EXACT DESCRIPTION OF SIR-BASS EPIDEMICS ON 1D LATTICES

GADI FIBICH AND SAMUEL NORDMANN*

Department of Applied Mathematics
Tel Aviv University, Israel

(Communicated by Hirokazu Ninomiya)

ABSTRACT. This paper is devoted to the study of a stochastic epidemiological model which is a variant of the SIR model to which we add an extra factor in the transition rate from susceptible to infected accounting for the inflow of infection due to immigration or environmental sources of infection. This factor yields the formation of new clusters of infections, without having to specify a priori and explicitly their date and place of appearance.

We establish an exact deterministic description for such stochastic processes on 1D lattices (finite lines, semi-infinite lines, infinite lines) by showing that the probability of infection at a given point in space and time can be obtained as the solution of a deterministic ODE system on the lattice. Our results allow stochastic initial conditions and arbitrary spatio-temporal heterogeneities on the parameters.

We then apply our results to some concrete situations and obtain useful qualitative results and explicit formulae on the macroscopic dynamics and also the local temporal behavior of each individual. In particular, we provide a fine analysis of some aspects of cluster formation through the study of patient-zero problems and the effects of time-varying point sources.

Finally, we show that the space-discrete model gives rise to new space-continuous models, which are either ODEs or PDEs, depending on the rescaling regime assumed on the parameters.

1. Introduction.

1.1. **Goal of the paper.** This paper is devoted to the study of a stochastic epidemic model on 1D (directed or undirected) lattices with a general epidemic source term and arbitrary heterogeneity on the parameters and initial conditions. Our model is a variant of the stochastic SIR model, originally introduced in its deterministic version by Kermack and McKendrick [30], which forms the framework of almost any model used in epidemiology. Despite the extensive literature on the SIR model (see, e.g., the books of Bailey [5], Anderson and May [1], Mollison [34] and the review of Heathcote [24]), some fundamental and practical questions remain open. In the current context of the propagation of the COVID-19, an in-depth study of this model and its variants is a major challenge.

In addition to the usual mechanism of contagion occurring in epidemics, we also assume the presence of an epidemiological *source* term, qualified as "*external*

influences". This term acts as an additional mechanism that leads to the transition from susceptible to infected, which does not depend on the level of infection in the population, and accounts for the inflow of infection due to immigration or environmental sources of infection. In our framework, it is allowed to depend on space and time. From a modeling point of view, dealing with external influences is important since they yield the formation of new clusters of infections, without having to specify a priori and explicitly their date and place of appearance. Such external influences are not traditionally included in the SIR model and refer more to the Bass model [8], originally introduced to describe the diffusion of new products. In the marketing context, external influences account for the spontaneous adoptions of the product induced by advertising or mass media, while "social imitation" and "word of mouth" are formally regarded as a "contagion" mechanism. In turn, the population of *potential adopters* of a product (resp. *adopters*) are seen as formal analogs of the population of *susceptible* (resp. *infected*) in the epidemiology context. Although each modeling situation has its own specificities, we keep a unified approach to both phenomena. In what follows, we use more frequently the terminology borrowed from epidemiology, but our considerations can be transposed to the context of the adoption of a new product. Let us mention that the *recovery* of infected individuals operating in the SIR model is also relevant in the marketing context even though it has not been included in the Bass model until recently [17]. Similar models are also used in other social situations: propagation of riots [9], rumours [15], etc.

. The main goal of the present paper is to establish an *exact deterministic description* for such stochastic SIR-Bass epidemics on 1D lattices. Specifically, we show that the probability of infection at a given point in space and time can be obtained as the solution of a deterministic ODE system on the lattice. We then apply our results to some concrete situations and obtain useful qualitative results and explicit formulae. In particular, we provide a fine analysis of some aspects of cluster formation through the study of *patient-zero problems* and the effects of *time-varying point sources*.

For simplicity and a clear presentation of our results and methodology, we focus in this paper on 1D lattices (i.e., finite, semi-infinite, and infinite lines) where the infection only occurs from a node to its direct neighbors. Our approach can be adapted without difficulties to the case where the contact network is a tree (i.e., a graph with no cycles). Some of our ideas may also apply to a more general context, which will be the subject of future works.

As a counterpart to the strong assumption on the structure of the graph, we try to keep the assumptions on the parameters and initial conditions as broad as possible by allowing any type of spatial-temporal heterogeneity of the coefficients. Understanding the effect of such local heterogeneities, even for simple networks, is still a major challenge in epidemiology [35, Section 1].

We emphasize that our results provide an exact description of the process at the individual-scale, that is, we describe both the macroscopic dynamics of the epidemics and the local temporal behavior of each individual. This takes us one step further than the usual description at the population-scale through aggregate quantities such as the expected fraction of infected individuals in the population. Such a spatial-temporal description at the local scale is required for predicting population-level dynamics from individual-level observations, all the more in the situation of strong heterogeneities.

Outline. We give a brief state of the art of the mathematical epidemiology literature in Section 1.2, which allow us to present the context of our study, to motivate it, and to highlight the novelty of our results. Our mathematical framework and its link with classical models are presented in Sections 2 and 3.

We first present and prove our results dealing with 1D one-sided lattices in which the contagion can only occur in one direction. We focus in Section 4 on the Bass model when there is no recovery. Section 5 deals with the full SIR-Bass model. Our main results (Theorem 4.1, Theorem 5.1) establish that the probability of a certain node to be susceptible at some time satisfies a differential equation which is continuous in time and discrete in space. This provides an exact spatial-temporal deterministic description of the stochastic process, for any type of heterogeneity on the parameters and stochastic initial conditions. We then use these results to perform a further analysis and derive explicit formulae in several typical situations, such as the spatially homogeneous case (Proposition 1, Proposition 5), the patient-zero problem (Proposition 2, Proposition 7), and the case of a time-varying point sources (Proposition 4, Proposition 6). We also propose a further analysis on the speed of propagation of epidemics (Proposition 3) as well as a detailed comparison of our model with the classical aggregate SIR model (Section 5.3).

In Section 6, we extend our results to two-sided 1D lattices (corresponding to an undirected underlying graph) in which epidemics can propagate in both directions. We derive a new useful formula that expresses the probability of infection in a two-sided lattice in terms of the probabilities of infections in one-sided lattices. In Section 7, we discuss two different space rescalings giving rise to either a limiting space-continuous ODE or a limiting PDE coupling time and space. Finally, we give some concluding remarks in Section 8.

1.2. **Context and motivation.** To present the context of our work, we provide a brief overview of the mathematical study of epidemiological models. Most epidemiological models proposed in the literature fall into two broad categories: a deterministic approach using either ODEs [25] or PDEs of the reaction-diffusion type [13,36], and a stochastic approach using individual-centered Poisson processes (agent-based models) [1,5]. The deterministic and stochastic approaches have mostly developed in parallel in the literature. In general, deterministic models can be obtained as the "mean-field" limits [31, 39] of stochastic models. Mean-field models describe the average dynamics and typically apply to sufficiently large populations where stochastic fluctuations and individual variabilities are negligible [32]. They have the advantage of being more amenable to analysis and to give a clear idea of some qualitative aspects of epidemic dynamics. Deterministic models are particularly useful to study the spatial speed of propagation and the shape of the propagation front of epidemics [14]. On the other hand, stochastic models are capable of reproducing observed data way more accurately, both quantitatively and qualitatively, see, e.g., [5, Section 5.1 & 7] and [34]. The main reasons for this are, firstly, that stochasticity may have a dramatic influence on the overall dynamics since small populations are involved (for example during the outbreak of an epidemic) and, secondly, that the stochastic framework is readily well suited for the incorporation of strong heterogeneities or complex social structures which are known to be of tremendous importance in epidemics propagation.

The simplest models, referred to as *aggregate* or *compartmental* models, assume that the population is homogeneously mixed [5, Section 5.10], that is, the connectivity between two individuals is constant among all pairs of individuals. In this

situation, individuals are interchangeable and so the behavior of the system can be described by aggregate quantities (such as the expected fraction of infected individuals). The classical analysis of these models relies on strong symmetry assumptions on the parameters and initial conditions [31, 32].

However, many studies reveal that homogeneous mixing only holds in small groups but not in large populations, see e.g. [5, Section 5.10] and [29]. Intuitively, at the scale of a city, it does not make sense to assume that an individual interacts as much with members of his family as with a socially unrelated individual living on the other side of the city. The limited number of interactions of each individual tends to hinder the spread of infection and to cause the formation of localized clusters. For these reasons, the network within which social interactions occur has a dramatic impact on the epidemic propagation, and so aggregate models turn out to be inaccurate to describe the spatial-temporal dynamics of the epidemics.

A more elaborate approach is to consider that each individual interacts only with a relatively small subset of the population. The connectivity between individuals is usually encoded by a weighted graph, called *contact network* in the epidemiology context [3]. We point out that this formalism assumes that social interactions are fixed once and for all, which is not a problem when considering a sufficiently short epidemic episode (evolving networks have also been proposed, see [16] and references therein). Contact networks are typically sparse and may embed various social variables such as place of residence, age, profession, etc. The different classes of contact networks considered in the literature can be roughly classified according to their *clustering* level [4, 27] reflecting the extent to which the system is capable of generating and maintaining localized clusters of infections. For example, *lattice networks* [33, 37] regard the population as placed on a regular grid of points with spatially localized connections, yielding a high clustering level; on the contrary, *random networks* [2, 6, 12] have a low clustering level since connections between individuals occur randomly regardless of their locations. Lattices and random networks appear as two opposite extreme instances, each one focusing on certain aspects of the dynamics observed in epidemics and discarding others. Small-world or scale-free networks have been proposed as examples of more realistic configurations, combining both localized and rare long-range connections, which display intermediate clustering levels. For more details, we refer the reader to the reviews of Keeling & Eames [29], Britton [10] and Danon et. al. [11]. Let us emphasize here that lattices, especially 2D lattices, are the physically relevant contact networks when considering contagion phenomena that primarily depend on the physical distance between the infected and susceptible individuals, which happens to be the case in the propagation of some diseases such as Measles [23] or the diffusion of some products such as solar panels [22].

Most of the approaches developed for studying complex networks make strong assumptions of homogeneity and symmetry at the individual-scale and are mainly concerned with aggregate quantities (such as the total number of infected) which describe the dynamics at the population level but fail to precisely grasp the dynamics at the individual-scale. It remains nevertheless a major challenge to understand the local behavior and the effects of strong heterogeneity [35, Section 1]. For example, the study of the *patient-zero* problem (where the epidemics spread from a single infected individual at initial time) or the influence of a spatially localized source of infection gives significant insights into the spatial propagation of epidemic clusters.

Several approximation methods have been proposed to deal with complex networks with a high level of heterogeneity. The goal of these approaches is to provide deterministic approximations that go beyond the insufficient mean-field approximation, but that are still less computationally demanding and more amenable to analysis than the full stochastic models. The pair-approximation method [28] consists in formally neglecting the possible intercorrelation between the nodes induced by cycles in the contact network. Under this formal approximation, the dynamics can be described by a system of ODEs featuring aggregate quantities such as the average number of neighbors per node and the ratio of triples of nodes which form an interconnected triangle (this quantity reflects the level of clustering of the contact network). While this method offers great flexibility and results which are quite in agreement with numerical simulations of the stochastic process, its theoretical ground is far from being completely understood. Other approximations have been proposed [7,26], such as degree-based models, probability generating function (PGF) formalism, edge-based compartment modeling, but all these methods can be derived from the pair-approximation method [26].

Regarding theoretical results and exact deterministic descriptions of the stochastic process, most of the studies, if not all, assume a tree-like structure on the contact network at some level. If the contact network is a tree-graph (i.e. a loopless graph), it is well known that the pair-wise approximation is exact [38] and yields an exact description of the stochastic process, namely, the probabilities of infection satisfy a system of ODEs. This allows a fine theoretical analysis of the dynamics. Results are also available when the graph has a local tree-like structure (such as random *configuration* networks [11]) at the limit as the population size becomes large and under homogeneity assumptions [11,12].

As mentioned in the introduction, the present paper deals with 1D lattices, which are particular cases of tree-graphs, and for which exact equations are available. The novelty of our approach is in the combination of the following ingredients. Firstly, we study the SIR-Bass model, that is, we add to the classical SIR model an external source term accounting for a spontaneous inflow of infection; secondly we allow any heterogeneity on the coefficients and allow stochastic initial conditions; thirdly, we use the obtained exact deterministic description to perform a further analysis at a local scale and derive explicit formulae on some specific typical epidemiological instances; finally we show that the space-discrete model gives rise to new space-continuous models, which are either ODEs or PDEs, depending on the rescaling regime assumed on the parameters.

2. **Framework.** In this paper, we consider time-discrete stochastic processes that converge to time-continuous processes as the time step tends to zero. Discrete and continuous models each have their own advantages: the former can be simulated numerically directly, while the latter is more amenable for analysis.

Time-discrete process. Let us fix a time-step $0 < \Delta t \ll 1$ and consider random processes represented by a family of random variables $((x_k^n)_{k \in \mathcal{K}})_{n \in \mathbb{N}}$, where x_k^n represents the *state* of individual $k \in \mathcal{K}$ at time $t^n := n\Delta t$. The set $\mathcal{K}$ can be finite $\mathcal{K} = \{1, \ldots, K\}$, infinite $\mathcal{K} = \mathbb{Z}^d$, or semi-infinite $\mathcal{K} = \mathbb{N}$. The state of individual/node k at time t^n can be *susceptible* $(x_k^n = s)$, *infected* $(x_k^n = i)$, or *recovered*

$(x_k^n = r)$. We define the following events:

$$S_k^n \text{ is the event that } x_k^n = s,$$
$$I_k^n \text{ is the event that } x_k^n = i,$$
$$R_k^n \text{ is the event that } x_k^n = r,$$

and denote their probabilities by

$$[S_k^n] = \mathbb{P}(S_k^n), \qquad [I_k^n] = \mathbb{P}(I_k^n), \qquad [R_k^n] = \mathbb{P}(R_k^n).$$

The events S_k^n, I_k^n, R_k^n are disjoint and complementary, and so

$$[S_k^n] + [I_k^n] + [R_k^n] \equiv 1, \qquad \forall n \geq 0, \ k \in \mathcal{K}. \tag{1}$$

We assume that the stochastic dynamics in the time-interval

$$(t^n, t^{n+1}) = (n\Delta t, (n+1)\Delta t)$$

is governed by the following mechanisms:

- If node k is susceptible at t^n, it can become infected
    - by a source term ("external influence") at a rate of p_k^n,
    - and by contagion from infected individual i ("internal influence") at a rate of q_{ik}^n.

  These effects are additive.
- If a node k is infected at t^n, it can recover at a rate of r_k^n.
- If node k is recovered at t^n, it remains so for all later times.

The transition probabilities are thus given by

$$\mathbb{P}\left(I_k^{n+1}\big|\boldsymbol{X}^n\right) = \begin{cases} \Delta t \left(p_k^n + \sum_{i \in \mathcal{K}} q_{ik}^n \mathbb{1}_{I_i^n} \right), & \text{if } x_k^n = s, \\ 1 - r_k^n \Delta t, & \text{if } x_k^n = i, \\ 0, & \text{if } x_k^n = r, \end{cases} \tag{2a}$$

where $\boldsymbol{X}^n = (x_k^n)_{k \in \mathcal{K}}$ is the state of the network at time t^n,

$$\mathbb{1}_{I_i^n} = \begin{cases} 1, & \text{if } x_i^n = i, \\ 0, & \text{else}, \end{cases}$$

and by

$$\mathbb{P}\left(R_k^{n+1}\big|\boldsymbol{X}^n\right) = \begin{cases} 0, & \text{if } x_k^n = s, \\ r_k^n \Delta t, & \text{if } x_k^n = i, \\ 1, & \text{if } x_k^n = r. \end{cases} \tag{2b}$$

The parameters p_k^n, q_{ik}^n, r_k^n are nonnegative and may depend on space k and time n. We also assume that $\boldsymbol{X}^{n+1}$ is a family of random variables that are mutually independent when conditioned with respect to X^n. This way, the process is well defined. The matrix $(q_{ik})_\mathcal{K}$ represents the weighted directed edges of the contact graph between the nodes.

The state variables (x_k^n) are coupled through the sum $\sum_{i \in \mathcal{K}} q_{ik}^n \mathbb{1}_{I_i^n}$, see (2a). This term, which embeds the structure of the graph, has a central role in the analysis. As noted previously, in this paper we focus on the case of 1D (directed or undirected) lattices, where $q_{ik}^n \equiv 0$ when $i \neq k \pm 1$.

Initial conditions. The starting point of the stochastic dynamics (2) is the initial conditions

$$(x_k^0)_{k \in \mathcal{K}}, \qquad x_k^0 := x_k(t = 0),$$

which are assumed to be uncorrelated random variables taking values in $\{s, i, r\}$. Thus, in terms of the probabilities $[S_k^n]$, $[I_k^n]$ and $[R_k^n]$, we consider the initial conditions

$$\begin{cases} [S_k^0], [I_k^0], [R_k^0] \in [0, 1], & k \in \mathcal{K}, \\ [S_k^0] + [I_k^0] + [R_k^0] \equiv 1, & k \in \mathcal{K}, \\ x_k^0 \perp\!\!\!\perp x_l^0 & k \neq l \in \mathcal{K}. \end{cases} \tag{3}$$

In some applications, it is natural to consider *deterministic* (*pure*) initial conditions, where the initial state of each node is given in a deterministic way. This is a special case of (3), when x_k^0 is determined with probability 1 among the possible states $\{s, i, r\}$. For example, in the SIR model, it is standard to assume that the epidemics start from "patients zero", and so that initially, all nodes are susceptible, except for a few which are infected. In the Bass model, the standard assumption is that the whole population is susceptible (non-adopters) when the new product is first introduced into the market.

In this paper, we analyze the more general case (3) where the initial state of each node is given by a random variable. In addition to being more general from a mathematical point of view, dealing with stochastic initial conditions is important from an application point of view. For example, in the context of epidemiology, it may be that the only available information is the percentage of the population in each area which are infected. Similarly, in the context of the diffusion of a new product, sometimes only the fraction of the population that has adopted the product in each area is available, rather than a precise description of the state of each individual.

Time-continuous limit. As noted, the n superscript denotes a state variable evaluated at the discrete time $t^n = n\Delta t$. Throughout the paper, we assume the following:

Time-continuous limit assumptions. *The parameters p_k^n, q_k^n and r_k^n in (2) converge to some functions $p_k(t)$, $q_k(t)$ and $r_k(t)$ respectively, as $\Delta t \to 0$ and $n \to +\infty$ so that $t = n\Delta t$ is constant.*

. Then, we classically have that the time-discrete process (x_k^n) converges to a time-continuous Poisson-type process $(x_k(t))$. We use the same notations as in the time-discrete setting, by replacing the dependence on n with a dependence on t. Thus,

$$S_k(t) \text{ is the event that } x_k(t) = s,$$
$$I_k(t) \text{ is the event that } x_k(t) = i,$$
$$R_k(t) \text{ is the event that } x_k(t) = r,$$

and

$$[S_k](t) = \mathbb{P}(S_k(t)), \qquad [I_k](t) = \mathbb{P}(S_k(t)), \qquad [R_k](t) = \mathbb{P}(S_k(t)).$$

The time-continuous process is a Poisson-type process, defined as follows. The transition of individual k from *susceptible* to *infected* at time t occurs through an exponential law with the rate $\lambda_k^{s \to i}(t) = p_k(t) + \sum_{i \in \mathcal{K}} q_{ik}(t) \mathbb{1}_{I_i(t)}$, and from *infected* to *recovered* with the rate $\lambda_k^{i \to r}(t) = r_k(t)$.

3. **Aggregate models.** The aggregate version of the above model (2) is obtained when the parameters and initial conditions do not depend on k, which corresponds to the case where the underlying graph is homogeneous and complete. In this case, in the limit as the population size becomes infinite, the dynamics is governed by the aggregate SIR-Bass model [17] which is given by the system of ODEs

$$\begin{cases} S'(t) = -S(p + qI), \\ I'(t) = S(p + qI) - rI, \\ R'(t) = rI, \end{cases} \tag{4}$$

where S, I, and R represent the fraction of susceptible, infected, and recovered, respectively (or equivalently the probability that any node is susceptible, infected, or recovered respectively). Similarily to (1), we have that

$$S(t) + I(t) + R(t) \equiv 1, \qquad t \geq 0. \tag{5}$$

If we take $p = 0$ in (4), i.e., we assume that there is no source term, then the SIR-Bass model reduces to the classical aggregate SIR model (see e.g. [25])

$$\begin{cases} S'(t) = -qSI, \\ I'(t) = qSI - rI, \\ R'(t) = rI. \end{cases} \tag{6}$$

To have a non-trivial dynamics in (6), we need to assume the presence of patient(s) zero, i.e., $I(0) > 0$.

Alternatively, if we take $r = 0$ and $R(t = 0) = 0$ (i.e. assume that there is no recovery), then $R(t) \equiv 0$ and so by (5) we have $I = 1 - S$. Hence, (4) reduces to the original Bass model [8]:

$$I'(t) = (1 - I)(p + qI). \tag{7}$$

If we further assume that $I(0) = 0$, this leads to the Bass formula

$$I_{Bass}(t) = \frac{1 - e^{-(p+q)t}}{1 + \frac{q}{p}e^{-(p+q)t}}. \tag{8}$$

Finally, assuming no recovery ($r = R(t = 0) = 0$) and no source term ($p = 0$), the SIR-Bass model (4) reduces to the classical SI model

$$\begin{cases} S'(t) = -qSI, \\ I'(t) = qSI. \end{cases} \tag{9}$$

4. **Bass model on 1D one-sided lattices.** We begin our analysis with the case of a 1D one-sided lattice, that is, when

$$q_{ik} = 0 \qquad \text{if } i \neq k - 1. \tag{10}$$

Therefore, infection of node k at time-step n can be caused either by a source term at the rate p_k^n, or by contagion from its left neighbor at the rate q_k^n. Hence, (2a)

reduces to[1]

$$\mathbb{P}\left(I_k^{n+1}\big|\boldsymbol{X}^n\right) = \begin{cases} \left(p_k^n + q_k^n \mathbb{1}_{I_{k-1}^n}\right)\Delta t, & \text{if } x_k^n = s, \\ 1 - r_k^n \Delta t, & \text{if } x_k^n = i, \\ 0, & \text{if } x_k^n = r, \end{cases} \tag{12a}$$

and (2b) remains unchanged, that is:

$$\mathbb{P}\left(R_k^{n+1}\big|\boldsymbol{X}^n\right) = \begin{cases} 0, & \text{if } x_k^n = s, \\ r_k^n \Delta t, & \text{if } x_k^n = i, \\ 1, & \text{if } x_k^n = r, \end{cases} \tag{12b}$$

In this section, we focus on the Bass model case when there is no recovery, i.e., when

$$r_k^n \equiv 0 \qquad \text{and} \qquad [R_k^0] \equiv 0, \quad \forall k \in \mathcal{K}. \tag{13}$$

This implies that $[R_k^n] \equiv 0$ and so by (1),

$$[S_k^n] + [I_k^n] \equiv 1, \qquad \forall k, n. \tag{14}$$

Therefore, (12a) and (12b) reduce to

$$\mathbb{P}\left(I_k^{n+1}\big|\boldsymbol{X}^n\right) = \begin{cases} \Delta t\left(p_k^n + q_k^n \mathbb{1}_{I_{k-1}^n}\right), & \text{if } x_k^n = s, \\ 1, & \text{if } x_k^n = i. \end{cases} \tag{15}$$

4.1. **Deterministic description.** In what follows, we show that the probability $[S_k^n](t)$ of node k to be susceptible at time t^n satisfies a deterministic discrete PDE (here, the term "discrete" means that the time and space derivatives are replaced by finite differences). We then let $\Delta t \to 0$ to deduce that the probabilities $[S_k](t)$ satisfy a system of differential equations.

Theorem 4.1. *Assume that the nodes are placed on a 1D one-sided graph, see* (10), *that the initial conditions are stochastic and uncorrelated, see* (3), *and that there is no recovery, see* (13), *so that the stochastic dynamics are governed by* (15) *where* $\mathcal{K}$ *is either a finite, semi-infinite or infinite line. Then for all* $k \in \mathcal{K}$, *as* $\Delta t \to 0$, $[S_k]$ *satisfies the differential equation*

$$[S_k]'(t) + (p_k(t) + q_k(t))[S_k](t) - q_k(t)e^{-\int_0^t p_k(s)ds}[S_k^0][S_{k-1}](t) = 0, \tag{16}$$
$$[S_k](t = 0) = [S_k^0].$$

In particular, if $p \equiv p_k^n$ *and* $q \equiv p_k^n$ *are independent of* k *and* n, *then* (16) *reads*

$$[S_k]' + (p + q)[S_k] - qe^{-pt}[S_k^0][S_{k-1}] = 0, \qquad [S_k](t = 0) = [S_k^0]. \tag{17}$$

---

[1] In the case of the semi-infinite line $\mathcal{K} = \mathbb{N}$, the transition probability for the boundary node $k = 0$ is

$$\mathbb{P}\left(I_0^{n+1}\big|\boldsymbol{X}^n\right) = \begin{cases} p_0^n \Delta t, & \text{if } x_0^n = s, \\ 1 - r_0^n \Delta t, & \text{if } x_0^n = i, \\ 0, & \text{if } x_0^n = r, \end{cases} \tag{11}$$

Proof. A straightforward computation gives

$$\mathbb{P}(S_k^n) - \mathbb{P}(S_k^{n+1})$$
$$= \mathbb{P}(I_k^{n+1} \cap S_k^n) = \mathbb{P}(I_k^{n+1} \cap S_k^n \cap S_{k-1}^n) + \mathbb{P}(I_k^{n+1} \cap S_k^n \cap I_{k-1}^n)$$
$$= \underbrace{\mathbb{P}(I_k^{n+1} | S_k^n \cap S_{k-1}^n)}_{=p_k^n \Delta t} \mathbb{P}(S_k^n \cap S_{k-1}^n) + \underbrace{\mathbb{P}(I_k^{n+1} | S_k^n \cap I_{k-1}^n)}_{=(p_k^n + q_k^n)\Delta t} \mathbb{P}(S_k^n \cap I_{k-1}^n).$$

Since $\mathbb{P}(S_k^n \cap I_{k-1}^n) = \mathbb{P}(S_k^n) - \mathbb{P}(S_k^n \cap S_{k-1}^n)$, we obtain that

$$\frac{\mathbb{P}(S_k^{n+1}) - \mathbb{P}(S_k^n)}{\Delta t} + (p_k^n + q_k^n)\mathbb{P}(S_k^n) - q_k^n \mathbb{P}\left(S_k^n \cap S_{k-1}^n\right) = 0. \qquad (18)$$

In principle, we now have to deal with the term $\mathbb{P}(S_k^n \cap S_{k-1}^n)$ which involves the 2-marginals of the process through the joint laws of x_{k-1}^n and x_k^n. The 2-marginals, however, can be expressed through the 1-marginal $([S_k])_{k \in \mathcal{K}}$ as follows. Let us write

$$\mathbb{P}(S_k^n \cap S_{k-1}^n) = \mathbb{P}(S_k^n \cap S_k^0 \cap S_{k-1}^n) = \mathbb{P}(S_k^n | S_k^0 \cap S_{k-1}^n)\mathbb{P}(S_k^0 \cap S_{k-1}^n) \qquad (19)$$

and compute each term of the product separately. First, since the initial conditions are uncorrelated, the variables x_k^0 and x_{k-1}^n are independent, hence

$$\mathbb{P}(S_k^0 \cap S_{k-1}^n) = \mathbb{P}(S_k^0)\mathbb{P}(S_{k-1}^n). \qquad (20)$$

This identity follows from the fact that there is no path from k to $k-1$, therefore the computation of the probability $\mathbb{P}(S_{k-1}^n)$ through (15) does not involve x_k^0. Second, if k is not infected at initial time and $k-1$ is not infected by t^n, then k can only become infected through the source term p_k. In this case, node k remains susceptible during the time interval (t^n, t^{n+1}) with a probability $1 - p_k^n$, i.e.,

$$\mathbb{P}(S_k^n | S_k^0 \cap S_{k-1}^n) = \prod_{i=0}^{n-1} \left(1 - p_k^i \Delta t\right) = e^{-\sum_{i=0}^{n-1} p_k^i \Delta t}(1 + O(\Delta t)) \qquad (21)$$
$$= e^{-\int_0^t p_k(\cdot)}(1 + O(\Delta t)).$$

Note by the way that identities (20)-(21) does not hold if the graph is a circle (i.e. $\mathcal{K} = \mathbb{Z}/K\mathbb{Z}$) since, roughly speaking, node $k-1$ can become infected by an epidemic starting from k. Substituting (21) and (20) in (19), we derive the *closure identity*

$$\mathbb{P}(S_k^n \cap S_{k-1}^n) = e^{-\int_0^{t^n} p_k(s)ds}\mathbb{P}(S_k^0)\mathbb{P}(S_{k-1}^n)(1 + O(\Delta t)). \qquad (22)$$

Finally, plugging (22) into (18), one gets

$$\frac{\mathbb{P}(S_k^{n+1}) - \mathbb{P}(S_k^n)}{\Delta t} + (p_k^n + q_k^n)\mathbb{P}(S_k^n) - q_k^n e^{-\int_0^{t^n} p_k(s)ds}\mathbb{P}(S_k^0)\mathbb{P}(S_{k-1}^n) = O(\Delta t). \quad (23)$$

Letting $\Delta t \to 0$ gives the result. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\square$

Remark 1. The key element in the proof is the closure identity (22). This identity holds pointwise whenever a node is only influenced by a single node. Therefore, even if we drop the assumption that the graph is 1D and let it be general, relation (16) still holds pointwise at any node which is only influenced by a single node.

Remark 2. The initial conditions $[S_k^0]$ explicitly appear in ODE (16) for $[S_k]$, and not just in the initial conditions. As mentioned already, the case of deterministic initial conditions is a special case where $[S_k^0] \in \{0, 1\}$.

Theorem 4.1 provides a spatial-temporal deterministic description of the stochastic dynamics (15). Once compute the solution $[S_k](t)$ through equation (16), we can complete the description of the stochastic process, as follows: $[I_k]$ is deduced from (14), the second-order marginals can be computed using the closure identity (22), and any higher-order marginal can be computed using the same methodology.

We emphasize that equation (16) on $[S_k](t)$ is exact and does not result from any approximation (such as the mean-field approximation) or from the assumption that the population is infinite.

4.2. **Explicit solutions.** In the spatially-homogeneous case, one can obtain explicit solutions of (16):

Proposition 1 (Spatially-homogeneous Bass solution). *Assume the conditions of Theorem 4.1. If the initial condition $[S_k^0] \equiv [S^0]$ and the parameters $p_k(t) \equiv p(t), q_k(t) \equiv q(t)$ do not depend on space k (we allow, however, p and q to depend on time t), then $[S_k](t) = [S](t)$ does not depend on k, and is given by*

$$[S](t) = [S^0]e^{-\int_0^t (p(s)+q(s))ds+[S^0]\int_0^t q(s)e^{-\int_0^s p}ds}. \tag{24}$$

Proof. By translation symmetry, $[S_k](t)$ does not depend on k. Hence, equation (16) reads

$$[S]' + \left(p(t) + q(t)\left(1 - [S^0]e^{-\int_0^t p(s)ds} \right) \right)[S] = 0, \qquad [S](t=0) = [S^0].$$

Expression (24) follows from a straightforward integration of the above equation. $\square$

If, in addition to the assumptions of Proposition 1, p and q do not depend on t, then

$$[S](t) = \begin{cases} [S^0]e^{-(p+q)t+q[S^0]\frac{1-e^{-pt}}{p}}, & \text{if } p > 0, \\ [S^0]e^{-q(1-[S^0])t}, & \text{if } p = 0. \end{cases} \tag{25}$$

This expression was already obtained in [19] for $p > 0$ and $[S^0] \equiv 1$.

Another situation where an explicit solution can be calculated is that of a semi-infinite line when the boundary node is infected at $t = 0$ and all other nodes are initially susceptible:

Proposition 2 (Patient-zero Bass problem). *Assume the conditions of Theorem 4.1, let $p_k(t) \equiv p$ and $q_k(t) \equiv q$ be independent of k and t, and let the nodes be placed on the semi-inifinite line $\mathcal{K} = \{0, 1, 2, \dots\}$, such that*

$$x_0^0 = i, \qquad x_k^0 = s, \quad k = 1, 2, \dots$$

Then the solution $[S_k](t)$ of (17) is given by

$$[S_0](t) \equiv 0, \qquad [S_k](t) = e^{-(p+q)t} \sum_{l=0}^{k-1} \frac{\left(q\frac{1-e^{-pt}}{p} \right)^l}{l!}, \qquad k = 1, 2, \dots \tag{26}$$

In addition, if we denote by $N_K(t) = \sum_{k=1}^{K} \mathbb{1}_{x_k(t)=i}$ the total number of infected nodes among nodes $\{1, \dots, K\}$, then

$$\mathbb{E}[N_K^n] = K - e^{-(p+q)t} \sum_{l=1}^{K} l \frac{\left(q\frac{1-e^{-pt}}{p} \right)^{K-l}}{(K-l)!}. \tag{27}$$

Proof. Expression (26) can be verified by direct substitution in (17). Expression (27), follows from $\mathbb{E}[N_K^n] = \sum_{k=1}^{K}[I_k](t)$, (14), (26) and some elementary computations detailed in Appendix A.1   □

Let us briefly justify the importance of the patient-zero Bass problem considered in Proposition 2. When the source term p is much smaller than the contagion term q, the dynamics of the population can be formally described in two time scales: propagation of clusters through contagion in a fast time-scale, and spontaneous creation of new seeds in a slow time-scale [19]. The patient-zero Bass problem thus models the fast time-scale evolution of clusters.

Passing to the limit in (26), we have that

$$\lim_{k \to +\infty} [S_k](t) = e^{-(p+q)t + q\frac{1-e^{-pt}}{p}}, \qquad k = 1, 2, \ldots \tag{28}$$

This limit coincides with the expression for $[S_k]$ in the homogeneous case with $[S_k^0] \equiv 1$, see (25), since the effect of patient zero vanishes as $k \to +\infty$. Note also that, as $K \to +\infty$, the expected fraction of infected nodes is given by, see (27),

$$\lim_{K \to +\infty} \frac{\mathbb{E}[N_K^n]}{K} = 1 - e^{-(p+q)t + q\frac{1-e^{-pt}}{p}} \tag{29}$$

which expresses again the fact that the solution of the patient zero problem converges to the homogeneous solution given by (25) as $K \to +\infty$.

In Proposition 2, the epidemic is triggered by two distinct mechanisms: firstly, by contagion from "patient-zero" located at $k = 0$, secondly, by spontaneous infection from the source term p. Expression (26) indicates how these two effects combine nonlinearly. To focus on the effect of the source term, we may remove "patient zero" and assume that all the individuals on the half-line are initially susceptible. In this case, it can be shown that the probability $[S_k](t)$ coincides with the probability of a node to be susceptible on a finite circle of length k. Therefore, it can be computed explicitly, see [21] for more details. If, on the other hand, we only focus on the effect of contagion from "patient-zero", we may discard the source term by letting $p \to 0$ in (26). This gives the solution of the *patient-zero SI problem*

$$[S_k](t) = e^{-qt} \sum_{l=0}^{k-1} \frac{(qt)^{l-1}}{(l-1)!}, \qquad k = 1, 2, \ldots \tag{30}$$

By (30), for any $k \geq 1$ and $t > 0$, we have on the one hand that $S_k(t) > 0$, implying that the epidemics propagates instantaneously from "patient zero" to any other node. On the other hand, $\lim_{k \to +\infty} S_k(t) = 1$ for any $t > 0$ implying that the propagation occurs at a finite *asymptotic speed*. In fact, we can show that the asymptotic speed of propagation is q:

Proposition 3 (Speed of propagation)**.** *Assume the conditions of Proposition 2 and assume furthermore that $p = 0$ (SI patient-zero problem). Denote by $\overline{k}(t) = \sup\{k \geq 0 : x_k(t) = i\}$ the location of the propagation front of the epidemic. Then, for all $\alpha \in \mathbb{R}$,*

$$\lim_{t \to +\infty} \mathbb{P}\left(\overline{k}(t) \leq qt + \alpha\sqrt{t}\right) = \frac{1}{\sqrt{2\pi q}} \int_{-\infty}^{\alpha} e^{-\frac{\tau^2}{2q}} d\tau.$$

Proof. The process $\overline{k}(t)$ is equivalent to a standard Poisson process with parameter q. It is then classical that the mean value of $\overline{k}(t)$ is t. Then, by the central limit

theorem, we have that $\frac{\overline{k}(t)-t}{\sqrt{t}}$ converges in probability when $t \to +\infty$ to a Gaussian with mean 0 and variance $q\sqrt{t}$, hence the result. $\qquad \square$

Thus the epidemics propagate through space at an average speed q, such that the exact location of the front has a variance $q\sqrt{t}$ around its mean value. Therefore, we can make a formal analogy between the propagation of the epidemics and the propagation of particles whose density $u(t, x)$ is given by the advection-diffusion equation [40]

$$\partial_t u + q \partial_x u - \frac{q}{2} \partial_{xx}^2 u = 0.$$

This equation describes the evolution of a density of particles subject to advection towards $x > 0$ at speed q (resulting in propagation at average speed q) and to diffusion (i.e., each particle follows a Brownian motion) with diffusivity $\frac{q}{2}$ (resulting in a variance q for the location of the propagation front).

Finally, we compute an explicit solution for a time-varying source term located at the boundary of a semi-infinite line:

Proposition 4 (Time-varying point source). *Assume that the individuals are placed on a semi infinite line $\mathcal{K} = \{0, 1, 2, \dots\}$ and that all individuals are initially suscep- tible, i.e., $x_k^0 = s$ for $k \in \mathcal{K}$. Assume that $q_k(t) \equiv q > 0$ is constant, that $p_k(t) \equiv 0$ for all $k \geq 1$, and allow a single time-varying source term $p_0(t) \geq 0$ at $k = 0$. Then*

$$[S_0](t) = e^{-\int_0^t p_0(\tau)d\tau},$$

$$[S_k](t) = 1 - \int_0^t \frac{(q\tau)^{k-1}}{(k-1)!} q e^{-q\tau} \left(1 - e^{-\int_0^{t-\tau} p_0(\tau')d\tau'}\right) d\tau, \qquad k = 1, 2 \dots \tag{31}$$

In particular,

$$\lim_{t \to +\infty} [S_k](t) = e^{-\int_0^{+\infty} p_0(\tau)d\tau}, \qquad k = 0, 1, \dots \tag{32}$$

Proof. Under the above assumptions, equation (16) reduces to

$$[S_0]'(t) + p_0(t)[S_0](t) = 0, \qquad S_0(0) = 1, \tag{33}$$

and

$$[S_k]'(t) + q([S_k](t) - [S_{k-1}](t)) = 0, \qquad S_k(0) = 1, \qquad k = 1, 2 \dots \tag{34}$$

The expression of $[S_0](t)$ in (31) is obtain by a straightforward integration of (33). Using (14), equation (34) yields

$$[I_k]'(t) + q([I_k](t) - [I_{k-1}](t)) = 0, \qquad I_k(0) = 0, \qquad k = 1, 2 \dots$$

Hence,

$$[I_k](t) = q e^{-q \cdot} \star_t [I_{k-1}](\cdot), \qquad e^{-q \cdot} \star_t [I_{k-1}](\cdot) := \int_0^t e^{-q(t-\tau)}[I_{k-1}](\tau)d\tau. \tag{35}$$

By induction, we infer that

$$[I_k](t) = q^k \underbrace{e^{-q \cdot} \star_t \cdots \star_t e^{-q \cdot}}_{k \text{ terms}} \star_t [I_0](\cdot), \tag{36}$$

where we have used the associativity of the convolution product. Now,

$$[I_0](t) = 1 - [S_0](t) = 1 - e^{\int_0^t p_0(\tau)d\tau}. \tag{37}$$

In addition, a straightforward induction shows that

$$q^k \underbrace{e^{-q\cdot} \star_t \cdots \star_t e^{-q\cdot}}_{k \text{ terms}} = \frac{(qt)^{k-1}}{(k-1)!} q e^{-qt}. \tag{38}$$

Plugging (37) and (38) into (36) gives (31).

Let us now prove (32). For a fixed $k \in \{0, 1, \dots\}$, by the dominated convergence theorem, we have

$$\lim_{t \to +\infty} \int_0^t \frac{(q\tau)^{k-1}}{(k-1)!} q e^{-q\tau} \left(1 - e^{-\int_0^{t-\tau} p_0} \right) d\tau$$

$$= \left(1 - e^{-\int_0^{+\infty} p_0} \right) \int_0^{+\infty} \frac{(q\tau)^{k-1}}{(k-1)!} q e^{-q\tau} d\tau.$$

Since

$$\int_0^{+\infty} \frac{(q\tau)^{k-1}}{(k-1)!} q e^{-q\tau} d\tau = 1,$$

this concludes the proof of (32). □

The limit (32) does not depend on k and, in particular, coincides with the limit as $t \to +\infty$ of $[S_0](t) = e^{-\int_0^t p_0}$. This limit can be obtained formally by passing to the limit in (16). Note also that, for a fixed $t \geq 0$, as $k \to +\infty$ in (31), we get that $[S_k](t) \to 1$. This expresses the fact that, with probability 1, the source term located at $k = 0$ does not influence far-away nodes in finite time, since the epidemics propagates at a finite speed (see Proposition 3).

Note that by formally putting $p_0(t) := K\delta_{t=0}$ in (31), where $K > 0$ and $\delta_{t=0}$ denotes the Dirac mass centered at 0, then we recover the formula (26) with $p = 0$ for the patient-zero problem as $K \to +\infty$.

5. **SIR-Bass model on 1D one-sided lattices.** We now consider the more general 1D one-sided case, when recovery is not neglected, see (12).

5.1. **Deterministic desciption.** We begin by deriving from the discrete stochastic model (12) a system of differential equations for $[S_k](t)$:

Theorem 5.1. *Assume that the nodes are placed on a 1D one-sided lattice, see (10), that the stochastic dynamics are governed by (12), and that the initial conditions are stochastic and uncorrelated, see (3). For any $k \in \mathcal{K}$, if $q_k(t) > 0$ for all $t \geq 0$, then $[S_k](t)$ satisfies*

$$[S_k]' + \left(p_k + q_k + r_k \right)[S_k]$$

$$= q_k e^{-\int_0^t p_k} [S_k^0] \left([S_{k-1}] + [R_{k-1}^0] + \frac{r_k(0)}{q_k(0)} + \int_0^t \left(\frac{r_k}{q_k} \right)' (s) \frac{[S_k](s)}{[S_k^0]} ds \right),$$

$$[S_k](0) = [S_k^0], \tag{39a}$$

and $[R_k]$ satisfies

$$[R_k]'(t) - r_k \left(1 - [S_k] - [R_k] \right) = 0, \qquad [R_k](0) = [R_k^0]. \tag{39b}$$

Proof. In the following calculations, we frequently make use of the identities

$$\mathbb{P}(A \cap B) = \mathbb{P}(A|B)\mathbb{P}(B), \qquad \mathbb{P}(A_1 \cap A_2|B) = \mathbb{P}(A_1|A_2 \cap B)\mathbb{P}(A_2|B),$$

$$\mathbb{P}(A_1 \cap A_2|B)\mathbb{P}(B) = \mathbb{P}(A_1|A_2 \cap B)\mathbb{P}(A_2 \cap B). \tag{40}$$

We have that

$$
\begin{aligned}
\mathbb{P}\left(S_k^n\right) - \mathbb{P}\left(S_k^{n+1}\right) &= \mathbb{P}\left(I_k^{n+1} \cap S_k^n\right) \\
&= \mathbb{P}\left(I_k^{n+1} \cap S_k^n \cap S_{k-1}^n\right) + \mathbb{P}\left(I_k^{n+1} \cap S_k^n \cap I_{k-1}^n\right) + \mathbb{P}\left(I_k^{n+1} \cap S_k^n \cap R_{k-1}^n\right) \\
&= \underbrace{\mathbb{P}\left(I_k^{n+1}\big|S_k^n \cap S_{k-1}^n\right)}_{=p_k^n \Delta t} \mathbb{P}\left(S_k^n \cap S_{k-1}^n\right) + \underbrace{\mathbb{P}\left(I_k^{n+1}\big|S_k^n \cap I_{k-1}^n\right)}_{=(p_k^n+q_k^n) \Delta t} \mathbb{P}\left(S_k^n \cap I_{k-1}^n\right) \\
&\quad + \underbrace{\mathbb{P}\left(I_k^{n+1}\big|S_k^n \cap R_{k-1}^n\right)}_{=p_k^n \Delta t} \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right).
\end{aligned}
$$

and so

$$
\frac{\mathbb{P}\left(S_k^{n+1}\right) - \mathbb{P}\left(S_k^n\right)}{\Delta t} + p_k^n \mathbb{P}(S_k^n) + q_k^n \mathbb{P}\left(S_k^n \cap I_{k-1}^n\right) = 0. \tag{41}
$$

Using the closure relation (22), we deduce that

$$
\begin{aligned}
\mathbb{P}\left(S_k^n \cap I_{k-1}^n\right) &= \mathbb{P}(S_k^n) - \mathbb{P}\left(S_k^n \cap S_{k-1}^n\right) - \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) \\
&= \mathbb{P}(S_k^n) - e^{-\int_0^{t^n} p_k(s)ds} \mathbb{P}(S_k^0)\mathbb{P}(S_{k-1}^n)(1 + O(\Delta t)) - \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right).
\end{aligned} \tag{42}
$$

Substituting this expression into (41), we deduce

$$
\begin{aligned}
\frac{\mathbb{P}\left(S_k^{n+1}\right) - \mathbb{P}\left(S_k^n\right)}{\Delta t} &+ (p_k^n + q_k^n)\mathbb{P}(S_k^n) \\
&- q_k^n e^{-\int_0^{t^n} p_k(s)ds} \mathbb{P}(S_k^0)\mathbb{P}(S_{k-1}^n) - q_k^n \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) = O(\Delta t).
\end{aligned} \tag{43}
$$

Letting $\Delta t \to 0$, we finally deduce

$$
\frac{d[S_k]}{dt}(t) + (p_k(t) + q_k(t))[S_k] - q_k e^{-\int_0^t p_k}[S_k^0][S_{k-1}] - q_k[S_k \cap R_{k-1}](t) = 0. \tag{44}
$$

Next, let us derive an equation for $[S_k \cap R_{k-1}]$:

$$
\begin{aligned}
\mathbb{P}\left(S_k^{n+1} \cap R_{k-1}^{n+1}\right) &- \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) \\
&= \mathbb{P}\left(S_k^{n+1} \cap R_{k-1}^{n+1}\right) - \mathbb{P}\left(S_k^{n+1} \cap R_{k-1}^n\right) + \mathbb{P}\left(S_k^{n+1} \cap R_{k-1}^n\right) - \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) \\
&= \mathbb{P}\left(S_k^{n+1} \cap I_{k-1}^n \cap R_{k-1}^{n+1}\right) - \mathbb{P}\left(I_k^{n+1} \cap S_k^n \cap R_{k-1}^n\right) \\
&= \underbrace{\mathbb{P}\left(R_{k-1}^{n+1}\big|S_k^{n+1} \cap I_{k-1}^n\right)}_{=r_k^n \Delta t} \mathbb{P}\left(S_k^{n+1} \cap I_{k-1}^n\right) - \underbrace{\mathbb{P}\left(I_k^{n+1}\big|S_k^n \cap R_{k-1}^n\right)}_{=p_k^n \Delta t} \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) \\
&= r_k^n \Delta t \mathbb{P}\left(S_k^{n+1} \cap I_{k-1}^n\right) - p_k^n \Delta t \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) \\
&= r_k^n \Delta t \mathbb{P}\left(S_k^n \cap I_{k-1}^n\right) - p_k^n \Delta t \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) \\
&\quad + r_k^n \Delta t \left[\mathbb{P}\left(S_k^{n+1} \cap I_{k-1}^n\right) - \mathbb{P}\left(S_k^n \cap I_{k-1}^n\right)\right] \\
&= r_k^n \Delta t \left[\mathbb{P}\left(S_k^n\right) - \mathbb{P}\left(S_k^n \cap S_{k-1}^n\right) - \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right)\right] - p_k^n \Delta t \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) \\
&\quad - r_k^n \Delta t \mathbb{P}\left(I_k^{n+1} \cap S_k^n \cap I_{k-1}^n\right)
\end{aligned}
$$

Using (22) and the relation

$$
\begin{aligned}
\mathbb{P}\left(\overline{S_k^{n+1}} \cap S_k^n \cap I_{k-1}^n\right) &= \mathbb{P}\left(I_k^{n+1} \cap S_k^n \cap I_{k-1}^n\right) \\
&= \mathbb{P}\left(I_k^{n+1}\big|S_k^n \cap I_{k-1}^n\right) \mathbb{P}\left(S_k^n \cap I_{k-1}^n\right) \\
&= (p_k^n + q_k^n)\Delta t \mathbb{P}\left(S_k^n \cap I_{k-1}^n\right),
\end{aligned}
$$

we deduce that

$$\frac{\mathbb{P}\left(S_k^{n+1} \cap R_{k-1}^{n+1}\right) - \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right)}{\Delta t} - r_k^n \mathbb{P}\left(S_k^n\right) + r_k^n \left(1 - p_k^n \Delta t\right)^n \mathbb{P}(S_k^0) \mathbb{P}\left(S_{k-1}^n\right)$$
$$+ (p_k^n + r_k^n) \mathbb{P}\left(S_k^n \cap R_{k-1}^n\right) = -r_k^n (p_k^n + q_k^n) \Delta t \mathbb{P}\left(S_k^n \cap I_{k-1}^n\right).$$
$$(45)$$

Letting $\Delta t \to 0$, we deduce

$$\frac{d[S_k \cap R_{k-1}]}{dt}(t) - r_k(t)[S_k](t) + r_k e^{-\int_0^t p(s)ds}[S_k^0][S_{k-1}] + (p_k + r_k)[S_k \cap R_{k-1}] = 0.$$
$$(46)$$

At this stage, equations (44) and (46) form an autonomous system of ODEs governing the evolution of the family of probabilities $\{[S_k], [S_k \cap R_{k-1}]\}_{k \in \mathcal{K}}$. This system can actually be reduced to an autonomous system involving only the probabilities $\{[S_k]\}_{k \in \mathcal{K}}$. Let us set

$$w_k(t) := \left(\frac{r_k(t)}{q_k(t)}[S_k](t) + [S_k \cap R_{k-1}](t)\right) e^{\int_0^t p_k}.$$

Using equations (44) and (46) satisfied by $[S_k]$ and $[S_k \cap R_{k-1}]$, we find that w_k satisfies satisfies

$$w_k'(t) = \left(\frac{r_k}{q_k}\right)'(t)[S_k](t).$$

Therefore,

$$w_k(t) = w_k^0 + \int_0^t \left(\frac{r_k}{q_k}\right)'(s)[S_k](s)ds,$$

and so

$$[S_k \cap R_{k-1}] = -\frac{r_k}{q_k}[S_k] + e^{-\int_0^t p_k}\left(\frac{r_k(0)}{q_k(0)}[S_k^0] + [S_k^0][R_{k-1}^0] + \int_0^t \left(\frac{r_k}{q_k}\right)'(s)[S_k](s)ds\right).$$
$$(47)$$

Substituting the above expression in (43), we find (39a).

To derive equation (39b) for $[R_k](t)$, we note that

$$\mathbb{P}\left(R_k^{n+1}\right) - \mathbb{P}\left(R_k^n\right) = \mathbb{P}\left(R_k^{n+1} \cap I_k^n\right) = \mathbb{P}\left(R_k^{n+1}|I_k^n\right) \mathbb{P}\left(I_k^n\right) = r_k^n \Delta t \mathbb{P}\left(I_k^n\right).$$

Therefore,

$$\frac{\mathbb{P}\left(R_k^{n+1}\right) - \mathbb{P}\left(R_k^n\right)}{\Delta t} - r_k^n \mathbb{P}\left(I_k^n\right) = 0.$$

Using (1) and letting $\Delta t \to 0$ in the above equality gives (39b). $\square$

Remark 3. Equations (39a) form a system of integrodifferential equation that only involves the probabilities $\{[S_k](t)\}_{t \geq 0, k \in \mathcal{K}}$. If the ratio $\frac{r_k}{q_k}$ does not depend on t, it boils down to the system of ODEs

$$[S_k]' + \left(p_k + q_k + r_k\right)[S_k] = e^{-\int_0^t p_k}[S_k^0]\left(q_k[S_{k-1}] + q_k[R_{k-1}^0] + r_k\right), \quad [S_k](0) = [S_k^0].$$
$$(48)$$

An apparently different system of ODEs for $[S_k]$ and $[R_k]$ was derived in [17] for the case of the deterministic initial conditions $[S_k^0] \equiv 1$. In that system, however, the equation for $[S_k]$ was not decoupled from that for $[R_k]$.

Theorem 5.1 gives a complete deterministic description of the stochastic process, since once $[S_k]$ and $[R_k]$ are computed through equations (39a)-(39b), $[I_k]$ is deduced from (1).

5.2. **Explicit solutions.** We can solve equations (39a)-(39b) explicitely in the homogeneous case.

Proposition 5 (Spatially homogeneous solution). *Assume the conditions of Theorem 5.1, and assume further that the initial conditions $[S_k^0] \equiv [S^0]$ do not depend on the space variable k and that the parameters $p_k(t) \equiv p$, $q_k(t) = q > 0$ and $r_k(t)$ do not depend on the time variable t nor the space variable k. Then the solutions $[S_k] = [S]$ and $[R_k] = [R]$ of (39a)-(39b) do not depend on k, and are given by*

$$[S](t) = [S^0]e^{-(p+q+r)t+[S^0]q\frac{1-e^{-pt}}{p}} \left(1 + (r + q[R^0]) \int_0^t e^{(q+r)\tau - q[S^0]\frac{1-e^{-p\tau}}{p}} d\tau\right),$$
(49)

$$[R](t) = 1 - (1 - [R^0])e^{-rt} - re^{-rt} \int_0^t e^{r\tau}[S](\tau)d\tau.$$
(50)

Proof. By translation symmetry, $[S_k]$ and $[R_k]$ do not depend on k. The result follows from a straighforward integration of equations (39a)-(39b). □

Remark 4. Expression (49) was already obtained in [18, Lemma 4.2] for $[S^0] \equiv 1$ and $[R^0] \equiv 0$.

We can also explicitly solve the case of a time-dependent source term located at $k = 0$.

Proposition 6 (SIR-Bass time-varying point source). *Assume the conditions of Theorem 5.1, that the individuals are placed on a semi infinite line $\mathcal{K} = \{0, 1, 2, \dots\}$, that $q_k(t) \equiv q > 0$ is constant, that $p_k(t) \equiv 0$ for all $k > 0$, and allow a source term $p_0(t) \geq 0$ at $k = 0$. Then*

$$[S_0](t) = e^{-\int_0^t p_0},$$
$$[S_k](t) = 1 - \int_0^t \frac{(q\tau)^{k-1}}{(k-1)!}qe^{-(q+r)\tau}\left(1 - e^{-\int_0^{t-\tau} p_0}\right)d\tau, \qquad k = 1, 2 \dots$$
(51)

In particular,

$$\lim_{t \to +\infty}[S_k](t) = 1 - \left(\frac{q}{q+r}\right)^k \left(1 - e^{-\int_0^{+\infty} p_0}\right), \qquad k = 0, 1, 2, \dots$$
(52)

Proof. Under our assumptions, equation (39a) reduces to

$$[S_0]'(t) + p_0(t)[S_0](t) = 0, \qquad [S_0](0) = 1,$$

and

$$[S_k]'(t) + q([S_k](t) - [S_{k-1}](t)) + r([S_k](t) - 1) = 0, \qquad [S_k](0) = 0, \qquad k = 1, 2 \dots$$

Therefore, $[S_0](t) = e^{-\int_0^t p_0}$ and it can be verified by direct substitution that the solution for $k \geq 1$ is given by

$$[S_k](t) = e^{-(q+r)t} + e^{-(q+r)\cdot} \star_t (r + q[S_{k-1}](\cdot)),$$
(53)

where

$$f(\cdot) \star_t g(\cdot) := \int_0^t f(t - \tau)g(\tau)d\tau.$$

Note also that the constant 1 is a fixed point of the recursion relationship (53), namely,

$$1 = e^{-(q+r)t} + e^{-(q+r)\cdot} \star_t (r + q),$$

Therefore, for all $k > 0$,

$$[S_k](t) - 1 = qe^{-(q+r)\cdot} \star_t ([S_{k-1}](\cdot) - 1).$$

By induction, we infer

$$[S_k](t) = 1 + q^k \underbrace{e^{-(q+r)\cdot} \star_t \cdots \star_t e^{-(q+r)\cdot}}_{k \text{ terms}} \star_t ([S_0](\cdot) - 1),$$

where we have used the associativity of the convolution product. A straightforward induction shows that

$$q^k \underbrace{e^{-(q+r)\cdot} \star_t \cdots \star_t e^{-(q+r)\cdot}}_{k \text{ terms}} = \frac{(qt)^{k-1}}{(k-1)!} qe^{-(q+r)t}.$$

Therefore

$$[S_k](t) = 1 + \left(\tau \mapsto \frac{(q\tau)^{k-1}}{(k-1)!} qe^{-(q+r)\tau} \right) \star_t ([S_0](\cdot) - 1).$$

Using $[S_0](t) = e^{-\int_0^t p_0}$, we deduce that

$$[S_k](t) = 1 + \left(\tau \mapsto \frac{(q\tau)^{k-1}}{(k-1)!} qe^{-(q+r)\tau} \right) \star_t \left(\tau \mapsto e^{-\int_0^\tau p_0} - 1 \right).$$

which proves (51).

Let us now prove (52). For a fixed $k \in \{0, 1, \dots\}$, by the dominated convergence theorem,

$$\lim_{t \to +\infty} \int_0^t \frac{(q\tau)^{k-1}}{(k-1)!} qe^{-(q+r)\tau} \left(1 - e^{-\int_0^{t-\tau} p_0} \right) d\tau$$

$$= \left(1 - e^{-\int_0^{+\infty} p_0} \right) \int_0^{+\infty} \frac{(q\tau)^{k-1}}{(k-1)!} qe^{-(q+r)\tau} d\tau.$$

Using the relation

$$\int_0^{+\infty} \frac{(q\tau)^{k-1}}{(k-1)!} qe^{-(q+r)\tau} d\tau = \left(\frac{q}{q+r} \right)^k$$

completes the proof of (52).   □

Thus, the limit (52) depends on space k if and only if $r > 0$. Note also that, for a fixed $t \geq 0$, as $k \to +\infty$ in (51), we get that $[S_k](t) \to 1$. This expresses again the fact that a source term located at $k = 0$ cannot influence far nodes in finite time.

From (52), we observe that

$$\frac{1 - \lim_{t \to +\infty} [S_{k+1}](t)}{1 - \lim_{t \to +\infty} [S_k](t)} = \frac{q}{q+r}. \tag{54}$$

This expression is actually equal to the probability that node $k + 1$ will eventually become infected knowing that node k has been infected.

5.3. SIR model on 1D one-sided lattices. When $p_k^n \equiv 0$, i.e., when there are no source terms, the model (12) reduces to the SIR model on a 1D one-sided lattice. Assuming that the parameters and initial conditions are homogeneous (they do not depend on time nor space), if we let $p \to 0$ in (49)-(50), then $\frac{1-e^{-pt}}{p} \to t$, and so we obtain

$$[S](t) = \frac{[S^0]}{r + q(1 - [S^0])}\left(r + q[R^0] + q[I^0]e^{-\left(r+q(1-[S^0])\right)t}\right),$$

and

$$[R](t) = 1 - [S^0]\frac{r + q[R^0]}{r + q(1 - [S^0])} - \frac{[I^0]}{1 - [S^0]}e^{-rt}$$
$$+ \frac{r[S^0][I^0]}{r + q(1 - [S^0])(1 - [S^0])}e^{-(r+q(1-[S^0]))t}.$$

Therefore,

$$[S^\infty] := \lim_{t \to +\infty}[S](t) = [S^0]\frac{r + q[R^0]}{r + q(1 - [S^0])}, \tag{55}$$

and

$$[R^\infty] := \lim_{t \to +\infty}[R](t) = 1 - [S^0]\frac{r + q[R^0]}{r + q(1 - [S^0])} = 1 - [S^\infty],$$

and so, by (1), we deduce that $[I^\infty] := \lim_{t \to +\infty}[I](t) = 0$.

The SIR model on a 1D one-sided lattice (12) satisfies a threshold phenomenon for the initial outbreak of an epidemic, namely, an initially small number of infected ($S^0 \approx 1$, $0 < I^0 \ll 1$, $R^0 = 0$) triggers an epidemic if and only if

$$q - r > 0. \tag{56}$$

Indeed, under the same conditions as in Proposition 5, if we further assume that $p = 0$ and $[R^0] = 0$, using (1), (39a) and (39b), we find that $[I]$ satisfies

$$[I]' - \left(r + q(1 - [S^0])\right)(1 - [I]) + r[S^0] + r[I] = 0,$$

i.e.,

$$[I]' - (r + q)[I^0] + \left(2r + q[I^0]\right)[I] = 0.$$

If we linearize this equation around $0 < [I] \approx [I^0] \ll 1$, we find that $[I]$ increases in small time if and only if (56) holds.

Let us now solve explicitly the "patient-zero" problem for the SIR model:

Proposition 7 (Patient-zero SIR problem). *Assume that $p_k^n \equiv 0$, $q_k^n \equiv q$ and $r_k^n \equiv r$ in (12), that the nodes are placed on a semi-inifinite line $\mathcal{K} = \{0, 1, 2, \dots\}$ and that, initially,*

$$x_{k=0}^0 = i, \qquad x_k^0 = s, \quad k = 1, 2, \dots \tag{57}$$

Then, the solution of (39a) is given by

$$[S_0](t) \equiv 0,$$

$$[S_k](t) = 1 - \left(\frac{q}{q+r}\right)^k\left(1 - e^{-(q+r)t}\sum_{l=0}^{k-1}\frac{((q+r)t)^l}{l!}\right), \qquad k = 1, 2, \dots \tag{58}$$

If we denote by $N_K(t) = \sum_{k=1}^{K} \mathbb{1}_{x_k(t)=s}$ the total number of susceptibles on a finite line of length $K + 1 > 0$, then

$$\mathbb{E}[N_K^n] = K - \frac{q}{r}\left(1 - \left(\frac{q}{q+r}\right)^K\right) - \frac{q+r}{r}e^{-(q+r)t}\sum_{l=1}^{K}\frac{(qt)^{K-l}}{(K-l)!}\left(1 - \left(\frac{q}{q+r}\right)^{l+1}\right).$$

$$(59)$$

Proof. See Appendix A.2. □

We see from (58) that $\lim_{k\to+\infty}[S_k(t)] = 1$, i.e., that the effect of patient-zero vanishes as we move far away from $k = 0$.

It is instructive to compare the SIR model on a 1D-one sided lattice to the classical aggregate SIR model (6). Since (6) corresponds to a complete graph whereas (12) corresponds to a structured sparse graph, this comparison reveals the effect of the network structure on the dynamics.

- First, we emphasize that equation (39a) for $[S_k]$ does not include $[I_k]$ and $[R_k]$, or higher-order marginals such as $[S_k \cap S_{k-1}]$. This property is unique to 1D one-sided graphs. Indeed, for the aggregate SIR model (6), it is not possible to derive an equation on S which does not involve I and R.

  Let us mention that the system (4), and therefore also (6) obtained from (4) by taking $p = 0$, can also be reduced to a single equation on $u(t) := \int_0^t I$. Indeed, integrating the equation for S in (4), we obtain $S(t) = S^0 e^{-pt-q\int_0^t I}$. Injecting this expression into the equation for I in (4), integrating the equation on $(0, t)$, we obtain

  $$u'(t) = I^0 + S^0\left(1 - e^{-pt-qu(t)}\right) - ru(t).$$

- The aggregate SIR model satisfies the same threshold condition (56) for the outbreak of an epidemic from a small initial number of infected, see [25].
- When (56) holds, the behavior after the outbreak of an epidemic can be quite different for the 1D one-sided case (12) and the aggregate model (6). In particular, a remarkable difference is the dependence of the final state on the initial conditions. From (55), we observe that
  1. $[S^\infty]$ is an increasing function of $[S^0]$: the more susceptible individuals at initial time, the more for large times,
  2. If $0 < [I^0], [R^0] \ll 1$ then $[S^\infty] \approx [S^0] \approx 1$.

  In contrast, for the aggregate SIR model (6), we have that
  1. $[S^\infty]$ is a *nonincreasing* function $[S^0]$, see [25],
  2. $[I^0], [R^0] \ll 1$ does not imply that $[S^\infty] \approx [S^0] \approx 1$.

6. **SIR-Bass model on 1D two-sided lattices.** We now allow both the left and the right neighbors to spread the infection. We denote by $q_k^{L,n} := q_{k-1,k}$ and $q_k^{R,n} := q_{k+1,k}$ the influence of the left and right neighbors on node k, respectively, and we assume that

$$q_{ik} = 0 \qquad \text{if } |i-k| \neq 1. \tag{60}$$

Hence, (2a) becomes

$$\mathbb{P}\left(I_k^{n+1}\big|X^n\right) = \begin{cases} \Delta t\left(p_k^n + q_k^{L,n}\mathbb{1}_{I_{k-1}^n} + q_k^{R,n}\mathbb{1}_{I_{k+1}^n}\right), & \text{if } x_k^n = s, \\ 1 - r_k^n\Delta t, & \text{if } x_k^n = i, \\ 0, & \text{if } x_k^n = r, \end{cases} \tag{61a}$$

The transition rates (2b) from i to r remain unchanged, that is:

$$\mathbb{P}\left(R_k^{n+1}\big|X^n\right) = \begin{cases} 0, & \text{if } x_k^n = s, \\ r_k^n \Delta t, & \text{if } x_k^n = i, \\ 1, & \text{if } x_k^n = r, \end{cases} \tag{61b}$$

6.1. Deterministic description.

Theorem 6.1. *Assume that the individuals are placed on a 1D two-sided lattice $\mathcal{K}$ (which is either a finite, semi-infinite, or infinite line), see (60), that the dynamics are governed by (61), and that the initial conditions are uncorrelated, see (3). Let p_k^n, q_k^n and r_k^n converge as $\Delta t \to 0$ to $p_k(t)$, $q_k(t)$ and $r_k(t)$. Then for all $k \in \mathcal{K}$,*

$$[S_k](t) = \begin{cases} \dfrac{[S_k^L](t)[S_k^R](t)}{[S_k^0]e^{-\int_0^t p_k(\cdot)}}, & \text{if } [S_k^0] > 0, \\ 0, & \text{if } [S_k^0] = 0, \end{cases} \tag{62a}$$

where $[S_k^R](t)$ (resp. $[S_k^L](t)$) is the probability that node k is susceptible by time t if we discard the influences of the left neighbors by setting $q_k^{L,n} \equiv 0$ in (61a) (resp. the influences of the right neighbors by setting $q_k^{R,n} \equiv 0$). In addition, $[R_k(t)]$ is given by

$$[R_k]'(t) - r_k(t)\big(1 - [S_k] - [R_k]\big) = 0. \tag{62b}$$

Proof. Let us first introduce some notations. We denote by $[S_k|q_k^L = 0]$ the probability of the event $\tilde{S}_k(t) := \{\tilde{x}_k(t) = s\}$ where $\tilde{x}_k$ is the process defined identically as x_k but in which we put $q_k^L \equiv 0$ in (61a). We use the name notation if S_k is replaced by any event or intersection of events, and if $q_k^{L,n} = 0$ is replaced by other conditions on the parameters. For example, with these notations, we have $\left[S_k^R\right] = \left[S_k|(q_k^L)_{k\in\mathcal{K}} \equiv 0\right]$.

First, notice that

$$\begin{aligned} [S_{k-1} \cap S_k] &= \left[S_{k-1} \cap S_k|q_{k-1}^R = q_k^L = 0\right] \\ &= \left[S_{k-1}|q_{k-1}^R = q_k^L = 0\right]\left[S_k|q_{k-1}^R = q_k^L = 0\right] \\ &= \left[S_{k-1}|q_{k-1}^R = 0\right]\left[S_k|q_k^L = 0\right] \\ &= \left[S_{k-1}|(q_k^R)_{k\in\mathcal{K}} \equiv 0\right]\left[S_k|(q_k^L)_{k\in\mathcal{K}} \equiv 0\right]. \end{aligned}$$

We deduce that

$$[S_{k-1} \cap S_k] = \left[S_{k-1}^L\right]\left[S_k^R\right]. \tag{63}$$

Similarly, we have

$$\begin{aligned} [S_{k-1} \cap S_k \cap S_{k+1}] &= \left[S_{k-1} \cap S_k \cap S_{k+1}|q_{k-1}^R = q_k^L = q_k^R = q_{k+1}^L = 0\right] \\ &= \left[S_{k-1}|q_{k-1}^R = 0\right]\left[S_k|q_k^L = q_k^R = 0\right]\left[S_{k+1}|q_{k+1}^L = 0\right] \\ &= \left[S_{k-1}|(q_k^R)_{k\in\mathcal{K}} \equiv 0\right][S_k^0]e^{-\int_0^t p}\left[S_{k+1}|(q_k^L)_{k\in\mathcal{K}} \equiv 0\right], \end{aligned}$$

and so

$$[S_{k-1} \cap S_k \cap S_{k+1}] = \left[S_{k-1}^L\right]\left[S_{k+1}^R\right][S_k^0]e^{-\int_0^t p}. \tag{64}$$

Relations (63) and (64) can also be proved using the *indifference principle* of [21].

The last ingredient in our proof is a *spatial Markovian property* establishing that the state of nodes $k-1$ and $k+1$ are independent if conditioned with respect to S_k. This is where the assumption on the graph (namely that it is 1D with no cylces) comes into play.

Lemma 6.2.

$$x^n_{k-1} \perp\!\!\!\perp x^n_{k+1} | S^n_k. \tag{65}$$

This standard result is proved separately in a slightly more general form, see Theorem A.1 in appendix A.3. Theorem 6.2 can also be deduced from the funnel node theorem [20].

Using Theorem 6.2, we have that

$$[S_{k-1} \cap S_k \cap S_{k+1}] = [S_{k-1} \cap S_{k+1}|S_k][S_k] = [S_{k-1}|S_k][S_{k+1}|S_k][S_k]$$
$$= \frac{[S_{k-1} \cap S_k][S_k \cap S_{k+1}]}{[S_k]}. \tag{66}$$

Substituting (63) and (64) in the right and left hand-side of the above equation yields equation (62a) on $[S_k]$.

Equation (62b) on $[R_k]$ is obtained similarly to equation (39b) in Theorem 5.1. □

Remark 5. In the case without recovery, $([S_k])_{k\in\mathcal{K}}$ and $([S_k \cap S_{k-1}])_{k\in\mathcal{K}}$ satisfy a closed system of ODEs. More precisely, under the same conditions as in Theorem 6.1 and if we further assume that there is no recovery (13), then

$$[S_k]'(t) + (p_k + q^R_k + q^L_k)[S_k] - e^{-\int_0^t p_k}[S^0_k](q^R_k[S_k \cap S_{k+1}] + q^L_k[S_k \cap S_{k-1}]) = 0, \tag{67a}$$

and

$$[S_k \cap S_{k-1}]'(t) = -(p_k + p_{k-1} + q^L_{k-1} + q^R_k)[S_k \cap S_{k-1}]$$
$$+ q^R_k \frac{[S_{k-1} \cap S_k][S_k \cap S_{k+1}]}{[S_k]} + q^L_{k-1} \frac{[S_{k-2} \cap S_{k-1}][S_{k-1} \cap S_k]}{[S_{k-1}]}. \tag{67b}$$

To prove this, first recall that from Theorem 5.1 we have

$$[S^R_k]' + (p_k + q^R_k + r_k)[S^R_k] - e^{-\int_0^t p_k}[S^0_k](q^R_k[S^R_{k+1}] + q^R_k[R^0_{k+1}] + r_k) = 0, \tag{68}$$

$$[S^L_k]' + (p_k + q^L_k + r_k)[S^L_k] - e^{-\int_0^t p_k}[S^0_k](q^L_k[S^L_{k-1}] + q^L_k[R^0_{k-1}] + r_k) = 0 \tag{69}$$

Then, we derive (67) from the identities (62a) and (63). Note also that

$$\frac{[S_{k-1} \cap S_k][S_k \cap S_{k+1}]}{[S_k]} = e^{-\int_0^t p_k}[S^0_k][S^L_{k-1}][S^R_{k+1}]. \tag{70}$$

6.2. **Explicit solutions.** Using Theorem 6.1, the results of the previous sections dealing with the spatially homogeneous case, the patient zero problem, and time-varying point sources can be adapted to the case of a two-sided 1D lattice. We now briefly present the obtained explicit formula. For simplicity, we focus on the Bass model, i.e., we assume that there is no recovery.

Corollary 1 (Spatially-homogeneous two-sided Bass solution)**.** *Assume the conditions of Theorem 6.1. If the initial condition $[S^0_k] \equiv [S^0]$ and the parameters $p_k(t) \equiv p(t), q^R_k(t) \equiv q^R(t), q^L_k(t) \equiv q^L(t)$ do not depend on space k, then $[S_k](t) = [S](t)$ does not depend on k, and is given by*

$$[S](t) = [S^0]e^{-\int_0^t (p(s)+q^R(s)+q^L(s))ds + [S^0]\int_0^t (q^R(s)+q^L(s))e^{-\int_0^s p}ds}. \tag{71}$$

Proof. The result is proved by a direct application of Theorem 6.1 and (24). □

Note that expression (71) coincides with (24) in the one-sided case by setting $q(t) := q^R(t) + q^L(t)$.

We now turn to the patient-zero problem.

Corollary 2 (Two-sided patient-zero Bass problem). *Assume the conditions of Theorem 6.1, let $r_k(t) \equiv 0$, $p_k(t) \equiv p$, $q_k^R(t) \equiv q^R$, $q_k^L(t) \equiv q^L$, be independent of k and t, and let the nodes be placed on a semi-inifinite line $\mathcal{K} = \mathbb{N}$. Assume that, initially, we place a patient zero at $k = 0$, i.e.,*

$$[I_0^0] = 1, \qquad [S_k^0] = 1, \quad k = 1, 2, \ldots$$

Then, $[S_k]$ is given by

$$[S_k](t) = e^{-(p+q^R+q^L)t+q^R\frac{1-e^{pt}}{p}} \sum_{l=0}^{k-1} \frac{\left(q^L \frac{1-e^{-pt}}{p}\right)^l}{l!}, \qquad \forall k = 1, 2, \ldots$$

Proof. Note that $[S_k^R]$ is given by the solution (24) of the one-sided Bass homogeneous problem with $q = q^R$, and that $[S_k^L]$ is given by the solution (26) of the one-sided Bass patient-zero problem with $q = q^L$. We conclude by using (62a). $\square$

We now present the solution in the case of a time-varying point source.

Corollary 3 (Time-varying point source). *Assume the conditions of Theorem 6.1, that the individuals are placed on a semi infinite line $\mathcal{K} = \{0, 1, 2, \ldots\}$ and that all individuals are initially susceptible, i.e., $x_k^0 = s$ for $k \in \mathcal{K}$. Assume that $q_k(t) \equiv q > 0$ is constant, that $p_k(t) \equiv 0$ for all $k \geq 1$, and allow a single time-varying source term $p_0(t) \geq 0$ at $k = 0$. Then*

$$[S_k](t) \equiv [S_k^L](t), \tag{72}$$

where $[S_k^L]$ is given by (31).

Proof. The proof follows from (62a), the fact that $[S_k^R] \equiv 1$ and that $[S_k^L]$ is given by (31). $\square$

7. **Space-continuous limits.** In this section, we discuss the space-continuous models that can be obtained after a space rescaling in our space-discrete process. Since we have shown in Theorem 6.1 that 1D two-sided lattices reduce to 1D one-sided lattices, we focus on the latter case to make the presentation clearer.

Equation (39a) can be rewritten as

$$[S_k]' + q_k e^{-\int_0^t p_k}[S_k^0]([S_k] - [S_{k-1}]) + \left(p_k + q_k \left(1 - e^{-\int_0^t p_k}[S_k^0]\right) + r_k\right)[S_k]$$
$$= q_k e^{-\int_0^t p_k}[S_k^0]\left([R_{k-1}^0] + \frac{r_k(0)}{q_k(0)} + \int_0^t \left(\frac{r_k}{q_k}\right)'(s)\frac{[S_k](s)}{[S_k^0]}ds\right). \tag{73}$$

Since the term $[S_k] - [S_{k-1}]$ is a discrete spatial derivative, this suggests that we can derive a space-continuous PDE by a proper space rescaling. To do that, assume that the individuals are placed on the inifinite line $\mathcal{K} = \mathbb{Z}$, fix $0 < \Delta x \ll 1$, and proceed to the space-rescaling

$$[S](t, x) := \left[S_{k=\lfloor\frac{x}{\Delta x}\rfloor}\right](t), \qquad \forall x \in \Delta x \, \mathbb{Z}.$$

Similarily, we define $p(t, x) = p_{k=\lfloor\frac{x}{\Delta x}\rfloor}(t)$, $q(t, x) = q_{k=\lfloor\frac{x}{\Delta x}\rfloor}(t)$, $r(t, x) = r_{k=\lfloor\frac{x}{\Delta x}\rfloor}(t)$, $[S^0](x) = \left[S_{k=\frac{x}{\Delta x}}^0\right]$, $[R^0](x) = \left[R_{k=\frac{x}{\Delta x}}^0\right]$.

The limit of (73) as $\Delta x \to 0$ depends on some additional assumptions on how the parameters rescale. In the present paper, we shall only discuss the formal passage to the limit and leave the rigorous proof to future works.

7.1. **Limiting integrodifferential ODE.** First, let us make the following assumptions:

Rescaling Assumption 1. the parameters $p_k(t)$, $q_k(t)$, $r_k(t)$ and the initial conditions $[S_k^0]$, $[R_k^0]$ converge as $\Delta x \to 0$ locally uniformly in C^1 in time and space to some smooth functions $p(t, x)$, $q(t, x) > 0$, $r(t, x)$, $[S^0(x)]$, and $[R^0(x)]$, respectively.

Under these assumptions, we have that $[S_k] - [S_{k-1}] = O(\Delta x)$ and so

$$[S](t, x) - [S](t, x - \Delta x) = O(\Delta x). \tag{74}$$

Hence, $S(t, x; \Delta x)$ converges to the solution of the (integrodifferential) ODE

$$\begin{cases} \partial_t [S](t, x) + \left[p(t, x) + q(t, x)\left(1 - [S^0](x)e^{-\int_0^t p(\cdot, x)}\right) + r(t, x) \right][S] \\ \quad = q(t, x)e^{-\int_0^t p(\cdot, x)}[S^0](x)\left([R_{k-1}^0] + \frac{r(0, x)}{q(0, x)} + \int_0^t \partial_t \left[\frac{r}{q} \right](s, x) \frac{[S](s, x)}{[S^0](x)} ds \right), \\ S(0, x) = [S^0(x)]. \end{cases} \tag{75}$$

In particular, if p, q, and r do not depend on time (but are allowed to depend on space), we can integrate the above ODE and get the explicit expression

$$[S](t, x) = [S^0](x)e^{-(p+q+r)t + [S^0]q\frac{1-e^{-pt}}{p}}\left(1 + (r + q[R^0]) \int_0^t e^{(q+r)\tau - [S^0]q\frac{1-e^{-p\tau}}{p}} d\tau \right). \tag{76}$$

Note that the solution (76) is equal pointwise to the spatially-uniform solution (49). Intuitively, this is because under Rescaling Assumptions 1, the spatial heterogeneity becomes locally uniform as $\Delta x \to 0$ and there are infinitely many nodes between two distinct points $x, x' \in \mathbb{R}$. Therefore, the solution can be computed at each point as a spatially constant solution.

The solution (76) is thus spatially homogeneous in the short spatial scale $O(dx)$, but inhomogeneous on the long spatial scale $x = O(1)$. This multiple-scale property occurs in many situations, e.g., the slowly-varying amplitude approximation in optics.

The convergence of the solution of (73) to the solution of the limiting ODE (75) as $\Delta x \to 0$ under *Rescaling Assumptions 1* is illustrated in Figure 1.

Remark 6. Solution (76) can be derived if the initial condition S^0, I^0, R^0 and the parameters p, q, r are only piecewise smooth. It allows us to include discontinuous coefficients (we can then solve the equation in each of the continuous piece separately). This may be interesting for several problems:

- The Riemann problem where the initial condition is a step function $[S^0](x) = 0$ when $x < 0$ and $[S^0](x) = 1$ when $x > 0$.
- The gap problem, when $q(x)$ and $p(x)$ are spatially compactly supported or periodic
- The time-control (or "optimal duration campaign" problem) when p and q are piecewise constant with respect to time and change value at some time $t_0 > 0$.

7.2. **Limiting PDE..** The former rescaling gives a limiting ODE for which space can be viewed as a parameter since there is no derivative with respect to the variable x. In other words, there is no spatial coupling in the system as $\Delta x \to 0$, as expressed in (74). Intuitively, this is because the influence of any node on its right neighbors propagates at the speed $c(t, x) := [S^0]qe^{-\int_0^t p}\Delta x$, see (73). Thus, $c(t, x) \to 0$ as
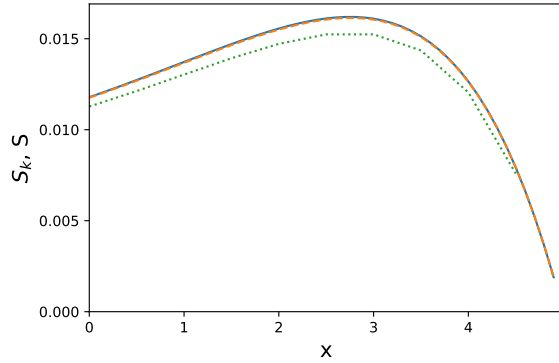
FIGURE 1. Convergence to the (integrodifferential) ODE limit (75) under *Rescaling Assumptions 1*. Snapshot at $t = 2$ of $[S_k(t)]$ on a segment $k = 0, \ldots, \lfloor \frac{5}{\Delta x} \rfloor$ so that $x \in (0, 5)$. Blue solid line represents the explicit solution (76) of the limiting ODE (75) ; Green dotted line and orange dashed line represent the solution $[S_k](t)$ of (73) for $\Delta x = 0.5$ and $\Delta x = 0.1$ respectively. Choice of parameters and initial conditions : $p(x) = 1 - \frac{x}{5}$, $q(x) = 5 + x$, $r(x) = 2 - \frac{2x}{5}$, $[S^0](x) = 0.2 - \frac{x}{25}$, $[R^0](x) = 0.2 + \frac{3x}{50}$

$\Delta x \to 0$ and so the limiting equation is spatially decoupled. Therefore, in order to get a limiting equation that includes space derivative, one has to rescale q by a factor of $1/\Delta x$, namely, to assume that $q = \frac{\tilde{q}}{\Delta x}$ for a function $\tilde{q}$ independent of Δx (or which converges locally uniformly as $\Delta x \to 0$). This way, the term $q \left([S](t, x) - [S](t, x - \Delta x) \right)$ is of order 1 and converges to $\tilde{q} \nabla_x [S]$ as $\Delta x \to 0$.

If we assume, however, that the contagion coefficient becomes large as the space-step Δx vanishes, i.e., that q is of order $1/\Delta x$, we also need to rescale the other parameters. Indeed, we see from equation (73) that, in order not to get the singular limit $S(t, x) \equiv 0$ as $\Delta x \to 0$ and $q = O(1/\Delta x)$, we need the term $1 - [S^0]e^{-pt}$ to be of order $O(\Delta x)$ as $\Delta x \to 0$. To fullfill thoses conditions, we make the following

Rescaling Assumption 2. Assume that $q_k(t) > 0$ and that the ratio $\frac{r_k(t)}{q_k(t)}$ does not depend on time. Further assume that the parameters and initial conditions scale as

$$q = \frac{\tilde{q}(t, x; \Delta x)}{\Delta x} > 0, \quad p = \tilde{p}(t, x; \Delta x)\Delta x, \quad r = \tilde{r},$$
$$[I^0(x)] = [\tilde{I}^0(x; \Delta x)]\Delta x, \quad [R^0(x)] = [\tilde{R}^0(x; \Delta x)]\Delta x, \tag{77}$$

where the "tilde functions" are assumed to converge locally uniformly as $\Delta x \to 0$.

Since we assume that $\frac{r(t, x)}{q(t, x)}$ does not depend on time, equation (73) boils down to

$$[S_k]' + q_k e^{-\int_0^t p_k}[S_k^0]([S_k] - [S_{k-1}]) + \left(p_k + q_k \left(1 - e^{-\int_0^t p_k}[S_k^0] \right) + r_k \right)[S_k]$$
$$= e^{-\int_0^t p_k}[S_k^0] \left(q_k[R_{k-1}^0] + r_k \right). \tag{78}$$
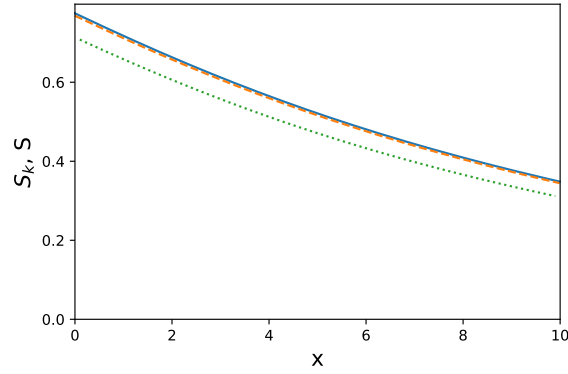
FIGURE 2. Convergence towards PDE limit (79) under *Rescaling Assumptions 2*. Snapshot at $t = 2$ of $[S_k(t)]$ on a segment $k = 0, \ldots, \lfloor \frac{10}{\Delta x} \rfloor$, with $\Delta x > 0$, so that $x \in (0, 10)$. Blue solid line represents the solution $S(t, x)$ of the limiting PDE (79) ; Green dotted line and orange dashed line represent the solution $[S_k](t)$ of (78) for $\Delta x = 0.1$ and $\Delta x = 0.01$ respectively. Choice of parameters and initial conditions : $\tilde{p}(x) = 0.1 + \frac{0.2x}{10}$, $\tilde{q}(x) = 1 + \frac{x}{10}$, $\tilde{r}(x) = 0.3 + \frac{0.5x}{10}$, $[\tilde{I}^0](x) = 0.2 + \frac{0.5x}{10}$, $[\tilde{R}^0](x) = 0.5 - \frac{0.3x}{10}$. The parameters p, q, r and initial conditions $[S^0]$, $[I^0]$, $[R^0]$ are then defined by the rescaling (77).

Under Rescaling Assumptions 2, the terms in (78) have the following limit for any fixed $t \geq 0$ as $\Delta x \to 0$:

$$[S^0] \sim 1$$

$$q[S^0]e^{-\int_0^t p}\big([S](t, x) - [S](t, x - \Delta x)\big) \sim \tilde{q}\partial_x[S](t, x),$$

$$p + q\big(1 - e^{-\int_0^t p}[S^0]\big) + r \sim \tilde{q}\big([\tilde{I}^0](x) + [\tilde{R}^0](x) + \int_0^t \tilde{p}\big) + \tilde{r},$$

$$e^{-\int_0^t p}[S^0]\big(q[R^0] + r\big) \sim \tilde{q}[\tilde{R}^0] + \tilde{r},$$

and so the limiting equation satisfied by $[S(t, x)]$ is

$$\begin{cases} \partial_t[S](t, x) + \tilde{q}(t, x)\partial_x[S] \\ \qquad + \left[\tilde{q}\left([\tilde{I}^0](x) + [\tilde{R}^0](x) + \int_0^t \tilde{p}(t, x)\right) + \tilde{r}(t, x)\right][S] = \tilde{q}[\tilde{R}^0] + \tilde{r}, \quad (79) \\ [S](0, x) = 1. \end{cases}$$

The convergence of the solution of (78) to the solution of the limiting PDE (79) as $\Delta x \to 0$ under *Rescaling Assumptions 2* is illustrated in Figure 2. The PDE (79) is solved numerically using a standard explicit finite-difference Euler scheme.

Equation (79) can be solved using the method of the caracteristics. For example, assuming for simplicity that there is no recovery (i.e., $[\tilde{R}^0] \equiv \tilde{r} \equiv 0$), no source term (i.e., $p \equiv 0$) and that $\tilde{q}(t, x) \equiv \tilde{q}$ is contant, the solution reads

$$S(t, x) = e^{-\int_{x - \tilde{q}t}^x [\tilde{I}^0](s)ds}. \tag{80}$$

7.3. **Discussion.** The two space-continuous models proposed above are obtained by different rescalings, each of which corresponding to a different modeling situation:

- The (integrodifferential) ODE limit (75) discussed in Section 7.1 corresponds to the case where the contagion between two consecutive nodes is driven by their discrete distance on the graph rather than by their metric distance on the line. In this case, the propagation of the epidemic propagates at finite speed q on the graph. When the rescaling factor Δx goes to 0, any interval on the real line contains infinitely many nodes, therefore the epidemics do not propagate on the line in finite time. This explains why the obtained limiting (integrodifferential) ODE (75) is spatially decoupled. The (integrodifferential) ODE limit (75) is relevant from the modeling point of view when the contagion phenomenon occurs at a small space-scale.

- In contrast, the PDE limit (79) obtained in Section 7.2 corresponds to a contagion phenomenon that is driven by the metric distance between nodes rather than on their discrete distance on the graph. In this case, the propagation of the epidemic occurs at a finite speed $\tilde{q}$ on the line, as expressed by the explicit form of the solution in (80). It is then natural to consider the parameter p and initial conditions $[I^0]$ and $[R^0]$ as space-densities, as suggested by the Rescaling Assumptions (77). The PDE limit (79) has the advantage to qualitatively reproduce the spatial propagation at finite speed occurring in the discrete model (73). This PDE rescaling is also relevant when the contagion rate outweighs the source term (i.e., $p \ll q$).

8. **Final remarks.** In the present paper, we establish an *exact deterministic description* of stochastic SIR-Bass epidemics on 1D lattices by showing that the probability of infection at a given point in space and time can be obtained as the solution of a deterministic ODE system on the lattice. Our framework allows any type of heterogeneity on the parameters and initial conditions. In addition, our results precisely describe the spatio-temporal dynamics at a local scale, which is still a major challenge in epidemiology.

Our results focus on 1D lattices (i.e., finite, semi-infinite, and infinite lines) where the infection only occurs from a node to its direct neighbors. This is the first step towards more realistic contact networks, such as 2D lattices, scale-free networks, small-world networks, as well as other types of networks that are relevant from the application point of view in the context of epidemiology or marketing. Most likely, an exact deterministic description such as the one obtained in this paper cannot be obtained for 2D lattices or other contact networks[2]. We believe however that the analytical tools introduced in the present paper will be useful for further analysis of epidemiological models in complex networks under the pair-approximation assumption [26].

---

[2]When the underlying undirected graph contain loops, such as in nD lattices for $n \geq 2$, it is not possible even for the Bass model to find a finite system of equations on the marginals as in Theorem 4.1, see [38].

REFERENCES

[1] R. Anderson and R. May, *Infectious Diseases of Humans: Dynamics and Control*, Oxford University Press, 1991.

[2] H. Andersson, Limit theorems for a random graph epidemic model, *The Annals of Applied Probability*, **8** (1998), 1331–1349.

[3] H. Andersson and T. Britton, Stochastic Epidemic Models and Their Statistical Analysis, volume 151. Springer-Verlag New York, 2000.

[4] J. Badham and R. Stocker, The impact of network clustering and assortativity on epidemic behaviour, *Theoretical Population Biology*, **77** (2010), 71–75.

[5] N. T. J. Bailey, *The Mathematical Theory of Infectious Diseases and its Applications*, Charles Griffin & Company Ltd, Bucks, 1975.

[6] F. Ball and D. Sirl, An SIR epidemic model on a population with random network and household structure, and several types of individuals, *Advances in Applied Probability*, **44** (2012), 63–86.

[7] S. Bansal, B. T. Grenfell and L. A. Meyers, When individual behaviour matters: Homogeneous and network models in epidemiology, *Journal of the Royal Society Interface*, **4** (2007), 879–891.

[8] F. M. Bass, A new product growth for model consumer durables, *Mathematical Models in Marketing*, 132 (1969), 351–353.

[9] H. Berestycki, J.-P. Nadal and N. Rodríguez, A model of riot dynamics: Shocks, diffusion, and thresholds, *Networks and Heterogeneous Media*, **10** (2015), 443–475.

[10] T. Britton, Stochastic epidemic models: A survey, *Mathematical Biosciences*, **225** (2010), 24–35.

[11] L. Danon, A. P. Ford, T. House, C. P. Jewell, M. J. Keeling, G. O. Roberts, J. V. Ross and M. C. Vernon, Networks and the epidemiology of infectious disease, *Interdisciplinary Perspectives on Infectious Diseases*, **2011** (2011), Article ID 284909.

[12] L. Decreusefond, J.-S. Dhersin, P. Moyal and V. C. Tran, Large graph limit for an SIR process in random network with heterogeneous connectivity, *The Annals of Applied Probability*, **22** (2012), 541–575.

[13] O. Diekmann, Limiting behaviour in an epidemic model, *Nonlinear Analysis: Theory, Methods & Applications*, **1** (1977), 459–470.

[14] O. Diekmann, Run for your life. A note on the asymptotic speed of propagation of an epidemic, *Journal of Differential Equations*, **33** (1979), 58–73.

[15] K. Dietz, Epidemics and Rumours: A Survey, *Journal of the Royal Statistical Society. Series A (General)*, **130** (1967), 505–528.

[16] J. Enright and R. R. Kao, Epidemics on dynamic networks, *Epidemics*, **24** (2018), 88–97.

[17] G. Fibich, Bass-SIR model for diffusion of new products in social networks, *Physical Review E*, **94** (2016), 32305, 5pp.

[18] G. Fibich, Diffusion of new products with recovering consumers, *Society for Industrial and Applied Mathematics*, **77** (2017), 1230–1247.

[19] G. Fibich and R. Gibori, Aggregate diffusion dynamics in agent-based models with a spatial structure, *Operations Reasearch*, **58** (2010), 1450–1468.

[20] G. Fibich and T. Levin, *Network Effects in the Discrete Bass Model*, Work in progress, 2021.

[21] G. Fibich, T. Levin and O. Yakir, Boundary effects in the discrete Bass model, *SIAM Journal on Applied Mathematics*, **79** (2019), 914–937.

[22] M. Graziano and K. Gillingham, Spatial patterns of solar photovoltaic system adoption: The influence of neighbors and the built environmentz, *Journal of Economic Geography*, **15** (2015), 815–839.

[23] B. T. Grenfell, O. N. Bjørnstad and J. Kappey, Travelling waves and spatial hierarchies in measles epidemics, *Nature*, **414** (2001), 716–723.

[24] H. Heathcote, The mathematics of infectious diseases, *SIAM Review*, **42** (2000), 599–653.

[25] H. W. Hethcote, Three basic epidemiological models, *Applied Mathematical Ecology (Trieste, 1986)*, 119–144, Biomathematics, 18, Springer, Berlin, 1989.

[26] T. House and M. J. Keeling, Insights from unifying modern approximations to infections on networks, *Journal of the Royal Society Interface*, **8** (2011), 67–73.

[27] M. Keeling, The implications of network structure for epidemic dynamics, *Theoretical Population Biology*, **67** (2005), 1–8.

[28] M. J. Keeling, The effects of local spatial structure on epidemiological invasions, *Proc. R. Soc. Long. B*, **266** (1999), 859–867.

[29] M. J. Keeling and K. T. Eames, Networks and epidemic models, *Journal of the Royal Society Interface*, **2** (2005), 295–307.

[30] W. Kermack and A. McKendrick, A contribution to the mathematical theory of epidemics, *Proceedings of the Royal Society A*, **115** 1927.

[31] T. G. Kurtz, Solutions of ordinary differential equations as limits of pure jump Markov processes, *Journal of Applied Probability*, **7** (1970), 49–58.

[32] T. G. Kurtz, Limit theorems for sequences of jump markov processes approximating ordinary, *Journal of Applied Probability*, **8** (1971), 344–356.

[33] H. Matsuda, N. Ogita, A. Sasaki and K. Sato, Statistical mechanics of population: The lattice lotka-volterra model, *Progress of Theoretical Physics*, **88** (1992), 1035–1049.

[34] D. Mollison, *Epidemic Models: Their structure and Relation to Data*, Cambridge University Press, 1995.

[35] L. Pellis, F. Ball, S. Bansal, K. Eames, T. House, V. Isham and P. Trapman, Eight challenges for network epidemic models, *Epidemics*, **10** (2015), 58–62.

[36] S. Ruan, Spatial-temporal dynamics in nonlocal epidemiological models, In *Mathematics for Life Science and Medicine*, Springer, (2007), 97–122.

[37] K. Sato, H. Matsuda and A. Sasaki, Pathogen invasion and host extinction in lattice structured populations, *Journal of Mathematical Biology*, **32** (1994), 251–268.

[38] K. J. Sharkey, I. Z. Kiss, R. R. Wilkinson and P. L. Simon, Exact equations for sir epidemics on tree graphs, *Bulletin of Mathematical Biology*, **77** (2015), 614–645.

[39] P. L. Simon and I. Z. Kiss, From exact stochastic to mean-field ODE models: A new approach to prove convergence results, *IMA Journal of Applied Mathematics*, **78** (2013), 945–964.

[40] S. A. Socolofsky and G. H. Jirka, *Advective Diffusion Equation (lecture notes),* 2004.

Appendix A. Appendix.

A.1. **Proof of Proposition 2.** Expression (26) can be verified by direct substitution in (17). To show (27), we first notice that $\mathbb{E}[N_K^n] = \sum_{k=1}^{K}[I_k](t)$. Using $[I_k](t) = 1 - [S_k](t)$ and expression (26), we compute

$$\mathbb{E}[N_K^n] = \sum_{k=1}^{K}\left(1 - e^{-(p+q)t}\sum_{l=0}^{k-1}\frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!}\right)$$

$$= K - e^{-(p+q)t}\sum_{k=1}^{K}\sum_{l=0}^{k-1}\frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!}$$

$$= K - e^{-(p+q)t}\sum_{l=0}^{K-1}\sum_{k=l+1}^{K}\frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!}$$

$$= K - e^{-(p+q)t}\sum_{l=0}^{K-1}(K-l)\frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!}.$$

We obtain (27) from the change of variable $l' = K - l$ in the above expression.

To show (29), note that above expression gives

$$\frac{\mathbb{E}[N_K^n]}{K} = 1 - e^{-(p+q)t}\sum_{l=0}^{K-1}\frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!} + \frac{1}{K}\sum_{l=0}^{K-1}l\frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!}.$$

Then, we conclude that, as $K \to +\infty$,

$$\sum_{l=0}^{K-1} \frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!} \sim e^{q\frac{1-e^{-pt}}{p}},$$

and that

$$\frac{1}{K}\sum_{l=0}^{K-1} l\frac{\left(q\frac{1-e^{-pt}}{p}\right)^l}{l!} = \frac{q\frac{1-e^{-pt}}{p}}{K}\underbrace{\sum_{l=1}^{K-1} \frac{\left(q\frac{1-e^{-pt}}{p}\right)^{l-1}}{(l-1)!}}_{=O(1)} \to 0.$$

A.2. **Proof of Proposition 7.** A straightforward substitution shows that expression (58) solves equation (39a) and satisfies the initial conditions (57).

To prove (59), first note that $\mathbb{E}[N_K^n] = \sum_{k=1}^{K}[S_k(t)]$, and so, using (58), we find

$\mathbb{E}[N_K^n]$

$$= \sum_{k=1}^{K}\left(1 - \left(\frac{q}{q+r}\right)^k\left(1 - e^{-(q+r)t}\sum_{l=0}^{k-1}\frac{((q+r)t)^l}{l!}\right)\right)$$

$$= K - \frac{q}{q+r}\frac{1-\left(\frac{q}{q+r}\right)^K}{1-\frac{q}{q+r}} - e^{-(q+r)t}\sum_{k=1}^{K}\sum_{l=0}^{k-1}\left(\frac{q}{q+r}\right)^k\frac{((q+r)t)^l}{l!}$$

$$= K - \frac{q}{r}\left(1 - \left(\frac{q}{q+r}\right)^K\right) - e^{-(q+r)t}\sum_{l=0}^{K-1}\frac{((q+r)t)^l}{l!}\left(\frac{q}{q+r}\right)^l\frac{1-\left(\frac{q}{q+r}\right)^{K-L+1}}{1-\frac{q}{q+r}}.$$

Then, we get (59) from a change of variable $l' = K - l$ in the above expression.

A.3. **Spatial Markovian property - proof of Theorem 6.2.** The proof of Theorem 6.1 relies on the spatial Markovian property stated in Theorem A.1. This section is devoted to the proof of this result. For the sake of generality, we actually prove the following more general statement.

Lemma A.1. *Assume that the individuals are placed on a 1D two-sided lattice (see (60)) and that the initial conditions are uncorrelated, see (3).*

For any integer $k \in \mathbb{Z}$, denote by $X_{k-}^n := (x_l^n)_{l<k}$ and $X_{k+}^n := (x_l^n)_{l>k}$ the left and right neighbors of k. We have that

$$x_{k-1}^n \perp\!\!\!\perp x_{k+1}^n | (x_k^m)_{0\leq m<n}, \qquad \forall k \in \mathbb{Z}, \tag{81}$$

meaning that the state of the left and right neighbors of k are independent under the conditioning with respect to the state of k for all preceeding times. In particular,

$$x_{k-1}^n \perp\!\!\!\perp x_{k+1}^n | S_k^n. \tag{82}$$

Proof. Let $(\tilde{x}_{k-1}^n, \tilde{x}_{k+1}^n) \in \{s, i, r\}^2$ be a possible realization of (x_{k-1}^n, x_{k+1}^n), and let us denote by $Y_-^n := \{x_{k-1}^n = \tilde{x}_{k-1}^n\}$ and $Y_+^n := \{x_{k+1}^n = \tilde{x}_{k+1}^n\}$ the events that x_{k+1}^n and x_{k-1}^n equal this realization respectively. For clarity, we also dentote by $X_k^n := (x_k^m)_{0\leq m<n}$.

Let us prove by induction on $n \geq 0$ that

$$\mathbb{P}\left(Y_-^n \cap Y_+^n | X_k^n\right) = \mathbb{P}\left(Y_-^n | X_k^n\right)\mathbb{P}\left(Y_+^n | X_k^n\right). \tag{83}$$

By assumption on the initial conditions, this property holds for $n = 0$. Assume that the above property holds for some $n \geq 0$. Then, conditionning with respect to $Y_-^n \cap Y_+^n$ and using (40), we have

$$\mathbb{P}\left(Y_-^{n+1} \cap Y_+^{n+1} | X_k^{n+1}\right) = \mathbb{P}\left(Y_-^{n+1} \cap Y_+^{n+1} | X_k^{n+1}, Y_-^n \cap Y_+^n\right) \mathbb{P}\left(Y_-^n \cap Y_+^n | X_k^{n+1}\right) \tag{84}$$

From the assumption that the $(x_k^{n+1})_{k \in \mathcal{K}}$ are all independent if we condition by $(x_k^n)_{k \in \mathcal{K}}$, we deduce

$$\mathbb{P}\left(Y_-^{n+1} \cap Y_+^{n+1} | X_k^{n+1}, Y_-^n \cap Y_+^n\right)$$
$$= \mathbb{P}\left(Y_-^{n+1} | X_k^{n+1}, Y_-^n \cap Y_+^n\right) \mathbb{P}\left(Y_+^{n+1} | X_k^{n+1}, Y_-^n \cap Y_+^n\right).$$

Moreover, from the assumption that $\mathcal{K}$ is a 1D two-side lattice, we further deduce that

$$\mathbb{P}\left(Y_-^{n+1} \cap Y_+^{n+1} | X_k^{n+1}, Y_-^n \cap Y_+^n\right) = \mathbb{P}\left(Y_-^{n+1} | X_k^{n+1}, Y_-^n\right) \mathbb{P}\left(Y_+^{n+1} | X_k^{n+1}, Y_+^n\right). \tag{85}$$

Besides, the induction assumption (83) implies that

$$\mathbb{P}\left(Y_-^n \cap Y_+^n | X_k^{n+1}\right) = \mathbb{P}\left(Y_-^n | X_k^{n+1}\right) \mathbb{P}\left(Y_+^n | X_k^{n+1}\right). \tag{86}$$

Injecting (85) and (86) into (84), we derive

$$\mathbb{P}\left(Y_-^{n+1} \cap Y_+^{n+1} | X_k^{n+1}\right)$$
$$= \mathbb{P}\left(Y_-^{n+1} | X_k^{n+1}, Y_-^n\right) \mathbb{P}\left(Y_-^n | X_k^{n+1}\right) \mathbb{P}\left(Y_+^{n+1} | X_k^{n+1}, Y_+^n\right) \mathbb{P}\left(Y_+^n | X_k^{n+1}\right)$$
$$= \mathbb{P}\left(Y_-^{n+1} | X_k^{n+1}\right) \mathbb{P}\left(Y_+^{n+1} | X_k^{n+1}\right),$$

which proves that (83) holds at the rank $n + 1$.

We have thus proved that (83) holds for all $n \geq 0$, from which we immediately deduce (81).

Property (82) is deduced as a particular instance of (81). $\qquad\square$

Remark 7. In constrast with Theorem A.1, we have that

$$x_{k-1}^n \not\perp\!\!\!\perp x_{k+1}^n | I_k^n.$$

The reason is that the condition I_k^n only prescribes the state of x_k^n at time t^n and not the states $(x_k^m)_{0 \leq m < n}$ at the preceeding times since it does not indicate at what time x_k becomes infected.

E-mail address: fibich@tau.ac.il
E-mail address: samnordmann@gmail.com