^*(1,n)$ is the initial state of $(X,Y)$, is defined by the probabilities

$$P(R_{exact,0} = j | (I(0), S(0), Y_1(0), Y_2(0)) = (1,n,y_1,y_2)), \quad 0 \leq j \leq n.$$  

Similarly to the case $(I(0), S(0)) = (1,n)$, we may define conditional probability mass functions $x_{(i,s,y_1,y_2)}$ of $R_{exact,0}$ in terms of probabilities $x_{(i,s,y_1,y_2)}(j) = P(R_{exact,0} = j | (I(0), S(0), Y_1(0), Y_2(0)) = (i,s,y_1,y_2))$, for $0 \leq j \leq s$ and phases $1 \leq y_k \leq L_k$ with $k \in \{1,2\}$, when initial basic states $(i,s) \in C(1,n)$ are observed.

For every basic state $(i,s) \in C(1,n)$ and values $0 \leq j \leq s$, we let $x_{(i,s)}(j)$ be the column vector of order $L_1L_2$ with entries $x_{(i,s)}(j)(y_1-1)L_2+y_2 = x_{(i,s,y_1,y_2)}(j)$ for $1 \leq y_k \leq L_k$ and $k \in \{1,2\}$. Then, by conditioning on the first transition of the augmented process $(X,Y)$, we notice that entries of $x_{(i,s)}(j)$ satisfy

$$-\left(\begin{array}{c}
\lambda_i^{-1}(1)\lambda_{i,s}C_0 \\
\lambda_i^{-1}(2)\mu_iD_0
\end{array}\right) + \left(\begin{array}{c}
y_0(i) \mathbf{e}_{L_1} \otimes \left(\lambda_i^{-1}(1)\lambda_{i,s}C_1\right) \\
\lambda_i^{-1}(2)\eta(i-1)\mathbf{e}_{L_2}
\end{array}\right)x_{(i,s)}(j)$$

$$+(1-\delta_{0,j}) \left(\begin{array}{c}
\left(\lambda_i^{-1}(1)\lambda_{i,s}C_0\right) \otimes \mathbf{I}_{L_2}
\end{array}\right)x_{(i,s-1)}(j-1)$$

$$+(1-\delta_{1,j}) \left(\begin{array}{c}
\left(\lambda_i^{-1}(1)\lambda_{i-1,s}C_1\right) \otimes \mathbf{I}_{L_2}
\end{array}\right)x_{(i+1,s-1)}(j),$$

for values $0 \leq j \leq s$ and $(i,s) \in C(1,n)$, where the rate $\lambda_{i,s}$ gives the individual rate at which a focal infective contacts with susceptibles, and $\tilde{\lambda}_{i-1,s}$ is the contact rate of the remaining $i-1$ non-focal infectives\(^3\). This means that the factor $\left(\begin{array}{c}
\lambda_i^{-1}(1)\lambda_{i,s}C_0 \\
\lambda_i^{-1}(2)\mu_iD_0
\end{array}\right)$ on the left side of (13) gives the total transition rate given the current state, whereas the terms on the right side correspond to the case $j = 0$ and the focal infective recovers without giving rise to a secondary infection; a non-focal infective recovers; the case $j > 0$ and the focal infective infects another individual; and the case $j < s$ and a non-focal infective infects another individual.

Starting from $x_{(i,0)}(0) = \mathbf{e}_{L_1} \mathbf{e}_{L_2}$ for $1 \leq i \leq 1+n$, the triangular system of equations (13) can be efficiently solved in the natural order $j \geq 0, j \leq s \leq n$ and $1 \leq i \leq 1+n-s$, and moments of $R_{exact,0}$ are then derived in an straightforward manner. At an invasion time (i.e., when a focal infective is introduced into a completely susceptible population of size $n$ at time $t = 0$), the role of $R_0$ may be replaced by the unconditional expected value

$$\overline{R}_{exact,0} = \sum_{j=1}^{n} j \eta^* x_{(1,n)}(j),$$

where $\eta^*$ is the vector of initial probabilities of the phase process $Y$ at the invasion time; see Appendix A for further details about the underlying MMAP governing $Y$. It should be noted that an important difference between $R_0$ and $\overline{R}_{exact,0}$ is that $R_0$ can take arbitrarily large values, whereas $\overline{R}_{exact,0} < n$ if the random variable $R_{exact,0}$ is defined at the invasion time; in addition, the values of $R_0$ overestimate the expected number of secondary cases generated by a focal infective before its infectious period expires.

\(^3\)In the special case $\lambda_{i,s} = \beta is$ with $\beta = (1+n)^{-1} \beta'$, they are given by $\lambda_{i,s}^* = (1+n)^{-1} \beta' s$ and $\tilde{\lambda}_{i-1,s} = (1+n)^{-1} \beta'(i-1)s$. 
3.3. The population transmission number $R_p$. The population transmission number $R_p$, as discussed in [14, Section 4], is an alternative measure of disease spread, and it counts the exact number of secondary cases produced by all currently infective individuals prior to the first removal. In a similar manner to $R_{\text{exact},0}$, the probability distribution of $R_p$ depends on the augmented state at time $t = 0$. Therefore, for a fixed augmented state $(i, s, y_1, y_2) \in C^*(1, n)$, we define the conditional probabilities $y_{(i, s, y_1, y_2)}(j)$ by

$$P(R_p = j | (I(0), S(0), Y_1(0), Y_2(0)) = (i, s, y_1, y_2)), \quad 0 \leq j \leq s,$$

from which we introduce the vectors $y_{(i, s)}(j)$ with entries $(y_{(i, s)}(j))(y_1-1)L_2+y_2 = y_{(i, s, y_1, y_2)}(j)$ for $1 \leq y_k \leq L_k$ and $k \in \{1, 2\}$.

An appeal to a first-step argument yields

$$-\left(\left(\hat{\lambda}^{-1}(1)\lambda_{i,s}C_0\right) \oplus \left(\hat{\lambda}^{-1}(2)\mu_iD_0\right)\right)y_{(i, s)}(j) = \delta_{0,j} \left(e_{L_1} \otimes \left(\hat{\lambda}^{-1}(2)\mu_iD_1e_{L_2}\right)\right) + (1 - \delta_{0,j}) \left(\left(\hat{\lambda}^{-1}(1)\lambda_{i,s}C_1\right) \oplus I_{L_2}\right)y_{(i+1, s-1)}(j - 1),$$

for values $0 \leq j \leq s$ and basic states $(i, s) \in C(1, n)$, which can be routinely solved by starting from $y_{(i, 0)}(0) = e_{L_1}L_2$. The unconditional expected value of $R_p$ is then evaluated as

$$R_p = \sum_{j=1}^{n} j \eta^* y_{(1, n)}(j),$$

when the interest is in an invasion time.

4. Numerical experiments and discussion. In our numerical examples, we focus on the general stochastic epidemic (Gani and Purdue [27]) at an invasion time, and we assume rates $\lambda_{i,s} = \beta is$ and $\mu_i = \eta_i$ with $\beta = (1 + n)^{-1}\beta'$, $\beta' > 0$, $\eta = 1.0$ and initial size $1+n \in \{10, 50, 100, 150\}$ of infectives and susceptibles. The aim is to study the dynamics of Markov-modulated versions of the basic SIR-model when the exponential regime on the underlying processes – governing a new infection and the removal of an infective – is replaced by other more general distributional assumptions. Nine scenarios (Table 2) are defined in terms of standard Poisson processes, and positively and negatively correlated Markovian events, which are appropriately specified from two predetermined MAPs (Appendix B) of order $L_1 = L_2 = 3$ with values $\rho_1 = 0.35016$ and $-0.35016$ for their correlation coefficients; in this setting, the basic SIR-model corresponds to scenario I.

In Figures 1-7, we illustrate the effects of the correlation structure in the underlying Markov-modulated streams on the mean number $E[S(T_{1,n})]$ of surviving susceptibles, the expected values of the exact reproduction number $R_{\text{exact},0}$ and the population transmission number $R_p$. For illustrative purposes, we let the vector $\eta^*$ of initial probabilities be the stationary probability vector $\eta(1) \otimes \eta(2)$ of the phase process $\gamma'$, regardless of the initial basic state $(i, s) \in C(1, n)$. It is important to note that, though there seems to be no reason for expecting the phase process to attain stationary before the disease dies out, the selection $\eta^* = \eta(1) \otimes \eta(2)$ has to be seen as a natural choice of initial phases, since it is inherently linked to the underlying characteristic matrices $(C_0, C_1)$ and $(D_0, D_1)$ and does not necessarily yield the uniform law of initial phases.

We display the expected values $E[S(T_{1,n})]$ (Figure 1), $R_{\text{exact},0}$ (Figures 2-4) and $R_p$ (Figures 5-7) as a function of the basic reproduction number $R_0 = \eta^{-1}\beta'$. It should be pointed out that, when a deterministic approach is adopted, the transition
between the epidemic and non-epidemic regimes happens at the point $R_0 = 1$ (equivalently $\beta' = \eta$); specifically, the size of the epidemic goes continuously to zero as $R_0$ approaches one from above and for $R_0 \leq 1$ there is no epidemic at all. We may explain this result by noting that if $\beta' \leq \eta$, the infectives recover faster than susceptible individuals become infected, so the disease will not get a toehold in the population and the number of infected individuals, which starts being small, decreases and the disease dies out instead of spreading out. Values $R_0 > 1$ reflect that each individual catching the disease would pass it on to $R_0$ others on average, each of them would pass it on to $R_0$ more and so forth, so that the number of new cases of the disease would be multiplied by the factor $R_0$ at
each round, thus growing exponentially. Therefore, increasing values of $R_0$ lead to decreasing values of the mean number $E[S(T_{1,n})]$ of survivors, which is graphically shown in Figure 1 for the basic SIR-model (scenario I), and two Markov-modulated versions with positively correlated MAPs governing new infections (scenario A) and recoveries (scenario C). However, it is observed that, for every fixed value of $R_0$, more susceptible individuals will survive if new infections are positively correlated, whereas positively correlated recoveries will imply smaller mean numbers of survivors. This means that the number of infections taking place during an outbreak is greater when recoveries are positively correlated (scenario C) than when a positively correlated MAP governs new infections (scenario A), being the basic SIR-model the intermediate case; more particularly, the basic SIR-model behaves closer to the dynamics of scenarios A and C when, in the deterministic setting, the epidemic dies out ($R_0 < 1$) and the disease is more virulent ($R_0 > 1$) in the long run, respectively.

Figures 2 and 5 are related to a positively correlated process governing new infections, and they allow us to carry out a comparative analysis between $R_0$ and the resulting values of $R_{exact,0}$ (Figure 2) and $R_p$ (Figure 5) under three assumptions for the removal of an infective: the exponential case (scenario A); the negatively correlated case (scenario E); and the positively correlated case (scenario G). We first observe that, as intuition tells us, $R_{exact,0}$ and $R_p$ increase with increasing values of $R_0$ (i.e., when the contact rate $\beta'$ increases, whereas the removal rate $\eta$ is constant) since susceptible individuals become infected faster than infectives recover. Magnitudes in Figures 2 and 5 show an important difference between the exact measurement provided by $R_{exact,0}$ and the population version $R_p$: more concretely, differences in magnitude become more apparent with increasing values of the initial size $n$ of susceptible individuals. We also notice that, for each fixed value of $R_0$, higher and smaller expected values of $R_{exact,0}$ and $R_p$ are always associated with scenarios G and A, respectively, regardless of $n$. In quantifying the spread of a disease when it is just starting out, this means that, for positively correlated infections, the maximum reproductive potential for an infectious disease is related to a positively correlated removal process, while the exponential assumption yields the minimum reproductive potential.

In Figures 3 and 6, the interest is in new infections generated by a Markovian stream with negative correlation, and we compare the value of $R_0$ with the expected values $R_{exact,0}$ (Figures 3) and $R_p$ (Figures 6) when the removal process amounts to the exponential regime (scenario B), the positively correlated case (scenario F), and the negatively correlated case (scenario H). In Figures 4 and 7, we compare the basic SIR-model (scenario I) with those models where removals are generated by Markovian streams with positive (scenario C) and negative (scenario D) correlation, when infections occur according to the exponential regime. Similarly to Figures 2 and 5, the expected values $R_{exact,0}$ and $R_p$ behave as an increasing function of $R_0$, and the maximum and minimum reproductive potentials for an infectious disease amount to removals governed by a positively correlated structure and the exponential regime, respectively. In agreement with Figures 2 and 5, negatively correlated removals lead us to an intermediate situation, where the expected values of $R_{exact,0}$ and $R_p$ are closer to those derived in the exponential case than under the assumption of positive correlation.
Figure 2. The expected value $\mathcal{R}_{exact,0}$ versus the basic reproduction number $\mathcal{R}_0$ for scenarios A, E and G, and initial size $N = 1 + n \in \{10, 50, 100, 150\}$.

Figure 3. The expected value $\mathcal{R}_{exact,0}$ versus the basic reproduction number $\mathcal{R}_0$ for scenarios B, F and H, and initial size $N = 1 + n \in \{10, 50, 100, 150\}$. 
Figure 4. The expected value $R_{\text{exact},0}$ versus the basic reproduction number $R_0$ for scenarios C, D and I, and initial size $N = 1 + n \in \{10, 50, 100, 150\}$

Figure 5. The expected value $R_p$ versus the basic reproduction number $R_0$ for scenarios A, E and G, and initial size $N = 1 + n \in \{10, 50, 100, 150\}$
Figure 6. The expected value $R_p$ versus the basic reproduction number $R_0$ for scenarios B, F and H, and initial size $N = 1 + n \in \{10, 50, 100, 150\}$

Numerical results (not reported here) with $n$ no greater than $10^3$ suggest that increasing sizes $n$ result in higher values of $R_{\text{exact,0}}$ and $R_p$, which can be reasonably explained by noting that removals occur at a constant rate and, on the contrary, infectious contacts take place more frequently when the initial number $n$ of susceptible individuals increases. More particularly, it is numerically seen that $\lim_{n \to \infty} R_{\text{exact,0}}$ is finite—with $\lim_{n \to \infty} R_{\text{exact,0}} < R_0$—and $\lim_{n \to \infty} R_p = R_0$ in the case of Markov-modulated infectious contacts and exponential removals (scenarios A and B), which is closely related to the limiting results $\lim_{n \to \infty} R_{\text{exact,0}} = R_0$ and $\lim_{n \to \infty} R_p = R_0$ suggested by Artalejo and López-Herrero [14, Sections 3 and 4] in the basic SIR-model (scenario I). Under the assumption of Markov-modulated removals (scenarios C-H), it appears that $\lim_{n \to \infty} R_{\text{exact,0}}$ and $\lim_{n \to \infty} R_p$ are always finite, but with $\lim_{n \to \infty} R_{\text{exact,0}} > R_0$ and $\lim_{n \to \infty} R_p > R_0$. These limiting results for Markov-modulated SIR-models illustrate that, similarly to the SIR-model [14], the existence of repeated contacts between infectives and any other individual becomes negligible when the initial number $n$ of susceptibles increases.

5. Conclusions. In this paper, we develop a stochastic modeling framework that incorporates correlation between events into the dynamics of the SIR-model (El Maroufy et al. [25], Gani and Purdue [27], Kermack and McKendrick [37], Neuts and Li [40]). We construct Markov-modulated versions of the basic SIR-model $X$ by allowing one to deal with non-exponential distributional assumptions on the contacts between an infective and a susceptible individual, and the random duration of an infectious period. We may view the resulting augmented Markovian process $(X,Y)$ in Section 3 as consisting of basic states $(i,s) \in C(m,n)$, describing the numbers of infectives and susceptibles in the basic SIR-model, and an auxiliary variable
Figure 7. The expected value $R_p$ versus the basic reproduction number $R_0$ for scenarios C, D and I, and initial size $N = 1 + n \in \{10, 50, 100, 150\}$

called phase, which is a useful artifice in creating Markovian structures that enable easy computations from analytical and numerical perspectives. In our approach, we construct a Markov-modulated model by scheduling events at a compartmental level, rather than according to the individual behavior of infectives and susceptibles. Specifically, we use two scaled streams of Markovian arrivals with suitably chosen level of correlation, in such a way that the augmented process $(\mathcal{X}, \mathcal{Y})$ is a regular, time-homogeneous CTMC exhibiting correlation between events, and the phase process $\mathcal{Y}$ amounts to a state-dependent version of an MMAP where the types of marked events correspond to either a new infection (i.e., $(z_1, z_2) = (1, -1)$) or the removal of an infective (i.e., $(z_1, z_2) = (-1, 0)$) in the basic SIR-model. The process $\mathcal{X}$ and its augmented version $(\mathcal{X}, \mathcal{Y})$ possess identical structural properties regarding the infinitesimal rates $\lambda_{i,s}$ and $\mu_i$, and the fundamental arrival rates $\bar{\lambda}_{(i,s)}(z_1, z_2)$ of type-$(z_1, z_2)$ marked events; see Eq. (6). The variability between SIR-models with distinct level of correlation is discussed in Section 3 in terms of extinction times, the final size of the epidemic, and the basic reproduction number $R_0$. More particularly, the role of $R_0$, which overestimates the reproductive potential in the SIR-model (Artalejo and López-Herrero [14]), is replaced by two alternative random measures –the exact reproduction number $R_{exact,0}$ and the population transmission number $R_p$– that overcome this difficulty and provide valuable insight. We emphasize that $R_{exact,0}$ and $R_p$ are random variables, rather than an expected value, and they can be defined either at the invasion time or at any time after the invasion. The key feature of $R_{exact,0}$ (respectively, $R_p$) is that, during the infection period of a focal infective (respectively, the elapsed time until the first recovery occurs), it eliminates the effect of repeated contacts since $R_{exact,0}$ (respectively, $R_p$) does not count the number of contacts taking place between the focal
infective (respectively, the population of infectives) and any contacted individual which has previously been infected. Consequently, important differences between $\mathcal{R}_0$ and the exact measurements provided by $\overline{R}_{\text{exact},0}$ and $\overline{R}_p$ are observed (Artalejo and López-Herrero [14], Diekmann et al. [24]) when the initial size $n$ of susceptibles is small. The expected values $\overline{R}_{\text{exact},0}$ and $\overline{R}_p$ may thus contribute to the design of control strategies for preventing major outbreaks, by means of vaccination, representing exact counterparts of the classical vaccination coverage levels; see [14, Section 5].

For illustrative purposes, in Section 4 we compare $\mathcal{R}_0$ with the expected values $E[S(T_{[1,n]})]$, $\overline{R}_{\text{exact},0}$ and $\overline{R}_p$ at an invasion time. It is worth noting that, irrespectively of concrete specifications for the correlation structure of the infectious process, exponential removals (scenarios A, B, I) yield lower bounds for the expectations of the exact number $R_{\text{exact},0}$ of secondary cases and its population counterpart $R_p$ when removals are negatively correlated (scenarios D, E, H), while upper bounds are linked to removal processes governed by Markovian streams with positive correlation (scenarios C, F, G). Therefore, if the basic reproduction number $\mathcal{R}_0$ – in the stochastic setting, its random versions $R_{\text{exact},0}$ and $R_p$ – is commonly the first descriptor to be estimated during the emergence of a new infectious disease, our examples in Section 4 reveal that we must supplement their meaning with the correlation structure of the underlying Markovian streams for a new infection and the removal of an infective.

It is clear that, similarly to the extensions in [23, 29, 45], Markov-modulated events in the SIR-model can be scheduled at an individual level. For instance, in the general stochastic epidemic (see e.g. Bailey [16, Chapter 6], Gani and Purdue [27]), we may let the Poisson process of rate $\beta > 0$–governing infectious contacts between a concrete infective and each susceptible individual–, and the exponential law with expected value $\eta^{-1}$ –governing the infectious period of this infective– be replaced by two independent MAPs with characteristic matrices $(C_0, C_1)$ of order $\tilde{L}_1$, and $(D_0, D_1)$ of order $\tilde{L}_2$, respectively. Instead of states $(i, s, y_1, y_2)$ with $1 \leq y_k \leq L_k$ and $k \in \{1, 2\}$ (Section 3), this results in augmented states $(i, s, \tilde{y}_1, \tilde{y}_2)$ with vectors $\tilde{y}_k = (\tilde{y}_{k,1}, ..., \tilde{y}_{k,k})$ containing the numbers $\tilde{y}_{k,k'}$ of contact processes ($k = 1$) and removal processes ($k = 2$) at phases $1 \leq k' \leq \tilde{L}_k$, for $k \in \{1, 2\}$, at an arbitrary time. Therefore, a set of $(\begin{pmatrix} i-1 \\ \tilde{L}_1-1 \end{pmatrix})(\begin{pmatrix} s-1 \\ \tilde{L}_2-1 \end{pmatrix})$ augmented states is obtained for each basic state $(i, s) \in \mathcal{C}(m, n)$–whereas a constant number $L_1L_2$ of augmented states is used in Section 3 for the process $(X, Y)$–, when Markov-modulated events in the SIR-model are scheduled at an individual level. We are not aware of any reported numerical experiments for this extension of the SIR-model, but we can point out that, based on preliminary numerical examples (not reported here) for numbers $\tilde{L}_1 = \tilde{L}_2 = 3$ of phases and small initial numbers $(m, n)$ of infectives and susceptibles, the storage cost in floating-point numbers becomes so high that numerical computations of $E[S(T_{[1,n],s})]$, $\overline{R}_{\text{exact},0}$ and $\overline{R}_p$ are rapidly prone to numerical instability.

Modeling the spread of a disease by using correlated inter-event times is interesting in its own right, but another important problem not discussed in this paper is that of estimation of parameters; in drawing inferences, a few of the main references on statistical methods for epidemics are Anderson and May [7], Andersson and Britton [8], and Becker [19], which should be appropriately supplemented with the recent monograph by Buchholz et al. [22] on parameter fitting of Markovian arrival processes. In using the augmented process $(X, Y)$ in order to approximate the above Markov-modulated SIR-model with events scheduled at an individual
level, the two phase approach (Buchholz et al. [22, Section 5.3]) can be generalized to jointly fit parameters of the characteristic matrices \((C_0, C_1)\) of order \(L_1\), and \((D_0, D_1)\) of order \(L_2\). In the two phase approach, a PH distribution of order \(L_k\) with representation \((\eta(1), C_k)\) if \(k = 1\), and \((\eta(2), D_0)\) if \(k = 2\), is first fitted (Buchholz et al. [22, Chapter 3]), and the matrices \(C_1\) and \(D_1\) are then constructed in such a way that the inter-event time distribution of the resulting MAPs remains unchanged. Current research on these statistical aspects will be reported shortly when the data trace of the epidemic process consists of inter-event times or, alternatively, the numbers of events in finite intervals are available and not the detailed inter-event times.

**Acknowledgments.** We thank two anonymous referees whose comments and suggestions led to improvements in the manuscript. The authors are members of the “Stochastic Modelling Group, UCM-910211”, which was supported by the Complutense University of Madrid and Banco de Santander, call 2014-GR3/14.

**Appendix A. Proof of Eqs. (6)-(8).** In this appendix, we prove (6)-(8) for a fixed basic state \((i, s) \in C(m, n)\).

To begin with, we recall that, for state \((i, s) \in C(m, n)\), the phase process \(\mathcal{Y}\) can be seen as an MMAP with characteristic matrices

\[
(C^0_{i,s} \oplus D^0_{i,s}, C^1_{i,s} \otimes I_{L_1}, I_{L_1} \otimes D^*_1_{i,s}),
\]

where matrices \(C^*_1_{i,s} \otimes I_{L_2} \text{ and } I_{L_1} \otimes D^*_1_{i,s}\) are related to arrivals of type-\((-1, -1)\) and type-\((-1, 0)\) marks, respectively, and transitions between phases (recording no arrival) are governed by \(C^*_{0_{i,s}} \oplus D^*_{0_{i,s}}\). Then, based on the general theory of MMAPs (He and Neuts [32], He [33, Section 2.5]), the fundamental arrival rate of type-\((z_1, z_2)\) marks with \((z_1, z_2) = (1, -1)\) is defined as

\[
\tilde{\lambda}_{(i,s)}(1, -1) = \eta(i,s)(1)C^*_{1_{i,s}}e_{L_1},
\]

where \(\eta(i,s)(1)\) is the stationary vector of a finite CTMC with infinitesimal generator \(C^*_{0_{i,s}} + C^*_{1_{i,s}}\), that is, the entries of \(\eta(i,s)(1)\) are uniquely determined as the solution of \(\eta(i,s)(1)(C^*_{0_{i,s}} + C^*_{1_{i,s}}) = 0_{L_1}^T\) and \(\eta(i,s)(1)e_{L_1} = 1\). Since \(\tilde{\lambda}(1)C^*_{k_{i,s}} = \lambda_{i,s}C_k\) for \(k \in \{1, 2\}\), the system of linear equations \(\eta(i,s)(1)(C^*_{0_{i,s}} + C^*_{1_{i,s}}) = 0_{L_1}^T\) is equivalent to \(\eta(i,s)(1)(C_0 + C_1) = 0_{L_1}^T\), from which it follows that \(\eta(i,s)(1) = \eta(1)\), where \(\eta(1)\) denotes the stationary vector of the infinitesimal generator \(C_0 + C_1\). As a result, it is easily seen that \(\tilde{\lambda}_{(i,s)}(1, -1) = \lambda_{i,s}\), since \(\eta(1)C_1e_{L_1} = \tilde{\lambda}(1)\). The proof of (6) for type-\((z_1, z_2)\) marks with \((z_1, z_2) = (1, 0)\) is similarly derived, and is thus omitted.

The fundamental arrival rate \(\tilde{\lambda}_{(i,s)}\) of the MMAP has the form

\[
\tilde{\lambda}_{(i,s)} = \pi(i,s)\left(C^*_{1_{i,s}} \oplus D^*_{1_{i,s}}\right)e_{L_1L_2},
\]

where \(\pi(i,s)\) is the stationary vector of a finite CTMC with infinitesimal generator \(C^*_{1_{i,s}} \oplus D^*_{1_{i,s}}\) with 

\[
C^*_{1_{i,s}} = C^*_{0_{i,s}} + C^*_{1_{i,s}} \text{ and } D^*_{1_{i,s}} = D^*_{0_{i,s}} + D^*_{1_{i,s}}.
\]

By direct substitution, we may verify that \(\pi(i,s) = \eta(1) \otimes \eta(2)\), where \(\eta(2)\) is the stationary vector of the infinitesimal generator \(D_0 + D_1\). This results in

\[
\tilde{\lambda}_{(i,s)} = \eta(1)C^*_{1_{i,s}}e_{L_1} + \eta(2)D^*_{1_{i,s}}e_{L_2},
\]

which completes the proof of (7) since \(\eta(1)C^*_{1_{i,s}}e_{L_1} = \lambda_{i,s}\) and \(\eta(2)D^*_{1_{i,s}}e_{L_2} = \mu_i\).
The proof of (8) is based on the following expression for the correlation coefficient between $X_1$ and $X_k$ in an MAP of order $L_1$ with characteristic matrices $(C_0, C_1)$:

$$\rho_k = \frac{\theta(-C_0^{-1})(I_{L_1} - e_{L_1}(\theta)P^k(-C_0^{-1})e_{L_1})}{\theta(-C_0^{-1})(2I_{L_1} - e_{L_1}(\theta)(-C_0^{-1})e_{L_1})}, \quad (15)$$

where $P = (-C_0^{-1})C_1$ is the one-step transition probability matrix of the embedded Markov chain at arrival instants, and $\theta$ is its stationary vector; see, for example, the reference [11, Section 2.3.2]. An alternative expression for $\rho_k$ can be readily derived by noting that $\theta = \hat{\lambda}^{-1}(1)\eta(1)C_1$, which leads us to

$$\theta(-C_0^{-1})(I_{L_1} - e_{L_1}(\theta)P^k(-C_0^{-1})e_{L_1} + \hat{\lambda}^{-2}(1)$$

$$\theta(-C_0^{-1})(2I_{L_1} - e_{L_1}(\theta)(-C_0^{-1})e_{L_1} = 2\hat{\lambda}^{-1}(1)\eta(1)(-C_0^{-1})e_{L_1} - \hat{\lambda}^{-2}(1), \quad (17)$$

since $\eta(1)C_1 = -\eta(1)C_0$. For the scaled matrices $(C_{0,(i,s)}, C_{1,(i,s)})$, the resulting expression for $\rho_{k,(i,s)}$ has the form (15) with the numerator replaced by

$$\lambda_{i,s}^{-1}\eta(1) \left( \left( -\lambda_{i,s}^{-1}\hat{\lambda}(1)C_0^{-1} \right) \hat{\lambda}^{-1}(1)\lambda_{i,s}C_1 \right)^k \left( -\lambda_{i,s}^{-1}\hat{\lambda}(1)C_0^{-1} \right) e_{L_1} - \lambda_{i,s}^{-2} \quad (18)$$

and the denominator replaced by

$$2\lambda_{i,s}^{-1}\eta(1) \left( -\lambda_{i,s}^{-1}\hat{\lambda}(1)C_0^{-1} \right) e_{L_1} - \lambda_{i,s}^{-2}. \quad (19)$$

Then, Eq. (8) is easily derived from (15)-(19).

For the sake of completeness, we point out that, for every basic state $(i, s) \in \mathcal{C}(m, n)$, the matrix $(\hat{\lambda}^{-1}(1)\lambda_{i,s}C_0) \oplus (\hat{\lambda}^{-1}(2)\mu_iD_0)$ on the left side of Eqs. (13)-(14) is related to the phase process $\hat{\lambda}$. Note that it corresponds to the sub-matrix $C_{0,(i,s)} \oplus D_{0,(i,s)}$ in the underlying MMAP and, consequently, it governs transitions between joint phases $(y_1, y_2)$, recording no arrival of type $(z_1, z_2)$ with $(z_1, z_2) \in \{(1, 1), (-1, 0)\}$. Then, the matrix $-(\hat{\lambda}^{-1}(1)\lambda_{i,s}C_0) \oplus (\hat{\lambda}^{-1}(2)\mu_iD_0)$ is non-singular and non-negative since the fundamental arrival rate $\hat{\lambda}_{i,s}$ of the MMAP is strictly positive. More particularly, its elements record expected total times spent in the phase $(y_1', y_2')$, starting from an initial phase $(y_1, y_2)$, before the arrival of a type $(z_1, z_2)$ marked event with $(z_1, z_2) \in \{(1, 1), (-1, 0)\}$ is observed.

Appendix B. Markovian arrivals with positive and negative correlation. In our numerical experiments in Section 4, eight scenarios (Table 2) are defined from two predetermined MAPs with positively and negatively correlated events. The characteristic matrices of these MAPs and the resulting values of $\rho_1$ (Eq. (15)) are specified as follows:

(i) Markovian arrivals with positive correlation. We consider the rate matrices

$$E_0^+ = \begin{pmatrix} -0.87478 & 0 & 0.0 \\ 0.0 & -0.87478 & 0.0 \\ 0.0 & 0.0 & -94.76811 \end{pmatrix},$$

$$E_1^+ = \begin{pmatrix} 0.08748 & 0.0 & 0.28730 \\ 0.78730 & 0.0 & 0.08748 \\ 7.28985 & 0.0 & 87.47826 \end{pmatrix},$$

which yield the value $\rho_1 = 0.35016$. 
(ii) Markovian arrivals with negative correlation. In this case, we deal with

\[
E_0^- = \begin{pmatrix}
-0.87478 & 0.68353 & 0.0 \\
0.0 & -0.87478 & 0.0 \\
0.0 & 0.0 & -94.76811
\end{pmatrix},
\]

\[
E_1^- = \begin{pmatrix}
0.10377 & 0.0 & 0.08748 \\
0.08748 & 0.0 & 0.78730 \\
87.47826 & 0.0 & 7.28985
\end{pmatrix}.
\]

Then, we have that \( \rho_1 = -0.35016 \).

REFERENCES

[1] L. J. S. Allen, An Introduction to Stochastic Processes with Applications to Biology, Second edition. CRC Press, Boca Raton, FL, 2011.
[2] L. J. S. Allen, An Introduction to Mathematical Biology, Pearson Education, Inc., New Jersey, 2007.
[3] L. J. S. Allen, An introduction to stochastic epidemic models, in Mathematical Epidemiology (eds. F. Brauer, P. van der Driessche and J. Wu), Lecture Notes in Mathematics, Vol. 1945, Springer-Verlag, Berlin Heidelberg, (2008), 81–130.
[4] L. J. S. Allen and A. M. Burgin, Comparison of determistic and stochastic SIS and SIR models in discrete time, Mathematical Biosciences, 163 (2000), 1–33.
[5] E. Almaraz, A. Gómez-Corral and M. T. Rodríguez-Bernal, On the time to reach a critical number of infections in epidemic models with infective and susceptible immigrants, BioSystems, 144 (2016), 68–77.
[6] J. Amador and J. R. Artalejo, Modeling computer virus with the BSDE approach, Computer Networks, 57 (2013), 302–316.
[7] R. M. Anderson and R. M. May, Infectious Diseases of Humans; Dynamics and Control, Oxford University Press, Oxford, 1991.
[8] H. Andersson and T. Britton, Stochastic Epidemic Models and Their Statistical Analysis, Lecture Notes in Statistics, Vol. 151, Springer-Verlag, New York, 2000.
[9] J. R. Artalejo, A. Economou and M. J. López-Herrero, On the number of recovered individuals in the SIS and SIR stochastic epidemic models, Mathematical Biosciences, 228 (2010), 45–55.
[10] J. R. Artalejo and A. Gómez-Corral, A state-dependent Markov-modulated mechanism for generating events and stochastic models, Mathematical Methods in the Applied Sciences, 33 (2010), 1342–1349.
[11] J. R. Artalejo, A. Gómez-Corral and Q. M. He, Markovian arrivals in stochastic modeling: A survey and some new results, SORT – Statistics and Operations Research Transactions, 34 (2010), 101–144.
[12] J. R. Artalejo and M. J. López-Herrero, Quasi-stationary and ratio of expectations distributions: A comparative study, Journal of Theoretical Biology, 266 (2010), 264–274.
[13] J. R. Artalejo and M. J. López-Herrero, The SIS and SIR stochastic epidemic models: A maximum entropy approach, Theoretical Population Biology, 80 (2011), 256–264.
[14] J. R. Artalejo and M. J. López-Herrero, On the exact measure of disease spread in stochastic epidemic models, Bulletin of Mathematical Biology, 75 (2013), 1031–1050.
[15] J. R. Artalejo and M. J. López-Herrero, Stochastic epidemic models: New behavioral indicators of the disease spreading, Applied Mathematical Modeling, 38 (2014), 4371–4387.
[16] N. T. L. Bailey, The Mathematical Theory of Infectious Diseases and its Applications, Charles Griffin and Company, London, 1975.
[17] F. G. Ball, A unified approach to the distribution of total size and total area under the trajectory of infectives in epidemic models, Journal of Applied Probability, 18 (1986), 289–310.
[18] F. Ball and P. Neal, A general model for stochastic SIR epidemics with two levels of mixing, Mathematical Biosciences, 180 (2002), 73–102.
[19] N. G. Becker, Analysis of Infectious Disease Data, Chapman and Hall, London, 1989.
[20] A. J. Black and J. V. Ross, Computation of epidemic final size distributions, Journal of Theoretical Biology, 367 (2015), 159–165.
[21] T. Britton, Stochastic epidemic models: A survey, Mathematical Biosciences, 225 (2010), 24–35.
[22] P. Buchholz, J. Kriege and I. Felko, Input Modeling with Phase-Type Distributions and Markov Models, Theory and Applications, Springer, Dordrecht, 2014.

[23] D. Clancy, SIR epidemic models with general infectious period distribution, Statistics and Probability Letters, 85 (2014), 1–5.

[24] O. Diekmann, M. C. M. de Jong and J. A. J. Metz, A deterministic epidemic model taking account of repeated contacts between the same individuals, Journal of Applied Probability, 35 (1989), 448–462.

[25] H. El Maroufy, L. Omari and Z. Taib, Transition probabilities for generalized SIR epidemic model, Stochastic Models, 28 (2012), 15–28.

[26] H. El Maroufy, D. Kiouach and Z. Taib, Final outcome probabilities for SIR epidemic model, Communications in Statistics–Theory and Methods, 45 (2016), 2426–2437.

[27] J. Gani and P. Purdue, Matrix-geometric methods for the general stochastic epidemic, Mathematical Medicine and Biology, 1 (1984), 333–342.

[28] A. Gómez-Corral and M. López García, Modeling host-parasitoid interactions with correlated events, Applied Mathematical Modelling, 37 (2014), 5452–5463.

[29] A. Gómez-Corral and M. López-García, On SIR epidemic models with generally distributed infectious periods: Number of secondary cases and probability of infection, International Journal of Biomathematics, 10 (2017), 1750024 (13 pages).

[30] L. F. Gordillo, S. A. Marion, A. Martin-Löf and P. E. Greenwood, Bimodal epidemic size distributions for near-critical SIR with vaccination, Bulletin of Mathematical Biology, 70 (2008), 589–602.

[31] J. Grasman, Stochastic epidemics: The expected duration of the endemic period in higher dimensional models, Mathematical Biosciences, 152 (1998), 13–27.

[32] Q. M. He and M. F. Neuts, Markov chains with marked transitions, Stochastic Processes and Their Applications, 74 (1998), 37–52.

[33] Q. M. He, Fundamentals of Matrix-Analytic Methods, Springer, New York, 2014.

[34] V. Isham, Stochastic models for epidemics with special reference to AIDS, The Annals of Applied Probability, 3 (1993), 1–27.

[35] V. Isham, Stochastic models for epidemics, in Celebrating Statistics: Papers in Honour of Sir David Cox on his 80th Birthday (eds. A.C. Davison, Y. Dodge and N. Wermuth), Oxford University Press, Oxford, 33 (2005), 27–54.

[36] M. J. Keeling and P. Rohani, Modeling Infectious Diseases in Humans and Animals, Princeton University Press, Princeton, 2008.

[37] W. O. Kermack and A. G. McKendrick, A contribution to the mathematical theory of epidemics, Proceedings of the Royal Society of London A, 115 (1927), 700–721.

[38] G. Latouche and V. Ramaswami, Introduction to Matrix Analytic Methods in Stochastic Modelling, ASA-SIAM, Philadelphia, 1999.

[39] M. F. Neuts, Matrix-geometric Solutions in Stochastic Models: An Algorithmic Approach, The Johns Hopkins University Press, Baltimore, 1981.

[40] M. F. Neuts and J. M. Li, An algorithmic study of SIR stochastic epidemic models, in Athens Conference on Applied Probability and Time Series Analysis, Lecture Notes in Statistics, Vol. 114, Springer, New York, (1996), 295–306.

[41] P. Picard and C. Lefèvre, A unified analysis of the final size and severity distribution in collective Reed-Frost epidemic processes, Advances in Applied Probability, 22 (1990), 269–294.

[42] J. V. Ross, Invasion of infectious diseases in finite homogeneous populations, Journal of Theoretical Biology, 289 (2011), 83–87.

[43] I. W. Saunders, A model for myxomatosis, Mathematical Biosciences, 48 (1980), 1–15.

[44] T. Sellke, On the asymptotic distribution of the size of a stochastic epidemic, Journal of Applied Probability, 20 (1983), 390–394.

[45] G. Streftaris and G. J. Gibson, Non-exponential tolerance to infection in epidemic systems—modeling, inference, and assessment, Biostatistics, 13 (2012), 580–593.

Received June 2016; revised April 2018.

E-mail address: ealmarazluengo@mat.ucm.es
E-mail address: antonio.gomez@icmat.es