Davide Secchi · Martin Neumann
*Editors*

# Agent-Based Simulation of Organizational Behavior

## New Frontiers of Social Science Research

Springer

# Agent-Based Simulation of Organizational Behavior

raffaello.seri@uninsubria.it

Davide Secchi • Martin Neumann
Editors

# Agent-Based Simulation of Organizational Behavior

## New Frontiers of Social Science Research

Springer

*Editors*
Davide Secchi
COMAC Research Cluster
Department of Language
and Communication
University of Southern Denmark
Slagelse, Denmark

Martin Neumann
Institute for information systems
Koblenz University
Koblenz, Germany

The Editors gratefully acknowledge the support of the Society for the Study of Artificial Intelligence and Simulation of Behavior (AISB) in developing this volume.

ISBN 978-3-319-18152-3 ISBN 978-3-319-18153-0 (eBook)
DOI 10.1007/978-3-319-18153-0

Library of Congress Control Number: 2015945340

Springer Cham Heidelberg New York Dordrecht London


Printed on acid-free paper

Springer International Publishing AG Switzerland is part of Springer Science+Business Media (www.springer.com)

# Chapter 13
# Analytical Approaches to Agent-Based Models

**Raffaello Seri**

**Abstract** The aim of this article is to present an approach to the analysis of simple systems composed of a large number of units in interaction. Suppose to have a large number of agents belonging to a finite number of different groups: as the agents randomly interact with each other, they move from a group to another as a result of the interaction. The object of interest is the stochastic process describing the number of agents in each group. As this is generally intractable, it has been proposed in the literature to approximate it in several ways. We review these approximations and we illustrate them with reference to a version of the epidemic model. The tools presented in the paper should be considered as a complement rather than as a substitute of the classical analysis of ABMs through simulation.

**Keywords** Individual-based models • Markov processes • Differential equations • Diffusion approximation • Central limit theorem

## 13.1 Introduction

The aim of this paper is to provide an introduction to the approximation of a class of models that may be of some interest in the study of organizations.

The models we consider here describe the evolution over time of a population composed of similar individuals moving from one to another of $d$ mutually exclusive categories. Models of this class are sometimes called *compartmental* (see, e.g., Matis & Kiffe, 2000) as they represent transitions of individuals between compartments. From another perspective, the models we consider belong to the class of *individual-based* models. Some authors consider individual-based and agent-based as synonyms (see, e.g., Railsback & Grimm, 2011), while others reserve the term individual-based for models in which rules of behavior are formulated in probabilistic terms at the individual level (see, e.g., Black & McKane, 2012, p. 338). This requires that some simplifying assumptions are needed in order to allow an

R. Seri (✉)
Dipartimento di Economia, Università degli Studi dell'Insubria,
Via Monte Generoso 71, 21100 Varese, Italy
e-mail: raffaello.seri@uninsubria.it



D. Secchi, M. Neumann (eds.), *Agent-Based Simulation of Organizational Behavior*, DOI 10.1007/978-3-319-18153-0_13

affordable analysis: usually each individual can be in only one of a finite number of states, and members of each state are supposed to be identical in every other respect.[1] These hypotheses are not necessary in agent-based models, in which the fact that the solution is obtained through simulation allows the researcher to consider more complex rules of behavior and continuous attributes. The price to pay for this freedom is that results thus obtained are only numerical. Even if we recognize these advantages of agent-based models, we claim that the techniques that we are going to present may still serve several purposes. First, in simplified contexts they could be seen as direct alternatives to the computer-intensive simulations of agent-based models. Second, they could be used as a preliminary step in the analysis of agent-based models, in order to obtain some hints concerning the behavior of the system. Third, they could serve as auxiliary simulation methods for parts of an agent-based simulation.

In practice, the models we consider are described by density dependent jump Markov processes with time homogeneous transition intensities (see, e.g., Kurtz, 1981, Chap. 8 and Ethier & Kurtz, 1986, Chap. 11). As the explicit analysis of these models may be daunting, we discuss some approximations that have been proposed in the literature. Their study has been pioneered by Feller (1951) and they have been developed by Norman (1968, 1972, 1974a, 1974b), Kurtz (1970, 1971, 1976, 1978, 1980, 1981, 1983) and Barbour (1972, 1974). Here we just consider the simplest situation, without any attempt to cover more advanced topics, such as spatial issues, age dependence, time inhomogeneity or dependence upon the whole past of the process. Moreover, we limit ourselves to introduce results already present in the literature.

We consider a series of models indexed by a number $N$, that can be the total population size, the area or the volume occupied by the population or any other indicator. For $N$ large, the behavior of the model can be approximated by what happens in the limiting situation in which $N$ is infinity. It turns out that this amounts to approximate these Markov processes through ordinary and stochastic differential equations. Clearly, the interest of the approximations is that the limiting behavior is simpler than what happens for finite $N$. These approximations have been used to describe abundances in biological populations (see the reviews in Pollett, 2001; Black & McKane, 2012), quantities of reactants in chemical reactions (see, e.g., Kurtz, 1972) and stochastic process algebra models in computer science (see, e.g., the introductory treatment in Bortolussi, Hillston, Latella, & Massink, 2013), among others. It is important to note here that the discrete time case is covered in Bortolussi et al. (2013, Sect. 2) and Challenger, Fanelli, and McKane (2014), while an alternative approach based on the so-called master equation can be found in Goutsias and Jenkinson (2013).

The theoretical results will be illustrated using an example of information spread in a fixed population. The model is simplistic in several respects. First, the structure of the model is the simplest possible, with only two compartments and a transition

[1]In the following models, discrete differences between individuals can be accounted for by adequately expanding the number of compartments and by varying the transition intensities.

between them. This is justified by the fact that the model has only an expository purpose. However, more reasonable models could be obtained, e.g., supposing that the population is stratified in several mutually exclusive groups, each one with a different exposure to the information and a different probability of passing it to someone else. Second, we suppose very simple interaction mechanisms between individuals in different states. Indeed, the present form, in which interaction terms are bilinear in the cardinalities of the two interacting subgroups, has a long history dating back at least to Lotka and Volterra (note that, while in Lotka, 1925, p. 89 and Volterra, 1962, pp. 119–120, the form of the interaction is justified as an approximation to the true one, in Volterra, 1931, p. 14 and, especially, in Volterra, 1962, pp. 9–10, pp. 119–120, a probabilistic interpretation is provided). However, more complex interactions can be considered, the price to pay being an increased complexity in the study of the system (by the way, the irrealistic form of interactions in the Lotka–Volterra model is the rationale that led Gause and Kolmogorov to introduce their variants of the predator-prey system, see, e.g., Sigmund, 2007). Third, the transition intensities between states are supposed to be time homogeneous, i.e. the rate at which individuals move between states does not depend explicitly upon time; moreover, the intensities do not depend upon the past of the process but only upon its present value. We maintain both hypotheses throughout the whole paper, but we remark that they can be relaxed using the results in Kurtz (1983). Fourth, the network modeling the agent interactions has no topological structure (see, e.g., Centola, 2010; Hirshman, Charles, & Carley, 2011; Zhang & Wu, 2012; Wang, Tao, Xie, & Yi, 2013; Plikynas & Masteika, 2014; however, see the Introduction in Collet, Dai Pra, & Sartori, 2010 for a justification of mean-field interactions without topological structure in social sciences).

Now we introduce the notation used in the following. The symbols $\mathbb{Z}$, $\mathbb{R}$, and $\mathbb{R}_+$ denote respectively the set of integer (positive and negative), real, and nonnegative real numbers. Vectors are always supposed to be column vectors and indicated with bold letters. For a vector $\mathbf{x}$, $x_i$ denotes its $i$–th element and $|\mathbf{x}|$ is the sum of the absolute values of the elements of $\mathbf{x}$. The superscript $\mathsf{T}$, as in $\mathbf{x}^{\mathsf{T}}$, indicates that the transpose of $\mathbf{x}$ is taken. Capital letters usually indicate random variables. Whenever needed, we will indicate derivatives of $X$ with respect to time as $\dot{X}$, while we reserve the prime symbol ($X'$, $X''$, $X'''$) for indicating approximations of $X$. Differentials of a variable $x$ are indicated as $\mathrm{d}x$ and derivatives of a function $f$ with respect to an argument $x$ are written as $\frac{\partial f}{\partial x}$. The quantities corresponding to finite values of $N$ will be indexed, whenever possible, by a superscript $(N)$.

As concerns the structure of the paper, in Sect. 13.2 we introduce the process specified in terms of transitions between compartments and of density dependent transition intensities. In Sect. 13.3 we present the first deterministic approximation through an ordinary differential equation. Section 13.4 contains two different stochastic results; the first one approximates directly the process with a stochastic differential equation (see Sect. 13.4.1), the second one shows that the scaled deviations of the original process from the deterministic process of Sect. 13.3 behaves for large $N$ as a Gaussian process (see Sect. 13.4.2). At last, Appendix contains a non-technical discussion of the conditions under which the results hold.

## 13.2 The Original Process

We consider a process $\left\{\hat{\mathbf{X}}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ such that, for any instant of time $t \geq 0$, $\hat{\mathbf{X}}_t^{(N)}$ is a vector of size $d$ with integer coordinates; more formally, we say that $\hat{\mathbf{X}}_t^{(N)}$ takes its values in $\mathbb{Z}^d$. Each coordinate of the vector $\hat{\mathbf{X}}_t^{(N)}$ corresponds univocally to one of the possible states or compartments of the model, and its value measures the number of individuals in that state in time $t$. The process moves in continuous time from a point of $\mathbb{Z}^d$, say $\mathbf{k}$, occupied in $t$, to another point, say $\mathbf{k}+\boldsymbol{\ell}$, occupied in $t+s$, with $s > 0$. As already explained in the introduction, the process $\left\{\hat{\mathbf{X}}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ is indexed by a number $N$, that can be integer (e.g., the size of the population) or real (e.g., the area in which the population dwells).

**Exercise 1 (News Diffusion Model).** The model that we are going to analyze is (a version of) the simple epidemic model. Consider the spread of a piece of news in a population of $N$ individuals. Let $\hat{S}_t$ and $\hat{I}_t$ be, respectively, the number of susceptibles (people that have not yet been reached by the news) and infected (i.e. people that have been reached by the news) at time $t$. Clearly, $\hat{S}_t + \hat{I}_t = N$ for any $t > 0$. Therefore, $\left[\hat{S}_t, \hat{I}_t\right]^\mathsf{T}$ will be equal to $\mathbf{k} = [k_1, k_2]^\mathsf{T}$ with $k_1 + k_2 = N$ and $0 \leq k_1 \leq N$. No other values are allowed for $\left[\hat{S}_t, \hat{I}_t\right]^\mathsf{T}$. On the other hand, $\left[\hat{S}_{t+s}, \hat{I}_{t+s}\right]^\mathsf{T}$ takes the value $\mathbf{k}+\boldsymbol{\ell}$. Two facts should be clear. First, the first element of $\boldsymbol{\ell}$ should be negative while the second should be positive, as people can become aware of the news but cannot do the reverse. Second, all values of $\boldsymbol{\ell}$ should be of the form $\boldsymbol{\ell} = [-\ell, +\ell]^\mathsf{T}$, otherwise the elements of $\mathbf{k}+\boldsymbol{\ell}$ fail to sum to $N$. □

We suppose that $\left\{\hat{\mathbf{X}}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ is a *Markov process*, i.e. a stochastic process whose state in $t+s$ depends on the past before $t$ only through the state occupied in $t$ (see, e.g., Ethier & Kurtz, 1986, Sect. 4.1 or Karlin & Taylor, 1975, Chap. 4, for definitions). A Markov process can be described by its *transition probability*, i.e. the probability that the process $\left\{\hat{\mathbf{X}}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ starting from the value $\mathbf{k}$ in $t$ reaches the value $\mathbf{k}+\boldsymbol{\ell}$ in $t+s$:

$$\mathbb{P}\left\{\hat{\mathbf{X}}_{t+s}^{(N)} = \mathbf{k}+\boldsymbol{\ell} \,\middle|\, \hat{\mathbf{X}}_t^{(N)} = \mathbf{k}\right\},$$

for any $t \geq 0$ and $s > 0$, and where $\mathbf{k}$ and $\boldsymbol{\ell}$ should in general respect some constraints. In the following, it will be particularly useful to consider what happens when $s = \mathrm{d}t$. In this case, we introduce the so-called *transition intensities* $q_{\mathbf{k},\mathbf{k}+\boldsymbol{\ell}}^{(N)}$ for $\mathbf{k}, \boldsymbol{\ell} \in \mathbb{Z}^d$, namely the quantities defined as:

$$\mathbb{P}\left\{\hat{\mathbf{X}}_{t+\mathrm{d}t}^{(N)} = \mathbf{k}+\boldsymbol{\ell} \,\middle|\, \hat{\mathbf{X}}_t^{(N)} = \mathbf{k}\right\} = q_{\mathbf{k},\mathbf{k}+\boldsymbol{\ell}}^{(N)} \cdot \mathrm{d}t + o(\mathrm{d}t)$$

raffaello.seri@uninsubria.it

or

$$\lim_{\mathrm{d}t \downarrow 0} \frac{\mathbb{P}\left\{\hat{\mathbf{X}}_{t+\mathrm{d}t}^{(N)} = \mathbf{k} + \boldsymbol{\ell} \,\middle|\, \hat{\mathbf{X}}_t^{(N)} = \mathbf{k}\right\}}{\mathrm{d}t} = q_{\mathbf{k},\mathbf{k}+\boldsymbol{\ell}}^{(N)}.$$

Therefore, a transition intensity in $t$ is the limit, as $s \downarrow 0$, of the transition probability between $t$ and $t+s$ divided by the length of the time period $s$. It measures the instantaneous probability that a jump of size $\boldsymbol{\ell}$ takes place immediately after $t$, for a process starting from $\mathbf{k}$ in $t$. It is often the case that transition intensities are null for large values of $|\boldsymbol{\ell}|$ and for some combinations of $\mathbf{k}$ and $\boldsymbol{\ell}$. As these intensities do not depend on $t$, the transition probabilities are called *time homogeneous* or *stationary*.

In the following we will further suppose that $\left\{\hat{\mathbf{X}}_t^{(N)}\right\}_{t \in \mathbb{R}_+}$ is *density dependent* (see Ethier & Kurtz, 1986, Chap. 11 for definition and examples), i.e. that its transition intensities $q_{\mathbf{k},\mathbf{k}+\boldsymbol{\ell}}^{(N)}$ depend on $\mathbf{k}$ only through the ratio $\mathbf{k}/N$. In particular, this means that the transition intensity between the states $\mathbf{k}$ and $\mathbf{k}+\boldsymbol{\ell}$ takes the following form:

$$q_{\mathbf{k},\mathbf{k}+\boldsymbol{\ell}}^{(N)} = N \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right)$$

for a function $\beta_{\boldsymbol{\ell}}$ indexed by the jump size $\boldsymbol{\ell}$, with $\mathbf{k}, \boldsymbol{\ell} \in \mathbb{Z}^d$. The requirement of density dependence implies that the transition intensity increases proportionally to the index $N$ and depends on the state of the process $\mathbf{k}$ only through its density $\mathbf{k}/N$. Because of density dependence, we are led to consider $\left\{\mathbf{X}_t^{(N)}\right\}_{t \in \mathbb{R}_+}$, defined by $\mathbf{X}_t^{(N)} = \hat{\mathbf{X}}_t^{(N)}/N$. In this case, we have:

$$\lim_{\mathrm{d}t \downarrow 0} \frac{\mathbb{P}\left\{\mathbf{X}_{t+\mathrm{d}t}^{(N)} = \frac{\mathbf{k}}{N} + \frac{\boldsymbol{\ell}}{N} \,\middle|\, \mathbf{X}_t^{(N)} = \frac{\mathbf{k}}{N}\right\}}{\mathrm{d}t} = N \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right). \tag{13.1}$$

**Exercise 2 (News Diffusion Model—Continued).** In each infinitesimal interval $\mathrm{d}t$, the number of contacts between susceptibles and infected can be assumed to be proportional to $\hat{S}_t \cdot \hat{I}_t$. The probability that two or more contacts take place in $\mathrm{d}t$ is $o(\mathrm{d}t)$. When such a contact takes place, we suppose that the probability that the news is transmitted from the infected to the susceptible is fixed and independent of everything else. Let $\hat{\mathbf{X}}_t^{(N)} = \left[\hat{S}_t, \hat{I}_t\right]^{\mathsf{T}}$. Therefore:

$$\mathbb{P}\left\{\begin{bmatrix}\hat{S}_{t+\mathrm{d}t}\\ \hat{I}_{t+\mathrm{d}t}\end{bmatrix} = \begin{bmatrix}\hat{s}_t - 1\\ \hat{i}_t + 1\end{bmatrix} \,\middle|\, \begin{bmatrix}\hat{S}_t\\ \hat{I}_t\end{bmatrix} = \begin{bmatrix}\hat{s}_t\\ \hat{i}_t\end{bmatrix}\right\} = p \cdot \frac{\hat{s}_t \cdot \hat{i}_t}{N} \cdot \mathrm{d}t + o(\mathrm{d}t).$$

In d$t$ values of $\boldsymbol{\ell}$ different from $\boldsymbol{\ell} = [-1,+1]^{\mathsf{T}}$ yield transitions of such a low probability to be $o\,(\mathrm{d}t)$. This means that:

$$\beta_{[-1,+1]^{\mathsf{T}}}\left([x_1,x_2]^{\mathsf{T}}\right) = p \cdot x_1 x_2.$$

This simple example lends itself to a further remark. As $\hat{S}_t + \hat{I}_t = N$ and $N$ is fixed, $\hat{S}_t$ is known when $\hat{I}_t$ is and we can also identify $\hat{\mathbf{X}}_t^{(N)} = \hat{I}_t$. In this case $\mathbf{k} = k$ with $0 \le k \le N$ and $\boldsymbol{\ell} = 1$:

$$\mathbb{P}\left\{\hat{I}_{t+\mathrm{d}t} = \hat{i}_t + 1 \,\middle|\, \hat{I}_t = \hat{i}_t\right\} = p \cdot \frac{\left(N - \hat{i}_t\right)\hat{i}_t}{N} \cdot \mathrm{d}t + o\,(\mathrm{d}t)\,, \tag{13.2}$$

while other values of $\boldsymbol{\ell}$ yield 0 (or $o\,(\mathrm{d}t)$) transition intensities. In the following we will use the definitions $S_t = \hat{S}_t/N$ and $I_t = \hat{I}_t/N$. Equation (13.2) becomes:

$$\mathbb{P}\left\{I_{t+\mathrm{d}t} = i_t + \frac{1}{N}\,|I_t = i_t\right\} = N \cdot p \cdot (1 - i_t)\, i_t \cdot \mathrm{d}t + o\,(\mathrm{d}t)\,,$$

where $\beta_1\,(x) = p \cdot x\,(1-x)$. In Fig. 13.1, a trajectory of $\left\{\hat{I}_t\right\}_{t\in\mathbb{R}_+}$ and $\{I_t\}_{t\in\mathbb{R}_+}$ with $p = 0.01$, $N = 20$ and $\hat{I}_0 = 2$, is reproduced as a step function. □

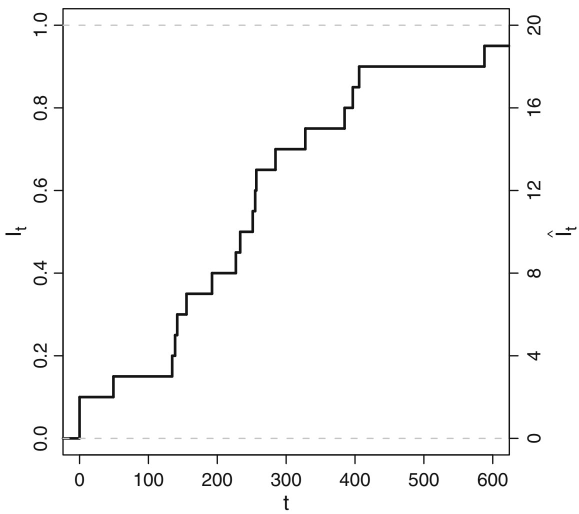


**Fig. 13.1** Graph of a trajectory of $\left\{\hat{I}_t\right\}_{t\in\mathbb{R}_+}$ and $\{I_t\}_{t\in\mathbb{R}_+}$ with $p = 0.01$, $N = 20$ and $\hat{I}_0 = 2$

Now we try to understand heuristically what happens to $\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)}$ when $N$ diverges. Equation (13.1) leads to:

$$\mathbb{P}\left\{\mathbf{X}_{t+\mathrm{d}t}^{(N)} = \frac{\mathbf{k}}{N} + \frac{\boldsymbol{\ell}}{N} \middle| \mathbf{X}_t^{(N)} = \frac{\mathbf{k}}{N}\right\} = \mathbb{P}\left\{\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)} = \frac{\boldsymbol{\ell}}{N} \middle| \mathbf{X}_t^{(N)} = \frac{\mathbf{k}}{N}\right\}$$
$$\simeq N \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right) \cdot \mathrm{d}t.$$

This means that, for any $\boldsymbol{\ell} \in \mathbb{Z}^d$, the random variable $\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)}$ will take the value $\frac{\boldsymbol{\ell}}{N}$ with probability approximately equal to $N \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right) \cdot \mathrm{d}t$. Therefore its mean and variance are approximately:

$$\mathbb{E}\left[\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)} \middle| \mathbf{X}_t^{(N)} = \frac{\mathbf{k}}{N}\right] \simeq \sum_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}}{N} \cdot N \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right) \cdot \mathrm{d}t = \sum_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right) \cdot \mathrm{d}t$$

and:

$$\mathbb{V}\left[\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)} \middle| \mathbf{X}_t^{(N)} = \frac{\mathbf{k}}{N}\right] \simeq \sum_{\boldsymbol{\ell}} \frac{\boldsymbol{\ell}\boldsymbol{\ell}^{\mathsf{T}}}{N^2} \cdot N \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right) \cdot \mathrm{d}t = \frac{1}{N} \cdot \sum_{\boldsymbol{\ell}} \boldsymbol{\ell}\boldsymbol{\ell}^{\mathsf{T}} \cdot \beta_{\boldsymbol{\ell}}\left(\frac{\mathbf{k}}{N}\right) \cdot \mathrm{d}t.$$

This shows that, when $\mathbf{X}_t^{(N)} = \mathbf{x}$, $\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)}$ approximately behaves, on average, as $\sum_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \beta_{\boldsymbol{\ell}}(\mathbf{x}) \cdot \mathrm{d}t$, and that the variance of $\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)}$ around its mean decreases as $N^{-1}$. This means that, when $N$ is very large, $\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)}$ is well approximated by $\sum_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \beta_{\boldsymbol{\ell}}(\mathbf{x}) \cdot \mathrm{d}t$, a fact that is the object of Sect. 13.3. Moreover, in Sect. 13.4, we will show that also the deviation between $\mathbf{X}_{t+\mathrm{d}t}^{(N)} - \mathbf{X}_t^{(N)}$ and $\sum_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \beta_{\boldsymbol{\ell}}(\mathbf{x}) \cdot \mathrm{d}t$ can be studied and used to improve the previous approximation.

## 13.3 The Deterministic Limit

In this section we show that, when $N$ is large enough, $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ can be approximated by a deterministic process $\left\{\mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$ (see Appendix for conditions).

Define the function:

$$\mathbf{f}(\mathbf{x}) := \sum_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \beta_{\boldsymbol{\ell}}(\mathbf{x})$$

where the sum is extended over all possible values of $\boldsymbol{\ell}$. As $N \to \infty$, under some additional conditions that will be detailed in Appendix, $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ converges to the deterministic process $\left\{\mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$ defined by:

$$\mathbf{X}'_t = \mathbf{X}'_0 + \int_0^t \mathbf{f}\left(\mathbf{X}'_s\right) \mathrm{d}s, \qquad t \geq 0.$$

Now, using this formula for $\mathbf{X}'_{t+\mathrm{d}t}$ and $\mathbf{X}'_t$, this process can be written as:

$$\begin{aligned}\mathbf{X}'_{t+\mathrm{d}t} - \mathbf{X}'_t &= \int_0^{t+\mathrm{d}t} \mathbf{f}\left(\mathbf{X}'_s\right) \mathrm{d}s - \int_0^t \mathbf{f}\left(\mathbf{X}'_s\right) \mathrm{d}s \\ &= \int_t^{t+\mathrm{d}t} \mathbf{f}\left(\mathbf{X}'_s\right) \mathrm{d}s = \mathbf{f}\left(\mathbf{X}'_t\right) \mathrm{d}t, \quad t \geq 0,\end{aligned}$$

or, using the equality $\frac{\mathbf{X}'_{t+\mathrm{d}t} - \mathbf{X}'_t}{\mathrm{d}t} = \dot{\mathbf{X}}'_t$, equivalently as:

$$\dot{\mathbf{X}}'_t = \mathbf{f}\left(\mathbf{X}'_t\right), \qquad t \geq 0$$

or:

$$\mathrm{d}\mathbf{X}'_t = \mathbf{f}\left(\mathbf{X}'_t\right) \cdot \mathrm{d}t, \qquad t \geq 0. \tag{13.3}$$

**Exercise 3 (News Diffusion Model—Continued).** In the first version of the news diffusion model (see Exercise 2):

$$\mathbf{f}(\mathbf{x}) = \sum_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \beta_{\boldsymbol{\ell}}(\mathbf{x}) = \begin{bmatrix} -1 \\ +1 \end{bmatrix} \cdot p \cdot x_1 x_2$$

and:

$$\begin{bmatrix} \dot{S}'_t \\ \dot{I}'_t \end{bmatrix} = \begin{bmatrix} -p \cdot S'_t I'_t \\ +p \cdot S'_t I'_t \end{bmatrix}, \qquad t \geq 0.$$

In the second rewriting of the model (see Exercise 2), we get:

$$f(x) = \sum_{\ell} \ell \cdot \beta_{\ell}(x) = p \cdot x(1 - x).$$

Therefore, the corresponding differential equation is:

$$\dot{I}'_t = p \cdot I'_t \left(1 - I'_t\right), \qquad t \geq 0$$

or

$$\mathrm{d}I'_t = p \cdot I'_t \left(1 - I'_t\right) \mathrm{d}t, \qquad t \geq 0. \tag{13.4}$$

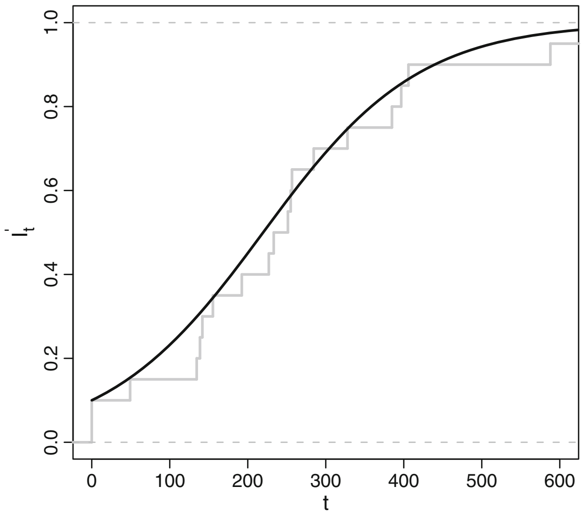


**Fig. 13.2** Graph of $\{I'_t\}_{t\in\mathbb{R}_+}$ (in *black*) over a trajectory of $\{I_t\}_{t\in\mathbb{R}_+}$ (in *grey*)

It is possible to see that the two models are indeed the same. By the way, this model has a closed form solution. Supposing that $I'_0 = i_0$, the solution is:

$$I'_t = \frac{\exp(p\cdot t)}{\frac{1-i_0}{i_0} + \exp(p\cdot t)}.$$

When $t = 0$, we have $I'_0 = i_0$ while, when $t \to \infty$, $\lim_{t\to\infty} I'_t = 1$. Moreover, the curve $t \mapsto I'_t$ is increasing. In Fig. 13.2, the deterministic approximation $\{I'_t\}_{t\in\mathbb{R}_+}$, corresponding to $p = 0.01$ and $i_0 = 0.1$, is reproduced in black over the previous trajectory of $\{I_t\}_{t\in\mathbb{R}_+}$, in grey. In Fig. 13.3, the difference between the trajectory of the original process and its deterministic approximation, $\{I_t - I'_t\}_{t\in\mathbb{R}_+}$, is displayed. □

As shown in the figures, the process $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ deviates from its deterministic approximation $\{\mathbf{X}'_t\}_{t\in\mathbb{R}_+}$. The new stochastic process $\left\{\mathbf{X}_t^{(N)} - \mathbf{X}'_t\right\}_{t\in\mathbb{R}_+}$ is characterized by fluctuations that decrease when $N$ increases. In particular, it can be shown that:[2]

$$\mathbf{X}_t^{(N)} = \mathbf{X}'_t + O_{\mathbb{P}}\left(\frac{1}{\sqrt{N}}\right). \tag{13.5}$$

[2]We write that $X_n = O_{\mathbb{P}}(a_n)$ where $n$ is an index diverging to infinity if, for any $\varepsilon > 0$, there exists a finite $M > 0$ such that $\mathbb{P}(|X_n/a_n| > M) < \varepsilon$ for any $n$ large enough.

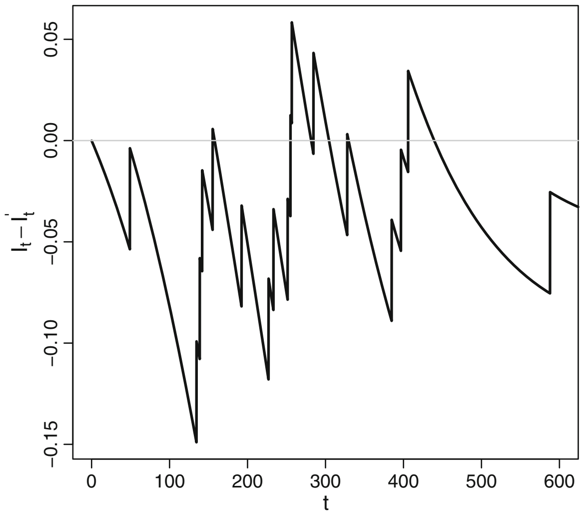


**Fig. 13.3** Graph of a trajectory of $\{I_t - I_t'\}_{t\in\mathbb{R}_+}$

In terms of the original process, we have:

$$\hat{\mathbf{X}}_t^{(N)} = N \cdot \mathbf{X}_t' + O_{\mathbb{P}}\left(\sqrt{N}\right).$$

In the next section we will see that the $O_{\mathbb{P}}\left(\frac{1}{\sqrt{N}}\right)$ term in (13.5) provides a refinement to this approximation.

## 13.4 The Stochastic Limit

The previous result states that, when $N$ is large enough, the process $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ converges to a deterministic process $\{\mathbf{X}_t'\}_{t\in\mathbb{R}_+}$ expressed as a differential equation. The results that we are going to present in this section describe the fluctuations of the process $\left\{\mathbf{X}_t^{(N)} - \mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$ for large values of $N$.

In the literature on approximations for density dependent Markov processes, two different kinds of stochastic results are considered. In the first one, often called *diffusion approximation*, $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ is directly approximated through a diffusion. In the second one, one approximates the process $\left\{\mathbf{V}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$, where:

$$\mathbf{V}_t^{(N)} := \sqrt{N}\left(\mathbf{X}_t^{(N)} - \mathbf{X}_t'\right),$$

given by the scaled fluctuations of the stochastic process $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ around the deterministic evolution $\left\{\mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$, through a Gaussian process $\left\{\mathbf{V}_t\right\}_{t\in\mathbb{R}_+}$. This goes under the name of *Central Limit Theorem approximation*.

### *13.4.1 The Diffusion Approximation*

Let us start from the diffusion approximation. In this case, $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ is approximated by the Gaussian process $\left\{\mathbf{X}_t''\right\}_{t\in\mathbb{R}_+}$ (see Appendix for conditions) described by the following stochastic differential equation (SDE) or Itô diffusion:[3]

$$\mathrm{d}\mathbf{X}_t'' = \mathbf{f}\left(\mathbf{X}_t''\right)\mathrm{d}t + \frac{1}{\sqrt{N}}\sum_{\boldsymbol{\ell}}\boldsymbol{\ell}\cdot\sqrt{\beta_{\boldsymbol{\ell}}\left(\mathbf{X}_t''\right)}\cdot\mathrm{d}W_{\boldsymbol{\ell},t}, \qquad t\geq 0,$$

where the processes $\left\{W_{\boldsymbol{\ell},t}\right\}_{t\in\mathbb{R}_+}$ are independent Brownian motions, each one associated with a value of $\boldsymbol{\ell}$. Remark that $\sum_{\boldsymbol{\ell}}\boldsymbol{\ell}\cdot\sqrt{\beta_{\boldsymbol{\ell}}(\mathbf{x})}\cdot\mathrm{d}W_{\boldsymbol{\ell},t}$ is a Gaussian random vector with $\mathbf{0}$ mean and variance:

$$\mathbb{V}\left(\sum_{\boldsymbol{\ell}}\boldsymbol{\ell}\cdot\sqrt{\beta_{\boldsymbol{\ell}}(\mathbf{x})}\cdot\mathrm{d}W_{\boldsymbol{\ell},t}\,|\mathbf{x}\right) = \mathrm{d}t\cdot\sum_{\boldsymbol{\ell}}\boldsymbol{\ell}\boldsymbol{\ell}^{\mathsf{T}}\cdot\beta_{\boldsymbol{\ell}}(\mathbf{x}). \tag{13.6}$$

The limit of the process $\left\{\mathbf{X}_t''\right\}_{t\in\mathbb{R}_+}$ for large $N$ is exactly $\left\{\mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$.

In integral terms, the process can be written as:

$$\mathbf{X}_t'' = \mathbf{X}_0'' + \int_0^t \mathbf{f}\left(\mathbf{X}_s''\right)\mathrm{d}s + \frac{1}{\sqrt{N}}\sum_{\boldsymbol{\ell}}\boldsymbol{\ell}\cdot\int_0^t\sqrt{\beta_{\boldsymbol{\ell}}\left(\mathbf{X}_s''\right)}\cdot\mathrm{d}W_{\boldsymbol{\ell},s}, \qquad t\geq 0.$$

**Exercise 4 (News Diffusion Model—Continued).** The diffusion equation is:

$$\mathrm{d}I_t'' = p\cdot I_t''\left(1-I_t''\right)\mathrm{d}t + \frac{1}{\sqrt{N}}\cdot\sqrt{p\cdot I_t''\left(1-I_t''\right)}\cdot\mathrm{d}W_t, \quad t\geq 0.$$

The stochastic part of the equation has variance:

$$\mathbb{V}\left(\frac{1}{\sqrt{N}}\cdot\sqrt{p\cdot i_t(1-i_t)}\cdot\mathrm{d}W_t\,|i_t\right) = \frac{1}{N}\cdot p\cdot i_t(1-i_t)\cdot\mathrm{d}t.$$

[3] We follow here the Kunrei-shiki romanization convention, instead of the more common Hepburn romanization Itō, because Itô himself used the first one in several publications.

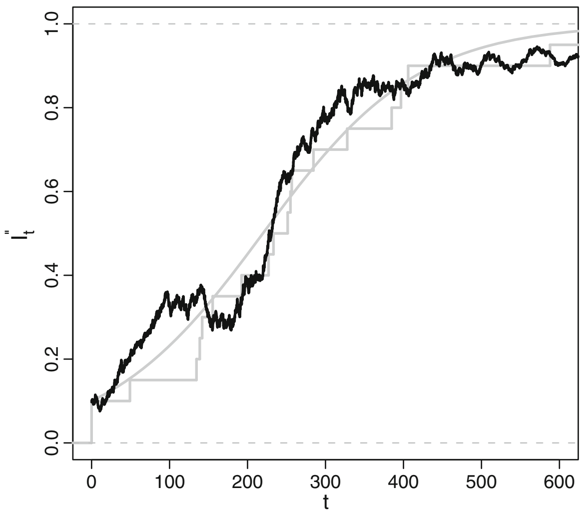


**Fig. 13.4** Graph of a trajectory of $\{I''_t\}_{t\in\mathbb{R}_+}$ (in *black*), over $\{I'_t\}_{t\in\mathbb{R}_+}$ and a trajectory of $\{I_t\}_{t\in\mathbb{R}_+}$ (both in *grey*)

This implies that the process is heteroskedastic, i.e. its variance depends on $t$. In Fig. 13.4, we reproduce a trajectory of $\{I''_t\}_{t\in\mathbb{R}_+}$, in black, over the deterministic approximation $\{I'_t\}_{t\in\mathbb{R}_+}$ and the previous trajectory of $\{I_t\}_{t\in\mathbb{R}_+}$, both in grey. From the graph there seems to be no particular similarity between $\{I''_t\}_{t\in\mathbb{R}_+}$ and $\{I_t\}_{t\in\mathbb{R}_+}$: we will pursue this point at the end of Sect. 13.4.2, showing that, for large enough $N$, the paths $\{I''_t\}_{t\in\mathbb{R}_+}$ and $\{I_t\}_{t\in\mathbb{R}_+}$ look similar in distribution. □

As concerns the precision of the approximation, for any process $\{\mathbf{X}_t\}_{t\in\mathbb{R}_+}$ it is possible to find a process $\{\mathbf{X}''_t\}_{t\in\mathbb{R}_+}$ such that (see Appendix for references):

$$\mathbf{X}_t^{(N)} = \mathbf{X}''_t + O\left(\frac{\ln N}{N}\right).$$

## 13.4.2 *The Central Limit Theorem Approximation*

As briefly explained above, this result considers the process $\left\{\mathbf{V}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$, where:

$$\mathbf{V}_t^{(N)} := \sqrt{N}\left(\mathbf{X}_t^{(N)} - \mathbf{X}'_t\right).$$

By (13.5), we expect $\mathbf{V}_t^{(N)}$ to be $O_{\mathbb{P}}(1)$.

Consider $\partial \mathbf{f}$, the matrix of partial derivatives of $\mathbf{f}$, defined as:

$$[\partial \mathbf{f}(\mathbf{x})]_{i,j} = \frac{\partial f_i(\mathbf{x})}{\partial x_j}$$

where $f_i$ is the $i$–th element of the vector of functions $\mathbf{f}$ and $x_j$ is the $j$–th element of $\mathbf{x}$. We define a Gaussian process $\{\mathbf{V}_t\}_{t\in\mathbb{R}_+}$ through the SDE:

$$\mathrm{d}\mathbf{V}_t = \partial \mathbf{f}\left(\mathbf{X}_t'\right) \cdot \mathbf{V}_t \mathrm{d}t + \sum_{\boldsymbol{\ell}} \boldsymbol{\ell} \cdot \sqrt{\beta_{\boldsymbol{\ell}}\left(\mathbf{X}_t'\right)} \cdot \mathrm{d}W_{\boldsymbol{\ell},t}, \qquad t \geq 0, \tag{13.7}$$

where the processes $\{W_{\boldsymbol{\ell},t}\}_{t\in\mathbb{R}_+}$ are independent Brownian motions, each one associated with a value of $\boldsymbol{\ell}$. Remark that $\left\{\mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$ is deterministic and therefore, in this SDE, both the drift and the diffusion coefficients are known in advance.

Under certain regularity conditions (see Appendix), we have:

$$\mathbf{V}_t^{(N)} \to_{\mathcal{D}} \mathbf{V}_t, \quad t \geq 0, \tag{13.8}$$

where the subscript on the arrow denotes convergence in distribution. This means that the fluctuations of $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ around $\left\{\mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$, opportunely scaled, behave as the Gaussian process $\{\mathbf{V}_t\}_{t\in\mathbb{R}_+}$.

**Exercise 5 (News Diffusion Model—Continued).** We get:

$$\partial f(x) = p(1-2x).$$

Therefore, the corresponding diffusion equation is:

$$\begin{aligned}\mathrm{d}V_t &= \partial f\left(I_t'\right) \cdot V_t \mathrm{d}t + \sqrt{\beta_1\left(I_t'\right)} \cdot \mathrm{d}W_t \\ &= p\left(1-2I_t'\right) \cdot V_t \mathrm{d}t + \sqrt{p \cdot I_t'\left(1-I_t'\right)} \cdot \mathrm{d}W_t.\end{aligned} \tag{13.9}$$

In Fig. 13.5, a trajectory $\{V_t\}_{t\in\mathbb{R}_+}$, in black, is plotted against the trajectory of $\left\{V_t^{(N)}\right\}_{t\in\mathbb{R}_+} = \left\{\sqrt{N}\cdot\left(I_t - I_t'\right)\right\}_{t\in\mathbb{R}_+}$, in grey, already displayed in Fig. 13.3 with a different scaling. □

Now, from the definition $\mathbf{V}_t^{(N)} := \sqrt{N}\left(\mathbf{X}_t^{(N)} - \mathbf{X}_t'\right)$ and the approximate result $\mathbf{V}_t^{(N)} \simeq \mathbf{V}_t$, valid in distribution for large $N$, we get:

$$\sqrt{N}\left(\mathbf{X}_t^{(N)} - \mathbf{X}_t'\right) \simeq \mathbf{V}_t,$$

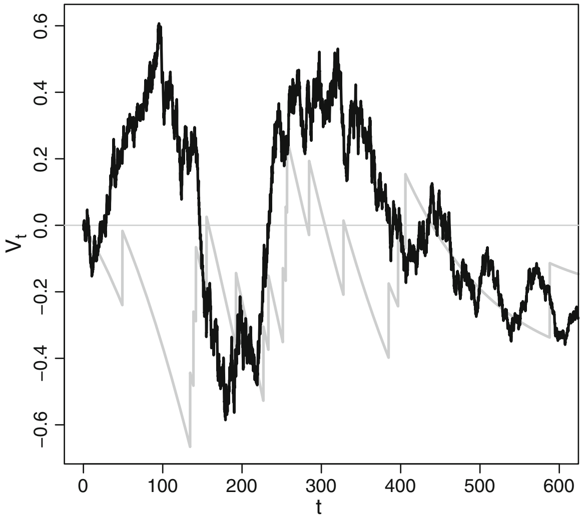


**Fig. 13.5** Graph of a trajectory of $\{V_t\}_{t\in\mathbb{R}_+}$ (in *black*) and of one of $\left\{\sqrt{N}\cdot\left(I_t - I'_t\right)\right\}_{t\in\mathbb{R}_+}$ (in *grey*)

$$\mathbf{X}_t^{(N)} \simeq \mathbf{X}'_t + \frac{1}{\sqrt{N}}\cdot\mathbf{V}_t. \tag{13.10}$$

The right-hand side of the last line leads us to consider the process $\left\{\mathbf{X}'''_t\right\}_{t\in\mathbb{R}_+}$, defined through the equality $\mathbf{X}'''_t := \mathbf{X}'_t + \frac{1}{\sqrt{N}}\cdot\mathbf{V}_t$. This process is an approximation to $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$. In differential terms, $\left\{\mathbf{X}'''_t\right\}_{t\in\mathbb{R}_+}$ is defined by:

$$\begin{aligned}
\mathrm{d}\mathbf{X}'''_t &= \mathrm{d}\mathbf{X}'_t + \frac{1}{\sqrt{N}}\cdot\mathrm{d}\mathbf{V}_t \\
&= \mathbf{f}\left(\mathbf{X}'_t\right)\cdot\mathrm{d}t + \frac{1}{\sqrt{N}}\cdot\partial\mathbf{f}\left(\mathbf{X}'_t\right)\cdot\mathbf{V}_t\cdot\mathrm{d}t + \frac{1}{\sqrt{N}}\cdot\sum_{\ell}\boldsymbol{\ell}\cdot\sqrt{\beta_{\ell}\left(\mathbf{X}'_t\right)}\cdot\mathrm{d}W_{\ell,t}, \qquad t\geq 0
\end{aligned}$$

where we have used (13.3) and (13.7). The replacement $\mathbf{V}_t = \sqrt{N}\cdot\left(\mathbf{X}'''_t - \mathbf{X}'_t\right)$ leads us to:

$$\begin{aligned}
\mathrm{d}\mathbf{X}'''_t &= \left\{\mathbf{f}\left(\mathbf{X}'_t\right) + \frac{1}{\sqrt{N}}\cdot\partial\mathbf{f}\left(\mathbf{X}'_t\right)\cdot\mathbf{V}_t\right\}\cdot\mathrm{d}t + \frac{1}{\sqrt{N}}\cdot\sum_{\ell}\boldsymbol{\ell}\cdot\sqrt{\beta_{\ell}\left(\mathbf{X}'_t\right)}\cdot\mathrm{d}W_{\ell,t} \\
&= \left\{\mathbf{f}\left(\mathbf{X}'_t\right) + \partial\mathbf{f}\left(\mathbf{X}'_t\right)\cdot\left[\mathbf{X}'''_t - \mathbf{X}'_t\right]\right\}\cdot\mathrm{d}t + \frac{1}{\sqrt{N}}\cdot\sum_{\ell}\boldsymbol{\ell}\cdot\sqrt{\beta_{\ell}\left(\mathbf{X}'_t\right)}\cdot\mathrm{d}W_{\ell,t}, \qquad t\geq 0.
\end{aligned}$$

The differences with respect to $\left\{\mathbf{X}''_t\right\}_{t\in\mathbb{R}_+}$ are the more complex form of the drift coefficient and the fact that the diffusion coefficient depends on the process

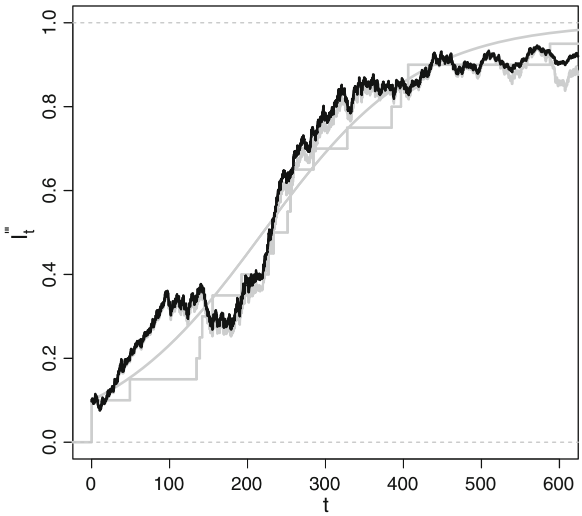


**Fig. 13.6** Graph of a trajectory of $\{I_t'''\}_{t\in\mathbb{R}_+}$ (in *black*), over trajectories of $\{I_t''\}_{t\in\mathbb{R}_+}$, $\{I_t\}_{t\in\mathbb{R}_+}$ and over $\{I_t'\}_{t\in\mathbb{R}_+}$ (all in *grey*)

$\{\mathbf{X}_t'\}_{t\in\mathbb{R}_+}$. As this process is a deterministic function of $t$, the diffusion coefficient behaves as if a forcing is applied.

**Exercise 6 (News Diffusion Model—Continued).** Starting from (13.10) and replacing in it the formulas (13.4) and (13.9) for $\mathrm{d}I_t'$ and $\mathrm{d}V_t$, we get:

$$\begin{aligned}\mathrm{d}I_t''' &= \mathrm{d}I_t' + \frac{1}{\sqrt{N}} \cdot \mathrm{d}V_t \\ &= p \cdot I_t'\left(1 - I_t'\right)\mathrm{d}t + \frac{1}{\sqrt{N}} \cdot \left\{p\left(1 - 2I_t'\right) \cdot V_t \mathrm{d}t + \sqrt{p \cdot I_t'\left(1 - I_t'\right)} \cdot \mathrm{d}W_t\right\} \\ &= p \cdot \left\{\left(1 - 2I_t'\right) \cdot I_t''' + \left(I_t'\right)^2\right\} \cdot \mathrm{d}t + \frac{1}{\sqrt{N}} \cdot \sqrt{p \cdot I_t'\left(1 - I_t'\right)} \cdot \mathrm{d}W_t, \quad t \geq 0.\end{aligned}$$

In Fig. 13.6, we plot a trajectory of $\{I_t'''\}_{t\in\mathbb{R}_+}$, in black, over the previous trajectories of $\{I_t''\}_{t\in\mathbb{R}_+}$, $\{I_t\}_{t\in\mathbb{R}_+}$ and over $\{I_t'\}_{t\in\mathbb{R}_+}$, all in grey. In order to ensure comparability, both $\{I_t'''\}_{t\in\mathbb{R}_+}$ and $\{I_t''\}_{t\in\mathbb{R}_+}$ have been based on the same Brownian motion path $\{W_t\}_{t\in\mathbb{R}_+}$. □

There is little to choose between $\{\mathbf{X}_t''\}_{t\in\mathbb{R}_+}$ and $\{\mathbf{X}_t'''\}_{t\in\mathbb{R}_+}$ as concerns the precision of the approximation. Indeed, for any process $\left\{\mathbf{V}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ it is possible to find a process $\{\mathbf{V}_t\}_{t\in\mathbb{R}_+}$ such that (see Appendix for references):

$$\sqrt{N} \cdot \left(\mathbf{X}_t^{(N)} - \mathbf{X}_t'\right) = \mathbf{V}_t + O\left(\frac{\ln N}{\sqrt{N}}\right).$$

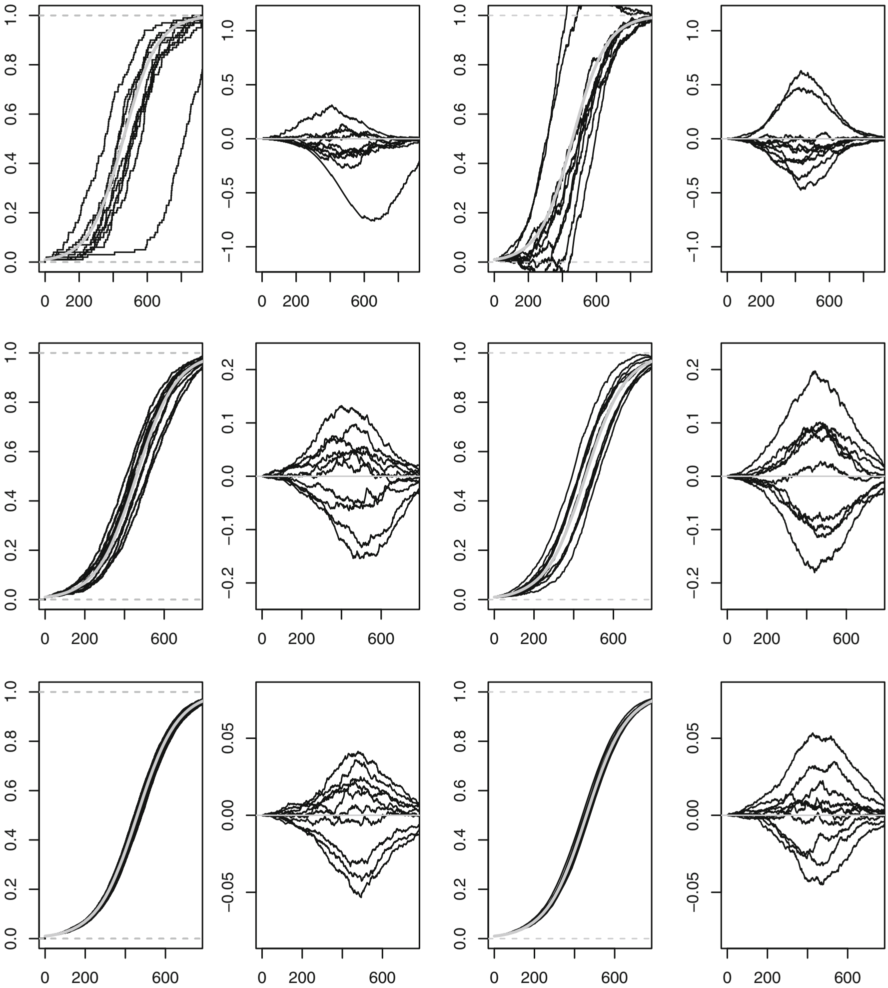


**Fig. 13.7** Plot of 10 paths of $\{I_t\}_{t\in\mathbb{R}_+}$, $\{I_t - I'_t\}_{t\in\mathbb{R}_+}$, $\{I'''_t\}_{t\in\mathbb{R}_+}$ and $\{I'''_t - I'_t\}_{t\in\mathbb{R}_+}$ (in columns, from *left to right*) for $N = 100, 1{,}000, 10{,}000$ (in rows, from *top to bottom*)

This implies that, for any process $\{\mathbf{X}_t\}_{t\in\mathbb{R}_+}$ it is possible to find a process $\{\mathbf{X}'''_t\}_{t\in\mathbb{R}_+}$ such that:

$$\mathbf{X}_t^{(N)} = \mathbf{X}'''_t + O\left(\frac{\ln N}{N}\right).$$

**Exercise 7 (News Diffusion Model—Continued).** In Fig. 13.7, we, respectively, represent some paths of $\{I_t\}_{t\in\mathbb{R}_+}$, $\{I_t - I'_t\}_{t\in\mathbb{R}_+}$, $\{I'''_t\}_{t\in\mathbb{R}_+}$ and $\{I'''_t - I'_t\}_{t\in\mathbb{R}_+}$, for three values of $N$. The parameters are $p = 0.01$ and $i_0 = 0.01$ for all the graphs.

Each row corresponds to a different value of $N$, namely, from top to bottom, 100, 1,000, and 10,000. The first column contains the graphs of 10 realizations of $\{I_t\}_{t\in\mathbb{R}_+}$, in black, as well as the curve $\{I'_t\}_{t\in\mathbb{R}_+}$, in grey. The second column illustrates the behavior of $\{I_t - I'_t\}_{t\in\mathbb{R}_+}$, displaying the deviations between each one of the 10 paths of the previous column and the deterministic approximation $\{I'_t\}_{t\in\mathbb{R}_+}$. The third column contains 10 realizations of the central limit theorem approximation $\{I'''_t\}_{t\in\mathbb{R}_+}$, in black, and the curve $\{I'_t\}_{t\in\mathbb{R}_+}$, in grey. The fourth column contains the differences $\{I'''_t - I'_t\}_{t\in\mathbb{R}_+}$. We do not plot $\{I''_t\}_{t\in\mathbb{R}_+}$ because the graphs in which this process replaces $\{I'''_t\}_{t\in\mathbb{R}_+}$ are undistinguishable with respect to these ones. The rationale of the graph is that the processes in the third (fourth) column should be approximations of the ones in the first (second) column. We have depicted several realizations in each subplot, because the approximation holds only in distribution and, as such, one realization would be insufficient to illustrate how its quality increases when passing from small $N$ (i.e., the first row) to large $N$ (i.e., the third row). Indeed, it is apparent from the graph that, for $N = 100$, the agreement between the distribution of the centered point process $\{I_t - I'_t\}_{t\in\mathbb{R}_+}$ and that of the Gaussian process $\{I'''_t - I'_t\}_{t\in\mathbb{R}_+}$ is not particularly good; this fact is witnessed by the different appearance of the curves $\{I_t\}_{t\in\mathbb{R}_+}$ and $\{I'''_t\}_{t\in\mathbb{R}_+}$. For $N = 1{,}000$, the agreement is clearly much better, while for $N = 10{,}000$ the two sets of curves are indistinguishable. □

## 13.5 Conclusions

In this paper, we have presented some probabilistic results that can be useful to approximate analytically a class of intrinsically stochastic individual- or agent-based models. With respect to classical agent-based models whose behavior is studied through simulation, the present approach is not able to deal with arbitrarily complex rules of behavior and often requires simplified assumptions. However, we believe that the methods presented here can still be helpful in the analysis of models customarily approached through simulation. Up to our knowledge, the most lucid example of this interaction is the analysis, performed in Galán and Izquierdo (2005), of the Norms and Metanorms models introduced in Axelrod (1986). This example shows how much insight can be gained when the mathematical approach is used as a supplement of simulations.

## Appendix: Technical Conditions

In this appendix, we discuss the technical conditions under which the results stated above hold true.

raffaello.seri@uninsubria.it

As concerns the deterministic approximation of Sect. 13.3, we follow Theorem 8.1 in Kurtz (1981) (similar results are Theorem 3.1 in Norman, 1968; Theorem (3.1) in Kurtz, 1970; Theorem 8.1.1 in Norman, 1972; Theorem (2.1) in Kurtz, 1976; Theorem 2.2 in Kurtz, 1978; Theorem (2.16) in Kurtz, 1980; Theorem 2.1 in Chap. 11 in Ethier & Kurtz, 1986).

Let $K \subset E$ be a bounded and closed (i.e., compact) set. The first condition requires that, for each $K$:

$$\sum_{\boldsymbol{\ell}} |\boldsymbol{\ell}| \cdot \sup_{\mathbf{x}\in K} \beta_{\boldsymbol{\ell}} (\mathbf{x}) < \infty.$$

The second condition requires that, for any $K$, there exists $M_K$ such that:

$$|\mathbf{f}(\mathbf{x}) - \mathbf{f}(\mathbf{y})| \leq M_K \cdot |\mathbf{x} - \mathbf{y}|, \quad \mathbf{x}, \mathbf{y} \in K.$$

At last, we require that the initial condition of the original process converges to the one of the deterministic one, i.e. $\lim_{N\to\infty} \mathbf{X}_0^{(N)} = \mathbf{x}_0$. By the way, under these conditions, the convergence of $\left\{\mathbf{X}_t^{(N)}\right\}_{t\in\mathbb{R}_+}$ to $\left\{\mathbf{X}_t'\right\}_{t\in\mathbb{R}_+}$ is uniform for $t$ belonging to bounded subsets of $\mathbb{R}_+$.

**Exercise 8 (News Diffusion Model—Continued).** Consider the epidemic model seen in Exercise 1 in the second rewriting. Using the fact that $x \in [0, 1]$, it is possible to see that $x(1-x) \leq \frac{1}{4}$. Therefore, we have:

$$\sum_{\ell} |\ell| \cdot \sup_{x\in K} \beta_\ell (x) = p \cdot \sup_{x\in K} x(1-x) \leq \frac{p}{4} < \infty.$$

As concerns the second hypothesis, we have:

$$\begin{aligned} |f(x) - f(y)| &= p \cdot |x(1-x) - y(1-y)| \\ &\leq p \cdot \sup_{z\in[x,y]} \left| \frac{\partial [z(1-z)]}{\partial z} \right| \cdot |x-y| \\ &= p \cdot \sup_{z\in[x,y]} |1-2z| \cdot |x-y| \leq p \cdot |x-y| \end{aligned}$$

where the second step derives from the mean value theorem. At last, we have supposed that $I_0 = i_0$ so that the initial condition is trivially verified. □

The diffusion approximation of Sect. 13.4.1 holds under the following conditions (this is Theorem 8.4 in Kurtz, 1981; see Theorem (3.13) in Kurtz, 1976; Theorem 3.3 in Kurtz, 1978; Theorem 2.1 in Kurtz, 1983; Theorem 3.1 in Chap. 11 in Ethier & Kurtz, 1986 for alternative or more general conditions):

- for any index $\boldsymbol{\ell}$ but a finite number, $\beta_{\boldsymbol{\ell}}(\mathbf{x}) \equiv 0$;
- for any index $\boldsymbol{\ell}$, $\overline{\beta}_{\boldsymbol{\ell}} = \sup_{\mathbf{x}} \beta_{\boldsymbol{\ell}}(\mathbf{x}) < +\infty$;

- there exists $M > 0$ such that:

$$\left|\beta_{\boldsymbol{\ell}}\left(\mathbf{x}\right)-\beta_{\boldsymbol{\ell}}\left(\mathbf{y}\right)\right| \leq M \cdot \overline{\beta}_{\boldsymbol{\ell}} \cdot \left|\mathbf{x}-\mathbf{y}\right|;$$

- there exists $M > 0$ such that:

$$\left|\mathbf{f}\left(\mathbf{x}\right)-\mathbf{f}\left(\mathbf{y}\right)\right| \leq M \cdot \left|\mathbf{x}-\mathbf{y}\right|.$$

The rate on the approximation of $\{\mathbf{X}_t\}_{t\in\mathbb{R}_+}$ through $\left\{\mathbf{X}_t''\right\}_{t\in\mathbb{R}_+}$ at the end of Sect. 13.4.1 can be found in Theorem (3.13) in Kurtz (1976), Theorem 3.3 in Kurtz (1978), Theorem 8.4 in Kurtz (1981) and Theorem 3.1 in Chap. 11 in Ethier and Kurtz (1986). By the way, the coupling is uniform over bounded intervals of the real line.

**Exercise 9 (News Diffusion Model—Continued).** There exists only one index $\ell$, i.e. $\ell = 1$, for which $\beta_\ell \not\equiv 0$. For this index, $\overline{\beta}_1 = p \cdot \sup_{x\in[0,1]} x\left(1-x\right) = p/4 < +\infty$. Now, from Exercise 8:

$$\left|\beta_1\left(x\right)-\beta_1\left(y\right)\right| \leq p \cdot \left|x-y\right|,$$

i.e. one can take $M = 4$. On the other hand, always from Exercise 8:

$$\left|f\left(x\right)-f\left(y\right)\right| = \left|\beta_1\left(x\right)-\beta_1\left(y\right)\right| \leq p \cdot \left|x-y\right|,$$

i.e. one can take $M = p$. Therefore, any $M \geq \max\{4, p\}$ respects the conditions.

□

The convergence in Sect. 13.4.2 holds under the following conditions (these are the ones stated in Theorem 8.2 in Kurtz, 1981; for related results, see Theorem 1.1 in Norman, 1968; Theorem (3.5) in Kurtz, 1971; Theorem 8.1.1 in Norman, 1972; Theorem 1 in Barbour, 1974; Theorem (2.3) in Kurtz, 1976; Theorem 2 in Allain, 1976a; Theorem 4.4 in Kurtz, 1978; Theorem 2.2 in Kurtz, 1983; Theorem 2.3 in Chap. 11 in Ethier & Kurtz, 1986):

- for each bounded closed set $K$, we have:

$$\sum_{\boldsymbol{\ell}} \left|\boldsymbol{\ell}\right|^2 \sup_{\mathbf{x}\in K} \beta_{\boldsymbol{\ell}}\left(\mathbf{x}\right) < \infty;$$

- the functions $\partial\mathbf{f}$ and $\beta_{\boldsymbol{\ell}}$, for each $\boldsymbol{\ell}$, are continuous;
- the initial conditions converge in such a way that $\lim_{N\to\infty} \sqrt{N}\left|\mathbf{X}_0^{(N)}-\mathbf{x}_0\right| = \mathbf{0}$.

Versions of this result holding uniformly for $t > 0$ have been stated in Theorem 3.2 (ii) in Norman (1974b), Theorem 1 in Norman (1974a), Theorem (2.7) in Kurtz (1976) and Theorem 8.5 in Kurtz (1981). Berry–Esséen-type theorems can be found in Theorem 1 in Barbour (1974), Theorem (2.5) in Kurtz (1976), Allain (1976b), Corollary 4.5 in Kurtz (1978) and Chapters 5 and 6 in Alm (1978).

The rate on the approximation of $\{\mathbf{X}_t\}_{t\in\mathbb{R}_+}$ through $\left\{\mathbf{X}_t'''\right\}_{t\in\mathbb{R}_+}$ at the end of Sect. 13.4.2 is uniform over bounded subsets of the real line and can be found in Theorem 4.4 in Kurtz (1978) and in Theorem 3.2 and following remarks in Chap. 11 in Ethier and Kurtz (1986).

**Exercise 10 (News Diffusion Model—Continued).** Reasoning as in Exercise 8, we have:

$$\sum_{\ell} |\ell|^2 \cdot \sup_{x\in K} \beta_\ell(x) \leq \frac{p}{4} < \infty.$$

As concerns $\partial f(x) = p \cdot (1 - 2x)$ and $\beta_1(x) = p \cdot x(1 - x)$, they are clearly continuous. □

## References


Allain, M.-F. (1976). Approximation par un processus de diffusion des oscillations, autour d'une valeur moyenne, d'une suite de processus de Markov de saut pur. *Comptes rendus de l'Académie des sciences Paris Séries A-B, 282*(16), Aiii, A891–A894.

Allain, M.-F. (1976). Étude de la vitesse de convergence d'une suite de processus de Markov de saut pur. *Comptes rendus de l'Académie des sciences Paris Séries A-B, 282*(17), Aiii, A1015–A1018.

Alm, S. E. (1978). On the rate of convergence in diffusion approximation of jump Markov processes. Report 8, Uppsala University, Department of Mathematics.

Axelrod, R. (1986). An evolutionary approach to norms. *American Political Science Review, 80*(4), 1095–1111.

Barbour, A. D. (1972). The principle of the diffusion of arbitrary constants. *Journal of Applied Probability, 9*, 519–541.

Barbour, A. D. (1974). On a functional central limit theorem for Markov population processes. *Advances in Applied Probability, 6*, 21–39.

Black, A. J., & McKane, A. J. (2012). Stochastic formulation of ecological models and their applications. *Trends in Ecology & Evolution, 27*(6), 337–345.

Bortolussi, L., Hillston, J., Latella, D., & M. Massink (2013). Continuous approximation of collective system behaviour: A tutorial. *Performance Evaluation, 70*, 317–349.

Centola, D. (2010). The spread of behavior in an online social network experiment. *Science, 329*(5996), 1194–1197.

Challenger, J. D., Fanelli, D., & McKane, A. J. (2014). The theory of individual based discrete-time processes. *The Journal of Statistical Physics, 156*(1), 131–155.

Collet, F., Dai Pra, P., & Sartori, E. (2010). A simple mean field model for social interactions: Dynamics, fluctuations, criticality. *The Journal of Statistical Physics, 139*(5), 820–858.

Ethier, S. N., & Kurtz, T. G. (1986). *Markov processes. Characterization and convergence.* Wiley Series in Probability and Mathematical Statistics: Probability and Mathematical Statistics. New York: John Wiley & Sons, Inc.

Feller, W. (1951). Diffusion processes in genetics. In J. Neyman (Ed.), *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability* (pp. 227–246). Berkeley: University of California Press.

Galán, J. M., & Izquierdo, L. R. (2005). Appearances can be deceiving: Lessons learned re-implementing Axelrod's 'evolutionary approach to norms'. *Journal of Artificial Societies and Social Simulation, 8*(3), 2.

Goutsias, J., & Jenkinson G. (2013). Markovian dynamics on complex reaction networks. *Physics Reports, 529*(2), 199–264.

Hirshman, B. R., Charles, J. St., & Carley, K. M. (2011). Leaving us in tiers: can homophily be used to generate tiering effects? *Computational & Mathematical Organization Theory, 17*(4), 318–343.

Karlin, S., & Taylor, H. M. (1975). *A first course in stochastic processes*, 2nd ed. New York/London: Academic Press.

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

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

Kurtz, T. G. (1972). The relationship between stochastic and deterministic models for chemical reactions. *The Journal of Chemical Physics, 57*(7), 2976–2978.

Kurtz, T. G. (1976). Limit theorems and diffusion approximations for density dependent Markov chains. *Mathematical Programming Studies, 5*, 67–78. Stochastic systems: Modeling, identification and optimization, I (Proc. Sympos., Univ. Kentucky, Lexington).

Kurtz, T. G. (1977/1978). Strong approximation theorems for density dependent Markov chains. *Stochastic Processes and their Applications, 6*(3), 223–240.

Kurtz, T. G. (1980). Relationships between stochastic and deterministic population models. In *Biological growth and spread (Proc. Conf., Heidelberg, 1979)*. Lecture Notes in Biomathematics (Vol. 38, pp. 449–467). Berlin/New York: Springer.

Kurtz, T. G. (1981). *Approximation of population processes*. CBMS-NSF Regional Conference Series in Applied Mathematics (Vol. 36). Philadelphia: Society for Industrial and Applied Mathematics (SIAM).

Kurtz, T. G. (1983). Gaussian approximations for Markov chains and counting processes. In *Proceedings of the 44th session of the International Statistical Institute, Vol. 1 (Madrid, 1983)* (Vol. 50, pp. 361–376). With a discussion in Vol. 3, pp. 237–248.

Lotka, A. J. (1925). *Elements of Physical Biology*. Baltimore: Williams & Wilkins Company.

Matis, J. H., & Kiffe, T. R. (2000). *Stochastic Population Models: A Compartmental Perspective*. Lecture Notes in Statistics. New York: Springer.

Norman, M. F. (1968). Slow learning. *British Journal of Mathematical and Statistical Psychology, 21*, 141–159.

Norman, M. F. (1972). *Markov processes and learning models*. Mathematics in Science and Engineering (Vol. 84). New York/London: Academic Press.

Norman, M. F. (1974a). A central limit theorem for Markov processes that move by small steps. *The Annals of Probability, 2*, 1065–1074.

Norman, M. F. (1974b). Markovian learning processes. *SIAM Review, 16*, 143–162.

Plikynas, D., & Masteika, S. (2014). Agent-based nonlocal social systems: Neurodynamic oscillations approach. In G. Meiselwitz (Ed.), *Social Computing and Social Media. Lecture Notes in Computer Science* (Vol. 8531, pp. 253–264). New York: Springer International Publishing.

Pollett, P. K. (2001). Diffusion approximations for ecological models. In F. Ghassemi, P. Whetton, R. Little & M. Littleboy (Eds.), *MODSIM 2001 International Congress on Modelling and Simulation* (pp. 843–848). Townsville: Modelling and Simulation Society of Australia and New Zealand Inc.

Railsback, S. F., & Grimm, V. (2011). *Agent-based and individual-based modeling: A practical introduction*. Princeton: Princeton University Press.

Sigmund, K. (2007). Kolmogorov and population dynamics. In *Kolmogorov's heritage in mathematics* (pp. 177–186). Berlin: Springer.

Volterra, V. (1931). *Leçons sur la théorie mathématique de la lutte pour la vie*. Cahiers scientifiques, Fascicule VII. Paris: Gauthier-Villars.

Volterra, V. (1962). *Opere matematiche. Memorie e Note. Volume quinto: 1926–1940*. Accademia nazionale dei Lincei.

Wang, X., Tao, H., Xie, Z., & Yi, D. (2013). Mining social networks using wave propagation. *Computational & Mathematical Organization Theory, 19*(4), 569–579.

Zhang, Y., & Wu, Y. (2012). How behaviors spread in dynamic social networks. *Computational & Mathematical Organization Theory, 18*(4), 419–444.

# Index





D. Secchi, M. Neumann (eds.), *Agent-Based Simulation of Organizational Behavior*, DOI 10.1007/978-3-319-18153-0