Stochastic simplicial contagion model

Daniel Hernández Serrano[a], Javier Villarroel[c], Juan Hernández-Serrano[b], Ángel Tocino[a]

[a] *Departamento de Matemáticas and Instituto Universitario de Física Fundamental y Matemáticas (IUFFyM), Universidad de Salamanca, Salamanca, Spain*
[b] *Department of Network Engineering, Universitat Politècnica de Catalunya, Barcelona, Spain.*
[c] *Departamento de Estadística and Instituto Universitario de Física Fundamental y Matemáticas (IUFFyM), Universidad de Salamanca, Salamanca, Spain*

Abstract

We propose a stochastic model that describes of epidemics over simplicial complex networks (SSCM) in which higher-order unforeseen or random interactions may occur. Its dynamics obeys a stochastic differential equation (SDE) based on the mean field approach of the simplicial social contagion model. In this stochastic regime the only possible equilibrium state is the origin. We give conditions to guarantee global stability and hence that the disease dies out. We partition the parameter space into the instability, the bi-stability and the globally asymptotically stable regions described in terms of appropriate epidemiological parameters. These regimes codify whether the disease will, can or will not disappear. We also present empirical results obtained by running different simulations of the SSCM over several real-world simplicial networks and over a synthetically generated one, which validate the theoretical results presented.

Keywords: complex networks, simplicial complexes, mathematical epidemiology, contagion models, stochastic differential systems, stochastic stability, SIS model, basic reproductive number
2020 MSC: 05C82, 55U10, 60H10, 60H35, 90B15

1. Introduction

The theory of complex networks ([6, 5]) has become a main tool to analyze complex systems appearing in the physical, social and biological sciences. In particular, they have been used to describe the spread of diseases ([2, 49, 59, 15, 29]) where the transmission between individuals is assumed to occur through pairwise interactions between infectious and susceptible agents (simple contagion processes); see also the recent works [3, 7, 30, 67].

Email addresses: dani@usal.es (Daniel Hernández Serrano), javier@usal.es (Javier Villarroel), j.hernandez@upc.edu (Juan Hernández-Serrano), bacon@usal.es (Ángel Tocino)

However, very often physical systems cannot be represented as traditional graphs and may have an internal structure that reflects interactions between several agents at the same time, group interactions or n-ary relationships between vertices of the network. Such complex systems can be modeled with hypergraphs or with simplicial complexes ([39, 8, 58, 11]); their application in the study of higher-order structures in complex systems have produced numerous advances in a wide variety of domains; see, for instance, [9, 10, 43, 51, 48, 25, 26] for social systems, [18, 19, 63, 64] for biomolecular systems or [14, 21, 41, 50, 54] for brain networks.

Classical epidemiological studies appear in the seminal work of Kermack and McKendrick [33] as deterministic compartmental systems described by ordinary differential equations ([28]). Recently, the study of epidemics spreading starts to be taken into account when the contact topology ceases to be a graph and becomes a hypergraph or a simplicial complex. Novel ideas for the analysis of complex contagion processes on multilayer networks appear also in [67, 23]. See [35] for a thorough review and general framework.

In [30, 3, 7], social contagion (compartmental and deterministic) models for structures where higher-order interactions might occur are proposed, and these remarkable results have resulted in the development of new insights into problems of propagation or diffusion processes in networks where group interactions exist ([45, 13, 40, 4]).

Real social networks are inherently random, a fact that the basic Kermack and McKendrick model does not capture, and which recent studies are now trying to incorporate by appropriate modifications of the dynamics. The crucial role that stochasticity plays in a variety of physical situations was recognized in the seminal work of Einstein, Smoluchowski and Langevin to describe Brownian motion, which in turn paved the way for the introduction of stochastic differential equations (SDE) in the groundbreaking work of Kiyosi Itô. The importance of SDE's is now widely recognized and is a most active research areas in the last decades. These techniques have become ubiquitous in fields like mathematical finance where, after the pioneering works of Black and Scholes and Merton [12] and Vasicek [60], the basic models for the evolution of stock prices, interest rates and and currency exchange rates are all assumed to evolve via SDE. However, they also find ample use in other fields like biology (to develop optimal fishing policies) and population dynamics (and extinction times). Here too stochastic versions of differential equations have been considered for decades. In epidemiology, different stochastic versions of the SIR model have been considered in [24, 17, 37, 42, 57, 31, 1, 44, 66], and a local stability analysis of the disease free equilibrium state is studied using Lyapunov functions. In [55] a systematic way to obtain sufficient conditions for the mean square stability stochastic stability of the linearized problem of a SIR model is presented. For other applications of SDE's see also [16, 61] (Physics) , [20, 47] (Biology), and [36, 38, 52, 65] (Finance).

Following this approach, in this paper we model the evolution of the epidemic by a stochastic SIS-type model described by a certain Ito's SDE, cf. (5). It generalizes the model considered in [24], to situations where simultaneous

contagious are possible. We expect that this equation − which is based in the mean field approach of the simplicial SIS model of [30]− gives a good approximation to the evolution based on real networks. Both from a mathematical and an epidemiological perspective, the key problem is the determination of conditions that guarantee that the disease dies out.

As far as we know, the randomness in higher-order complex systems studied so far is only captured when higher-order contagion models are run on real (or synthetically generated) networks (where an infected agent, or group of infected agents, infect a susceptible agent with a certain probability) or in other modeling situations such as the microscopic Markov-chain approaches. In them, the probabilities of infection are constant with respect to time.

We propose a different stochastic process making these higher-order probabilities dependent on time: for each instant of time, the probability β of infection (and its relative higher-order generalizations) will be modified to be a time dependent random variable $\beta + \sigma\xi_t$ where ξ_t is a white Gaussian noise and σ denotes the intensity with which the white noise is applied. With this we intend to model random situations such as fortuitous and unexpected encounters of a susceptible individual with one or more infected individuals. The randomness is represented in the generation of the white Gaussian noise (which will follow a standard normal distribution), ξ_t, and the intensity σ with which it is applied could also be interpreted as a risk factor associated with a specific scenario (school, bar, soccer stadium,...). This could help to establish measures to contain the spread of a disease depending on the unforeseen multiple interactions that may occur in different places.

In addition, and in order to validate these theoretical results from an empirical point of view, we propose a stochastic simplicial contagion model (SSCM) based on the simplicial contagion model (SCM) of [30]. We run the SSCM over several real-world simplicial networks coming from different domains (and over synthetically generated ones) to carry out our analysis of the dynamics of evolution in higher-order simplicial social structures. We allow for the contagion probabilities to vary randomly at every time step. This feature fits well with several real-life situations, such as unforeseen encounters of susceptible and one or more infected individuals, or the possibility that the contagion probability may increase, say, in certain forums. These notes therefore expands the results of [30] to a stochastic case. We have made available Python code for our SSCM at [27].

The paper is organized as follows. In section 2, we review the basic results of simplicial complexes and a simplicial SIS-type model. Section 3 presents the first theoretical result: a SDE that governs the dynamics of the spread of an epidemic over a simplicial network and the study of global stability conditions. This is related to several mathematical concepts of interest, like equilibrium points and the concept of barrier points. We use them to show that− as expected− motion is limited to $(0, 1)$. However, while in a deterministic situation both disease free and endemic equilibrium points may coexist, stochasticity destroys endemic equilibrium points, so only the origin remains as a possible equilibrium point.

Section 4 is devoted to fully characterize the stability regions in terms of

new epidemiological parameters for the case on which the rate of higher-order infection is proportional to the rate of standard individual infection. The main result in this regard is quite technical: theorem (3) gives sharp conditions to guarantee that the origin in (5) is globally asymptotically stable and hence with probability one the disease dies out. We then polish this result and set it down in terms of parameters with epidemiological meaning, the basic reproduction numbers. This allows us to partition the parameter space into the instability, the bi-stability and the globally asymptotically stable (g.a.s.) regimes, according to whether eventually the disease *will, can or will not* disappear. Figures 1 and 3 show these regions in the deterministic and stochastic cases. The case where the rate of recovery is zero is remarkable, since here the evolution arising from (5) is such that both 0 and 1 are two different equilibrium points, that is, both of them may be simultaneously stable. Different sample paths obtained solving (5) with Runge-Kutta numerical methods illustrate the possibilities.

In section 6 we perform a study of disease transmission as a complex network and show different empirical results; concretely, a SSCM is defined and ran over different types of real-world and synthetically generated simplicial networks. We compare our techniques with the original one of [30]. We also relate the network evolution and that arising from the SDE (5) of previous sections. We find similar propagation dynamics for both the real-world and the synthetic simplicial networks, which also show a remarkable similarity with that predicted by the SDE in all different regimes. We include a section 7 with conclusions and two appendices with technical results.

2. The simplicial SIS model.

To analyze the evolution of the disease transmission in terms of a simplicial SIS model, let us first recall some basic notions of simplicial complexes (we refer to [46, 22, 21] for a detailed exposition). Given a finite set of distinct points $\{v_0, v_1, \ldots, v_N\}$, $N \in \mathbb{N}$, which we call vertices, a q-simplex is a subset of vertices $\sigma^{(q)} = \{v_0, v_1, \ldots, v_q\}$, where $q \leq n$ is called the dimension of $\sigma^{(q)}$. For $p \leq q$, a $p-$face of a q-simplex $\sigma^{(q)}$ is a subset of $p+1$ vertices of $\sigma^{(q)}$. A simplicial complex K is a collection of simplices such that if σ is a simplex in K, then all the faces of σ are also in K, and any non-empty intersection of any two simplices of K is a face of each of them. Simplices which are not faces of any other simplex are called facets and the dimension of a simplicial complex, $\dim K$, is the maximum among the dimensions of its simplices.

We consider a population, composed of $N \in \mathbb{N}$ individuals and susceptible to a certain disease; we assume that the population remains constant through time and it is divided into two classes: susceptible and infected. Individuals v_i in this population can be thought of as the vertices of a certain simplicial complex K of dimension D, having $N \in \mathbb{N}$ vertices.

Following [30], we consider a mean-field (MF) approach under homogeneous mixing assumption to describe the dynamics of a simplicial SIS model:

$$\frac{dx(t)}{dt} = \sum_{h=1}^{D} \beta_h \langle k_h \rangle \big(x(t)\big)^h \big(1 - x(t)\big) - \mu x(t)\,, \tag{1}$$

where

- The basic observable $x(t)$ represents the density of infected individuals at time t (so that $1 - x(t)$ represents the density susceptible individuals).
- μ is the recovery probability.
- $\{\beta_1, \ldots, \beta_D\}$ are the infection probabilities, where β_h represents the probability for a susceptible node v_i, which belongs to the h-simplex

  $$\sigma^{(h)} = \{v_i, v_{j_0}, v_{j_1}, \ldots, v_{j_{h-1}}\}\,,$$

  to be infected by the $(h-1)$-face

  $$\tau^{(h-1)} = \{v_{j_0}, v_{j_1}, \ldots, v_{j_{h-1}}\}$$

  under the assumption that all vertices v_{j_k} of $\tau^{(h-1)}$ are infectious.
- k_h is a random variable counting the number of h-simplices a node belongs to and $\langle k_h \rangle$ stands for its statistical mean.

Remark 1. *The generalised simplicial degree k_h coincides with the (h,h)-upper simplicial degree of [25, 26], where several notions of lower, upper and of a general adjacency (and their associated simplicial degrees) valid for any dimensional comparison between simplices are defined.*

In the sequel we restrict to the $D = 2$ case, i.e. assume that the simplicial complex K is bounded to dimension 2, since this situation is already substantial enough. For ease of notation let us rename $\beta \equiv \beta_1 \langle k_1 \rangle$, $\beta_\Delta \equiv \beta_2 \langle k_2 \rangle$. To avoid trivialities we assume $\beta > 0$, $\beta_\Delta > 0$ and $\mu > 0$. Equation (1) is transformed into:

$$\frac{dx(t)}{dt} = x\big(-\beta_\Delta x^2 + (\beta_\Delta - \beta)x + (\beta - \mu)\big) = -\beta_\Delta x(x - x_+)(x - x_-)\,, \tag{2}$$

where abusing notation we write $x \equiv x(t)$ and define $R_0 = \frac{\beta}{\mu}$, $R_{0\Delta} = \frac{\beta_\Delta}{\mu}$ and

$$x_\pm = \frac{R_{0\Delta} - R_0 \pm \sqrt{(R_0 + R_{0\Delta})^2 - 4R_{0\Delta}}}{2R_{0\Delta}} \tag{3}$$

Before turning to the study of the stochastic case, and for the sake of completeness, we summarize stability results for the deterministic case.

Definition 1. Let $x(t; x_0)$ denote the solution of(2) with initial condition $x(0; x_0) = x_0$. The origin is *locally* asymptotically stable if whenever the *initial condition x_0 is close enough* to the origin then so it is $x(t; x_0)$ and besides it holds that

$$\lim_{t \to \infty} x(t; x_0) = 0 \tag{4}$$

The origin is *globally* asymptotically stable if (4) holds for all x_0. We use l.a.s. for "locally asymptotically stable" and g.a.s. for "globally asymptotically stable".

The *possible* limit states are the points 0 and $x_\pm$−provided they exist in $[0,1]$. We call them the disease free and, respectively, endemic equilibrium points (e.p.). Their classification is given in terms of only two parameters $(R_0, R_{0\Delta})$. To this end we divide the parameter space $S := \{(R_0, R_{0\Delta}), R_0 > 0, R_{0\Delta} > 0\}$ in the following 4 sub-regions, see Figure 1:

1. $S_1 := \{(R_0, R_{0\Delta}) : R_0 < 1, R_{0\Delta} < 1, \text{ and } (R_{0\Delta} + R_0)^2 - 4R_{0\Delta} \geq 0\}\}$
2. $S_2 := \{(R_0, R_{0\Delta}) : \ (R_{0\Delta} + R_0)^2 - 4R_{0\Delta} < 0\}$
3. $S_3 := \{(R_0, R_{0\Delta}) : R_0 < 1, R_{0\Delta} > 1, \text{ and } (R_{0\Delta} + R_0)^2 - 4R_{0\Delta} \geq 0\}\}$.
4. $S_4 := \{(R_0, R_{0\Delta}) : R_0 \geq 1\}$

Here, a complete classification of the limit points $\lim_{t\to\infty} x(t)$ is possible using classical Lyapunov theory (see also [30] for a different but equivalent approach). We have

Proposition 1. *(Stability)*

1. *If $x_0 \in [0,1]$ and $x(t)$ evolves via (2) then $x(t; x_0) \in [0,1]$.*
2. *If $R_0 < 1$ then the origin is a stable equilibrium point.*
   Besides the origin is globally stable– and the disease disappears– if parameters satisfy $(R_0, R_{0\Delta}) \in S_1 \cup S_2$. Here endemic e.p. do not exist.
   If $(R_0, R_{0\Delta}) \in S_3$ there are two stable e.p., namely 0 and x_+ and one unstable, x_-, with $0 \leq x_- \leq x_+ \leq 1$.
   In this case the disease may disappear or become endemic. The initial condition $x_0 = x_-$ marks the threshold between these regimes.
   Concretely, we have

$$\lim_{t\to\infty} x(t) = \begin{cases} 0, \ \textit{if } (R_0, R_{0\Delta}) \in S_1 \cup S_2 \\ 0, \ \textit{if } (R_0, R_{0\Delta}) \in S_3, \ \textit{and } x_0 \in [0, x_-) \\ x_+, 0, \ \textit{if } (R_0, R_{0\Delta}) \in S_3, \ \textit{and } x_0 \in (x_-, 1] \end{cases}$$

3. *If $R_0 \geq 1$ then $x_- \leq 0 \leq x_+ \leq 1$, and x_+ is a stable equilibrium point, the endemic state, while 0 is unstable. Hence $\lim_{t\to\infty} x(t; x_0) = x_+$ for all $x_0 > 0$.*

Remark 2. We say that $S_1 \cup S_2$ is the g.a.s. regions, S_3 *the bi-stable region while S_4 the unstable region.* Figure (1) illustrates the situation (see also [30]). Note that, by increasing R_0 any point $(R_0, R_{0\Delta})$ belonging to the interior of $(S_1 \cup S_2) \cap \{R_{0\Delta} \geq 1\}$ will pass to the bi-stable regime before reaching the unstable one. The special tri-critical point $p_T = (1,1)$ marks the boundary of these three regions. Any perturbation of p_T may result in going to a globally stable, a bi-stable or to an unstable state.

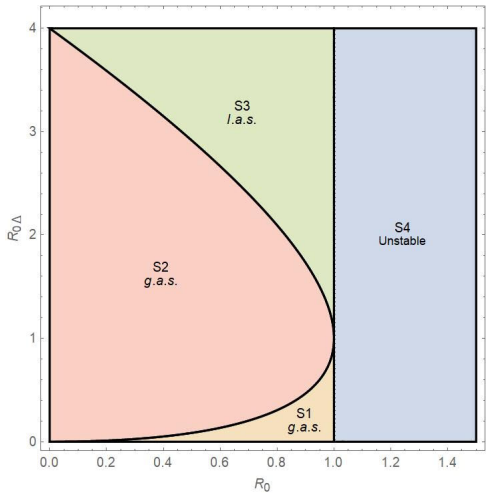

Figure 1: Parameter space $p = (R_0, R_{0\Delta})$ (horizontal and vertical axis respectively) and stability regions for the simplicial SIS deterministic model. Here, the disease disappears in the g.a.s. regions $p \in S_1 \cup S_2$; the disease becomes endemic if $p \in S_4$; or the disease sometimes disappears and sometimes becomes endemic in the bi-stability region S_3.

3. The stochastic differential equation for the epidemic evolution.

The dynamics given by (2) assumes tacitly that at every instant of time t the proportion of infected individuals satisfy certain rules *with certainty*. For example, (2) states that the number of infected individuals that overcome the illness in $[t, t+h]$ is $\mu N x(t)h + o(h)$. The crucial role that randomness plays in real world epidemics has been overlooked so far, where the best one can do is to talk in terms of probabilities. To this end we model the evolution of a epidemic over a simplicial network by a certain Itô stochastic differential equation (ISDE) which responds to the following assumptions

- the population has only two compartments, susceptible and infected.
- Simultaneous contact with one or two sources of activation occur, whose intensity is measured, respectively, by β and β_Δ. Simultaneous contact with three or more individuals do not occur.
- Contagious rates [1] are no longer fixed: there exist as stochastic factors which modify the contagions rates per unit of time of equation (2) as

  $$\beta\Delta t \to \beta\Delta t + \sigma\Delta W(t),\ \beta_\Delta\Delta t \to \beta_\Delta\Delta t + \sigma_\Delta\Delta W(t),$$

  where $W(t)$ is a Brownian motion, $\beta, \beta_\Delta, \sigma, \sigma_\Delta, \mu$ are positive model constants and σ and σ_Δ measure the intensity of the stochastic perturbation

[1] We recall that if β_1, β_2 are the basic contagious rates then $\beta \equiv \beta_1\langle k_1\rangle$, $\beta_\Delta \equiv \beta_2\langle k_2\rangle$

It follows from the above that we restrict ourselves to a $D = 2$ dimensional simplicial complex. Besides the proportion of infected individuals at time t, $x(t)$, evolves via the SDE

$$dx(t) = a(x(t))dt + b(x(t))dW(t), \tag{5}$$

where

$$a(x) = (\beta - \mu)x + (\beta_\Delta - \beta)x^2 - \beta_\Delta x^3; \quad b(x) = \sigma x + (\sigma_\Delta - \sigma)x^2 - \sigma_\Delta x^3 \tag{6}$$

are, respectively, the drift and the diffusion coefficients of the Ito process.

We expect that (6) constitutes a good approximation to the dynamics of epidemics, which generalizes the ISDE considered in [24], to a case where simultaneous contagious are possible, namely $\beta_\Delta > 0, \sigma_\Delta > 0$. We denote as $x(t; x_0)$ the solution to (5)-(6) satisfying the initial condition $x(0; x_0) = x_0$.

3.1. Implications of stochasticity.

When stochasticity is introduced equilibrium points (e.p.) are defined as the common solutions to $a(x) = b(x) = 0$. If $\sigma > 0, \mu > 0$, the only such solution is $x = 0$, so it appears that *endemic equilibrium point do not exist.* This contrasts with the deterministic case, where both disease free and endemic equilibrium points may coexist. We next study sufficient conditions for the stability of the origin. Actually, the study is far more complex in this case: while equilibrium points, can still be divided into unstable, locally stable and globally stable, the concept of barrier points is also relevant; these may be cataloged as attracting or not, and attainable or not. Besides, limit states, if they exist, can not be classified in terms of epidemiological parameters and initial state x_0.

The first result establishes several properties of the system evolving via equation (5).

Proposition 2. *If $\sigma, \sigma_\Delta, \mu$ are positive, the only critical point of the system is $x = 0$.*

Proof. We note the factorization, see (3),

$$b(x) = -\sigma_\Delta x(x-1)(x + \sigma/\sigma_\Delta), \quad a(x) = -\beta_\Delta x(x - x_-)(x - x_+).$$

Hence, x solves $a(x) = b(x) = 0$ if and only if $x = 0$. □

In the sequel we focus on the stability of the origin $x = 0$.

Definition 2. The origin is *stochastically unstable* if for any $\epsilon > 0$

$$\lim_{x_0 \to 0} \Pr(\sup_{t<\infty} |x(t; x_0)| \geq \epsilon) > 0. \tag{7}$$

The origin is *locally* stochastically asymptotically stable (l.a.s.) if it is not unstable and it holds that

$$\lim_{x_0 \to 0} \Pr(\lim_{t \to \infty} x(t; x_0) = 0) = 1. \tag{8}$$

Namely, initial conditions x_0 "close enough" to the origin have a high probability to end up in 0.

It turns out that stability can be pin down in terms of one number.

Theorem 1. *(Instability and local stability)*

If $J_0 := \frac{\beta}{\mu+\sigma^2/2} < 1$ *then the origin is a locally asymptotically stable e.p. for the equation* (5)-(6).

If $J_0 \geq 1$ *then the origin is an unstable e.p. for the equation* (5)-(6) *and* $\lim_{t\to\infty} x(t;x_0) \neq 0, \forall x_0 \in (0,1]$ *with probability one. Actually, for any* $x_0 \in (0,1]$, $\lim_{t\to\infty} x(t;x_0)$ *can not converge to any* $\hat{x} \in [0,1]$

Proof. The proof of the theorem appears in Appendix A. □

Remark 3. If we accept that the epidemic evolution is well described by (5), then the parameter $J_0 \in [0,\infty)$ could be interpreted as an effective *basic reproductive number*, that controls the basic stability: if $J_0 \geq 1$ the disease never disappears with probability one. However, if $J_0 < 1$ all that can be said is that *there is positive probability that the disease disappears.* Figures (4) and (6) below illustrate this fact. This holds regardless the value of β_Δ. However, even in this case an endemic regime, with statistical variations in time, may set in− much the same as in the unstable case (see Figures (5), (9) and (10) below).

Recall that if $J_0 < 1$ the origin is stable; however, this is not strong enough to guarantee that the disease eventually disappears. For this to be the case we need the following stronger condition.

Definition 3. The origin is *globally* stochastically asymptotically stable (g.a.s.) if it is stochastically asymptotically stable and for any initial condition $x_0 \in [0,1)$ it is:

$$\Pr\big(\lim_{t\to\infty} x(t;x_0) = 0\big) = 1\,. \tag{9}$$

In this case the *disease dies out no matter what the initial conditions are.*

We study when condition (9) holds in terms of the values of parameters. In principle, there are 5 free parameters and hence a parameter space

$$S := \{(\beta,\beta_\Delta,\sigma,\sigma_\Delta,\mu) : \beta>0, \beta_\Delta>0, \sigma>0, \sigma_\Delta>0, \mu>0\} = \mathbb{R}^5_+.$$

The problem of determining when the disease dies out regardless the initial conditions is non-trivial. Hence, we shall first consider some particular cases where the situation simplifies.

4. Proportional stochastic case: global stability conditions.

Here we consider the case when

$$\sigma_\Delta/\beta_\Delta = \sigma/\beta \tag{10}$$

We note that this condition is natural as it guarantees that both sources of noise (operating on β and β_Δ) have the same intensity. In this case the parameter

space is reduced to the set of positive points: $p = (\sigma, \beta, \beta_\Delta, \mu) \in \mathbb{R}^4$ where $\sigma, \beta, \beta_\Delta, \mu$ are strictly positive.

The following result precises the *boundaries* of the process– where motion takes place– and also its nature. We use Feller's classification (see [32]). Let τ be the first exit time from $(0,1)$. Then 0 is an attracting barrier if

$$\Pr\big(\lim_{t\to\tau} x(t;x_0) = 0\big) > 0,\ \forall x_0 \in (0,1) \tag{11}$$

and attainable if, additionally, $\Pr(\tau < \infty) > 0$.

Theorem 2. *Consider the SDE* (5)-(6).

1. *Suppose condition* (10) *is satisfied. Let* p *be the point in parameter space* $p = (\sigma, \beta, \beta_\Delta, \mu)$. *If* $\beta - \mu - \sigma^2/2 < 0$ *then* 0 *is an attracting, but not attainable, barrier. Otherwise* 0 *is a repelling (i.e., non-attracting) barrier. If* $\mu > 0$ *then* 1 *is a repelling barrier (the case* $\mu = 0$ *is considered briefly below).*
2. $(0,1)$ *is an invariant set of the dynamics. Concretely, if* $x_0 \in (0,1)$, *then* $x(t;x_0) \in (0,1)$ *for all* t.

Proof. From Proposition 2, $a(x) = b(x) = 0$ if and only if $x = 0$.

Analysis of this, shows then that around the origin the Feller scale function behaves as

$$\varphi(x) = x^{2(\mu-\beta)/\sigma^2} 0(1) + o(1)$$

while around $x = 1$ we have that $\varphi(x) \approx e^{2\mu/(\sigma_\Delta(x-1))} + o(x-1)$. The nature of the boundaries can then be discerned from this and Feller's theory ([32]). It follows that $(0,1)$ must be an invariant set of the dynamics, namely whenever $x_0 \in (0,1)$ then $x(t;x_0) \in (0,1)$ for all t. □

Thus when $J_0 < 1$, the origin 0 is an attracting, non-attainable, barrier, and also a locally stable e.p. We now give sharp conditions to guarantee that the origin is g.a.s.

Theorem 3. *Suppose condition* (10) *is satisfied. Let* p *the point parameter space* $p \equiv (\beta, \beta_\Delta, \sigma, \mu)$ *and denote* $\ell := \frac{(\beta_\Delta+\beta)^2}{4\beta_\Delta}$. *The origin in* (5) *is globally asymptotically stable if parameters satisfy*

1. *either* $\sigma > \sqrt{2\mu}$ *and* $p \in G_1$.
2. *or,* $\sigma \le \sqrt{2\mu}$ *and*

$$p \in G := G_1 \cup G_2 \cup G_3, \tag{12}$$

where we define the regions G_1, G_2, G_3 *as*

$$G_1 := \{p : \beta < \sqrt{2\mu}\sigma\} \tag{13}$$

$$G_2 := \{p : \beta \ge \beta_\Delta;\ \sigma^2 \le \beta < \sigma^2/2 + \mu;\ \beta \ge \sqrt{2\mu}\sigma\} \tag{14}$$

$$G_3 := \{p : \beta_\Delta > \beta;\ 2(\ell - \mu) < \frac{\ell^2\sigma^2}{\beta^2} < \ell;\ \beta \ge \sqrt{2\mu}\sigma\} \tag{15}$$

Remark 4. 1. *The proof of this result (which is quite technical) is given in Appendix B.*
2. *Equation* (12) *holds true in all cases. Nevertheles, if* $\sigma > \sqrt{2\mu}$ *the sets* G_2 *and* G_3 *vanish:* $G_2 \cup G_3 = \emptyset$ *and* (12) *simplifies.*
3. *While not strictly necessary, the inclusion of the condition* $\beta \geq \sqrt{2\mu}\sigma$ *results in having disjoint regions* G_1, G_2 *and* G_3.

Remark 5. *The condition* $\beta \geq \beta_\Delta; \sigma^2 \leq \beta < \sigma^2/2 + \mu$ *for g.a.s. for the full simplicial SIS system was already given in [56].*

We next clarify further the previous results and consider several particular cases.

First, if we let $\beta_\Delta \to 0, \sigma_\Delta \to 0$ we recover the SDE considered in [24] describing the *random SIS model, without simplicial interactions.* By letting $\beta_\Delta \to 0$ in Theorem (3), and noting that then $\ell \to \infty$ and $G_3 \to \emptyset$, we can polish the results of [24] and characterize global stability as follows

Corollary 1. *Suppose condition* $\beta_\Delta = \sigma_\Delta = 0$ *is satisfied and* $x(t)$ *evolves via*

$$dx(t) = \big((\beta - \mu)x - \beta x^2\big)\, dt + \sigma x\big(1 - x\big)dW(t). \tag{16}$$

The origin is globally asymptotically stable if parameters satisfy

1. *Either* $\beta < \sqrt{2\mu}\sigma$,
2. *or* $\sigma^2 \leq \beta < \sigma^2/2 + \mu$.

Alternatively, 0 *is g.a.s. if*

1. *either* $\sigma > \sqrt{2\mu}$ *and* $\beta < \sqrt{2\mu}\sigma$,
2. *or,* $\sigma \leq \sqrt{2\mu}$ *and* $\beta < \sigma^2/2 + \mu$.

Remark 6. *The condition* $\sigma^2 \leq \beta < \sigma^2/2 + \mu$ *is given in [24]. It is void when* $\sigma > \sqrt{2\mu}$. *The novel insight here is that* $\beta < \sqrt{2\mu}\sigma$ *also guarantees g.a.s.*

Remark 7. *To recover the deterministic SIS system we let* $\sigma \to 0, \sigma_\Delta \to 0$. *Here*

$$G_1 \to \emptyset, \quad G_2 \to \{p : \beta_\Delta \leq \beta < \mu\}, \quad G_3 \to \{p : \beta_\Delta > \beta; \ell < \mu\}, \tag{17}$$

which are the deterministic regions of Proposition 1 and [30].

4.1. Epidemiological meaningful parameters

Here we recover the general case when $\beta_\Delta > 0, \sigma_\Delta > 0$. Stability can be described in terms of the following set of parameters $\eta, J_0, J_{0\Delta}$

$$\eta := \frac{\sigma^2}{2\mu}, J_0 := \frac{\beta}{\mu + \sigma^2/2}; J_{0\Delta} := \frac{\beta_\Delta}{\mu + \sigma^2/2}$$

which have an epidemiological meaning as they are positive and could be interpreted as contagious rates. Concretely, J_0 *measures the average number of*

people to whom one single individual with the disease, will eventually spread the disease. The average number of people to whom a pair can jointly transmit the disease is $J_{0\Delta}$. Define also

$$J_0^* := \frac{2\sqrt{\eta}}{1+\eta}$$

In terms of these parameters Theorem 3 reads

Theorem 4. *Suppose condition* (10) *is satisfied. The origin in* (5) *is globally asymptotically stable (and hence disease dies out w.p. 1) if and only if parameters satisfy*

1. *either* $\eta \geq 1$ *and* $J_0 < J_0^*$,
2. *or* $\eta < 1$ *and* $(J_0, J_{0\Delta}) \in G := G_1 \cup G_2 \cup G_3$ *where*

$$G_1 = \{(J_0, J_{0\Delta}) : J_0 < J_0^*\} \tag{18}$$

$$G_2 = \{(J_0, J_{0\Delta}) : J_0^* \leq J_0 < 1, \quad J_{0\Delta} \leq J_0\} \tag{19}$$

$$G_3 = \{J_{0\Delta} > J_0\} \cap D_2^- \tag{20}$$

and D_2^- *is the interior of the curve*

$$D_2 = \{(J_0, J_{0\Delta}) : \ \big(1+\eta - 4J_{0\Delta}(J_{0\Delta}+J_0)^{-2}\big)4J_{0\Delta}J_0^2 = \eta(J_{0\Delta}+J_0)^2\}. \tag{21}$$

Figures 2 and 3 show the situation. Note how when $\eta < 1$, G is made up of the unit square, the strip $J_0 < 1$ and the interior of the "belly" D_2^-.

Remark 8. Hence, the parameter space can be partitioned as follows: $\{J_0 < 1\}$ is the stability region while the complimentary set $\{J_0 \geq 1\}$ is the unstability region. The former can be partitioned into two sub-regions: $G_1 \cup G_2 \cup G_3$, the g.a.s. regions, where the disease dies out: $\Pr(x_\infty = 0) = 1$; besides $\{J_0 < 1\} - \big(G_1 \cup G_2 \cup G_3\big)$ is the bi-stable region where the disease may die out: $0 < \Pr(x_\infty = 0) < 1$.

Table 1 summarizes all possibilities

4.2. The case when $\mu = 0$

Recall that in this paper we assume a strictly positive recovery rate $\mu > 0$. Nevertheless the case with $\mu = 0$, corresponding to a SI model, is also of interest. In a deterministic situation such evolution is asymptotically trivial: $\lim_{t\to\infty} x(t; x_0) = 1, \forall x_0 \neq 0$. However, when stochasticity is introduced the situation is more complex. Here (5) has two critical points: $x = 0$ (disease free) and $\hat{x} = 1$ and evolutions starting from $x_0 > 0$ need not be asymptotic to $x(\infty) = 1$.

Proposition 3.

1. *If* $\mu = 0$, (5) *has two critical points:* $x = 0$ *(disease free) and* $\hat{x} = 1$.

J_0	$p = (J_0, J_{0\Delta})$	0	$\Pr(x(\infty) = 0)$
≥ 1	–	Unstable	0
< 1	$\notin G$	Locally stable	> 0
$\in [J_0^*, 1)$	$\in G_2 \cup G_3$	Globally stable	1
$< J_0^*$		Globally stable	1

Table 1: Stability regions of the stochastic system.

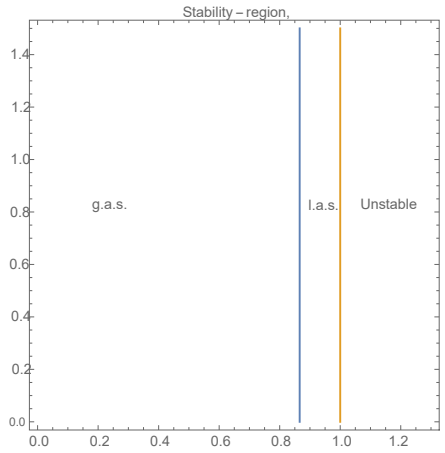

Figure 2: The region of global stochastic stability $G_1 = \{0 \leq J_0 < 2\sqrt{\eta}/(1+\eta)\}$ when condition (10) holds and $\eta = \sigma^2/(2\mu) = 3.5 > 1$. Horizontal and vertical axis represent, respectively, the values of J_0 and $J_{0\Delta}$.

2. 1 *is always an attractive non-attainable barrier and a locally asymptotically stable e.p.*
3. *If* $J_0 := \frac{2\beta}{\sigma^2} < 1$ *the origin is an attractive non-attainable barrier and a locally asymptotically stable e.p. for equation* (5).
4. *If* $J_0 \geq 1$ *the origin is repelling barrier and a locally asymptotically unstable point. In this case there are two different stable equilibrium points.*

We skip the proof but note that if $\mu = 0$, then

$$b(x) = -\sigma_\Delta x(x-1)(x + \sigma/\sigma_\Delta) \text{ and } a(x) = -\beta_\Delta x(x-1)(x + \beta/\beta_\Delta).$$

It develops from the above that if $x_0 \neq 0$, then $\lim_{t\to\infty} x(t; x_0)$ is either 0 or 1, i.e. a Bernoulli's binary variable such that

$$\Pr\big(\lim_{t\to\infty} x(t;x_0) = 1\big) = 1 - \Pr\big(\lim_{t\to\infty} x(t;x_0) = 0\big) = \begin{cases} 1 & \text{if } J_0 \geq 1 \\ < 1 & \text{if } J_0 < 1 \end{cases}$$

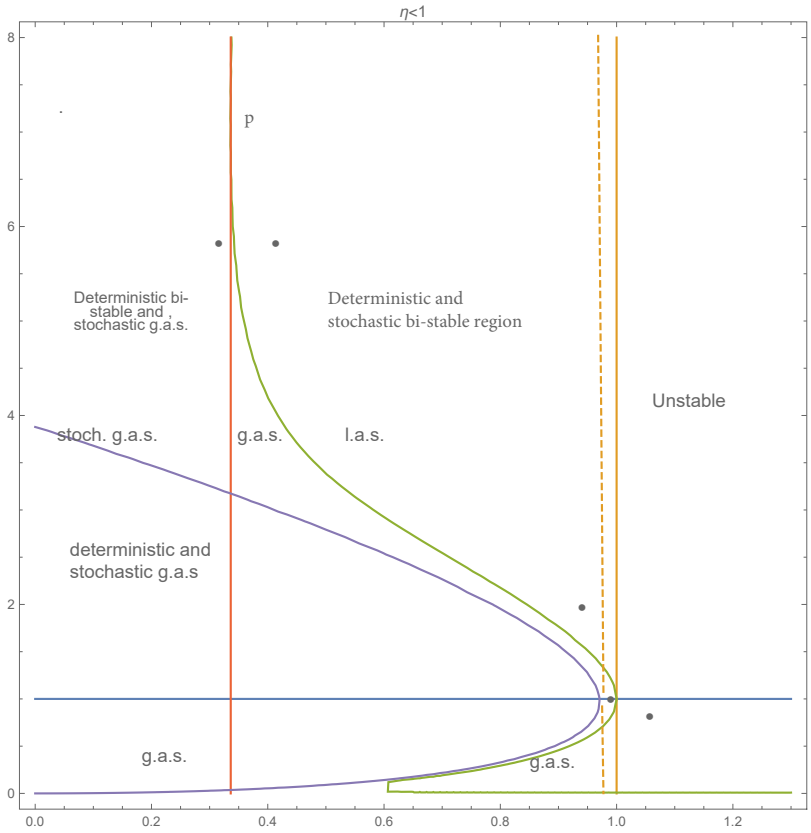

Figure 3: The regions of global and local stability in the stochastic case when $\mu = 1, \eta = \sigma^2/2 = 0.03$. Horizontal and vertical axis represents, respectively, the values of J_0 and $J_{0\Delta}$. Note also the biestable region when stability is only local (l.a.s.). Dots indicate value of the parameters for which we provide a sample path obtained from (5). We also include a comparison with the stability regions in the deterministic case, where $\mu = 1, \sigma = 0$, whose boundaries are the blue curve and the dashed vertical line

4.3. Numerical simulations using a Runge Kutta method

To illustrate the above results, different sample paths are produced solving (5) (under the condition (10)) with Runge-Kutta numerical methods.

Notice that how− except for the last case− the l.a.s. condition $R_0 < 1$ or $J_0 < 1$, is satisfied (for both deterministic and stochastic cases), but not that of g.a.s. Despite this, note how the limit *needs not to be the disease free e.p.*. Notice also the great variability of the possible evolutions.

Figure 9 considers the situation when $J_0 = 0.6 < 1$ so two different stable equilibrium points coexist. Notice how different initial conditions have different limits, the stable e.p. $x = 0, x = 1$

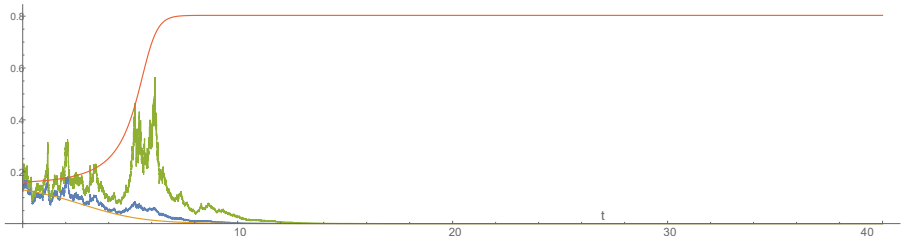

Figure 4: Sample paths for equation (5) under two different initial conditions $x_0 = 0.12$ and $x_0 = 0.15$. They have been obtained when $J_0 = 0.254902 < 0.277297 \equiv J_0^*, J_{0\Delta} = 5.88, \eta = 0.04$ ($\mu = 1, \sigma = 0.2$) We have also plotted the corresponding deterministic paths that follow evolving via (2) for large time. This corresponds to a region of (stochastic) g.a.s and deterministic bi-stability where $\lim_{t\to\infty} x(t) = 0$ or x_+ (the endemic state) .

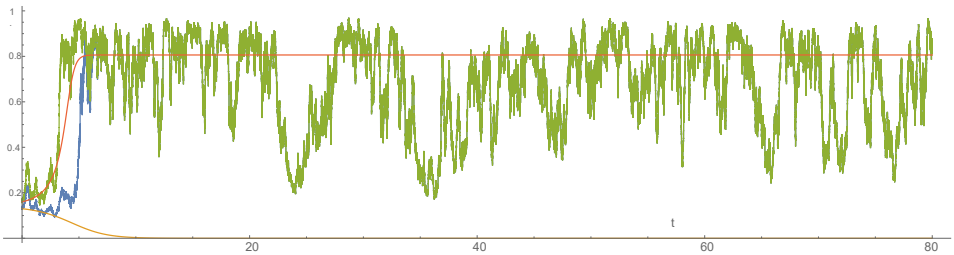

Figure 5: We take $J_0 = 0.317 > J_0^* \equiv 0.183, J_{0\Delta} = 5.95, \sigma = 0.13$. We have plotted the stochastic and also the deterministic paths that follow evolving via (2) for large time $t = 80$. This corresponds to a bi-stability region for both stochastic and deterministic cases. Note how in the former the disease-free regime never sets in.

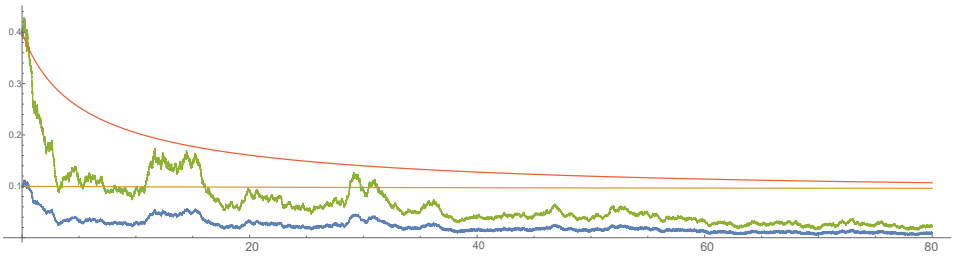

Figure 6: We take $J_0 = 0.99, J_{0\Delta} = 0.98, \sigma = 0.2$. This corresponds to a deterministic bi-stability region, and stochastic g.a.s. region.

5. The general case.

Study of g.a.s. conditions in a general situation where condition (10) needs not to hold is quite involved. Nevertheless, we give here conditions in this regard.

In this case the bulk of Proposition 4 and Theorem 6 (see Appendix B) still hold and g.a.s. is determined in terms of the existence of roots to a quartic function, cf. (22). However, it is not easy to determine from here the complete

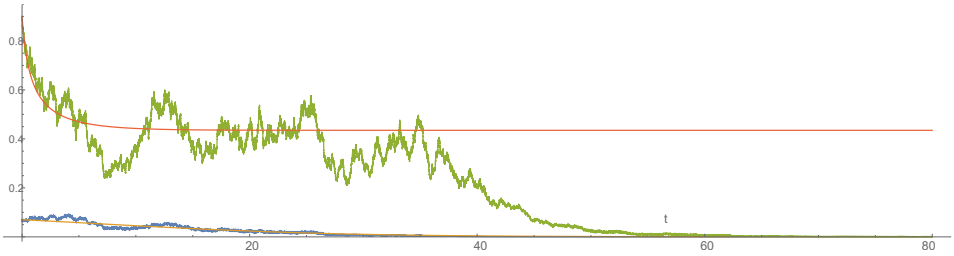

Figure 7: Sample paths corresponding to a region of (stochastic) bi-stability (l.a.s.) region, just on the outside of the "belly": $J_0 = 0.88, J_{0\Delta} = 1.96, \sigma = 0.2$. Here, if $x_0 << 1$ the solution converges to zero with a high probability.

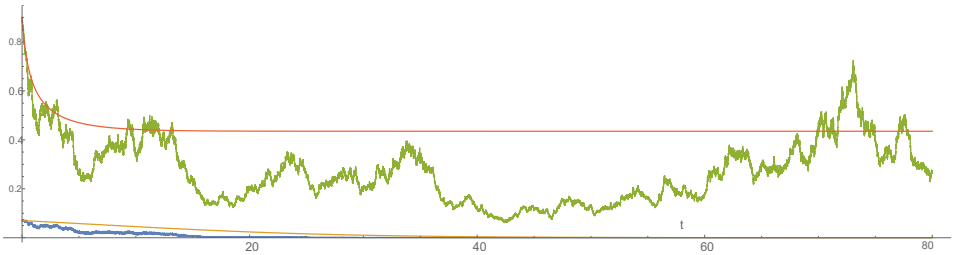

Figure 8: We take $J_0 = 0.88, J_{0\Delta} = 1.96, \sigma = 0.2$. This corresponds to a bi-stability (l.a.s.) region, just on the outside of the "belly". Note how different initial conditions have given rise to different evolutions: for one of them a disease free regime sets in very quickly while it never seems to be the case for the other.

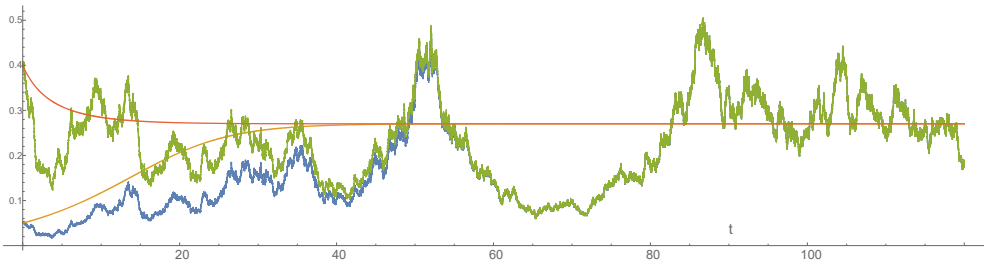

Figure 9: We take $J_0 = 1.08, J_{0\Delta} = 0.98, \sigma = 0.2$. This corresponds to unstability regions in both deterministic and stochastic cases: $p \in S_4$ so the disease becomes endemic under the deterministic evolution.

stability region in a neat way. We have

Theorem 5. *Suppose that there exists $r > 0$ such that the quartic function*

$$Q(x;r) := (\mu - \beta) + (\beta - \beta_\Delta)x + \beta_\Delta x^2 + \frac{1-r}{2}\left(\sigma + (\sigma_\Delta - \sigma)x - \sigma_\Delta x^2\right)^2 \quad (22)$$

has no real roots on $[0,1]$. Then the origin in (5) is globally stochastically asymptotically stable. In particular, the origin is g.a.s. if parameters satisfy

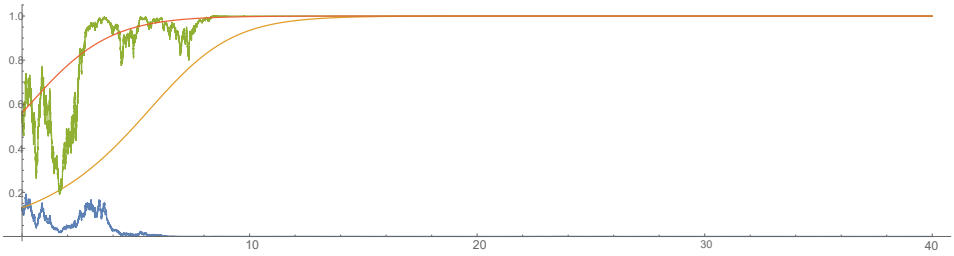

Figure 10: We have taken $\mu = 0, \beta = \beta_\Delta = 0.3, \sigma = 1$ giving $J_0 = 0.6 < 1$. Here $\mu = 0$, so under the stochastic evolution there are two stable e.p. $x = 0, x = 1$.

one of the conditions

- C1
$$\beta \leq \beta_\Delta \sigma/\sigma_\Delta < \sqrt{2\mu}\sigma$$
- C2
$$\beta_\Delta \sigma/\sigma_\Delta \leq \beta < \mu + \beta_\Delta \sigma/\sigma_\Delta - (1/2)(\beta_\Delta/\sigma_\Delta)^2$$
- C3
$$\sigma^2 \leq \beta_\Delta \sigma/\sigma_\Delta \leq \beta < \sigma^2/2 + \mu, \text{ and } \sigma \geq \sigma_\Delta$$
- C4
$$\sigma^2 \leq \beta \leq \beta_\Delta \sigma/\sigma_\Delta < \sigma^2/2 + \mu, \text{ and } \sigma \geq \sigma_\Delta$$
- C5: $\sigma \leq \sigma_\Delta$, and
$$(\sigma + \sigma_\Delta)^2 \leq 4\beta_\Delta \sigma/\sigma_\Delta \leq 4\beta, (\sigma + \sigma_\Delta)^2\big(\beta_\Delta - (\sigma + \sigma_\Delta)^2/8\big) \leq 4\mu\sigma_\Delta^2$$
- C6: $\sigma \leq /\sigma_\Delta, \beta_\Delta \sigma \geq \beta\sigma_\Delta$ and
$$(\sigma + \sigma_\Delta)^4 - 8\beta_\Delta(\sigma + \sigma_\Delta)^2 + 32\mu\sigma_\Delta^2 + \beta_\Delta \sigma\sigma_\Delta > 32\sigma_\Delta^2\beta$$
- C7 (deterministic region): $\beta_\Delta \leq \beta \leq \mu$
- C8 (deterministic region): $(\beta + \beta_\Delta)^2 - 4\mu\beta_\Delta < 0$

Remark 9.

1. *The above conditions are sufficient but not necessary conditions for g.a.s. We note that they need not be excluding.*
2. *When $\beta = \beta_\Delta \sigma/\sigma_\Delta$ they are both sufficient and necessary. In this case both C1 and C2 reduce to $\beta < \sqrt{2\mu}\sigma$. Besides C3 and C4 simplify to $\sigma^2 \leq \beta < \sigma^2/2 + \mu$, and $\beta \geq \beta_\Delta$. Finally, C5 and C6 give the region G_3 of* (15)*. The deterministic regions defined by conditions C7 and C8 are subregions of $G_1 \cup G_2 \cup G_3$. Hence theorem* (3) *is recovered.*
3. *The idea od the proof is similar to that of theorem* (3) *and we skip it.*

6. Stochastic simplicial contagion model and empirical results.

6.1. The stochastic simplicial contagion model

We propose a stochastic simplicial contagion model (SSCM) which we will use to run simulations in real-world and synthetically generated simplicial networks. With it, we aim to model situations in which the probability of infection of a susceptible individual, defined by the different higher-order channels of the simplicial contagion model (SCM) of [30], could be perturbed at each instant of time by random situations, such as unforeseen or unplanned encounters of that susceptible individual with one or more infected individuals.

We model this type of situation using a single white Gaussian noise ξ_t following a standard normal distribution to perturb both infection parameters β_1 and β_2 with different intensities σ_1 and σ_2. In other words, the SSCM is based on the SCM of [30] and consists of replacing their infection parameters β_1 and β_2 by the time-dependent random variables $\beta_1 + \sigma_1\xi_t$ and $\beta_2 + \sigma_2\xi_t$.

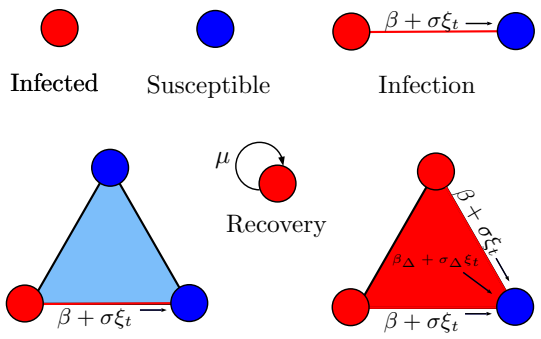

Figure 11: Channels of infection of the stochastic simplicial contagion model (SSCM)

.

In order to be able to simulate all possible channels of infection given in Figure 11, for every simulation we first parse the selected dataset to obtain a simplicial complex of dimension $D = 2$, with its nodes, edges (connections between two nodes) and triangles (connections between three nodes). In [27], we have published and made publicly available parsers for the cliques analyzed in [30], the real-world datasets provided in [9, 25], and for randomly generated simplicial complexes (which are detailed in Section 6.3). The code for the SSCM described below is also available at the very same repository. It is based on the one provided in [30] but modified to make it support larger datasets and higher connection degrees. It also supports our stochastic perturbation of the contagion probabilities.

Before running the SCCM, a simulation must be prepared. As explained above, input parameters for a simulation of our SSCM requires: a recovery probability μ; infection probabilities β_1 and β_2, and their corresponding intensities of the stochastic perturbation σ_1 and σ_2. Moreover, for every simulation we

have to set an initial density x_0 of infected nodes, which results in $\lfloor x_0+0.5 \rfloor$ randomly chosen nodes set as infected. Obviously, the rest of nodes are considered to be susceptible of being infected.

Once the simulation is ready, in each time step, we generate a single Gaussian noise ξ_t following a standard normal distribution. For every susceptible node we run through the list of its neighbors; in case a neighbor is infected, the susceptible node can be infected with probability $\beta_1 + \sigma_1 \xi_t$. In addition, we run through the list of its triangles; in case the triangle is infected (the other two nodes are infected), the susceptible node can be infected with probability $\beta_2 + \sigma_2 \xi_t$. For the sake of protocol efficiency, once a susceptible node is infected we stop processing potential infections for that node; for a node cannot be infected more than once. When the time-step simulation for each susceptible node is over, for every node that was infected at the beginning of the time-step execution, we recover it with probability μ, and finally, the next time step can be executed. The simulation stops when no infected nodes are left or when a maximum predefined amount of time steps is performed.

6.2. The SSCM on real world simplicial datasets.

In [30] numerical simulations were performed on both real and synthetically generated simplicial networks. For the simulations on real networks, they used four datasets of face-to-face interactions collected from social contexts such as a workplace, a conference, a hospital or a high school. From the aggregated simplicial data, and given that the study is done for simplicial complexes of maximum dimension $D = 2$, they consider the 2- and 3-cliques and weight them according to their frequency of occurrence (forgetting the other higher-dimensional simplices); they then retain with 20% of the simplices that occur more often. To reduce the finite size effects, they use an augmented version of the simplicial data in which they duplicate five times the list of the resulting facets and use this as input dataset for the simplicial contagion model. As mentioned there, it would be interesting to investigate the simplicial contagion model on simplicial complexes with heterogeneous generalized degree distribution.

Following this suggestion, in this paper we will use instead some of the real-world simplicial networks of [9, 25] to carry out our analysis of the dynamics of evolution in empirical social simplicial structures. These datasets (publicly available at `https://github.com/arbenson/ScHoLP-Data`) are already collected as simplices bounded to a maximum of 25 nodes, and they are obtained from real world data of different domains. The following list summarizes the datasets in terms of the data that is relevant to our study, i.e., the datasets' simplices and nodes.

contact-high-school, contact-primary-school. Nodes are people. Simplices are sets of people that were connected during 20-second intervals.

email-Eu. A simplex is an email. The nodes of a simplex are the recipient addresses along with the sender one.

NDC-classes. Each simplex corresponds to a national drug code (NDC) for a drug. Nodes are the classes applied to that drug (NDC-classes) or the substances that make up that drug (NDC-substances).

	nodes	simplices	distinct simplices	unordered distinct simplices	facets	% facets / simp	% facets / u.d.simp.
contact-high-school	327	172,035	7,937	7,818	4,862	2.83%	62.19%
contact-primary-school	242	106,879	12,799	12,704	8,010	7.49%	63.05%
NDC-classes	1,161	49,724	1,222	1,088	563	1.13%	51.75%
email-Eu	998	234,760	25,791	25,027	8,102	3.45%	32.37%

Table 2: Summary statistics obtained from the analysis of the datasets. Column 'simplices' correspond to the data available in the datasets; column 'distinct simplices' represents the number of simplices where duplicates are discarded; the 'unordered distinct simplices' are counted by deleting the 'distinct simplices' that have the same set of vertices (regardless of their order) than another distinct simplex; and the 'facets' are calculated by removing the unordered distinct simplices that were already faces of another unordered distinct simplex.

Table 2 shows the summary statistics of the dataset and Table 3 shows the facet size statistics (recall that if a facet is given by a simplex $\sigma^{(q)}$, then its size is $s = q + 1$). Arithmetic mean Sum of values of a data set divided by number of values $(1 + 2 + 2 + 3 + 4 + 7 + 9)/7$.

Table 3: Facet size statistics. We show the maximum, average $\langle s \rangle$, median s_m and mode s_p of the facet.

	max s	$\langle s \rangle$	s_m	s_p
contact-high-school	5	2.42	2	2
contact-primary-school	5	2.57	3	3
NDC-classes	24	5.54	4	2
email-Eu	25	3.82	2	2

Remark 10. *Note how the mean almost doubles, or triples, the mode in several cases; this deviation can be put down to the fact that for real-world simplicial networks we expect having scale-free tails, and hence a large number of outliers, which pull away the mean from the mode.*

In addition, and in order to compare with the results of the deterministic simplicial social contagion model of [30], we have included in the study one of the datasets they use: InVS15 (contacts in a workplace).

The generalized degree distributions of the real world datasets associated with both the classical node degree of a node v, k_1 and the node-to-triangle upper simplicial degree k_2 (the number of 2-simplices a vertex belongs to) are given in Figures 12, 13 and 14. Note the heterogeneity of the distributions.

When running the SSCM, we start with an initial density x_0 of infectious nodes. The simulation stops if an absorbing state is reached, otherwise it is stopped at a fixed number of time-steps if an endemic stationary state is achieved. Each trajectory represents an actual temporal evolution of the densities of infectious nodes with a particular value of x_0 and we plot them by

Table 4: We show the average classical node degree and the average node-to-triangle degree of the datasets.

	$\langle k_1 \rangle$	$\langle k_2 \rangle$
contact-high-school	35.58	21.74
contact-primary-school	68.74	63.71
NDC-classes	10.72	81.34
email-Eu	58.71	482.78
InVS15	24.42	6.92

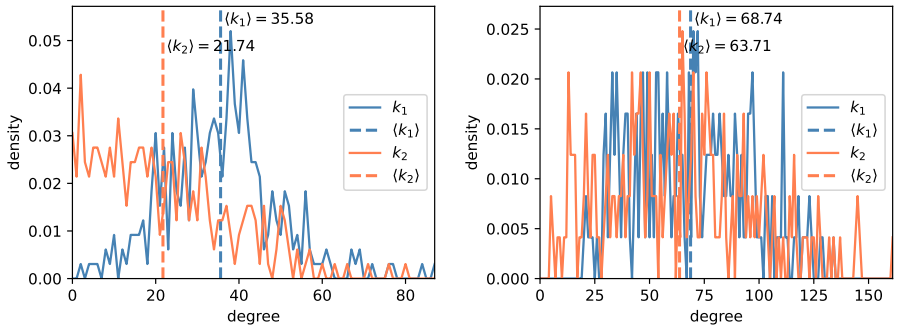

Figure 12: Linear plot of the degree distribution for the real-world simplicial datasets contact-high-school (left) and contact-primary-school (right). Vertical dashed lines represents the statistical means $\langle k_1 \rangle$ and $\langle k_2 \rangle$

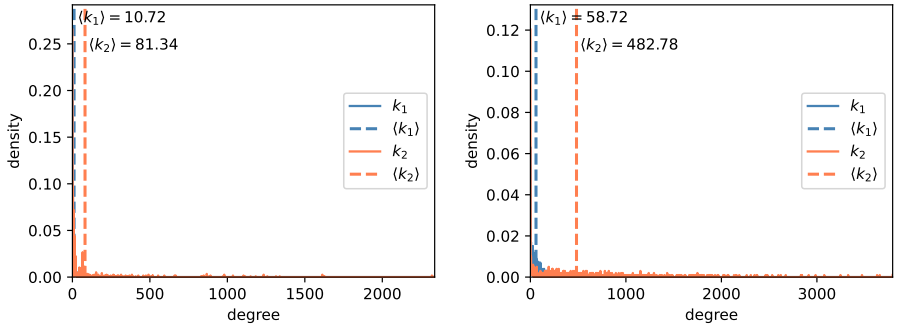

Figure 13: Linear plot of the degree distribution for the real-world simplicial datasets NDC-classes (left) and email-Eu (right).

different colored curves. We perform different simulations corresponding to different ranges of the infectivity parameters.

In order to make these notes clearer and more concise, we detail below the simulations for the contact-high-school and email-Eu simplicial networks and include the rest of the simulations in the Appendix C.

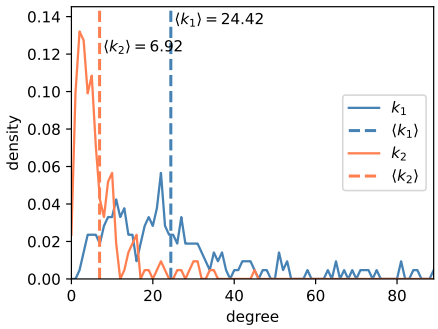

Figure 14: Linear plot of the degree distribution for the real-world simplicial dataset InVS15.

6.2.1. Simulations for the region G_1.

We fix the parameters

$$\mu = 1\,,\ \sigma = 0.2\,,\ J_0 = 0.2549\,,\ J_{0\Delta} = 5.88\,. \tag{23}$$

We have that $\eta = 0.02 < 1$ and since $J_0 < J_0^* = 0.2772$ we are indeed in region G_1 (see equation (18) and Figure 3). Thus, by Theorem 3 the origin of equation (5) has to be globally asymptotically stable, a fact confirmed by the graphics on the left in Figure 15 (see also Figures C.27 and C.28 in Appendix C), where different realizations eventually hit 0, so the cloud of trajectories becomes thinner over time.

However, letting $\sigma_1 = 0$ one escapes the stochastic g.a.s. region to enter a deterministic bi-stable region. Indeed, here one has that $R_0 = 0.2599$ and $R_{0\Delta} = 5.9975$, within the bistability region S_3 (see Figure 1); according to Proposition 1 (see also [30]) in this regime both the disease-free state and the endemic state can coexist. This is confirmed by the graphics on the right of Figure 15 (see also Figures C.27 and C.28 in Appendix C) obtained running the SCM of [30]).

This simulation suggests that the introduction of noise in a contagion process can stabilize the possible endemic solution of a deterministic system.

For the following datasets, we will slightly modify the parameters by increasing only the intensity, which allows us to observe the stochastic effect in fewer time-steps. We fix parameters:

$$\mu = 1\,,\ \sigma = 1.4\,,\ J_0 = 0.2549\,,\ J_{0\Delta} = 5.88\,. \tag{24}$$

We have that $\eta = 0.98 < 1$ and since $J_0 < J_0^* = 0.99$ we are again in region G_1. As shown in Figure 16 (see also Figure C.29 in Appendix C), we observe the same behavior as in the previous datasets.

Note that if we increase the intensity, the probability that the epidemic disappears increases and the trajectories would fall faster, which is consistent with the fact that by increasing σ we are making η larger, and therefore approaching the Figure 2, in which the stochastic bistability region (l.a.s.) is much smaller.

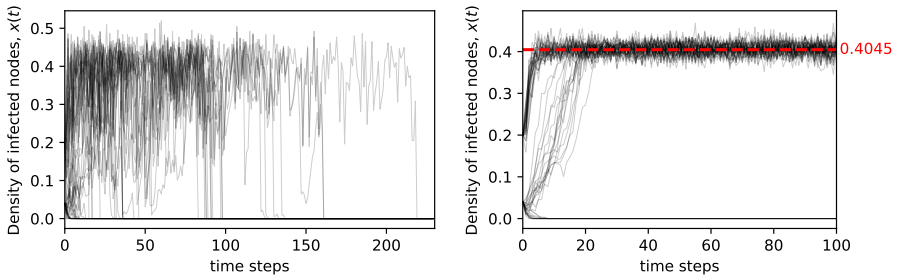

Figure 15: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-high-school. 30 trajectories are plotted for $x_0 = 0.04$ and 30 for $x_0 = 0.2$. Parameter values are in the region G_1.

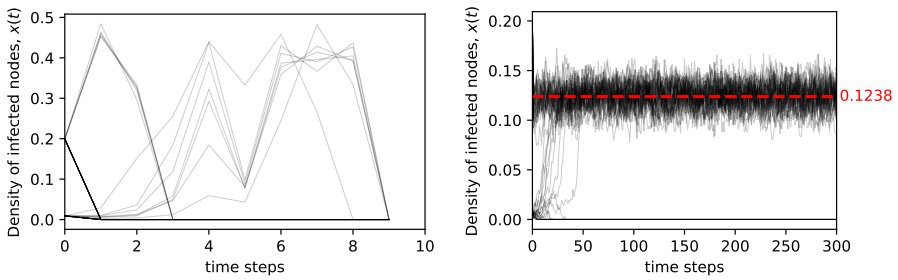

Figure 16: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network email-Eu. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.2$. Parameter values are in the region G_1.

Let us also notice that letting $\sigma_1 = 0$ in equation (5), the stochasticity disappears and a deterministic evolution results. This contrasts with the network evolution where running the SCM always provides a stochastic evolution, since at very step contagious happen in a a random way− even though we refer formally to the case in which the intensities are zero as a deterministic case. This fact can be observed in the right graphics of Figure 15 (see also Figures C.27 and C.28 in Appendix C), which display the typical sharp wiggles of a stochastic process; in particular limit states are random and a certain percentage of trajectories starting from $x_0 = 0.04$ end up in the endemic state. However, in this deterministic case, and when living in a deterministic bistable region, no trajectory is observed to fall once it reaches an endemic state and the reason is that since the probability of contagion is constant over time for the deterministic system, although there might be a probability that the disease disappears (that some trajectory would fall) when running the SCM on a dataset, it is almost negligible. As we will see in the next example, this is the main difference between the stochastic bistable region and the bistable deterministic one.

6.2.2. Simulations for the stochastic bistable region.

We fix parameters:

$$\mu = 1\,,\ \sigma = 0.2\,,\ J_0 = 0.88\,,\ J_{0\Delta} = 1.96\,. \tag{25}$$

We have that $\eta = 0.02 < 1$ and the parameter values are in region $D_2^+ \cap \{J_0 < 1\}$ and thus we are in stochastic bistable region (see equation (21), Figure 3 and Table 1). In the left graphic of Figure 17 it is seen that 4000 time steps are not enough for the epidemic to disappear in the stochastic case, while in the graphic on the left of Figure 19 it had practically disappeared. Notice also that in the left graphic of Figure C.30 of Appendix C we see that the disease has disappeared for the SSCM and that it has taken more time steps to do so than in left graphic the Figure C.32 of Appendix C.

For the left graphic of Figure 18 (see also Figure C.31 in Appendix C) we still see a stochastic bistable regime.

Making $\sigma_1 = 0$ we have $R_0 = 0.8976$ and $R_{0\Delta} = 1.9991$ and then we are in the deterministic bistable region S_3 (see Figure 2), in which both the disease-free state and the endemic state can coexist (depending on the initial number of infected of the system). This is confirmed by the graphics on the right of Figures 17and 18 (see also Figures C.30, C.31 in Appendix C).

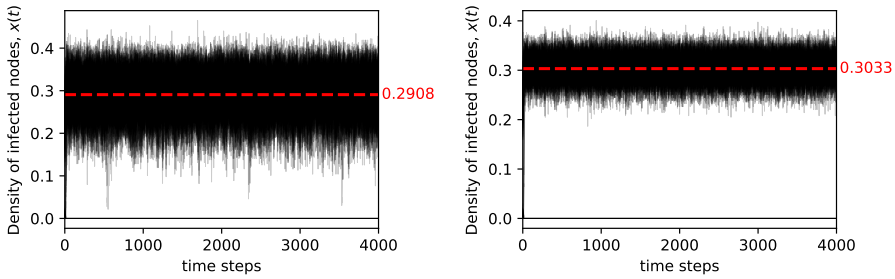

Figure 17: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-high-school. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.2$. Parameter values are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

Now let's increase the intensity to observe the stochastic effect at a faster rate. We fix the parameters with:

$$\mu = 1\,,\ \sigma = 0.41\,,\ J_0 = 0.88\,,\ J_{0\Delta} = 1.96\,. \tag{26}$$

We have that $\eta = 0.08 < 1$ and the parameter values live in the region $D_2^+ \cap \{J_0 < 1\}$ (see (21) and Figure 3), which is a stochastic bistable region. Hence any initial configuration may wind up in 0 with a rather high probability, as Figures 7 and 8 suggest. This is observed too in graphics on the left of Figures 19 and 20 (see also Figures C.32, C.33 in Appendix C) where most of the trajectories starting from $x_0 = 0.2$ disappear in finite time.

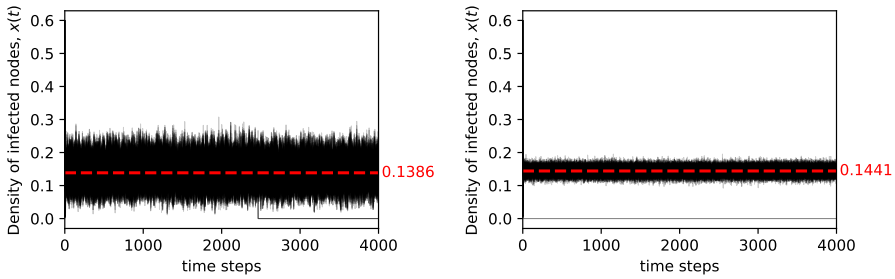

Figure 18: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network email-Eu. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.6$. Parameter values are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

Making $\sigma_1 = 0$ we have $R_0 = 0.9539$ and $R_{0\Delta} = 2.1247$ and then we are in the deterministic bistable region S_3 (see Figure 1), in which both the disease-free state and the endemic state can coexist (depending on the initial number of infected nodes of the system). This is confirmed by the graphics on the right of Figures 19and 20 (see also Figures C.32, C.33 in Appendix C).

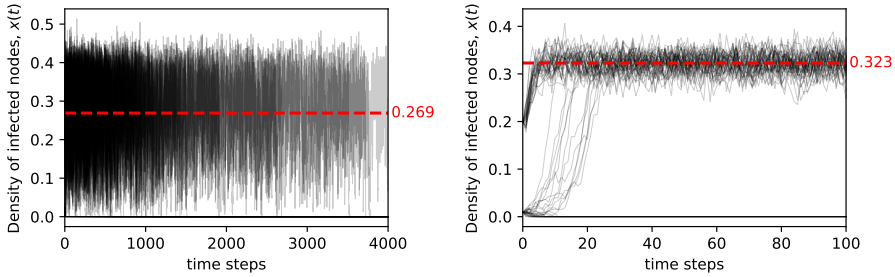

Figure 19: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-high-school. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.2$. Parameter values are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

Note in particular the drastic change in the dynamic behavior of the email-Eu dataset as the intensity σ increases from 0.2 to 0.4. For the latter case, the probability of the disease disappearing has increased considerably, as shown by the left graphic of the Figure 20 (compare with Figure 18 left).

6.2.3. Simulations for the unstable region.

We fix the parameters

$$\mu = 0.5\,,\ \sigma = 0.2\,,\ J_0 = 1.2\,,\ J_{0\Delta} = 1.2\,. \tag{27}$$

We have that $\eta = 0.04 < 1$ and the parameter values are in the unstable region $\{J_0 > 1\}$ of Figure 3, a fact confirmed by the left graphics in Figures 21 and 22

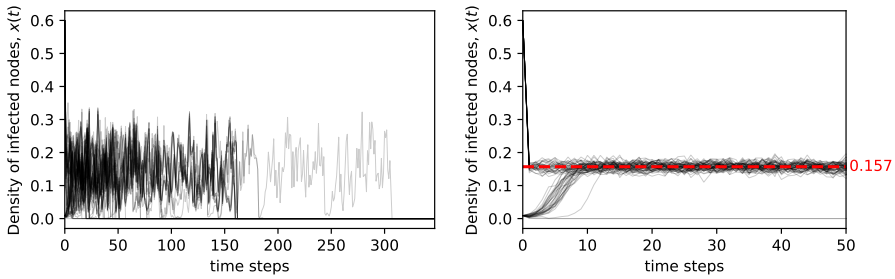

Figure 20: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network email-Eu. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.6$. Parameter values are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

(see also Figures C.35, C.36 andC.37). Making $\sigma_1 = 0$ we have $R_0 = 1.248$ and $R_{0\Delta} = 1.248$ and then we are in the deterministic unstable region, as confirmed by the graphics on the right of Figures 21 and 22 (see also Figures C.35, C.36 andC.37).

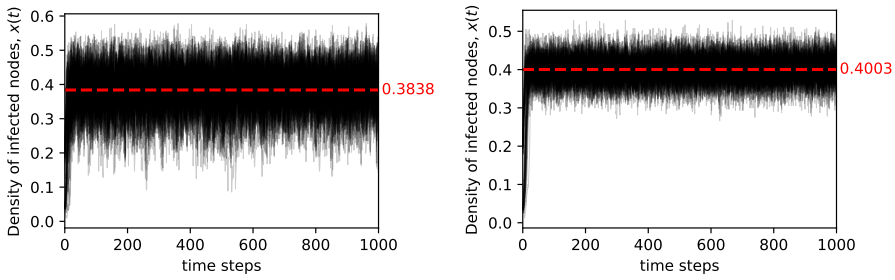

Figure 21: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-high-school. 30 trajectories are plotted for $x_0 = 0.04$ and 30 for $x_0 = 0.2$. Parameter values are in the unstable region.

6.3. The SSCM on a synthetic simplicial network.

We generate synthetic simplicial networks using the random simplicial complex (RSC) model of dimension $D = 2$ of [30]. That is, once we have $\langle k_1 \rangle, \langle k_2 \rangle$ a simplicial complex is generated using the inputs $(N, \{p_1, p_2\})$ where N is the number of vertices and p_i are the probabilities defined by:

$$\begin{aligned} p_1 &= \frac{\langle k_1 \rangle - 2\langle k_2 \rangle}{(N-1) - 2\langle k_2 \rangle} \\ p_2 &= \frac{\langle k_2 \rangle}{\binom{N-1}{2}} \end{aligned} \tag{28}$$

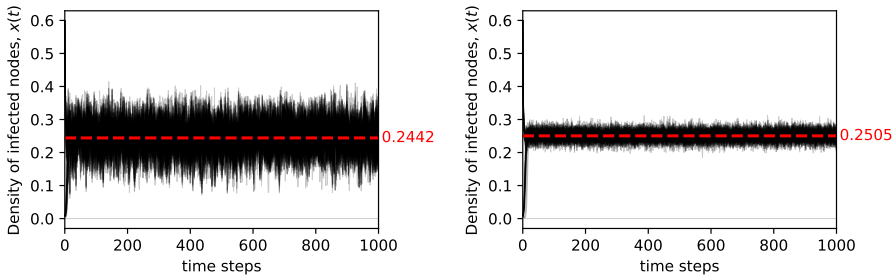

Figure 22: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network email-Eu. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.6$. Parameter values are in the unstable region.

Notice that in order for $0 \leq p_1 \leq 1$ we have one of the following constrains:

$$\begin{cases} \text{If } \langle k_2 \rangle \leq \frac{\langle k_1 \rangle}{2} \implies \langle k_2 \rangle \leq \frac{N-1}{2} \\ \text{If } \langle k_2 \rangle > \frac{\langle k_1 \rangle}{2} \implies \langle k_2 \rangle > \frac{N-1}{2} \end{cases} \quad \text{and} \quad \langle k_1 \rangle = N - 1$$

These constraints are satisfied for the synthetically generated simplicial networks used in [30] (where $\langle k_1 \rangle = 20$ and $\langle k_2 \rangle = 6$, degrees which are close to the ones of their real simplicial datasets). However, in many real networks the first constraint is not satisfied (such as all the datasets we use but InVS15, see Table 4) and in order to satisfy the second condition, we would have to generate a synthetic network with very few nodes, which would make the model simulation not sufficiently representative. It would therefore be interesting to study a simplicial network generation method in order to cover these cases.

Thus, in order to compare with the results obtained in the InVS15 dataset (also used in [30] for the SCM), we will generate a simplicial network using the RSC for $D = 2$ with the same average degrees ($\langle k_1 \rangle = 24.42$ and $\langle k_2 \rangle = 6.92$) as those given in InVs15 (which satisfy $\langle k_2 \rangle \leq \frac{\langle k_1 \rangle}{2}$). The resulting synthetic simplicial network has the degree distribution given by Figure 23.

Remark 11. *The general RSC model of order D allows to maintain the average node-to-h-simplex degrees (for $h \leq D$), and thus it permits to tune the expected local connectivity of the simplicial network. For the general D-order case we have that given the number of nodes $N \in \mathbb{N}$ and reals $\langle k_1 \rangle, \ldots, \langle k_D \rangle \in \mathbb{R}$ satisfying*

$$\langle k_i \rangle \leq \binom{N-1}{i}; \langle k_i \rangle \geq \sum_{j=i+1}^{D} j \langle k_j \rangle, i = 1, \ldots D-1, \tag{29}$$

we can synthetically generate a simplicial complex by using the RSC with inputs $(N, \{p_1, \ldots, p_D\})$ where N is the number of nodes and p_i are the probabilities

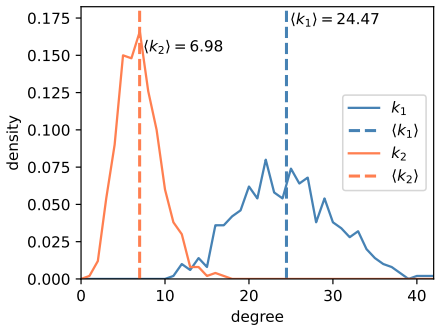

Figure 23: Linear plot of the degree distribution of a synthetic simplicial network generated by the RSC for $D = 2$.

defined by:

$$p_i = \frac{\langle k_i \rangle - \sum_{j=i+1}^{D} j\langle k_j \rangle}{\binom{N-1}{i} - \sum_{j=i+1}^{D} j\langle k_j \rangle} \qquad \textit{for } i = 1, \ldots, D-1.$$

$$p_D = \frac{\langle k_D \rangle}{\binom{N-1}{D}}$$

As in the previous section, we run the stochastic SCM on this randomly generated simplicial dataset and perform the different examples corresponding to different ranges of the infectivity parameters. We obtain similar results (see Figures 24, 25 and 26) to those of the real-world dataset InVS15 (with whose degrees we have generated this synthetic network) and also remarkable similar to those of the other real-world dataset of the previous subsection and with the ones predicted by the theoretical results. Notice that since our ISDE was based on a MF approach under homogeneous mixing assumption, these results could be expected to more closely approximate the theoretical ones.

6.3.1. Simulations for the region G_1.

We fix the parameters

$$\mu = 1\,,\ \sigma = 0.2\,,\ J_0 = 0.2549\,,\ J_{0\Delta} = 5.88\,. \tag{30}$$

We have that $\eta = 0.02 < 1$ and since $J_0 < J_0^* = 0.2772$ we are indeed in region G_1 (see equation (18) and Figure 3). This is a g.a.s. stochastic region and a deterministic bistable region as shown by Figure 24.

6.3.2. Simulations for the stochastic bistable region.

We fix parameters:

$$\mu = 1\,,\ \sigma = 0.57\,,\ J_0 = 0.7\,,\ J_{0\Delta} = 6\,. \tag{31}$$

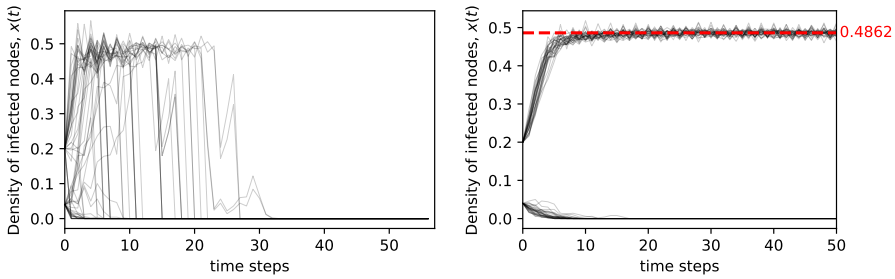

Figure 24: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the synthetic simplicial network. 30 trajectories are plotted for $x_0 = 0.04$ and 30 for $x_0 = 0.2$). Parameter values are in the unstable region.

We have that $\eta = 0.02 < 1$ and the parameter values are in region $D_2^+ \cap \{J_0 < 1\}$ (see 21 and Figure 3). This is a stochastic bistable region and a deterministic bistable region as validated by Figure 25.

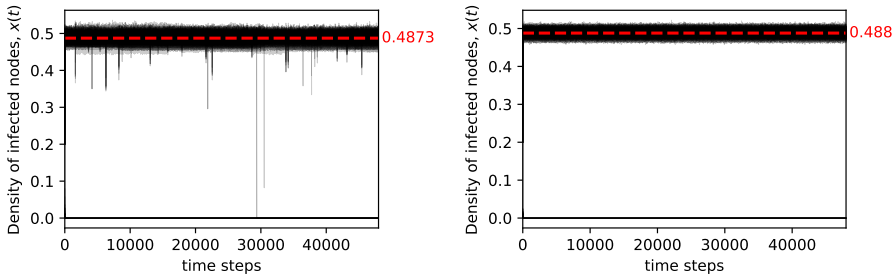

Figure 25: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the synthetic simplicial network. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.2$). Parameter values are in the unstable region.

6.3.3. Simulations for the unstable region.

We fix the parameters

$$\mu = 0.5\,,\ \sigma = 0.9\,,\ J_0 = 1.2\,,\ J_{0\Delta} = 1.2\,. \tag{32}$$

We have that $\eta = 0.81 < 1$ and the parameter values are in the unstable region $\{J_0 > 1\}$ of Figure 3. This unstable regime (both stochastic and deterministic) is corroborated by Figure 26.

7. Conclusions

We have presented a stochastic model of social contagion in simplicial networks based on the mean-field approach of [30]. We simulated the dynamics

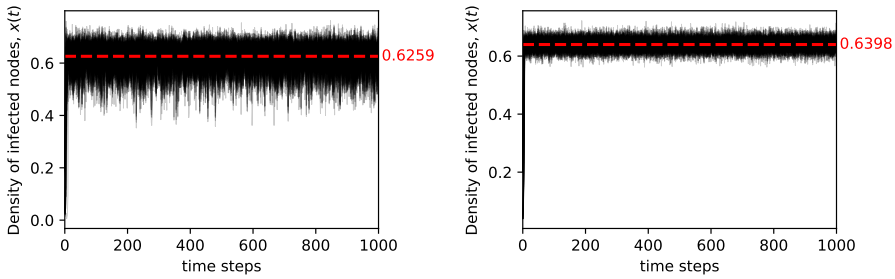

Figure 26: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the synthetic simplicial network. 30 trajectories are plotted for $x_0 = 0.04$ and 30 for $x_0 = 0.2$). Parameter values are in the unstable region.

of this network by a certain SDE and studied conditions that guarantee that the disease disappears in the long run. We have fully characterized the stability of the SDE in terms of new epidemiological meaningful parameters and show the existence of several new regions, among which we find two particularly significant ones: a stochastic bistable region (generalizing the deterministic bistable one but in which there exists a certain probability to be or not to be bistable) and a region which is bistable for the deterministic case but in which, nevertheless, the origin of the SDE is globally asymptotically stable. Moreover, we have introduced a stochastic simplicial contagion model (SSCM) and used to perform different simulations over both real-world (from different domains and following heterogeneous degree distributions) and synthetically generated datasets, obtaining similar results among them which are remarkable similar to those predicted by the numerical simulations (with a Runge Kutta method) of the SDE, and thus confirming theoretical results here presented. We have published the source code for running simulations of our SSCM with the analyzed datasets in [27].

With this we intend to model random situations such as fortuitous and unexpected encounters of a susceptible agent with one or more infected agents, where the randomness is represented in the generation of the white Gaussian noise and the intensity with which it is applied could also be interpreted as a risk factor associated with a specific scenario (school, bar, soccer stadium, etc.). We hope that this can contribute to improve and refine measures to contain the spread of epidemics over networks where unforeseen higher-order interactions may occur.

Acknowledgements

D.H.S. is supported by Grant PID2021-128665NB-I00 funded by MCIN/AEI/ 10.13039/501100011033 and, as appropriate, by “ERDF A way of making Europe”, Ministerio de Economía y Competitividad (Spain) under grant MTM2017-86042-P, the project STAMGAD 18.J445 / 463AC03 by Consejería de Edu-

cación (GIR, Junta de Castilla y León, Spain) and by Universidad de Salamanca (Spain) under project PIC2-2021-10. J.H.S. is supported by the Spanish Ministry of Science and Education under the project TCO-RISEBLOCK (PID2019-110224RB-I00), the European Union's H2020 Research and Innovation Programme under the Grant Agreement No. 871754 (i3-MARKET), and the Generalitat de Catalunya grant 2017-SGR-00782.

References

[1] L. Arnold, *Stochastic differential equations: theory and applications.*, John Wiley & Sons, New York, 1974.

[2] G.F. de Arruda, E. Cotto, T. P. Peixoto, F. A. Rodrigues, and Y. Moreno, *Disease Localization in Multilayer Networks*, Phys. Rev. X 7, 011014 (2017).

[3] G. F. de Arruda, G. Petri, and Y. Moreno, *Social Contagion Models on Hypergraphs*, Phys. Rev. Res. 2, 023032 (2020).

[4] G.F. de Arruda, M. Tizzani, and Y. Moreno, *Phase transitions and stability of dynamical processes on hypergraphs*, Commun. Phys. 4, no. 1 (2021) 24.

[5] A. Barabási and M. Pósfai, *Network Science*, Cambridge University Press, Cambridge, 2016.

[6] A. Barabási and A. Réka, *Statistical mechanics of complex networks*, Rev. Modern Phys. 74, (2002) 47-97.

[7] A. Barrat, G. Arruda, I. Iacopini and Y. Moreno, *Social contagion on higher-order structures*, arXiv preprint arXiv:2103.03709 (2021).

[8] F. Battiston, G. Cencetti, I. Iacopini, V. Latora, M. Lucas, A. Patania, J.-G. Young, and G. Petri, *Networks beyond Pairwise Interactions: Structure and Dynamics*, Phys. Rep. 874,(2020) 1-92.

[9] A.R. Benson, R. Adebe, M.T.Schaub, A. Jadbabaie and J. Kleinberg, *Simplicial closure and higher-order link prediction*, Proc. Natl. Acad. Sci. 115(48) (2018).

[10] G. Bianconi and A.P. Kartun-Giles, *Beyond the clustering coefficient: A topological analysis of node neighbourhoods in complex networks*, Chaos Solitons Fractals X, 1 (2019) 100004.

[11] C. Bick,, E. Gross, H. A. Harrington and M. T. Schaub, *What are higher-order networks?*, arXiv preprint arXiv:2104.11329 (2021).

[12] F Black and M. Scholes, *The Pricing of Options and Corporate Liabilities*, J. Pol. Econ., 81 (3), (1973) 637–654.

[13] G. Burgio, A. Arenas, S. Gómez, and J. T. Matamalas, *Network clique cover approximation to analyze complex contagions through group interactions*, arXiv preprint arXiv:2101.03618 (2021).

[14] M.K. Chung, H. Lee, A. DiChristofano, H. Ombao and V. Solo, *Exact topological inference of the resting-state brain networks in twins*, Net. Neurosci., 3(3) (2019) 674–694.

[15] R. Cowan and N. Jonard, *Network structure and the diffusion of knowledge*, J. Econ. Dyn. Control 28 (2004) 1557–1575.

[16] C. Dubi, and R. Atar, *Modeling neutron count distribution in a subcritical core by stochastic differential equations*, Ann. Nucl. Energy 111 (2018) 608-615.

[17] I. Dzhalladova, M. Růžičková, and V. Růžičková, *Stability of the zero solution of nonlinear differential equations under the influence of white noise*, Advan. Diff. Eqs., 143 (2015).

[18] E. Estrada and G.J. Ross, *Centralities in simplicial complexes. Applications to protein interaction networks*, J. Theoret. Biol. 438 (2018) 46–60.

[19] M. Gameiro, Y. Hiraoka, S. Izumi, M. Kramar, K. Mischaikow and V. Nanda, *A topological measurement of protein compressibility*, Japan J. Indust. Appl. Math 32 (2015) 1–17.

[20] Gard, T.C., *Stability for multispecies population models in random environments.*, Nonl. Anal. Theory, Meth. Appl., 10 (1986) 1411-1419.

[21] C. Giusti, R. Ghrist and S. Bassett, *Two's company, three (or more) is a simplex*, Journal of Computational Neuroscience 41 (1) (2015) 1–14.

[22] T.E. Goldberg, *Combinatorial Laplacians of Simplicial Complexes*, Annandale-on-Hudson, New York (2012).

[23] C. Granell, S. Gómez, and A. Arenas, *Dynamical Interplay between Awareness and Epidemic Spreading in Multiplex Networks*, Phys. Rev. Lett. 111, 128701 (2013).

[24] A. Gray, A., D. Greenhalgh, L. Hu, X. Mao, and J. Pan, *A stochastic differential equations sis epidemic model.*, SIAM J. Appl.Math., 71 (2011) 876-902.

[25] D. Hernández Serrano, J. Hernández Serrano and D. Sánchez Gómez, *Simplicial Degree in Complex Networks. Applications of Topological Data Analysis to Network Science.*, Chaos Solitons Fractals 137 (2020) 109839.

[26] D. Hernández Serrano and D. Sánchez Gómez, *Centrality measures in simplicial complexes: Applications of topological data analysis to network science* , Appl. Math. and Comp. 382 (2020) 125331.

[27] D. Hernández Serrano, J. Villarroel, J. Hernández-Serrano and A. Tocino, *Stochastic Simplicial Contagion Model*, GitHub repository. `https://github.com/juanelas/sscm`

[28] H. W. Hethcote, *The Mathematics of Infectious Diseases*, SIAM Rev. 42 (2000) 599-653.

[29] I. Iacopini, S. Milojević and V. Latora, *Network dynamics of innovation processes*, Phys. Rev. Lett. 120 (2018) 048301.

[30] I. Iacopini, G. Petri, A. Barrat and V. Latora, *Simplicial models of social contagion.*, Nat Commun 10, 2485 (2019).

[31] C. Ji, D. Jiang, and N. Shi, *The Behavior of an SIR Epidemic Model with Stochastic Perturbation*, Stoch. Anal. Appl., 30 (2012) 755–773.

[32] S. Karlin and H. Taylor *A second course in stochastic processes* Acad. Press, New York 1981.

[33] W. O. Kermack, and A. G. McKendrick, *Contribution to the mathematical theory of epidemics*, Proc. R. Soc. Lond. A 115 (1927) 700-721.

[34] R. Khasminskii, *Stochastic stability of differential equations*, 2nd. ed., Springer, Berlin, 2012.

[35] M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter,, *Multilayer networks*, J. Complex Networks 2, 203 (2014).

[36] P.E. Kloeden, and E. Platen, *Numerical Solution of Stochastic Differential Equations*, Springer-Verlag, Berlin, 1999.

[37] A. Lahrouz, L. Omari, and D. Kiouach, *Global analysis of a deterministic and stochastic nonlinear SIRS epidemic model*, Nonlinear Anal.: Model Control, 16 (2011) 59-76.

[38] A. Lahrouz, A. Settati, M. El Fatini, and A. Tridane, *The effect of a generalized nonlinear incidence rate on the stochastic SIS epidemic model*, Math. Meth. Appl. Sci. 44 (2021) 1137-1146.

[39] R. Lambiotte, M. Rosvall and I. Scholtes, *From networks to optimal higher-order models of complex systems.*, Nat. Phys. 15, no. 4 (2019) 313-320.

[40] N. W. Landry and J. G. Restrepo, *The Effect of Heterogeneity on Hypergraph Contagion Models*, Chaos 30, 103117 (2020).

[41] H. Lee, M.K. Chung, H. Kang, H. Choi, S. Ha, Y. Huh, E. Kim and D.S. Lee, *Coidentification of Group-Level Hole Structures in Brain Networks via Hodge Laplacian*, MICCAI 2019, LNCS 11767 (2019) 674–682.

[42] Q. Lu, *Stability of SIRS system with random perturbations*, Physica A 388 (2009) 3677-3686.

[43] S. Maletic and M. Rajkovic, *Combinatorial Laplacian and entropy of simplicial complexes associated with complex networks*, Eur. Phys. J. ST 212 (2012) 77–97.

[44] X. Mao, *Stochastic Differential Equations and Applications* , Horwood, Chichester, 2008.

[45] J. T. Matamalas, S. Gómez, and A. Arenas, *Abrupt Phase Transition of Epidemic Spreading in Simplicial Complexes*, Phys. Rev. Res. 2, (2020) 012049.

[46] J. R. Munkres, *Elements of algebraic topology*, Addison-Wesley Publishing Company, Menlo Park, CA, 1984.

[47] A. Oroji, M. Omar, and S. Yarahmadian, *An Ito stochastic differential equations model for the dynamics of the MCF-7 breast cancer cell line treated by radiotherapy*, J. Theor. Biol. 407 (2016) 128-137.

[48] A. Patania and P. Giovanni and F. Vaccarino, *The shape of collaborations*, EPJ Data Science 6 (2017) 18.

[49] R. Pastor-Satorras, C. Castellano, P. Van Mieghem and A. Vespignani, *Epidemic processes in complex networks*, Rev. Mod. Phys. 87 (2015) 925.

[50] G. Petri, P. Expert, F. Turkheimer, R. Carhart-Harris, D. Nutt, P.J. Hellyer and F. Vaccarino, *Homological scaffolds of brain functional networks*, J. R. Soc. Interface 11 (2014) 101.

[51] G. Petri, M. Scolamiero, I. Donato and F. Vaccarino, *Topological strata of weighted complex networks*, PLoS One 8 (2013) e66506.

[52] T.A. Phan, J.P. Tian, and B.X. Wang, *Dynamics of cholera epidemic models in fluctuating environments*, Stoch. Dyn. (Singap) 21 (2021) 2150011.

[53] Z. Schuss, *Theory and Applications of Stochastic Processes: An Analytical Approach*, Springer Verlag (2010)

[54] A.E. Sizemore, J.E. Phillips-Cremins, R. Ghrist and S. Bassett, *The importance of the whole: topological data analysis for the network neuroscientist*, Network Neuroscience 3(3) (2019) 656–673.

[55] A. Tocino and A. Martín del Rey, *Local stochastic stability of SIRS models without Lyapunov functions.*, Commun. Nonlinear Sci. Numer. Simul., 103 (2021) 105956.

[56] A. Tocino, D. Hernández Serrano, J. Hernández-Serrano and J. Villarroel, *A stochastic simplicial SIS model for complex networks*, submitted for publication (2022).

[57] E. Tornatore, S.M. Buccellato, and P. Vetro, *Stability of a stochastic SIR system*, Physica A 354 (2005) 111-126.

[58] L. Torres, A.S. Blevins, D.S. Bassett and T. Eliassi-Rad., *The why, how, and when of representations for complex systems*, arXiv preprint arXiv:2006.02870 (2020).

[59] T. W. Valente, *Network models of the diffusion of innovations*, Comp. Math. Org. Th. 2 (1996) 163–164.

[60] Vasicek, O., *An equilibrium characterization of the term structure*, J. Fin. Econ. 5 2 (1977) 177–188

[61] H. Verdejo, A. Awerkin, W. Kliemann, and C. Becker, *Modelling uncertainties in electrical power systems with stochastic differential equations*, Int. J. Electr. Power Energy Syst. 113 (2019) 322-332.

[62] D. Wang, Y. Zhao, J. Luo, and H. Leng, *Simplicial SIRS epidemic models with nonlinear incidence rates.*, Chaos 31, (2021) 053112.

[63] K. Xia and G.W. Wei, *Persistent homology analysis of protein structure, flexibility and folding*, Int. J. Numer. Methods Biomed. Eng. 30 (8) (2014) 814–844.

[64] K. Xia and G.W. Wei, *Multidimensional persistence in biomolecular data*, IJ. Comput. Chem. 36 (20) (2015) 1502–1520.

[65] R.M Xu, and R.W. Guo, *Pontryagin's Maximum Principle for Optimal Control Stochsatic SEIR models*, Complexity 2020 (2020) 6479087.

[66] Y. Zhao, D. Jiang, and D. O'Regan, *The extinction and persistence of the stochastic SIS epidemic model with vaccination*, Phys. A 392(20) (2013), 4916-4927

[67] Y. Zhuang, A. Arenas, and O. Yağan, *Clustering determines the dynamics of complex contagions in multiplex networks*, Phys. Rev. E 95 (2017) 012312.

Appendix A. Proof of Theorem 1.

Recall that the infinitesimal generator $\mathbf{L}$ is the differential operator

$$\mathbf{L} := a(x)\partial_x + \frac{b^2(x)}{2}\partial_{xx}. \tag{A.1}$$

Itô rule gives that the directional derivative along the trajectories of any function $V : \mathbb{R} \to \mathbb{R}$ is given by

$$\mathbb{E}(dV(x(t)) = \mathbb{E}\big(\mathbf{L}V(x(t))\big)dt.$$

It follows that if $\mathbf{L}V(x) < 0,$ for all x then

$$\frac{d}{dt}\mathbb{E}V(x(t)) < 0$$

and $V(x(t))$ is a supermartingale.

Here we consider the Lyapunov function $V(x) = -\log|x|$ which satisfies $\lim_{x\to 0} V(x) = \infty$. Hence, we have

$$\begin{aligned}\mathbf{L}V(x) &= -\big(\beta - \mu + (\beta_\Delta - \beta)x - \beta_\Delta x^2\big) + 1/2\big(\sigma + (\sigma_\Delta - \sigma)x - \sigma_\Delta x^2\big)^2\big)\\ &= -(\mu + \sigma^2/2)\big(J_0 - 1\big) + O(x)\end{aligned}$$

Thus, if $J_0 > 1$, $\mathbf{L}V(x)$ is definite negative in an appropriate neighbourhood of the origin.

If $J_0 = 1$ then, $\mathbf{L}V(0) = 0$; nevertheless it is easily seen that the function $b^2(x)$ is positive-definite around $x = 0$.

Hence in both cases we conclude the instability of the origin [1, 34, 44].

For the stable situation we consider the Lyapunov function $V(x) = |x|^\lambda$ for some appropriate $\lambda > 0$. Thus $V(x) \in \mathcal{K}$, the family of all continuous, positive-definite and decrescent functions. Here the directional derivative is

$$\mathbf{L}V(x) = \lambda|x|^\lambda\Big(\beta - \mu + \frac{\lambda - 1}{2}\sigma^2 + \xi(x)\Big)$$

where

$$\xi(x) := (\beta_\Delta - \beta)x - \beta_\Delta x^2 + \frac{\lambda - 1}{2}\big((\sigma_\Delta - \sigma)x - \sigma_\Delta x^2\big)\left(2\sigma + (\sigma_\Delta - \sigma)x - \sigma_\Delta x^2\right).$$

By continuity, for given $\epsilon > 0$ there exists a $\delta = \delta(\lambda, \epsilon)$ such that

$$\xi(x) \le \epsilon \text{ for } 0 < x \le \delta$$

and

$$\mathbf{L}V(x) \le \lambda|x|^\lambda\big(\epsilon + \lambda\sigma^2/2 + \beta - \mu - \sigma^2/2\big).$$

Hence, by choosing appropriately ϵ and λ it follows that $\mathbf{L}V(x) < 0$. This implies asymptotic (local) stability, cf. [1, 34, 44].

Appendix B. Proof of Theorem 3.

For the proof of the Theorem 3 we require the following partial results

Lemma 1. *For given $r > 0$, the quartic curve $Q(x;r)$ of (22) under condition (10) has no roots on $[0,1]$ if and only if there exists no common point to the sets*

$$\mathcal{I} = Im_{x\in[0,1]} L(x) \text{ and } \mathcal{Z} := \{z \in \mathbb{R} : H(z) \le 0\} \tag{B.1}$$

where we define $\nu := \sigma_\Delta/\beta_\Delta$,

$$H(z) = \mu + z + \frac{1-r}{2}\nu^2 z^2; \quad L(x) := -\beta + (\beta - \beta_\Delta)x + \beta_\Delta x^2. \tag{B.2}$$

Proof When (10) is satisfied, we can write

$$\begin{aligned} Q(x;r) &= \mu - \beta + (\beta - \beta_\Delta)x + \beta_\Delta x^2 + \frac{1-r}{2}\nu^2(-\beta + (\beta - \beta_\Delta)x + \beta_\Delta x^2)^2 \\ &= \mu + L(x) + \frac{1-r}{2}\nu^2 (L(x))^2 = H(L(x)). \end{aligned}$$

For $r = 1$, the roots $Q(x;1)$ are given by (3) and the stability study is reduced to that in the deterministic case. With no loss of generality we shall suppose that $r < 1$. Then H is a parable with branches opening upward. If there exists a point z^* satisfying $z^* \in \mathcal{I}$ and $z^* \in \mathcal{Z}$ then $z^* = L(x^*)$ where $x^* \in [0,1]$ and $H(z^*) \le 0$; hence

$$Q(x^*) = H(L(x^*)) = H(z^*) \le 0 \text{ with } x^* \in [0,1].$$

Since $Q(x^*) \le 0$ and $Q(1) = \mu > 0$, Q has at least one zero in $[0,1]$.

Reciprocally, if $Q(x^*) = 0$ for some $x^* \in [0,1]$ then $z^* := L(x^*)$ satisfies simultaneously $z^* \in \mathcal{I}$ and $z^* \in \mathcal{Z}$. c.q.d.

Proposition 4. *Let p be the point in parameter space $p = (\sigma, \beta, \beta_\Delta, \mu)$ and $\mu_r = \mu(1-r)$. Suppose condition (10) is satisfied. For given $r > 0$ the quartic $Q(x;r)$ of (22) has no roots on $[0,1]$ if and only if parameters satisfy*

$$p \in G_r := G_{1;r} \cup G_{2;r} \cup G_{3;r}, \tag{B.3}$$

where we define the regions

$$G_{1;r} := \{p : \beta < \sqrt{2\mu_r}\sigma\} \tag{B.4}$$

$$G_{2;r} := \{p : \ (1-r)\sigma^2 \le \beta; \beta_\Delta \le \beta < \mu + (\sigma^2/2)(1-r)\} \tag{B.5}$$

$$G_{3;r} := \{p : \beta_\Delta > \beta; 2(\ell - \mu) < (1-r)\frac{\ell^2\sigma^2}{\beta^2} < \ell\} \tag{B.6}$$

where

$$\ell := \frac{(\beta_\Delta + \beta)^2}{4\beta_\Delta}.$$

Proof We work out when does condition $\mathcal{I} \cap \mathcal{Z} = \emptyset$ hold. It follows from Lemma 1 (see equation (B.1)), that $\mathcal{I} \cap \mathcal{Z} = \emptyset$ holds if either

A. If $\beta < \sigma\sqrt{2\mu_r}$, since then $1 < 2\mu_r\nu^2$, which implies

$$H(z) > 0 \ \ \forall z, \text{ and } \mathcal{Z} := \emptyset.$$

B. Suppose $1 \geq 2\mu_r\nu^2$. In this case H has two negative roots $z_- < z_+$ such that

$$(1-r)\nu^2 z_\pm = -1 \pm \sqrt{1-2\mu_r\nu^2}.$$

Furthermore

$$z_- < z_+ < 0 \text{ and } \mathcal{Z} = [z_-, z_+] \neq \emptyset.$$

Besides, see equation(B.2), $\mathcal{I} := \mathrm{Im}_{x\in[0,1]} L(x) = [i, 0]$ where

$$i := \min_{x\in[0,1]} L(x) = \begin{cases} -\beta & \text{if } \beta_\Delta \leq \beta, \\ -\frac{(\beta_\Delta+\beta)^2}{4\tilde{\beta}} & \text{if } \beta_\Delta > \beta. \end{cases}$$

Hence the condition $\mathcal{I} \cap \mathcal{Z} = \emptyset$ holds if

$$i > z_+ = \frac{-1+\sqrt{1-2\mu_r\nu^2}}{(1-r)\nu^2}$$

This in turn yields

B.1 When $\beta_\Delta \leq \beta$, then

$$i > z_+ \Leftrightarrow (1-r)\sigma^2 \leq \beta < \mu + (\sigma^2/2)(1-r)$$

B.2 When $\beta_\Delta > \beta$, defining $\ell := \frac{(\beta_\Delta+\beta)^2}{4\tilde{\beta}}$ we see that $i > z_+$ if and only if

$$2(\ell - \mu) < (1-r)\ell^2\nu^2 < \ell$$

holds (we skip the details), and the proposition is proved.

We now state the key result, and prove theorem (3).

Theorem 6. *Suppose that there exists $r > 0$ such that the quartic function (22) has no real roots on $[0, 1]$. Then the origin in (5) is globally stochastically asymptotically stable.*

Besides global asymptotic stability holds if parameters belong to the regions of theorem (3).

Proof Consider the Lyapunov radially unbounded function $V(x) := |x|^r$ for some $r > 0$. Then

$$\mathbf{L}V(x) := r|x|^r\Big(\beta - \mu + (\beta_\Delta - \beta)x - \beta_\Delta x^2 + \frac{r-1}{2}\big(\sigma + (\sigma_\Delta - \sigma)x - \sigma_\Delta x^2\big)^2\Big)$$

Hence we can write $\mathbf{L}V(x) = -r|x|^r Q(x, r)$ where $Q(x, r)$ is the family of quartic polynomials (22). Hence if $Q(x; r) > 0$ for $x \in [0, 1]$ holds for some $r > 0$, $V(x)$

and $-\mathbf{L}V(x)$ are positive definite, which implies global asymptotic stability [34, 44].

Additionally, to connect with theorem (3) we note that if $Q(x;r_1) > 0$ for some r_1 then $Q(x;r_2) > 0$ for all $0 \leq r_2 < r_1$. Indeed, (see (B.2)) $H(z;r_2) > H(z;r_1)$ for all z. Hence

$$Q(x;r_2) = H(L(x);r_2) > H(L(x);r_1) > 0 \text{ for all } r_2 < r_1.$$

It follows that with no loss of generality *we can restrict to the case* $r = 0$. Thus stability holds if parameters satisfy the condition (B.3) of proposition (4) with $r = 0$. This gives the stability region of theorem (3)− upon redefinition so as to take disjoint components.

Besides, when $\sigma \geq \sqrt{2\mu}$ there are no points satisfying $\sigma^2 \leq \beta < \sigma^2/2 + \mu$ and hence $G_2^0 = \emptyset$ and similarly $G_{3,0} = \emptyset$.

The formulation of theorem (4) involves some further elaboration of the regions appearing, so as to obtain a neater formulation in terms of (i) a disjoint decomposition; and (ii) epidemiological parameters. □

Appendix C. The SSCM over more real simplicial networks.

- *Simulations for the region G_1.*

  We plot here the simulations of the SSCM over the real simplicial networks contact-primary-school, NDC-classes and InVS15 for values in the region $G1$. See Figures C.27, C.28 and C.29.

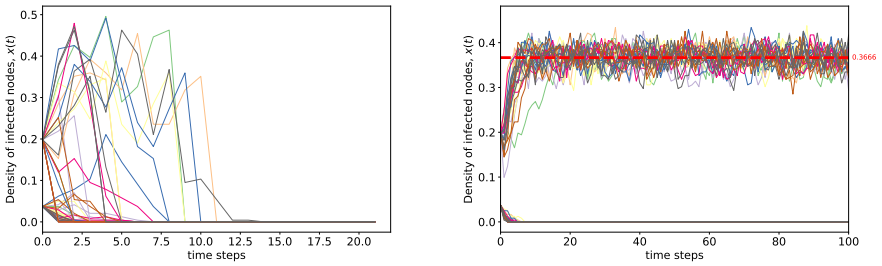

Figure C.27: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-primary-school. 30 trajectories are plotted for $x_0 = 0.04$ and 30 for $x_0 = 0.2$. Parameter values ($\mu = 1$, $\sigma = 0.2$, $J_0 = 0.2549$ and $J_{0\Delta} = 5.88$) are in the region G_1.

- *Simulations for the stochastic bistable region.*

  We plot here the simulations of the SSCM over the real simplicial networks contact-primary-school, NDC-classes and InVS15 for values in the stochastic bistable region. See Figures C.30, C.31, C.32, C.33 and C.34.

  For the InVS15 dataset we see in Figure C.34 that, choosing the parameters $\sigma = 0.57$, $\mu = 1$, $J_0 = 0.7$ and $J_{0\Delta} = 6$ (and running 48.000 time

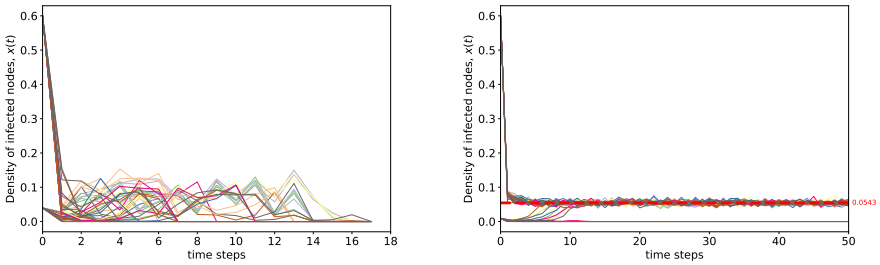

Figure C.28: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network NDC-classes. 30 trajectories are plotted for $x_0 = 0.04$ and 30 for $x_0 = 0.6$ for the left one and the same number for $x_0 = 0.01$ and $x_0 = 0.6$ in the right one. Parameter values ($\mu = 1$, $\sigma = 0.2$, $J_0 = 0.2549$ and $J_{0\Delta} = 5.88$) are in the region G_1.

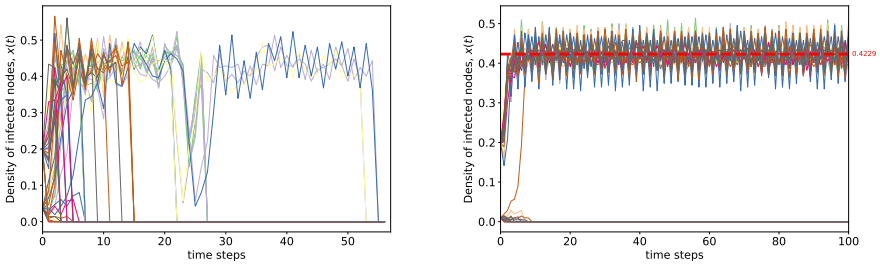

Figure C.29: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network InVS15. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.2$. Parameter values ($\mu = 1$, $\sigma = 1.44$, $J_0 = 0.2549$, $J_{0\Delta} = 5.88$) are in the region G_1.

steps), we are in a bistable stochastic regime (left graphic). However, and due to the stochastic nature of this regime, we observe in the figure that there is still a probability that a trajectory, which has apparently reached a state of endemic equilibrium, will fall to 0. The $\sigma = 0$ case (Figure C.34 right) produces, as expected, a deterministic bistable regime.

- *Simulations for the unstable region.*

  We plot here the simulations of the SSCM over the real simplicial networks contact-primary-school, NDC-classes and InVS15 for values in the unstable region. See Figures C.35, C.36 and C.37. For the InVS15 dataset we have used $\sigma = 0.9$ in order to see the stochastic perturbation faster.

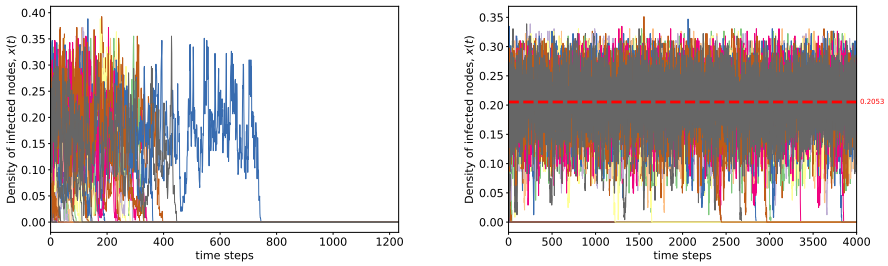

Figure C.30: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-primary-school. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.2$. Parameter values ($\mu = 1$, $\sigma = 0.2$, $J_0 = 0.88$, $J_{0\Delta} = 1.96$) are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

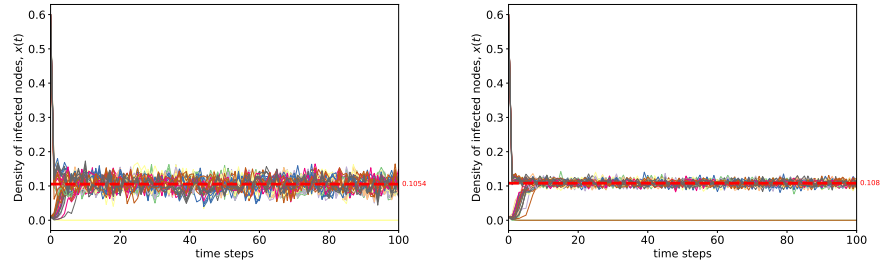

Figure C.31: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network NDC-classes. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.6$. Parameter values ($\mu = 1$, $\sigma = 0.2$, $J_0 = 0.88$, $J_{0\Delta} = 1.96$) are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

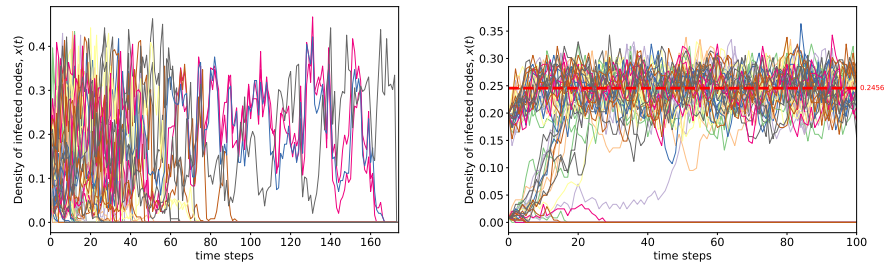

Figure C.32: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-primary-school. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.2$. Parameter values ($\mu = 1$, $\sigma = 0.41$, $J_0 = 0.88$, $J_{0\Delta} = 1.96$) are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

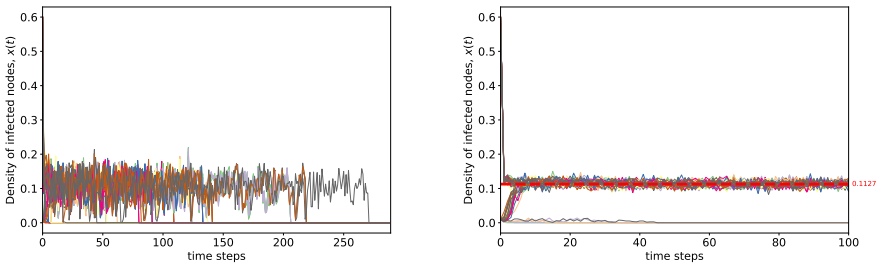

Figure C.33: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network NDC-classes. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.6$. Parameter values ($\mu = 1$, $\sigma = 0.41$, $J_0 = 0.88$, $J_{0\Delta} = 1.96$) are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

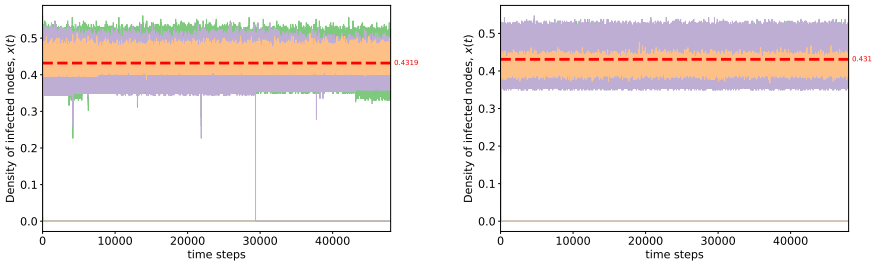

Figure C.34: 20 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network InVS15. 10 trajectories are plotted for $x_0 = 0.01$ and 10 for $x_0 = 0.6$. Parameter values are in the stochastic bistable region $D_2^+ \cap \{J_0 < 1\}$ of Figure 3).

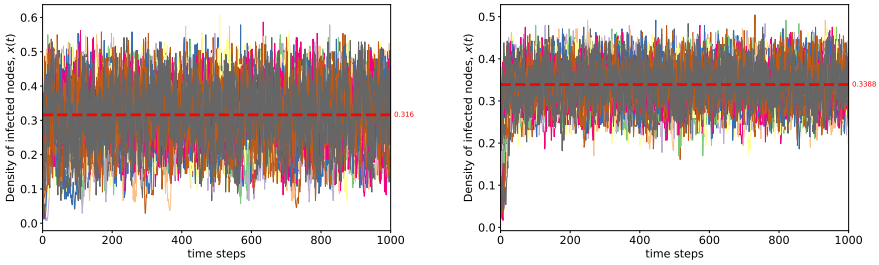

Figure C.35: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network contact-primary-school. 30 trajectories are plotted for $x_0 = 0.04$ and 30 for $x_0 = 0.2$. Parameter values ($\mu = 0.5$, $\sigma = 0.2$, $J_0 = 1.2$, $J_{0\Delta} = 1.2$) are in the unstable region.

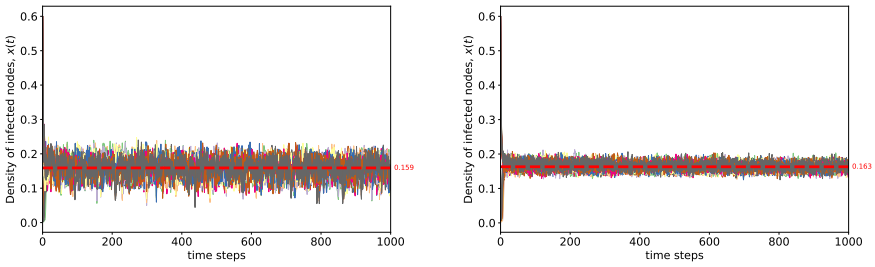

Figure C.36: 60 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network NDC-classes. 30 trajectories are plotted for $x_0 = 0.01$ and 30 for $x_0 = 0.6$. Parameter values ($\mu = 0.5$, $\sigma = 0.2$, $J_0 = 1.2$, $J_{0\Delta} = 1.2$) are in the unstable region.

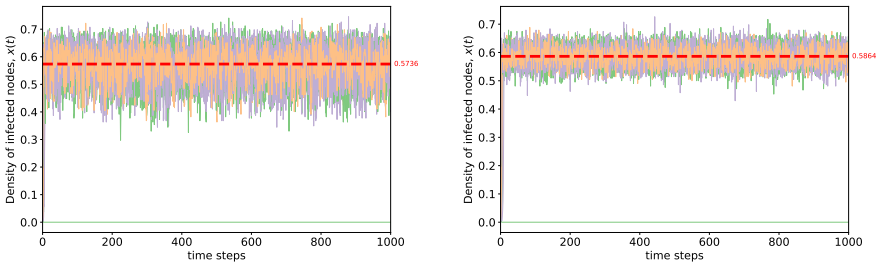

Figure C.37: 20 trajectories resulting from running the SSCM (left) and the SCM (right) on the real simplicial network InVS15. 10 trajectories are plotted for $x_0 = 0.01$ and 10 for $x_0 = 0.6$. Parameter values ($\mu = 0.5$, $\sigma = 0.2$, $J_0 = 1.2$, $J_{0\Delta} = 1.2$) are in the unstable region.