mathematics

Article

Stochastic Growth Models for the Spreading of Fake News

Antonio Di Crescenzo *,†, Paola Paraggio † and Serena Spina †

Dipartimento di Matematica, Università di Salerno, Via Giovanni Paolo II n.132, I-84084 Fisciano, SA, Italy; pparaggio@unisa.it (P.P.); sspina@unisa.it (S.S.)
* Correspondence: adicrescenzo@unisa.it
† These authors contributed equally to this work.

Abstract: The propagation of fake news in online social networks nowadays is becoming a critical issue. Consequently, many mathematical models have been proposed to mimic the related time evolution. In this work, we first consider a deterministic model that describes rumor propagation and can be viewed as an extended logistic model. In particular, we analyze the main features of the growth curve, such as the limit behavior, the inflection point, and the threshold-crossing-time, through fixed boundaries. Then, in order to study the stochastic counterparts of the model, we consider two different stochastic processes: a time non-homogeneous linear pure birth process and a lognormal diffusion process. The conditions under which the means of the processes are identical to the deterministic curve are discussed. The first-passage-time problem is also investigated both for the birth process and the lognormal diffusion process. Finally, in order to study the variability of the stochastic processes introduced so far, we perform a comparison between their variances.

Keywords: fake news; rumor propagation; growth model; birth processes; diffusion processes; first-passage-time

MSC: 60J85; 60J70

check for updates

Citation: Di Crescenzo, A.; Paraggio, P.; Spina, S. Stochastic Growth Models for the Spreading of Fake News. *Mathematics* **2023**, *11*, 3597. https://doi.org/10.3390/math11163597

Academic Editors: Alexander Zeifman, Victor Korolev and Alexander Sipin

Received: 26 July 2023
Revised: 16 August 2023
Accepted: 17 August 2023
Published: 19 August 2023

1. Introduction

Several types of real growth dynamics can be described by mathematical models. The most simple model is the Malthusian one in which the population size grows in time as an exponential function. Clearly, this model, in some instances, turns out not to be fully appropriate since it possesses an infinite limit value. Indeed, for long-term growth, it is necessary to take into account factors which can slow down or speed up the growth rate of the population. Aiming to describe these real situations, one can refer to so-called sigmoidal growth models, characterized by an initial slow growth followed by an explosion of an exponential-type, which finally flattens out to an equilibrium status, known as the carrying capacity.

Over the years, several sigmoidal curves have been introduced, such as Gompertz (see Tan [1]), Korf (introduced for the first time in Korf [2]), logistic (see, for instance, Di Crescenzo and Paraggio [3]), Bertalanffy–Richards (Richards [4]), and other generalizations of already existing models (as in Asadi et al. [5], Di Crescenzo and Spina [6], Di Crescenzo et al. [7,8]). The fields of possible applications of sigmoidal models are various and range from software reliability (see Erto et al. [9]) to biology (as in Brauer and Castillo-Chavez [10]) and economics (see, for example, Smirnov and Wang [11]).

Recently, S-shaped models have been used to model the spread of rumors in online social networks (see San Martìn et al. [12]). The attention stimulated by this topic arises from the global increase in social network usage and the ease of sending messages instantaneously. The growth of instantaneous communication has proved to be a fertile ground for the spread of fake news (see De Martino and Spina [13], Giorno and Spina [14], Figueira and Oliveira [15]).

By the term 'fake news', we generally mean false or misleading information. The techniques used to generate fake news have been the subject of many investigations, such as research reports by the RAND Corporation. Accordingly, in a broad sense, fake news can be characterized as follows:

(i) Fabrication: the invention of entirely false or misleading information,
(ii) Misappropriation: the misrepresentation of existing facts, events and people,
(iii) Deceptive identities: the employment of misleading source of information,
(iv) Obfuscation: the offering of multiple and contradictory accounts for the same event in order to confuse the audience,
(v) Conspiracy theories: the proposal of conspiracy plots related to real events/phenomena,
(vi) Selective use of facts: the selection of information in a manipulative way,
(vii) Rhetorical fallacies: reasoning which is logically invalid but cognitively effective,
(viii) Appeals to emotion/authority: the use of messages which elicit emotions.

The diffusion of fake news, whether intentional or accidental, causes disinformation which may be used for various aims, such as influencing public opinion, instigating hatred, damaging the image of particular states/companies, etc. Hence, studying the propagation of rumors and disinformation may be of great interest in order to design countermeasures and avoid potential impacts on society. The urgency of the matter has prompted various states, and also private companies, to invest in cybersecurity. In this context, a great deal of scientific research effort has been invested, especially in relation to the proposal of stochastic models with effective predictive capabilities (see, for instance, Abraham and Nair [16], Abimbola et al. [17], Paul and Zhang [18], Alandihallaj et al. [19]). In addition, the further need to optimize investments in cybersecurity also arises. This topic has been addressed in the work of Miaoui and Boudriga [20]. In particular, these authors propose a model that optimizes enterprise investments in cybersecurity using expected utility theory.

The development of suitable mathematical models capable of simulating the propagation of rumors is potentially of considerable value in making strategic choices (see, for example, Mahmoud [21], Kapsikar et al. [22], Ben Aissa et al. [23]). However, the development of stochastic models relating to the spread of information is not an easy task. Indeed, as noted in Raponi et al. [24], the propagation of fake news is a complex phenomenon influenced by several factors the identification and assessment of which is challenging. To overcome this difficulty, many models have been proposed in the literature that have been inspired by epidemiological models. However, although the two contexts have various similarities, the dissemination of news follows different rules than the diffusion of contagious diseases. Thus, a variety of models in stochastic environments has been developed which emphasize different aspects of interest. For instance, Esmaeeli and Sajadi [25] developed a sceptical rumor model for individuals located on a non-negative integer line, whereas a case illustrating the spread of fake news in a community of finite size was considered by Mahmoud [21]. Moreover, recent developments in stochastic rumor propagation modeling can be found in Jia and Cao [26] and Roy and Saha [27].

Nevertheless, although many attempts have been made, a model that includes all the properties of the real fake-news propagation phenomenon has not yet been reported. Bearing in mind the complexity of the problems connected to the phenomenon, the aim of this paper is twofold: (i) to study and enhance a recent growth model for rumor propagation, (ii) to build and study two stochastic processes that are able to describe the growth model itself in the presence of random fluctuations. In contrast to stochastic models treated in [21,25–27], which are based on spatial dynamics on suitable state-spaces and depend on network topologies, for analysis of point (ii), we focus on time-inhomogeneous settings involving suitable birth–death and diffusion processes whose means are identical to the growth model considered in point (i).

Usually, growth models are described by means of differential equations. In order to make them more realistic, it is possible to introduce a noise term, summarizing random fluctuations, in the differential equations (see Øksendal [28]) and to consider the resulting stochastic differential equations (as reported in Román-Román et al. [29] and Di Crescenzo et al. [7]). Other investigations have proposed introducing a random environment by considering special birth–death processes using an expected value which corresponds to the deterministic growth function (see Di Crescenzo and Spina [6], Di Crescenzo and Paraggio [3], Giorno and Nobile [30], Ricciardi [31]).

Hence, in the present work, we analyze both of the strategies stimulated by the above mentioned research lines.

A key concern with regard to disinformation and fake news that has recently emerged is the need for reassurance on the validity and quality of the news in the face of new pitfalls that can arise by use of the Web and from the use of artificial intelligence (AI). Major efforts are, therefore, needed to create a stable alliance between all stakeholders to promote, by any means, communication and awareness-raising activities aimed at all users so that they are able to recognize bad information and protect themselves from the dangers that can arise from it. Therefore, tools pertaining to AI can be fruitfully used to support the collective efforts of relevant institutions, web companies and communication professionals which are called upon to implement clear and shared actions to counter disinformation and the spread of fake news. Hence, identifying and extracting the most appropriate and significant features from information flows is one of the biggest challenges for AI-based detection. In this area, examples of recent contributions related to feature extraction and anomaly detection can be found in Khan et al. [32] and Arunnehru et al. [33].

We focus on the growth model proposed by San Martin et al. [12] for rumor propagation, postponing consideration of AI-based strategies for future work.

In this paper, our investigation is described along the following novel lines:

(i) analyzing some limit behaviors of the growth function, which is shown to be a suitable extension of the logistic curve,
(ii) studying the corresponding mean time in which a randomly chosen individual is reached by the rumor,
(iii) determining the initial specific growth rate, the inflection point, and other related quantities,
(iv) conducting a sensitivity analysis based on the perturbation on the parameters of the model,
(v) addressing the related threshold-crossing problem,
(vi) studying and comparing two different stochastic counterparts for the model based on suitable time-inhomogenous Markov processes, i.e., a linear birth–death process and a lognormal diffusion process.

Specifically, we provide conditions such that the mean values of these stochastic processes are identical to the growth curve, which allows modeling of the diffusion of rumors in the presence of random fluctuations. Moreover, we provide explicit expressions for various quantities of interest in applications, such as the conditional mean, the conditional variance, the index of dispersion, and the Fano factor. Finally, in order to investigate the variability of the two considered stochastic processes, we perform a comparison of their variances.

1.1. Relation with Epidemiological Models

Usually, propagation models for rumors are very similar to those used for the spread of infectious diseases. One of the best-known epidemiological models is the SI model, according to which the population is divided into two categories, i.e., susceptible and infected. In this case, the resulting growth curve describing the time evolution of those infected follows an exponential trend. In other recent studies, various generalizations of compartmental models have been introduced. For example, in Jin et al. [34] the authors consider the population to be divided into susceptible, exposed, infected (i.e., reached by the rumor) and skeptics (SEIZ model). Similarly, in [14] the population is divided into three classes: ignorant, spreader and stifler. In all the aforementioned studies, the model which mimics the dynamics of rumor spread is represented by a system of differential equations (one equation for any compartment). In other studies, researchers have focused on analysis of the time evolution of a rumor among the population by considering only one of the compartments into which the population is divided. This is the case as reported in [13], where the authors consider a differential equation describing the time evolution of infected individuals. In particular, the considered differential equation is a logistic one with a time-dependent growth rate. In the present paper, we consider an existing model which represents the fraction of the population reached by the rumor. This growth model embodies both the exponential and the logistic one. Indeed, they can be recovered by considering the special limit values of the parameters (cf. Section 2 below). Clearly, to have a more realistic representation of the rumor spread among individuals in a population, it may be interesting to consider a suitable compartmental model and its stochastic counterpart. This topic may be the subject of future investigations.

1.2. Plan of the Paper

The paper is organized in detail as follows: In Section 2, we study the main features of the deterministic model introduced in [12], such as the carrying capacity and the inflection point. A sensitivity analysis based on perturbation of the parameters and a study related to the problem of the first-crossing-time of the special threshold are also performed. Moreover, we analyze the expected time in which a randomly chosen individual is reached by the rumor. Then, in Section 3, we define a special time-inhomogeneous linear pure birth process having a mean which corresponds to the deterministic curve of Section 2. For this process, the transition probabilities, the moment-generating function, the variance, and some indexes of dispersion, are also determined. The first-passage-time problem of the pure birth process through constant boundaries is also addressed. Section 4 is devoted to description of a special lognormal diffusion process having the same mean as the pure birth process introduced in Section 3. The moments, the mode, and the quantiles of this process are also provided in closed form. Moreover, we study the first-passage-time problem by considering particular time-dependent boundaries in order to obtain an explicit expression for the corresponding probability density function. Finally, in order to provide a comparison between the stochastic processes introduced previously, since they possess the same mean, we investigate the ratio between their variances.

2. The Deterministic Model

In San Martín et al. [12], the authors propose a novel mathematical model to represent the spread of rumors in online social networks. It is assumed that the individuals are linked either by person-to-person relations or by belonging to the same group. Moreover, in the considered model, the population is divided into four categories: burned, sender, receiver and seed.

(i) A burned individual is defined as a person who knows the rumor. Note that when an individual becomes burned, he/she remains in this state until the end of rumor propagation.
(ii) A sender is a person who knows the message and texts it to his/her contacts.
(iii) A receiver is an individual who is reached by the rumor.
(iv) Finally, the seed is the set of all the individuals who know the rumor at the beginning of its propagation.

Note that the aforementioned categories are not mutually disjoint. For example, a sender is also a burned individual and a receiver may be also be an already burned individual.

In this work, we focus our attention on the function $F(t)$ that represents the fraction of burned individuals at the time $t \geq 0$. According to Equation (3) of [12] with $a = 1$ and $b = 0$, it is defined as follows

$$F(t) = \frac{C\exp((1+\varepsilon)t) - 1}{C\exp((1+\varepsilon)t) + \frac{2\varepsilon}{1-\varepsilon}}, \qquad t \geq 0, \tag{1}$$

where $C \geq 1$ and $-1 < \varepsilon < 1$. The positions performed in Equation (3) of [12] correspond to consider the original time scale ($a = 1$) and take 0 as the time origin ($b = 0$). According to the assumption specified in (i), from Equation (1), we have that the function $F(t)$ is monotone non-decreasing in $t \geq 0$, and satisfies $0 \leq F(t) \leq 1$ for all $t \geq 0$.

Remark 1. *From Equation (1), we have that the carrying capacity of the population, obtained as* $\lim_{t\to\infty} F(t)$*, is equal to 1. In other terms, the percentage of the population that will eventually be informed about the rumor is unity.*

Moreover, $F(t)$ is the solution of the following differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}F(t) = 2\varepsilon\left(\frac{1-\varepsilon}{2\varepsilon} + F(t)\right)(1 - F(t)), \qquad t \geq 0, \tag{2}$$

with the initial condition due to Equation (1) for $t = 0$, i.e.,

$$F(0) = \frac{C-1}{C + \frac{2\varepsilon}{1-\varepsilon}} \geq 0. \tag{3}$$

From the above formulas, we have that the parameter ε is involved both in the differential Equation (2) and the initial condition (3), whereas the parameter C is only linked to the initial size of the burned individuals $F(0)$. Clearly, a large value of C, or ε close to -1, correspond to a large initial size of burned individuals. Figure 1 illustrates the behavior of $F(0)$, which is decreasing in ε and increasing in C.

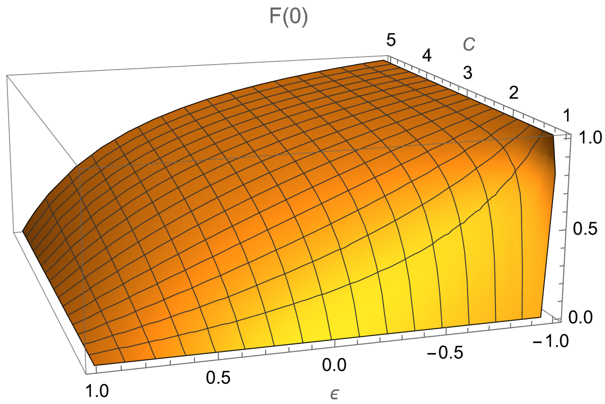

Figure 1. $F(0)$ given in (3) as a function of ε and C.

The parameters $\varepsilon \in (-1, 1)$ and $C \geq 1$ allow obtaining different kinds of growth, including the limit cases $\varepsilon \to -1$ and $\varepsilon \to 1$. Let us now examine some features.

Case no. 1: $\varepsilon \to -1$

If $\varepsilon \to -1$, then Equations (2) and (3) yield the differential problem

$$\begin{cases} \frac{\mathrm{d}}{\mathrm{d}t}F(t) = 2(1 - F(t))^2, & t \geq 0, \\ F(0) = 1 \end{cases}$$

with trivial constant solution

$$F(t) = 1, \qquad t \geq 0.$$

In this case, all the individuals already know the rumor since the beginning.

Case no. 2: $\varepsilon \to 0$

When $\varepsilon \to 0$, from (2) and (3), we have

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}F(t) = 1 - F(t), & t \geq 0, \\ F(0) = \frac{C-1}{C} \end{cases}$$

so that

$$F(t) = 1 - \frac{1}{C}e^{-t}, \qquad t \geq 0. \tag{4}$$

In this case, the spread of the rumor among the individuals grows in an (increasingly concave) exponential way.

Case no. 3: $\varepsilon \to 1$

If $\varepsilon \to 1$, then Equations (2) and (3) give the problem

$$\begin{cases} \dfrac{\mathrm{d}}{\mathrm{d}t}F(t) = 2(1 - F(t))F(t), & t \geq 0, \\ F(0) = 0 \end{cases}$$

which corresponds to the logistic model starting with a vanishing initial solution (for instance, cf. Remark 2.1 of Albano et al. [35]). In this case, the solution is trivially

$$F(t) = 0, \qquad t \geq 0,$$

so that, if no individuals know the rumor at the beginning, then it does not spread in the population.

Case no. 4: $C = 1$

If $C = 1$ then $F(0) = 0$, so that no individuals know the rumor at the beginning. However, in this case, Equation (1) is still a non-trivial function of t. Indeed, for $\varepsilon \in (-1, 1)$, the rumor can spread among the population, whereas when $\varepsilon \to 1$, the rumor cannot spread anymore. Moreover, for $C = 1$, from Equations (2) and (3), we have $\frac{d}{dt}F(t)\big|_{t=0} = 1 - \varepsilon$, so that, in this case, ε can be viewed as a reversed measure of the initial intensity of rumor spreading.

In Figure 2, some plots of the function $F(t)$ are provided for different choices of both the parameters C and ε. The first frame, for $C = 1$, confirms the remarks stated in Case no. 4, in particular that the initial intensity of rumor spreading is decreasing for $\varepsilon \in (-1, 1)$. This behavior is exhibited also for $C = 2$, as shown in in the second frame of Figure 2, where it is also evident that $F(0)$ is decreasing in ε due to (3).

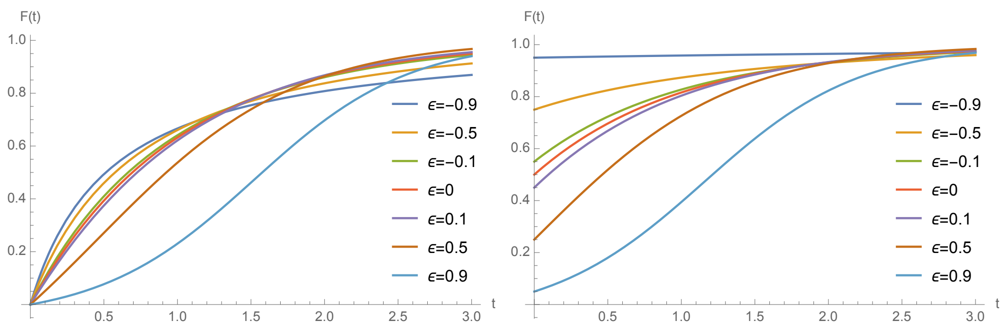

Figure 2. The function $F(t)$ given in (1) for $C = 1$ (**left**) and $C = 2$ (**right**), for various choices of ε.

Concerning the complexity of the model, it is not hard to see that the function (1) is $O(g(t))$, where $g(t) = 1 - \frac{1}{C}e^{-t}$, $t \geq 0$, corresponds to the growth function obtained in Equation (4) in the limit as $\varepsilon \to 0$.

Remark 2. *It is worth noting that $F(t)$ may be viewed as the distribution function of a random variable, say B, having support $[0, +\infty)$, which describes the instant in which an individual randomly chosen in the population is reached by the rumor. Clearly, due to (1) and (3), one has that B is absolutely continuous if, and only if, $C = 1$, whereas for $C > 1$, it is a mixed random variable, with an atom at 0 (corresponding to the individuals informed at time 0). The corresponding mean $\mathsf{E}(B)$ represents the expected time in which a randomly chosen individual is reached by the rumor, with*

$$\mathsf{E}(B) = \int_0^{+\infty} \mathsf{P}(B > t)\mathrm{d}t = \int_0^{+\infty} [1 - F(t)]\mathrm{d}t$$
$$= \begin{cases} \dfrac{1}{2\varepsilon} \log\left(1 + \dfrac{2\varepsilon}{C(1-\varepsilon)}\right), & \text{if } \varepsilon \in (-1, 0) \cup (0, 1), \\ \dfrac{1}{C}, & \text{if } \varepsilon = 0. \end{cases}$$

Note that $\mathsf{E}(B)$ is decreasing with respect to C; moreover, for $C = 1$, it is an even convex function in ε. These comments are confirmed by Figure 3, which shows some plots of $\mathsf{E}(B)$.

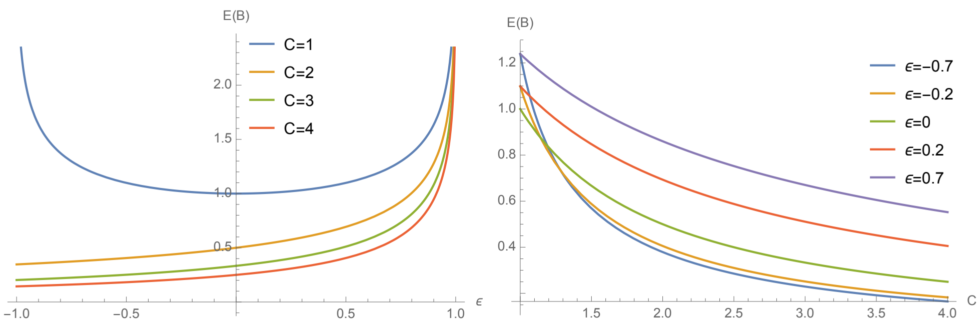

Figure 3. The expected value $\mathsf{E}(B)$ for various choices of C (**left**) and ε (**right**).

2.1. A Different Formulation

It has been pointed out in various recent investigations on population dynamics that problems concerning differential equations of the form (2) can, in some instances, be expressed in a different way (cf. [3,6–8]). In this vein, the present section is devoted to the determination of a different formulation of Equation (2), introducing a time-dependent growth rate. In more detail, in our case, the function (1) can also be viewed as a solution of the following differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}F(t) = \xi(t)F(t), \qquad t \geq 0, \tag{5}$$

where the time-dependent growth rate $\xi(t)$ has the following expression

$$\xi(t) = \frac{Ce^{(1+\varepsilon)t}(1+\varepsilon)^2}{(Ce^{(1+\varepsilon)t} - 1)(2\varepsilon + (1-\varepsilon)Ce^{(1+\varepsilon)t})}, \qquad t \geq 0. \tag{6}$$

We note that the growth rate (6) is decreasing in $t \geq 0$; this implies that the intensity of information spread gradually fades over time. Clearly, since the differential Equation (5) is a reformulation of the same problem described by Equations (2) and (3), the solution $F(t)$ is still expressed by (1).

Let us now introduce the function

$$f(\varepsilon) := \frac{C(1-\varepsilon)(1+\varepsilon)^2}{(C + 2\varepsilon - C\varepsilon)^2} \equiv \frac{\mathrm{d}}{\mathrm{d}t}F(t)\Big|_{t=0} = \xi(0)F(0). \tag{7}$$

This function can be viewed as the initial specific growth rate, since it describes the slope of the tangent to che curve $F(t)$ for $t = 0$. With respect to its behavior as a function of $\varepsilon \in (-1, 1)$, from (7), we have two cases:

(a) if $C = 1$, and, thus, $F(0) = 0$, then, $f(\varepsilon)$ is linearly decreasing, being $f(\varepsilon) = 1 - \varepsilon$; in this case, even if the initial size of burned individuals is 0, the population of burned individuals grows; such growth is more pronounced for small values of ε;

(b) if $C > 1$, and, thus, $F(0) > 0$, then, $f(\varepsilon)$ is not monotonic; indeed, $f(\varepsilon)$ is increasing for $\varepsilon \in (-1, \varepsilon_M)$ and is decreasing for $\varepsilon \in (\varepsilon_M, 1)$, where

$$\varepsilon_M = \begin{cases} \dfrac{-1 + 2C - \sqrt{C^2 + 6C - 7}}{C - 2}, & \text{if } C \in (1, 2) \cup (2, +\infty), \\ \dfrac{1}{3}, & \text{if } C = 2; \end{cases}$$

hence, at the beginning of the rumor propagation, the speed of the growth of burned individuals is increasing for small values of ε and decreasing for larger ε. See Figure 4 for some instances of the function $f(\varepsilon)$.

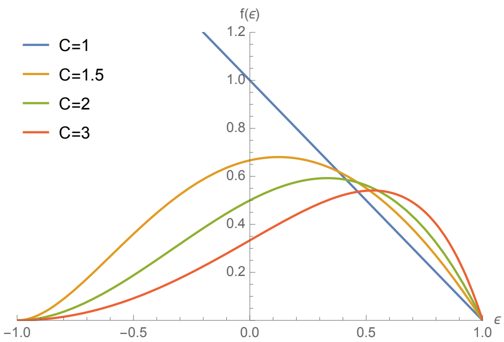

Figure 4. The function $f(\varepsilon)$, given in (7), for various choices of C.

2.2. The Inflection Point

Let us now focus on the inflection point of the function $F(t)$. In the context of population dynamics, the analysis of the inflection time is of great interest, since this point represents the instant at which the growth rate of the function is maximum. From (1), by evaluating the second derivative of $F(t)$, one obtains the following results. For $-1 < \varepsilon \leq 0$, the function $F(t)$ has downward concavity for any $t \geq 0$, whereas, if $0 < \varepsilon < 1$, the curve $F(t)$ is sigmoidal with an inflection point at

$$t_F = \frac{1}{1+\varepsilon} \log\left(\frac{2\varepsilon}{C(1-\varepsilon)}\right). \tag{8}$$

Note that $t_F \geq 0$ if, and only if, $\varepsilon > C/(2 + C)$. In this case, the size of the population at the inflection point t_F is given by

$$F(t_F) = \frac{3}{4} - \frac{1}{4\varepsilon} \geq 0.$$

As performed in similar studies, one can approximate the behavior of the curve in proximity to the inflection point by the tangent line. With this aim, let us now compute the maximum specific growth rate μ, which is defined as the slope of the tangent to the curve $F(t)$ at the inflection point t_F. Hence, if $\varepsilon > C/(2 + C)$, one has

$$\mu := \frac{\mathrm{d}}{\mathrm{d}t} F(t)\Big|_{t=t_F} = \frac{(1+\varepsilon)^2}{8\varepsilon},$$

so that the line tangent to the curve $F(t)$ at the inflection point has the following expression:

$$y = \frac{1}{8\varepsilon}\left[(1+\varepsilon)^2 t - 2(1-3\varepsilon) - (1+\varepsilon)\log\left(\frac{2\varepsilon}{C(1-\varepsilon)}\right)\right].$$

As a consequence, the lag time λ, which denotes the intercept with the t-axis of this tangent, is given by

$$\lambda = \frac{2(1-3\varepsilon)}{(1+\varepsilon)^2} + \frac{1}{1+\varepsilon}\log\left(\frac{2\varepsilon}{C(1-\varepsilon)}\right).$$

We note that the lag time represents the initial time of an ideal growth curve, which increases linearly at a constant rate given by the maximum specific growth rate and reaches the same size of the original population at the inflection point. In Figure 5, we show the function $F(t)$ and the corresponding tangent at the inflection point t_F for various choices of the parameters.

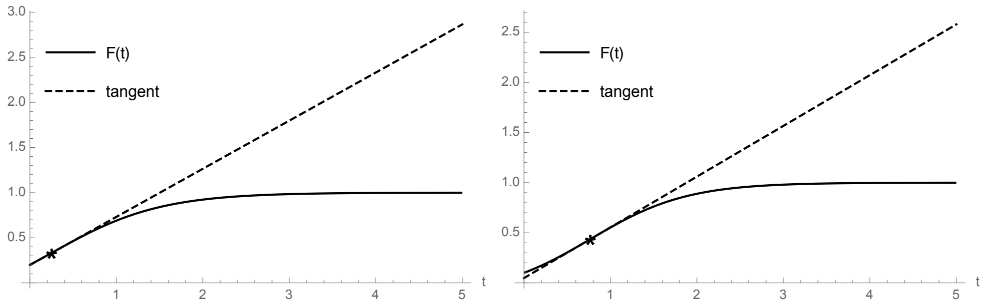

Figure 5. The function $F(t)$, with the line tangent to the curve at the inflection point for $C = 2$, $\varepsilon = 0.6$ (**left**) and $\varepsilon = 0.8$ (**right**). The inflection point is shown by the star.

2.3. Sensitivity Analysis

In this section, we analyze how the perturbations on the parameters $C > 1$ and $-1 < \varepsilon < 1$ involved in the model influence the behaviour of $F(t)$. In the following, we denote this function by F^ν to emphasize the dependence on a generic parameter ν. Specifically, starting from Equation (1), we expand $F^{\nu+\eta}$ in a Taylor series evaluated at ν, with $\eta > 0$, for $\nu = C$ and $\nu = \varepsilon$.

- Perturbation on C

$$F^{C+\eta} - F^C \approx \eta\,\frac{(1-\varepsilon^2)\,e^{(1+\varepsilon)t}}{\left[2\varepsilon + (1-\varepsilon)Ce^{(1+\varepsilon)t}\right]^2}.$$

The latter term is positive for all $t \geq 0$.

- Perturbation on ε

$$F^{\varepsilon+\eta} - F^\varepsilon \approx \eta\,\frac{2 - Ce^{(1+\varepsilon)t}[2 + (\varepsilon^2-1)t]}{\left[2\varepsilon + (1-\varepsilon)Ce^{(1+\varepsilon)t}\right]^2}. \tag{9}$$

The sign of the right-hand side of (9) is equal to that of the function

$$h(t) := 2 - Ce^{(1+\varepsilon)t}[2 + (\varepsilon^2-1)t], \qquad t \geq 0.$$

For $C > 1$, we have $h(0) = 2(1-C) < 0$, the first derivative $h'(t)$ is negative for all $t > 0$, and $\lim_{t\to+\infty} h(t) = -\infty$; therefore, from the continuity of h, we have $h(t) < 0$ for all $t \geq 0$. Hence, the right-hand side of (9) is negative for all $t \geq 0$.

As an example, in Figure 6, we show the curve $F(t)$ and the effect of the perturbation $\eta = 0.1$ on the parameter C (on the left) and on ε (on the right).

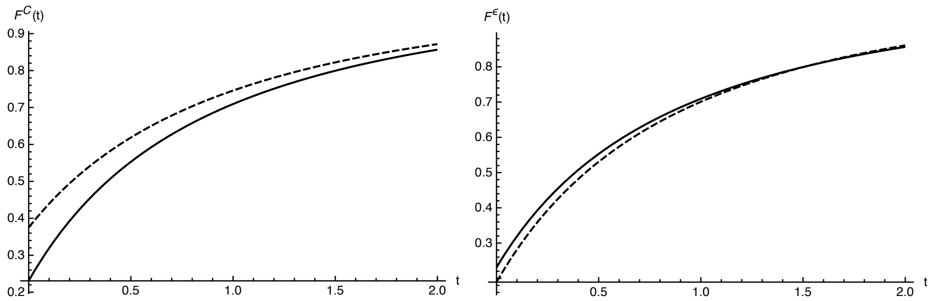

Figure 6. On the **left**, the curve $F(t)$ for $C = 1.1$ (solid) and $C = 1.2$ (dashed), with $\varepsilon = -0.5$. On the **right**, $F(t)$ is plotted for $\varepsilon = -0.5$ (solid) and $\varepsilon = -0.4$ (dashed), with $C = 1.1$.

2.4. Threshold Crossing Time

In many real contexts, it is of interest to study the time spent by the growth function below (or above) a specific threshold. Indeed, such a boundary may represent a critical value related to the dynamics of the modeled population. Hence, in this section, we focus on analysis of the first-crossing-time of the function $F(t)$ through a specific constant boundary representing a percentage $p \in (0, 1)$ of the whole population. In more detail, we consider θ defined as follows

$$\theta = \theta(p) := \inf\{t \geq 0 : F(t) = p\}. \tag{10}$$

Hence, by recalling (1), the equation $F(\theta) = p$ yields

$$\begin{aligned}\theta &= \frac{1}{1+\varepsilon} \log\left(\frac{1-\varepsilon+2p\varepsilon}{(1-\varepsilon)C(1-p)}\right) \\ &= t_F + \frac{1}{1+\varepsilon} \log\left(\frac{1-\varepsilon+2p\varepsilon}{2\varepsilon(1-p)}\right),\end{aligned}$$

where the last equality follows from (8). By Equation (1), we have that $F(t)$ is a continuous function. Hence, due to Remark 1, the existence of θ is guaranteed if p is larger than the initial proportion of burned individuals. Indeed, recalling (3), it follows that $\theta > 0$ if, and only if,

$$p > F(0) \equiv \frac{C-1}{C + \frac{2\varepsilon}{1-\varepsilon}} \geq 0,$$

and $\theta > t_F$ if, and only if, $p > \frac{3\varepsilon-1}{4\varepsilon}$. In Figure 7, the first-crossing-time θ is shown for different choices of the parameters. Note that θ is increasing both with respect to p and ε, with $\lim_{p\to 1} \theta(p) = +\infty$ and $\lim_{\varepsilon\to 1} \theta(p) = +\infty$.

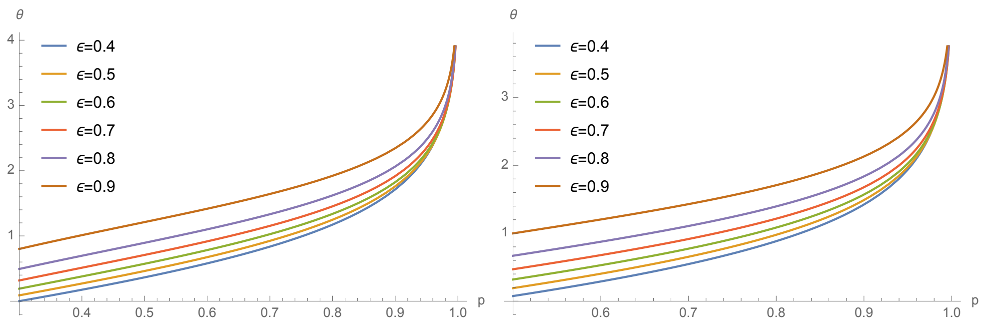

Figure 7. The first-crossing-time θ as a function of p for (**left**) $C = 2$, (**right**) $C = 3$.

The determination of $\theta(p)$ deserves attention since it represents the time required for the information to reach the percentage p of the population. In this respect, it is useful to investigate some cases for given choices of p. Due to (10), one has

$$\theta(p_2) - \theta(p_1) = \frac{1}{1+\varepsilon} \log \frac{(1-p_1)(1-\varepsilon+2p_2\varepsilon)}{(1-p_2)(1-\varepsilon+2p_1\varepsilon)},$$

so that, for instance,

$$\frac{\theta(3/4) - \theta(1/2)}{\theta(1/2) - \theta(1/4)} = \frac{\log(2+\varepsilon)}{\log(3/(2-\varepsilon))} > 1.$$

This implies that, whatever the value of ε, the time required for increasing the informed percentage from $1/2$ to $3/4$ is greater to that required for increasing the same from $1/4$ to $1/2$.

Remark 3. *We recall that the growth function $F(t)$, given in (1), represents the fraction of burned individuals in the population. Hence, assuming that the population size is $N > 1$, due to (1), the function*

$$\widetilde{F}(t) := N \cdot F(t) = N \cdot \frac{C \exp((1+\varepsilon)t) - 1}{C \exp((1+\varepsilon)t) + \frac{2\varepsilon}{1-\varepsilon}}, \qquad t \geq 0, \tag{11}$$

with $C \geq 1$ and $-1 < \varepsilon < 1$, denotes the (approximated) total number of burned individuals at time t. Making use of Equation (5), it is easy to see that the function $\widetilde{F}(t)$ satisfies the following Malthusian-type equation:

$$\frac{\mathrm{d}}{\mathrm{d}t}\widetilde{F}(t) = \xi(t)\widetilde{F}(t), \quad t \geq 0, \qquad \widetilde{F}(0) = N \cdot F(0), \tag{12}$$

where $\xi(t)$ is given in (6). Due to (11), the main difference between $F(t)$ and $\widetilde{F}(t)$ lays in the carrying capacity. Indeed, the carrying capacity for $F(t)$ is 1, whereas the carrying capacity for $\widetilde{F}(t)$ is equal to the population size N. Since N can be quite large, this will allow us to consider stochastic processes with infinite state-space as a stochastic counterpart of the considered growth model, as specified in Sections 3 and 4 below.

3. A Special Time-Inhomogeneous Linear Pure Birth Process

The introduction of stochasticity in growth equations can be performed in several ways. A classical approach in this framework is based on the variation of one or more parameters in the given model (see, for instance, the review article by Karim et al. [36] on logistic growth equations). However, the need for data-driven and applicable models in stochastic growth equations implies the criteria leading to stochastic processes whose mean value is identical to the underlying growth curve. Hence, aiming to introduce a stochastic counterpart of the model considered in (11), with $F(t)$ given in (1), we focus on a description of a continuous-time Markov chain having a countable state-space. The latter assumption is justified by the fact that the size of the population considered in Remark 3 may be large; consequently, the carrying capacity of $\widetilde{F}(t)$ may be large as well. For this reason, we refer to the growth function $\widetilde{F}(t)$ rather than $F(t)$. Moreover, in order to describe the effect of environmental perturbations that lead to fluctuations in the experimentally observed growth curves, hereafter, we introduce a suitable point process whose mean exhibits the same behavior shown by $\widetilde{F}(t)$. This approach has been followed in similar growth schemes. However, in contrast to cases in which the sample-paths of the relevant processes follow skip-free behavior (as seen e.g., in Section 4 of [6], Section 3 of [3], Section 5 of [5] and Sections 4 and 5 of [7]), we consider a process having non-decreasing sample-paths, so that the analogy with the strictly increasing function $\widetilde{F}(t)$ is more tight. Specifically, we refer to an inhomogeneous linear pure birth process $\{M(t); t \geq 0\}$ having state space $S := \{y, y+1, y+2, \dots\}$, for a fixed $y \in \mathbb{N}$, with birth rates

$$\lambda_n(t) := \lim_{h\to 0^+} \frac{1}{h} \mathsf{P}[M(t+h) = n+1 | M(t) = n] = n\lambda(t), \qquad n \in S. \tag{13}$$

Here, $\lambda(t)$ is a positive function, integrable on any interval $(0, t)$ for $t > 0$, that represents the individual birth rate. The probability of a single birth during an infinitesimal time-interval after time t is proportional to the current size of the population and to the time-dependent rate $\lambda(t)$ representing the individual birth rate at time t. In this context, $M(t)$ describes the number of burned individuals, i.e., the number of individuals reached by the rumor within $[0, t]$, and the births correspond to the individual burnings.

The transition probabilities of $M(t)$ are given by (see [1])

$$P_{y,x}(t) = \mathsf{P}[M(t) = x | M(0) = y] = \binom{x-1}{y-1} e^{-y\Lambda(t)} \left(1 - e^{-\Lambda(t)}\right)^{x-y}, \qquad t \geq 0,\ x \in S, \tag{14}$$

where

$$\Lambda(t) := \int_0^t \lambda(s)\mathrm{d}s, \qquad t \geq 0 \tag{15}$$

is the individual cumulative birth rate over $[0, t]$. In this case, the probability generating function has the following expression (see [1]), for any $0 \leq z \leq 1$ and $t \geq 0$,

$$G(z,t) = \sum_{x=0}^{+\infty} P_{yx}(t) z^x = \left\{1 - (z-1)[(z-1)\phi(t) - \psi(t)]^{-1}\right\}^y, \tag{16}$$

where

$$\psi(t) := \exp\left(-\int_0^t \lambda(\tau)\mathrm{d}\tau\right), \qquad \phi(t) := \int_0^t \lambda(\tau)\psi(\tau)\mathrm{d}\tau, \qquad t \geq 0. \tag{17}$$

We note that $\psi(t)$ and $\phi(t)$ are two auxiliary functions which allow expression of the probability generating function $G(z, t)$ in a more compact manner. Thanks to Equation (16), it is possible to show that

$$\begin{aligned} \mathsf{E}_y(t) &:= \mathsf{E}[M(t)|M(0) = y] = \frac{y}{\psi(t)}, \\ \mathsf{Var}_y(t) &:= \mathsf{Var}[M(t)|M(0) = y] = y\frac{\psi(t) + 2\phi(t) - 1}{\psi^2(t)}, \end{aligned} \tag{18}$$

for $t \geq 0$. The following proposition provides a necessary and sufficient condition so that the conditional mean of the process $M(t)$ equals the growth curve $\widetilde{F}(t)$ specified in (11).

Proposition 1. *The linear birth process $M(t)$ with transition rates specified in Equation (13) and initial value $M(0) = y$, has a conditional mean*

$$\mathsf{E}_y(t) = \widetilde{F}(t), \qquad t \geq 0$$

if, and only if,

$$\lambda(t) = \xi(t), \qquad t \geq 0 \tag{19}$$

where $\xi(t)$ is given in (6) for

$$C = \frac{2\varepsilon y + N - N\varepsilon}{(1-\varepsilon)(N-y)} \geq 1, \tag{20}$$

with $N \in S \setminus \{y\}$.

Proof. The result follows immediately considering that $\mathsf{E}_y(t)$ satisfies the differential equation

$$\frac{\mathrm{d}}{\mathrm{d}t}\mathsf{E}_y(t) = \lambda(t)\mathsf{E}_y(t), \qquad t \geq 0,$$

with $\mathsf{E}_y(0) = y$, and recalling Equation (12). □

With reference to Equation (11), we recall that N represents the size of the carrying capacity, i.e., the total number of individuals that will eventually be reached by the rumor, and y is the number of individuals who know the rumor at the beginning of the spread, i.e., at $t = 0$, with $y < N$. Instead, considering the linear birth process $M(t)$ with birth rates specified in Equation (13), we note that N represents the mean total number of individuals eventually reached by the rumor, i.e., $\lim_{t\to+\infty} \mathsf{E}_y(t) = N$, and $y = M(0)$ is the initial state of the process $M(t)$. Note that, differently from the deterministic growth model whose initial state (3) may be equal to 0, for the stochastic process $M(t)$, we have $M(0) = y > 0$.

In the following, we assume that the individual birth rate is fixed as specified in Equation (19). In this special case, $\lambda(t)$ is a decreasing and convex function that approaches 0 as $t \to +\infty$, as shown in Figure 8. Hence, due to (15), the function $\Lambda(t)$ can be expressed as follows:

$$\Lambda(t) = \log\left(\frac{N\left[(e^{(1+\varepsilon)t} - 1)(\varepsilon - 1)N + y(\varepsilon - 1 - 2\varepsilon e^{(1+\varepsilon)t})\right]}{y\left[2\varepsilon(y - N) + e^{(1+\varepsilon)t}((\varepsilon - 1)N - 2\varepsilon y)\right]}\right), \qquad t \geq 0. \tag{21}$$

Clearly, due to (14), Equation (21) allows expression of the transition probabilities of $M(t)$ in a closed form under the conditions specified in Proposition 1. Moreover, the function $\Lambda(t)$ has a finite limit when $t \to +\infty$, i.e.,

$$\lim_{t\to+\infty} \Lambda(t) = \log\frac{N}{y}.$$

This can be used in Equation (14) in order to obtain the asymptotic probabilities of $M(t)$, i.e., $\lim_{t\to+\infty} P_{y,x}(t)$. In Figure 9, we provide some plots of the probability $P_{y,x}(t)$ obtained by means of Equation (14).

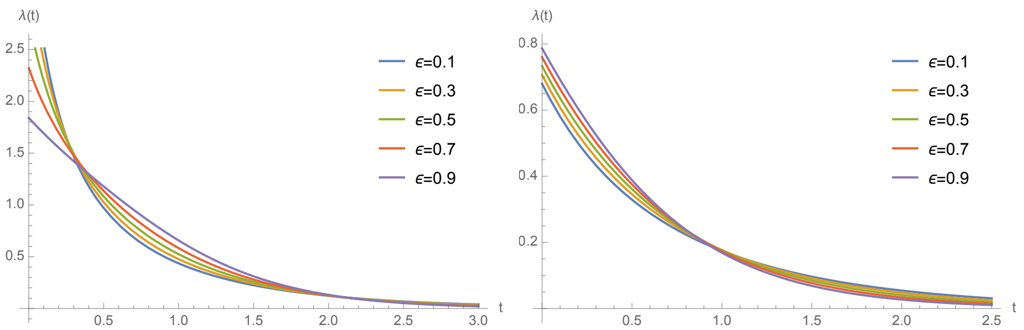

Figure 8. The individual birth rate $\lambda(t)$ as in Equation (19) for $N = 5$, **(left)** $y = 1$ and **(right)** $y = 3$.

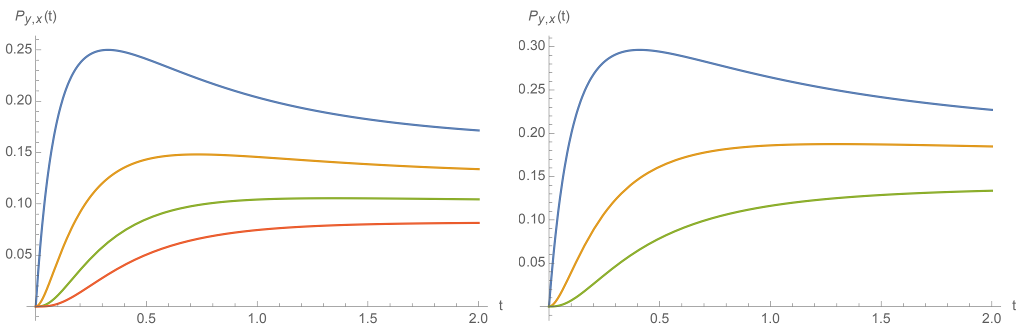

Figure 9. The transition probabilities $P_{y,x}(t)$ with the individual birth rate $\lambda(t)$ given in Equation (19), $N = 5$, **(left)** $\varepsilon = 0.3$, $y = 1$, $x = 2, 3, 4, 5$ (from top to bottom), **(right)** $\varepsilon = -0.5$, $y = 2$, $x = 3, 4, 5$ (from top to bottom).

The functions $\phi(t)$ and $\psi(t)$ are available in closed form thanks to Equation (17). This allows us to obtain an explicit expression for the conditional variance of $X(t)$, as shown in the following proposition.

Proposition 2. *The linear birth process $M(t)$ with transition rates (13) and with individual birth rate (19), with $\xi(t)$ and C given, respectively, in (6) and (20), has conditional variance given by*

$$\mathsf{Var}_y(t) = \frac{(e^{(1+\varepsilon)t}-1)N(N-y)}{y\big(2\varepsilon(N-y)+e^{(1+\varepsilon)t}(N-\varepsilon N+2\varepsilon y)\big)^2} \times \Big[(N(\varepsilon-1)-2\varepsilon y)\Big(e^{(1+\varepsilon)t}-1\Big)(\varepsilon-1)N + y(-1+\varepsilon-2\varepsilon e^{(1+\varepsilon)t})\Big], \qquad t \geq 0.$$

Proof. Determining $\psi(t)$ and $\phi(t)$ by means of (17), the result, thus, follows from Equation (18) after some calculations. □

Various plots of the conditional variance $\mathsf{Var}_y(t)$ are given in Figure 10.

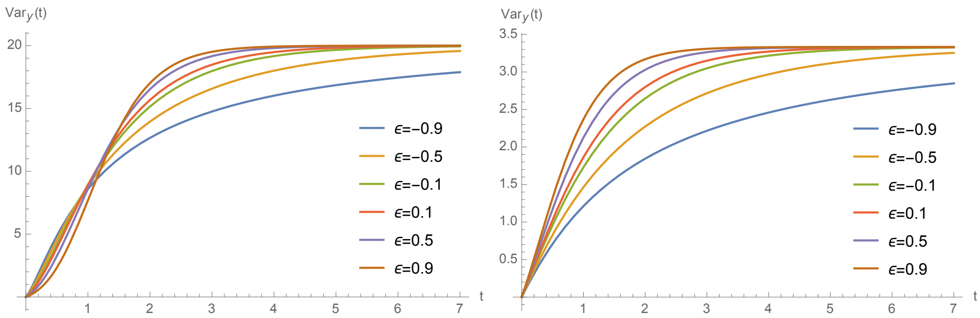

Figure 10. The conditional variance $\mathsf{Var}_y(t)$ with $\lambda(t)$ given in Equation (19), $N=5$, **(left)** $y=1$, **(right)** $y=3$.

Under the assumptions of Proposition 2, it is worth noting that the variance has a finite limit for $t \to +\infty$; indeed,

$$\lim_{t\to+\infty} \mathsf{Var}_y(t) = \frac{N(N-y)}{y}.$$

Since the conditional variance $\mathsf{Var}_y(t)$ has a finite limit as $t \to +\infty$, the corresponding conditional mean $E_y(t)$ is a significant index for the description of the birth process. Then, this property is found for other related quantities. Indeed, it is possible to obtain explicit expressions also for some indexes of dispersion for $M(t)$, such as the Fano factor and the coefficient of variation. In particular, when the conditions of Proposition 2 are satisfied, the Fano factor is given by

$$\mathsf{D}_y(t) := \frac{\mathsf{Var}_y(t)}{\mathsf{E}_y(t)} = \frac{N-y}{y\Big[1+\frac{(1+\varepsilon)N}{(e^{(1+\varepsilon)t}-1)(N-\varepsilon N+2\varepsilon y)}\Big]}, \qquad t \geq 0. \tag{22}$$

It is easy to show that $\mathsf{D}_y(t)$ is increasing with respect to t; its initial value is given by $\mathsf{D}_y(0) = 0$ and

$$\lim_{t\to+\infty} \mathsf{D}_y(t) = \frac{N}{y} - 1 > 0.$$

Hence, we have that

(i) if $y > N/2$, then the pure birth process $M(t)$ is underdispersed, i.e., $\mathsf{D}_y(t) < 1$ for any $t \geq 0$,

(ii) if $y < N/2$, then the pure birth process $M(t)$ is underdispersed for $t < \widetilde{t}$ with

$$\widetilde{t} := \frac{1}{1+\varepsilon} \log\left(\frac{(1+\varepsilon)Ny}{(N(1-\varepsilon)+2\varepsilon y)(N-2y)} + 1 \right),$$

and $M(t)$ is overdispersed for $t > \widetilde{t}$.

In Figure 11, some plots of the Fano factor are provided. Moreover, if the condition given in Proposition 2 is fulfilled, then the coefficient of variation $\mathsf{CV}_y(t) := \sqrt{\mathsf{Var}_y(t)}/\mathsf{E}_y(t)$ can be obtained in closed form for any $t \geq 0$. However, we omit the expression for brevity. Clearly, one has $CV_y(0) = 0$. Moreover, the following limit holds

$$\lim_{t\to+\infty} \mathsf{CV}_y(t) = \sqrt{\frac{N-y}{Ny}} > 0.$$

Some plots of the coefficient of variation $\mathsf{CV}_y(t)$ are provided in Figure 12.

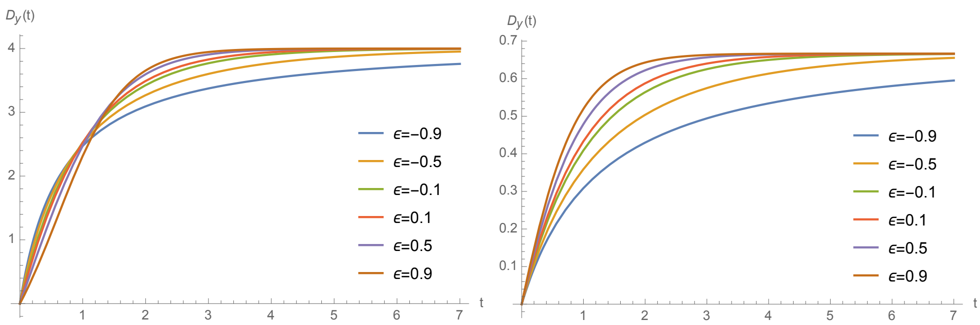

Figure 11. The Fano factor (22) with $\lambda(t)$ given in Equation (19), $N = 5$, **(left)** $y = 1$, **(right)** $y = 3$.

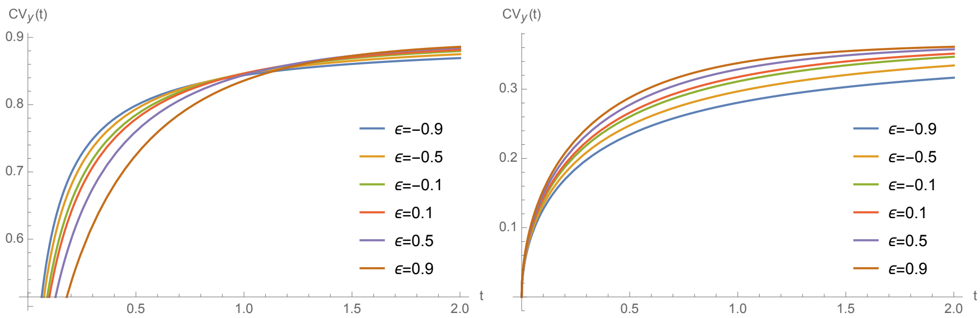

Figure 12. The coefficient of variation $\mathsf{CV}_y(t)$ with the birth rate $\lambda(t)$ given in Equation (19), $N = 5$, **(left)** $y = 1$, **(right)** $y = 3$.

Note that the conditional variance $\mathsf{Var}_y(t)$, the Fano factor $\mathsf{D}_y(t)$, and the coefficient of variation $\mathsf{CV}_y(t)$ are all increasing in t, and, for large times, they are increasing also with respect to ε. Hence, for large t and for an increasing number of initial burned individuals, the considered variability indexes attain large values. Finally, even though the expressions of the indexes obtained so far are quite cumbersome, it is worth noting that they are available in useful closed forms.

First-Passage-Time Problem

By analogy with the threshold crossing time problem considered in Section 2.4, in this section, we refer to the first-passage-time problem of the process $M(t)$. Considering the initial state $M(0) = y \in \mathbb{N}$, we fix a threshold $k \in \mathbb{N}$ with $k > y$. Consequently, the first-passage time of the process $X(t)$ through the boundary k is defined as follows

$$T_{y,k} := \inf\{t \geq 0 : M(t) = k\}, \qquad M(0) = y. \tag{23}$$

Thus, $T_{y,k}$ represents the first (random) instant in which k individuals are reached by the rumor, when the initial number of informed ones is $y > 0$. The probability density function of $T_{y,k}$ is denoted by $g_{y,k}(t)$. Since the sample paths of the pure birth process $M(t)$ are non-decreasing over the state space S, it is easy to show that, for $t > 0$,

$$g_{y,k}(t) = (k-1)\lambda(t)P_{y,k-1}(t) = (k-1)\lambda(t)\binom{k-2}{y-1}e^{-y\Lambda(t)}\left(1-e^{-\Lambda(t)}\right)^{k-1-y}, \quad (24)$$

with $k > y$, $\lambda(t)$ defined in Equation (19) and $\Lambda(t)$ given in Equation (15). In general, the first-passage time (23) is finite w.p. less than unity. However, from (24), we have

$$\lim_{N\to+\infty} \mathsf{P}(T_{y,k} < +\infty) = \lim_{N\to+\infty} \int_0^{+\infty} g_{y,k}(t)\mathrm{d}t = 1.$$

This is in accordance with the fact that the pure birth process with a countable state space is a suitable stochastic counterpart of the growth function $\widetilde{F}(t)$ when the population size N of the model (1) is large.

Moreover, we have that $P_{y,k-1}(0) = 0$, for $k-1 \neq y$. Hence, due to Equations (19) and (24), the initial value of the first-passage-time probability density function is given by

$$\lim_{t\to 0} g_{y,k}(t) = \begin{cases} y\,\lambda(0) = y\,\dfrac{C(1+\varepsilon)^2}{(C-1)(2\varepsilon+(1-\varepsilon)C)}, & \text{if } k = y+1 \\ 0, & \text{if } k > y+1. \end{cases}$$

Let $\mathbb{I}$ denote the indicator function. Some plots of the expected value $\eta_{y,k} := \mathsf{E}(T_{y,k} \cdot \mathbb{I}(T_{y,k} < +\infty))$ are given in Figure 13 for different values of the parameters. In this case, such an expectation is non-monotonic with respect to k, being increasing (decreasing) for small (large) values of k. Moreover, it is monotonic increasing with respect to ε for small values of k.

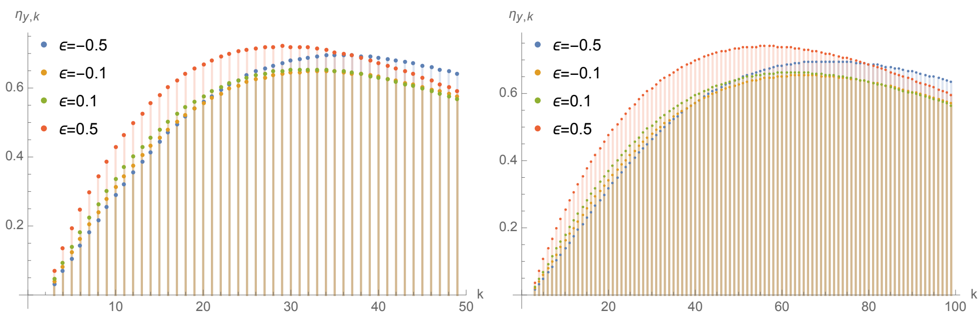

Figure 13. The expected value of the first passage time $T_{y,k}$ for $y = 2$, (**left**) $N = 50$ and (**right**) $N = 100$.

The discrete process considered in this section provides a suitable proposal for describing the growth of rumor spreads, thanks to the results of Proposition 1. However, when the population size N is very large, the information concerning the process $M(t)$ is not very manageable, so that the adoption of an alternative process is recommended. For instance, it is appropriate to consider a stochastic process with continuous state-space that, as seen for $M(t)$, possesses a mean which is identical to the growth curve. For this reason, in the following section, we focus on a diffusion process that will constitute an alternative tractable model for the description of the rumor spreads.

4. A Special Lognormal Diffusion Process

Let us consider a non-homogeneous diffusion process $\{X(t); t \geq 0\}$, with state-space $\mathbb{R}^+$ and infinitesimal moments

$$A_1(x,t) = \xi(t)x, \qquad A_2(x) = \sigma^2 x^2, \qquad t \geq 0, \quad x \in \mathbb{R}^+, \tag{25}$$

where $\xi(t)$ is defined in Equation (6), and $\sigma > 0$. The statistical properties of $X(t)$, including the infiniteness of its state-space, make it an ideal candidate for the description of growth phenomena in the presence of high environmental variability, in agreement with Remark 3. In addition, we note that the considered non-homogeneous diffusion process can be regarded as a diffusive approximation of a special birth–death process with quadratic transition rates (see Section 5.2 of [7]). With reference to (25), note that $X(t)$ is a lognormal diffusion process with a time-dependent drift, and it is the solution of the following stochastic differential equation

$$\mathrm{d}X(t) = \xi(t)X(t)\mathrm{d}t + \sigma X(t)\mathrm{d}W(t), \quad t > 0, \qquad X(0) = X_0,$$

where $W(t)$, $t \geq 0$, is a standard Wiener process independent from the initial condition $X(0) = X_0$. By means of Itô's formula, the resulting process can be expressed as follows

$$X(t) = X_0 \exp\big(\Xi(0,t) + \sigma W(t)\big), \qquad t \geq 0, \tag{26}$$

where, for $\xi(t)$ defined in (6), for $t \geq s$, one has

$$\begin{aligned}\Xi(s,t) :&= \int_s^t \xi(\tau)\mathrm{d}\tau - \frac{\sigma^2}{2}(t-s)\\ &= \log\left(\frac{2\varepsilon - Ce^{(1+\varepsilon)s}(\varepsilon-1)}{1-Ce^{(1+\varepsilon)s}}\right) - \log\left(\frac{2\varepsilon - Ce^{(1+\varepsilon)t}(\varepsilon-1)}{1-Ce^{(1+\varepsilon)t}}\right) - \frac{\sigma^2}{2}(t-s).\end{aligned} \tag{27}$$

In both cases, if X_0 has a lognormal distribution $\Lambda_1\big(\mu_0,\sigma_0^2\big)$, or if $\mathsf{P}(X_0 = x_0) = 1$, then, the random vector $(X(t_1),\ldots,X(t_n))^T$ has an n-dimensional lognormal distribution $\Lambda_n(\vec{\epsilon},\boldsymbol{\Sigma})$, where $\vec{\epsilon} = (\epsilon_1,\ldots,\epsilon_n)^T \in \mathbb{R}^n$ and $\boldsymbol{\Sigma} = \big(\sigma_{i,j}\big)_{i,j} \in \mathbb{R}^{n\times n}$ with

$$\epsilon_i = \mu_0 + \Xi(0,t_i), \qquad \sigma_{ij} = \sigma_0^2 + \sigma^2 \min(t_i,t_j), \quad i,j = 1,\ldots,n.$$

From the joint probability density function of $(X(t_1),\ldots,X(t_n))^T$, it is possible to obtain the transition probability density function of $X(t)$ given $X(s) = y$, for $0 \leq s < t$. In more detail, for $x,y \in \mathbb{R}^+$ and $0 \leq s < t$, one has

$$\begin{aligned}f(x,t|y,s) &:= \frac{\mathrm{d}}{\mathrm{d}t}\mathsf{P}(X(t) \leq x|X(s) = y)\\ &= \frac{1}{x\sqrt{2\pi\sigma^2(t-s)}}\exp\left(-\frac{(\log(x/y) - \Xi(s,t))^2}{2\sigma^2(t-s)}\right),\end{aligned} \tag{28}$$

where $\Xi(s,t)$ is defined in Equation (27). It is easy to note that $[X(t)|X(s) = y]$ follows a lognormal distribution with parameters $\log y + \Xi(s,t)$ and $\sigma^2(t-s)$, i.e.,

$$[X(t)|X(s) = y] \sim \Lambda_1\Big(\log y + \Xi(s,t), \sigma^2(t-s)\Big), \qquad s < t.$$

Since the conditional distribution of $X(t)$ is available in closed form, it is possible to obtain some of the most relevant characteristics of this process, as shown below. The conditional n-th moment of the process $X(t)$ given $X(s) = y \in \mathbb{R}^+$ for $0 \leq s < t$ is

$$\mathsf{E}(X(t)^n|X(s)=y)=y^n\left(\frac{\left(2\varepsilon-Ce^{(1+\varepsilon)s}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)t}\right)}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)s}\right)}\right)^n\exp\left(\frac{n(n-1)}{2}\sigma^2(t-s)\right),\tag{29}$$

for any $n\in\mathbb{N}$. Consequently, the unconditional n-th moment of the process $X(t)$, $t\geq 0$, can be expressed as follows

$$\mathsf{E}(X(t)^n)=(\mathsf{E}(X_0))^n\left(\frac{(2\varepsilon-C(\varepsilon-1))(1-Ce^{(1+\varepsilon)t})}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)(1-C)}\right)^n\exp\left(\frac{n(n-1)\sigma^2 t}{2}\right)\tag{30}$$

for any $n\in\mathbb{N}$. From Equations (29) and (30), one can obtain the expression of the conditional and the unconditional expected value of the process $X(t)$, i.e.,

$$\mathsf{E}(X(t)|X(s)=y)=y\,\frac{\left(2\varepsilon-Ce^{(1+\varepsilon)s}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)t}\right)}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)s}\right)},\qquad t\geq s,\tag{31}$$

and

$$\mathsf{E}(X(t))=\mathsf{E}(X_0)\,\frac{(2\varepsilon-C(\varepsilon-1))(1-Ce^{(1+\varepsilon)t})}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)(1-C)},\qquad t\geq 0,\tag{32}$$

respectively. Note that both the conditional mean (31) and the unconditional mean (32) have the same form of the function $\widetilde{F}(t)$ given in Equation (11), with C defined in Equation (20). Hence, the lognormal diffusion process $X(t)$, introduced in Equation (26), as the pure birth process considered in Section 3, has the conditional mean and the unconditional mean identical to the corresponding deterministic function $\widetilde{F}(t)$. This allows the process $X(t)$ to describe a continuous-time rumor spread subject to random fluctuations included in the infinitesimal variance $A_2(x)$ given in Equation (25). In this way, the mean of the process corresponds to the growth function $\widetilde{F}(t)$, but the sample paths of $X(t)$ are not strictly increasing. This instance is suitable to model real situations in which the diffusion of the rumor may endure abrupt slowdowns or accelerations over the time due to rough environmental perturbations.

Moreover, the conditional mode of the process $X(t)$ given $X(s)=y$, for $t\geq s\geq 0$ is given by

$$\mathsf{Mo}(X(t)|X(s)=y)=y\,\frac{\left(2\varepsilon-Ce^{(1+\varepsilon)s}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)t}\right)}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)s}\right)}\cdot\exp\left(-\sigma^2(t-s)\right),$$

whereas the unconditional mode is

$$\mathsf{Mo}(X(t))=e^{\mu_0}\,\frac{(2\varepsilon-C(\varepsilon-1))(1-Ce^{(1+\varepsilon)t})}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)(1-C)}\cdot\exp\left(-\sigma_0^2-\sigma^2 t\right),\qquad t\geq 0.$$

The α-quantiles of the process $X(t)$ can also be determined. In more detail, the conditional α-quantile for $t\geq s\geq 0$ is given by

$$\mathsf{C}_\alpha(X(t)|X(s)=y)=y\,\frac{\left(2\varepsilon-Ce^{(1+\varepsilon)s}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)t}\right)}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)\left(1-Ce^{(1+\varepsilon)s}\right)}\cdot\exp\left(z_\alpha\sqrt{\sigma^2(t-s)}\right),$$

and the unconditional α-quantile is

$$\mathsf{C}_\alpha(X(t))=e^{\mu_0}\,\frac{(2\varepsilon-C(\varepsilon-1))(1-Ce^{(1+\varepsilon)t})}{\left(2\varepsilon-Ce^{(1+\varepsilon)t}(\varepsilon-1)\right)(1-C)}\cdot\exp\left(z_\alpha\sqrt{\sigma_0^2+\sigma^2 t}\right),\qquad t\geq 0,$$

where z_α denotes the α-quantile of a standard normal random variable. In Figure 14, we provide some plots of the conditional expected value, of the conditional variance, of the conditional mode, and of the conditional coefficient of variation of $X(t)$ given $X(s) = y$. Note that the variance and the coefficient of variation are increasing with respect to t.

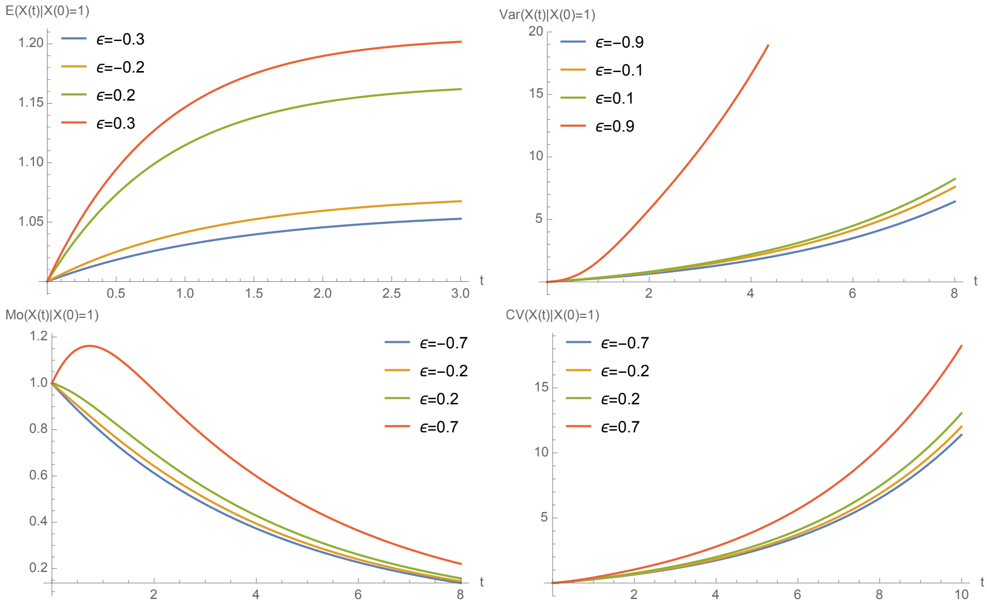

Figure 14. (**Top-left**) The conditional mean $\mathsf{E}(X(t)|X(s) = y)$, (**top-right**) the conditional variance $\mathsf{Var}(X(t)|X(s) = y)$, (**bottom-left**) the conditional mode $\mathsf{Mo}(X(t)|X(s) = y)$ and (**bottom-right**) the conditional coefficient of variation $\mathsf{CV}(X(t)|X(s) = y)$, for $s = 0$, $y = 1$, $C = 10$, $\sigma^2 = 0.25$ and different choices of ε.

4.1. First-Passage-Time Problem

This section is devoted to the study of the first-passage-time (FPT) of the diffusion process $X(t)$, defined in Equation (26), through special time-dependent boundaries. The relevance of the FPT problem for diffusion processes modeling growth phenomena is well-known in the literature. In this framework, for brevity, we limit ourselves to recalling the recent results obtained in this area in Albano et al. [37], where the FPT problem is faced for two stochastic forms of a general growth model in the presence of time-varying single or paired barriers.

Considering a continuous positive function $S(t)$, $t \geq 0$, by analogy with $T_{y,k}$ in (23), the FPT of the process $X(t)$ through the boundary $S(t) > 0$ conditional on the initial state x_0 is defined as

$$T_{x_0} := \begin{cases} \inf\{t \geq t_0 : X(t) > S(t)|X(t_0) = x_0\}, & x_0 < S(t_0) \\ \inf\{t \geq t_0 : X(t) < S(t)|X(t_0) = x_0\}, & x_0 > S(t_0). \end{cases}$$

Let us denote by $g(S(t), t|x_0, t_0)$ the probability density function of T_{x_0}. Determining an explicit expression for the density $g(S(t), t|x_0, t_0)$ is, in general, a hard task, since the function $g(S(t), t|x_0, t_0)$ is the solution of a Volterra integral equation (see for example Gutiérrez et al. [38]). However, it can be expressed in a closed form by considering special choices of the threshold $S(t)$. In more detail, by considering the results given in [38], if

$$S(t) = e^{A+Bt}\,\frac{2\varepsilon + Ce^{(1+\varepsilon)t}(1-\varepsilon)}{Ce^{(1+\varepsilon)t} - 1}, \qquad t \geq 0, \tag{33}$$

with $A, B \in \mathbb{R}$, then the FPT density is expressed in terms of the transition probability density function of $X(t)$, given in (28), i.e.,

$$\begin{aligned} g(S(t), t|x_0, t_0) &= f(S(t), t|x_0, t_0) S(t) \frac{\left|\log \frac{x_0}{S(t_0)}\right|}{t - t_0} \\ &= \frac{\left|\log \frac{x_0}{S(t_0)}\right|}{\sqrt{2\pi\sigma^2(t-t_0)^3}} \exp\left(-\frac{\left(\log \frac{S(t)}{x_0} - \Xi(t_0, t)\right)^2}{2\sigma^2(t-t_0)}\right), \qquad t \geq t_0 \geq 0, \end{aligned} \tag{34}$$

where $S(t_0) \neq x_0$ and $\Xi(s, t)$ are defined in Equation (27). Some examples of the threshold (33) and the density $g(S(t), t|x_0, t_0)$ given in (34) are plotted in Figure 15.

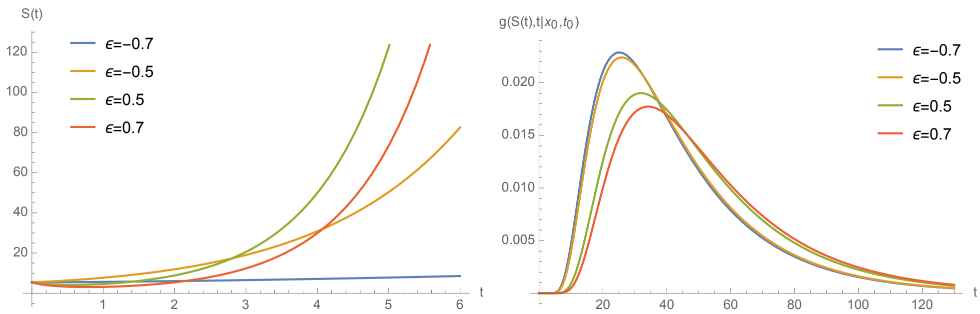

Figure 15. (**Left**) The boundary $S(t)$ and (**right**) the corresponding probability density function $g(S(t), t|x_0, t_0)$ for $A = 1$, $B = 0.5$, $C = 10$, $\sigma^2 = 0.25$, $t_0 = 0$, $N = 1000$ and different choices of ε.

4.2. Comparison between the Stochastic Growth Models

We note that the birth process and the diffusion process studied, respectively, in Sections 3 and 4, share the same mean. Hence, both processes are suitable to describe 'randomized' growth pertaining to the spread of a rumor among the members of a population. Consequently, it is appropriate to perform a comparison between the variances of the two processes in order to investigate their variability. To this end, we take into account the following ratio, for any $t \geq 0$

$$\begin{aligned} r(t) &:= \frac{\mathsf{Var}(X(t)|X(0) = y)}{\mathsf{Var}_y(t)} \\ &= \frac{(e^{\sigma^2 t} - 1)(\varepsilon - 1)^2 N y^3 \left(2\varepsilon(y - N) + e^{(1+\varepsilon)t}((\varepsilon - 1)N - 2\varepsilon y)\right)^2}{(e^{(1+\varepsilon)t} - 1)(N - y)((\varepsilon - 1)N - 2\varepsilon y)(-2N + y + \varepsilon y)^2} \\ &\times \frac{1}{\left((e^{(1+\varepsilon)t} - 1)(\varepsilon - 1)N + (\varepsilon - 1 - 2e^{(1+\varepsilon)t}\varepsilon)y\right)\left(2\varepsilon(y - N) + e^{(1+\varepsilon)t}(N(1+\varepsilon) - 2\varepsilon y)^2\right)}, \end{aligned} \tag{35}$$

where $\mathsf{Var}(X(t)|X(0) = y)$ is the conditional variance of the diffusion process (26) and $\mathsf{Var}_y(t)$ is the conditional variance of the birth process introduced in Section 3. As can be deduced from Figure 16, for large values of t, the ratio $r(t)$ is greater than 1. Hence, in the presence of the same parameters and the same initial values, the variance of the birth process, with birth rate $\lambda(t)$ given in Equation (19), is smaller than the variance of the lognormal diffusion process with infinitesimal moments (25). This leads to the conclusion that the considered birth process has less variability in the modeling of the rumor spreads.

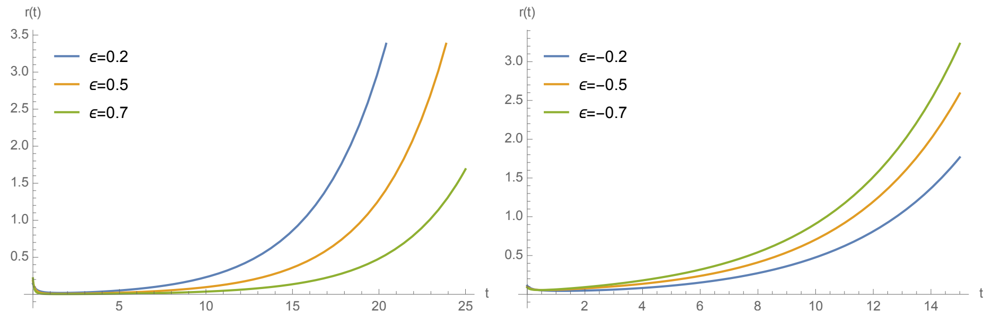

Figure 16. The ratio $r(t)$ given in (35) for $\sigma^2 = 0.25$, $N = 10$, $y = 2$ and different choices of ϵ.

5. Conclusions

The study of fake news propagation has become crucial, especially with reference to online social networks,where control mechanisms are very difficult to implement. Several attempts have been made to provide manageable functions describing the time evolution of rumor spread. We started by considering the model introduced by San Martìn et al. [12] and we studied it from a deterministic point of view. Then, we analyzed the behavior of the curve by making different choices of the parameters, showing the flexibility of the proposed model. The stochastic counterparts of the growth function were also considered. In detail, we introduced a time non-homogeneous linear pure birth process and a lognormal diffusion process with time-dependent drift. We analyzed the conditions under which the means of such processes correspond to the deterministic growth curve. The first-passage-time problem was also addressed. Furthermore, in order to analyze the variability of the processes, we performed a comparison between their variances, noting that, for large times, the variance of the birth process was smaller than the variance of the diffusion process.

In Remark 1, it was pointed out that the carrying capacity of the model (1) is unity. This can be viewed, in a sense, as a limitation of the growth model since, in various contexts, only a fraction of the population is eventually informed by the news. However, this feature makes the model particularly appropriate for describing the diffusion of news in closed communities or in restricted environments, for example, in homogeneous groups in social networks, where it is expected that all members will be reached by the rumor.

This study should be viewed as a first step to the construction of stochastic generalization of growth models for the spreading of fake news. Clearly, the present study can be improved in the future. For instance, it will be useful to adopt stochastic schemes similar to epidemiological models, such as SIR or SIS models, to describe not only the evolution of spreaders, but also of the other components of the population. Moreover, the investigation should be extended along the following lines:

(a) application of the considered models to real data for prediction purposes;
(b) new growth models for the percentage of burned, i.e., informed individuals, based on suitable compartmental models;
(c) new stochastic processes finalized to describe the diffusion of rumors subject to randomness;
(d) constructing more general birth–death processes to model the spread of rumors also in the presence of individuals prone to forgetting the rumors;
(e) adopting AI-based strategies for detecting disinformation and fake news.

The above sketched proposals can be the subject of future research.

Author Contributions: Conceptualization: A.D.C., P.P. and S.S.; investigation: A.D.C., P.P. and S.S.; methodology: A.D.C., P.P. and S.S. All authors have read and agreed to the published version of the manuscript.

Funding: This research was partially funded by MIUR-PRIN 2017, Project "Stochastic Models for Complex Systems" (no. 2017JFFHSH) and by INdAM-GNCS, Project "Modelli di shock basati sul processo di conteggio geometrico e applicazioni alla sopravvivenza"(CUP E55F22000270001). The authors are members of the group GNCS of INdAM (Istituto Nazionale di Alta Matematica).

Data Availability Statement: No new data were created or analyzed in this study. Data sharing is not applicable to this article.

Conflicts of Interest: The authors declare no conflict of interest. The funders had no role in the design of the study; in the collection, analyses, or interpretation of data; in the writing of the manuscript; or in the decision to publish the results.

References

1. Tan, W.Y. A stochastic Gompertz birth-death process. *Stat. Prob. Lett.* **1986**, *4*, 25–28. [CrossRef]
2. Korf, V. Prìspevek k matematickè formulaci vzrustovèho zàkona lesnìch porostu. Mathematical Definition of the Law of Stand Volume Growth. *Lesnickà Pràce* **1939**, *18*, 339–379.
3. Di Crescenzo, A.; Paraggio, P. Logistic growth described by birth-death and diffusion processes. *Mathematics* **2019**, *7*, 489. [CrossRef]
4. Richards, F.J. A flexible growth function for empirical use. *J. Exp. Bot.* **1959**, *10*, 290–301. [CrossRef]
5. Asadi, M.; Di Crescenzo, A.; Sajadi, F.A.; Spina, S. A generalized Gompertz growth model with applications and related birth-death processes. *Ric. Mat.* **2020**. [CrossRef]
6. Di Crescenzo, A.; Spina, S. Analysis of a growth model inspired by Gompertz and Korf laws, and an analogous birth-death process. *Math. Biosci.* **2016**, *282*, 121–134. [CrossRef] [PubMed]
7. Di Crescenzo, A.; Paraggio, P.; Román-Román, P.; Torres-Ruiz, F. Applications of the multi-sigmoidal deterministic and stochastic logistic models for plant dynamics. *Appl. Math. Model* **2021**, *92*, 884–904. [CrossRef]
8. Di Crescenzo, A.; Paraggio, P.; Román-Román, P.; Torres-Ruiz, F. Statistical analysis and first-passage-time applications of a lognormal diffusion process with multi-sigmoidal logistic mean. *Stat. Pap.* **2022**. [CrossRef]
9. Erto, P.; Giorgio, M.; Lepore, A. The generalized inflection s-shaped software reliability growth model. *IEEE Trans. Reliab.* **2020**, *69*, 228–244. [CrossRef]
10. Brauer, F.; Castilo-Chavez, C. *Mathematical Models in Population Biology and Epidemiology*; Springer: New York, NY, USA, 2010.
11. Smirnov, R.; Wang, K. In search of a new economic model determined by logistic growth. *Eur. J. Appl. Math.* **2020**, *31*, 339–368. [CrossRef]
12. San Martín, J.; Drubi, F.; Rodríguez Pérez, D. Uncritical polarized groups: The impact of spreading fake news as fact in social networks. *Math. Comput. Simul.* **2020**, *178*, 192–206. [CrossRef]
13. De Martino, G.; Spina, S. Exploiting the time-dynamics of news diffusion on the Internet through a generalized Susceptible—Infected model. *Phys. A Stat. Mech. Its Appl.* **2015**, *438*, 634–644. [CrossRef]
14. Giorno, V.; Spina, S. Rumor spreading models with random denials. *Phys. A Stat. Mech. Its Appl.* **2016**, *461*, 569–576. [CrossRef]
15. Figueira, A.; Oliveira, L. The current state of fake news: Challenges and opportunities. *Procedia Comput. Sci.* **2017**, *121*, 817–825. [CrossRef]
16. Abraham, S.; Nair, S. Cyber Security Analytics: A Stochastic Model for Security Quantification Using Absorbing Markov Chains. *J. Commun.* **2014**, *9*, 12. [CrossRef]
17. Abimbola, O.O.; Odunola, A.B.; Temitope, A.A.; Aderounmu, G.A.; Hamidja, K.B. An improved stochastic model for cybersecurity risk assessment. *Comput. Inf. Sci.* **2019**, *12*, 4. [CrossRef]
18. Paul, J.A.; Zhang, M. Decision support model for cybersecurity risk planning: A two-stage stochastic programming framework featuring firms, government, and attacker. *Eur. J. Oper. Res.* **2021**, *291*, 349–364. [CrossRef]
19. Alandihallaj, M.A.; Assadian, N.; Khorasani, K. Stochastic model predictive control-based countermeasure methodology for satellites against indirect kinetic cyber-attacks. *Int. J. Control* **2023**, *96*, 1895–1908. [CrossRef]
20. Miaoui, Y.; Boudriga, N. Enterprise security economics: A self-defense versus cyber-insurance dilemma. *Appl. Stoch. Model. Bus. Ind.* **2019**, *35*, 448–478. [CrossRef]
21. Mahmoud, H. A model for the spreading of fake news. *J. Appl. Probab.* **2020**, *57*, 332–342. [CrossRef]
22. Kapsikar, S.; Saha, I.; Agarwal, K.; Kavitha, V.; Zhu, Q. Controlling fake news by collective tagging: A branching process analysis. In Proceedings of the 2021 American Control Conference, New Orleans, LA, USA, 25–28 May 2021; pp. 1332–1337. [CrossRef]
23. Ben Aissa, A.; Abdalla, I.; Hussein, L.F.; Elhadad, A. A novel stochastic model for cybersecurity metric inspired by Markov Chain model and Attack Graphs. *Int. J. Sci. Technol. Res.* **2020**, *9*, 6329–6335. Available online: https://www.ijstr.org/research-paper-publishing.php?month=mar2020 (accessed on 13 July 2023).
24. Raponi, S.; Khalifa, Z.; Oligeri, G.; Di Pietro, R. Fake news propagation: A review of epidemic models, datasets, and insights. *ACM Trans. Web* **2022**, *16*, 12. [CrossRef]
25. Esmaeeli, N.; Sajadi, F.A. On the probability of rumour survival among sceptics. *J. Appl. Probab.* **2023**, *60*, 1096–1111. [CrossRef]
26. Jia, F.; Cao, C. Dynamic analysis of a stochastic rumor propagation model with regime switching. *Mathematics* **2021**, *9*, 3277. [CrossRef]

27. Roy, R.; Saha, K. How fast do rumors spread? *arXiv* **2023**, arXiv:2308.05940.
28. Øksendal, B. *Stochastic Differential Equations. An Introduction with Applications*, 6th ed.; Springer: Berlin/Heidelberg, Germany, 2003.
29. Román-Román, P.; Serrano-Pérez, J.J.; Torres-Ruiz, F. Some notes about infer- ence for the lognormal diffusion process with exogenous factors. *Mathematics* **2018**, *6*, 85. [CrossRef]
30. Giorno, V.; Nobile, A. Bell Polynomial Approach for Time-Inhomogeneous Linear Birth–Death Process with Immigration. *Mathematics* **2020**, *8*, 1123. [CrossRef]
31. Ricciardi, M. Stochastic population theory: Birth and death processes. In *Mathematical Ecology*; Hallam, T.G., Levin, S.A., Eds.; Springer: Berlin/Heidelberg, Germany, 1986; Volume 17.
32. Khan, A.A.; Nauman, M.A.; Shoaib, M.; Jahangir, R.; Alroobaea, R.; Alsafyani, M.; Binmahfoudh, A.; Wechtaisong, C. Crowd anomaly detection in video frames using fine-tuned AlexNet Model. *Electronics* **2022**, *11*, 3105. [CrossRef]
33. Arunnehru, J.; Thalapathiraj, S.; Dhanasekar, R.; Vijayaraja, L.; Kannadasan, R.; Khan, A.A.; Haq, M.A.; Alshehri, M.; Alwanain, M.I.; Keshta, I. Machine vision-based human action recognition using spatio-temporal motion features (STMF) with difference intensity distance group pattern (DIDGP). *Electronics* **2022**, *11*, 2363. [CrossRef]
34. Jin, F.; Dougherty, E.; Saraf, P.; Mi, P.; Cao, Y.; Ramakrishnan, N. Epidemiological modeling of news and rumors on Twitter. In Proceedings of the 7th SNA-KDD Workshop ’13 (SNA-KDD’13), Chicago, IL, USA, 11 August 2013.
35. Albano, G.; Giorno, V.; Román-Román, P.; Torres-Ruiz, F. Study of a general growth model. *Commun. Nonlinear Sci. Numer. Simul.* **2022**, *107*, 106100. [CrossRef]
36. Karim, M.A.U.; Aithal, V.; Bhowmick, A.R. Random variation in model parameters: A comprehensive review of stochastic logistic growth equation. *Ecol. Model.* **2023**, *484*, 110475. [CrossRef]
37. Albano, G.; Barrera, A.; Giorno, V.; Román-Román, P.; Torres-Ruiz, F. First passage and first exit times for diffusion processes related to a general growth curve. *Commun. Nonlinear Sci. Numer. Simul.* 2023, *online first*. [CrossRef]
38. Gutiérrez, R.; Ricciardi, L.M.; Román-Román, P.; Torres-Ruiz, F. First-passage-time densities for time-non-homogeneous diffusion processes. *J. Appl. Prob.* **1997**, *34*, 623–631. [CrossRef]