



# EQUATION-FREE MULTISCALE COMPUTATIONS IN SOCIAL NETWORKS: FROM AGENT-BASED MODELING TO COARSE-GRAINED STABILITY AND BIFURCATION ANALYSIS

A. C. TSOUMANIS\*, C. I. SIETTOS<sup>†</sup>,  
G. V. BAFAS\* and I. G. KEVREKIDIS<sup>‡</sup>

*\*School of Chemical Engineering,  
National Technical University of Athens,  
GR 157 80, Athens, Greece*

*<sup>†</sup>School of Applied Mathematics and Physical Sciences,  
National Technical University of Athens,  
GR 157 80, Athens, Greece*

*<sup>‡</sup>Department of Chemical Engineering,  
Program in Computational and Applied Mathematics,  
Department of Mathematics, Princeton University,  
Princeton, NJ 08544, USA*

Received June 9, 2009; Revised August 6, 2010

We focus on the “trijunction” between multiscale computations, bifurcation theory and social networks. In particular, we address how the Equation-Free approach, a recently developed computational framework, can be exploited to systematically extract coarse-grained, emergent dynamical information by bridging detailed, agent-based models of social interactions on networks, with macroscopic, systems-level, continuum numerical analysis tools. For our illustrations, we use a simple dynamic agent-based model describing the propagation of information between individuals interacting under mimesis in a social network with private and public information. We describe the rules governing the evolution of the agents’ emotional state dynamics and discover, through simulation, multiple stable stationary states as a function of the network topology. Using the Equation-Free approach we track the dependence of these stationary solutions on network parameters and quantify their stability in the form of coarse-grained bifurcation diagrams.

**Keywords:** Complex systems; multiscale computations; equation-free approach; networks; social systems; nonlinear dynamics; bifurcation analysis.

## 1. Introduction

Agent-based models constitute the state-of-the-art in contemporary computer simulation of many complex systems ranging from ecology [DeAngelis

*et al.*, 1994; Grimm, 1999; Bonabeau *et al.*, 2000] and epidemiology [Longini *et al.*, 1986; Eubank *et al.*, 2004; Ferguson *et al.*, 2005] to economics and financial markets [Raberto *et al.*, 2001; Iori, 2002;

---

<sup>†</sup>Author for correspondence

Wang & Zhang, 2005; Samanidou *et al.*, 2007; Liu *et al.*, 2008], and from traffic and supply-chain networks [Dijkstra *et al.*, 2001; Helbing, 2001; Julka *et al.*, 2002; Raney *et al.*, 2003; Balmer *et al.*, 2004] to biology [Levine *et al.*, 2001; Liu & Passino, 2004] and physiology [Bonabeau, 2002; Khulood & Raed, 2007].

On the other hand, networks are identified as a key feature of the structure of many real-world systems and recent years' research on the subject is a part of a broader trend towards research on the dynamics of complex systems [Bollobás, 1979; Strogatz, 2001; Callaway *et al.*, 2001; Newman *et al.*, 2002; Albert & Barabasi, 2002; Newman, 2003; Szabó & Fàth, 2007]. Within this context there has been an immense interest in studying social dynamics using detailed agent-based modeling on networks as they pertain to many complex systems such as epidemiological [Boots & Sasaki, 1999; Bogunà & Pastor-Satorras, 2002; Eubank *et al.*, 2004; Keeling & Eames, 2005; Keeling, 2005], economic and financial [Daniel *et al.*, 2002; Iori, 2002; Barberis, 2003; Ritter, 2003; Wang & Zhang, 2005; Palatella *et al.*, 2005; Omurtag & Sirovich, 2006; Hirshleifer *et al.*, 2006; Kim *et al.*, 2007; Boginski *et al.*, 2006; Reichardt & White, 2007; Hein *et al.*, 2008], opinion formation and voter dynamics [Clifford & Sudbury, 1973; Holley & Liggett, 1975; Steels & Baillie, 2000; Tessone & Toral, 2005; Castellano *et al.*, 2009; Lambiotte *et al.*, 2007a; Lambiotte *et al.*, 2007b], culture and language dynamics [Dunbar, 1993; Axelrod, 1997; Castellano *et al.*, 2000; Dorogovtsev & Mendes, 2001; Loreto & Steels, 2007], traffic and crowd flow design and management [Helbing & Molnár, 1995; Hughes, 2000; Helbing *et al.*, 2000; Parisi & Dorso, 2005], diffusion of news and innovations [Guardiola *et al.*, 2002; Llas *et al.*, 2003; Wu & Huberman, 2007]. The rapid technological and theoretical progress has dramatically enhanced our knowledge in modeling, simulating and analyzing such complex systems and we are getting better at it.

However due to the effectively complex, nonlinear, stochastic nature of such systems, their behavior at the macroscopic level — where usually the extensive analysis, design and optimization requirements arise — cannot be, most of the time, accurately modeled and analyzed in a straightforward manner. While one can try to write down a general master equation to describe the probabilistic time evolution of the macroscopic quantities, major problems arise in trying to find the closures required

to bridge the gap between the micro-scale of the individual-based interactions on the network and the macro-scale at which the systems-level analysis and design are required [Weidlich, 2002; Castellano *et al.*, 2009]. This lack of accurate, explicit macroscopic closures poses a major obstacle to the performance of important systems-level computational tasks, which rely on the availability of good low-order closed models in terms of a few macroscopic (coarse-grained) variables.

To date what is usually done with very detailed multi-agent-based descriptions of social networks is simple simulation: set up many initial conditions, for each initial condition create a large enough number of ensemble realizations, probably change some of the rules and then run the detailed dynamics for a long time to investigate how things like stock exchange indices or the spread rate of news change with parameters. However, this approach is both time consuming and inappropriate for systems-level tasks such as stability, bifurcation analysis or control design and optimization since efficient well-established computational tools for such analysis and design tasks require the availability of coarse-grained models in the form of ordinary and/or (integro)-partial differential-algebraic (PDAEs) equations in a closed form.

In this work we show how the so-called “Equation-Free” approach for multiscale computations [Theodoropoulos *et al.*, 2000; Gear *et al.*, 2002; Makeev *et al.*, 2002; Runborg *et al.*, 2002; Sietos *et al.*, 2003; Kevrekidis *et al.*, 2003; Kevrekidis *et al.*, 2004] can be exploited to systematically analyze the macroscopic behavior of agent-based simulations. For our illustrations we use a simple model describing the emotional state formation and propagation process under private and public information on a social network. Actually the model is a simple caricature of a society whose members are interested in some specific matter (e.g. a local election between two opponents, the trend to invest in a stock fund, etc.). We start with presenting how two important elements of the Equation-Free framework, namely the lifting and restricting operators, can be constructed to deal with certain features of the networks topology. We then construct the coarse-grained bifurcation diagrams with respect to network degree distribution and analyze the stability of the identified stationary solutions branches through coarse-grained eigen-computations.

The paper is organized as follows: in Sec. 2, we give a short description of the Equation-Free

approach and the procedure for deriving the lifting and restricting operators. In Sec. 3, we describe the agent-based model as well as the algorithm we used to create random connected networks with degree sequences sampled from a specific probability distribution. More details are given in the appendix. In Sec. 4, we present and discuss the simulation results and finally we conclude in Sec. 5.

## 2. The Equation-Free Framework for Complex/Multiscale Computations

The concept of the *coarse timestepper* is the key element in the Equation-Free framework [Theodoropoulos *et al.*, 2000; Gear *et al.*, 2002; Makeev *et al.*, 2002; Runborg *et al.*, 2002; Siettos *et al.*, 2003; Kevrekidis *et al.*, 2003; Kevrekidis *et al.*, 2004], an efficient methodology for extracting systems-level information from microscopic (detailed, fine scale, atomistic, agent-based) simulations. For the approach to be applicable we must know — or have good reason to believe — that predictive coarse-grained models of the detailed process can be written in terms of (can “close” with) a reduced number of “coarse-grained variables” or “observables”. Knowing such good variables through experience with the problem or through data-mining techniques [Cooley & Lohnes, 1971; Coifman *et al.*, 2005] is an important prerequisite. One might hypothesize, for example, that the average state (over several network realizations and over all nodes of a given degree) for each degree may constitute a good set of coarse-grained observables for our problem. Given this knowledge, the coarse timestepper consists of the following fundamental steps (see also Fig. 1):

- (a) Prescribe a coarse-grained initial condition (e.g. values of the average state for each connectivity degree in the network)  $u(t_0)$ ;
- (b) Transform the coarse-grained initial conditions through a lifting operator, say  $\mu$ , to consistent microscopic ensembles, i.e. get  $U(t_0) = \mu u(t_0)$ ;
- (c) Evolve these microscopic distributions in time with the detailed simulations and report their values after a short macroscopic time  $T$  (time-horizon). An appropriate choice of  $T$  can be estimated from the spectrum derived from the mean field equations linearization around their steady states. Matrix-free, iterative eigencomputations using the coarse timestepper are used to obtain the leading part of this spectrum, and

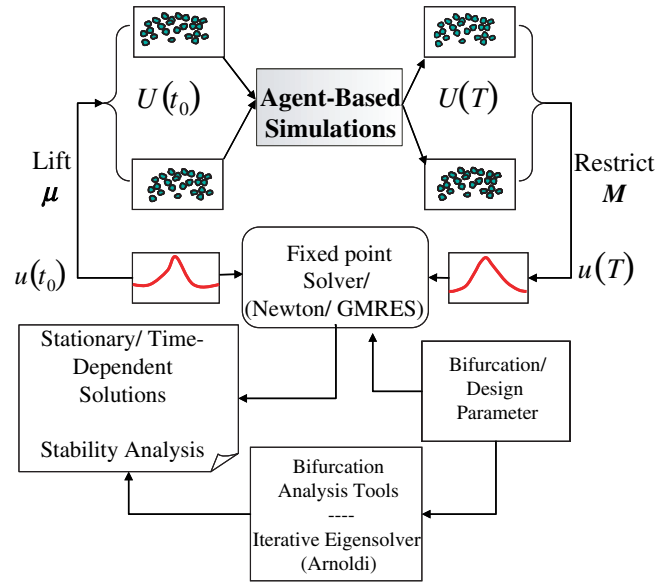


Fig. 1. A schematic of the Equation-Free approach.

search for a gap in it that will dictate the choice of  $T$  (for a more detailed discussion refer to [Siettos *et al.*, 2003; Kevrekidis *et al.*, 2003]).

- (d) Compute the corresponding coarse-grained variables by restricting the resulting microscopic distributions using a restricting operator, say  $M$ , i.e. compute  $u(T) = MU(T)$ .

The above procedure can be considered as a “black box” coarse timestepper of the coarse-grained system observables: a map that, using the detailed simulations, and given an initial coarse-grained state  $u$  at time  $t_0$  will report the result of the evolution after a given timehorizon  $T$ , i.e.:

$$u(t_0 + T) = \Phi_T(u(t_0), p),$$

where  $p$  denotes the systems parameters.

At this point one can

- (e) Utilize computational superstructures like the Newton’s or Newton-GMRES method [Kelley, 1995] as a shell “wrapped around” the coarse timestepper to compute and trace branches of coarse-grained equilibria or periodic solutions (even unstable ones) [Kavousanakis *et al.*, 2008]; other computational superstructures (other “wrappers”) e.g. iterative eigensolvers such as Arnoldi’s algorithm [Saad, 1992; Christodoulou & Scriven, 1998] can also be used to extract information about the stability of the coarse-grained system dynamics. For example,

coarse-grained (macroscopic) equilibria can be obtained as fixed points of the mapping  $\Phi_T$ :

$$u - \Phi_T(u, p) = 0 \quad (1)$$

The tracing of solution branches past turning points in one parameter space can be obtained by augmenting (1) with the linearized pseudo arc-length condition:

$$N(u, \mu, \mu) = \alpha \cdot (u - u_1) + \beta(p - p_1) - \Delta s = 0, \quad (2)$$

where

$$\alpha \equiv \frac{(u_1 - u_0)^T}{\Delta s}, \quad \beta \equiv \frac{(\mu_1 - \mu_0)}{\Delta s}$$

and  $\Delta s$  is the pseudo arc-length continuation step.  $(u_0, p_0)$  and  $(u_1, p_1)$ , represent two already computed solutions. The pseudo arc-length condition (2) constrains the consequent solution  $(u, p)$  to lie on a hyperplane perpendicular to the tangent of the bifurcation diagram at  $(u_1, p_1)$ , approximated through  $(a, \beta)$  and at a distance  $\Delta s$  from it. The computation of the fixed points can now be obtained using an iterative procedure like the Newton–Raphson technique. The procedure involves the iterative solution of the following linearized system:

$$\begin{bmatrix} I - \frac{\partial \Phi_T}{\partial u} & -\frac{\partial \Phi_T}{\partial p} \\ a & \beta \end{bmatrix} \begin{bmatrix} \delta u \\ \delta p \end{bmatrix} = \begin{bmatrix} u - \Phi_T(u, p) \\ N(u, p) \end{bmatrix} \quad (3)$$

Note that for the calculation of both the Jacobian  $\partial \Phi_T / \partial u$  and the derivative  $\partial \Phi_T / \partial p$ , no explicit macroscopic evolution equations are required. They can be approximated numerically by calling the black-box coarse timestepper at appropriately perturbed values of the corresponding unknowns  $(u, p)$ . Hence, the above framework enables the microscopic simulator to converge to both coarse-grained stable and unstable solutions and trace their locations, i.e. to fulfill tasks that the simulator was not explicitly designed for.

The main assumption behind the coarse timestepper is that a coarse-grained model for the microscopic-scale dynamics exists and closes in terms of (relatively few) coarse-grained variables which serve as the system observables. These are usually some low-order moments of the microscopically evolving distributions. The assumption about the existence of a coarse-grained model implies that the higher order moments of the microscopic

distributions become, relatively quickly over the macroscopic time scales, functionals of (become “slaved to”) the few lower ones. What a coarse timestepper does, in fact, is a model reduction “on demand”.

The above methodology essentially sidesteps the derivation of a closed form macroscopic description: relatively short and appropriately initialized short runs of the microscopic simulation provide the necessary closure (refer to [Gear *et al.*, 2002; Sietos *et al.*, 2003; Kevrekidis *et al.*, 2003] for more detailed discussions); when accurate macroscopic equations for a system of interest can be derived in closed form, one should work with them directly.

The construction algorithms of the lifting and restricting operators for networks that can be described by the degree distribution are introduced in the following section.

## 2.1. Lifting and restriction operators for a certain class of social networks under the equation-free approach

Here, we consider the case of undirected graphs for which their connectivity structure is described according to the degree distribution.

Let  $U(x)$  denote the distribution function of the microscopic variables over the set of microscopic coordinates  $x$ . The key idea underlying coarse-graining and model reduction is that certain “fast” system observables (e.g. higher-order moments of the evolving microscopic distributions  $U$ ) become rapidly (say, over a few characteristic interaction times in an individual-based simulation) slaved to the “slow” observables (e.g. the lower order moments)  $u$ . This functional relation can be stated in a general form as

$$u = \int h[U(x)] dx. \quad (4)$$

Assuming the existence of a separation of time scales (and an associated slaving) in the overall system dynamics, we can proceed to construct our restriction and lifting operators. We first start by describing the restriction procedure. For the sake of clarity of the presentation we will illustrate both procedures by considering the average value over the microscopic state.

Let the overall network ensemble assembled by  $M$  social networks,  $G_j = (V_j, E_j)$ , where  $1 \leq j \leq M$ .



We start by averaging the values of the microscopic variables  $U(x)$  over each degree and for all networks in the ensemble. We subsequently sort each individual according to its degree and we assign to it the average value of the microscopic variables. Having all individuals sorted and assigned the “new-averaged” values of the microscopic variable, we proceed to the computation of the coarse-grained variables as follows: let  $m$  be the number of coarse-grained variables and  $K$  be the total number of agents, over our entire network ensemble simulation; then each coarse-grained variable is defined as the average microscopic state over each  $K/m$  fraction of the overall agent population, sorted by the degree.

We define the superset of the vertices of all created networks,  $V = \bigcup_{j=1}^M V_j$ . Given  $V$  we proceed to construct the subsets of  $V$  containing only vertices of the same degree:  $v \in D_i \subseteq V \Leftrightarrow \deg(v) = i, \forall v \in V$ ; the cardinality of each such set is  $d_i = |D_i|$ .

Each agent  $v$  has a microscopic state  $U_v(x)$ . For each degree  $i$  we calculate the average value of the microscopic state  $\bar{U}_i(x)$  as

$$\bar{U}_i(x) = \sum_{v \in D_i} \frac{U_v(x)}{d_i} \quad (5)$$

The total agent population size is  $K = MN$ . We define a new set containing all the agents sorted by their degree  $S \subseteq V, S = \{v_1, v_2, \dots, v_K\}, \deg(v_i) \leq \deg(v_l), i < l$ .

If  $m$  is the number of the coarse-grained variables, each coarse-grained variable represents  $K/m$  agents. Thus the restriction of the microscopic distribution(s) to the coarse-grained variables (observables) reads:

$$\text{obs}_i = \sum_{q=(i-1)K/m+1}^{iK/m} \frac{\bar{U}_{\deg(v_q)}}{\frac{K}{m}}, \quad (6)$$

$$v_q \in S \quad \text{and} \quad 1 \leq i \leq m$$

The lifting operation consists of the inverse actions. The inputs to the lifting subroutine are the network parameter value (such as the probability  $p$  in the discrete geometric distribution) and the instantaneous values of the  $m$  coarse-grained variables. We first create a graph by sampling over the probability distribution. Then we sort its  $N$  agents based on their degree. Finally, we assign to each agent in the sorted list the microscopic state given by the corresponding coarse-grained variable value.

Consider the  $j$ th lifting attempt, giving rise to the social network  $G_j = (V_j, E_j)$  in our ensemble of networks. We then construct the set that contains all the nodes of the set  $V_j$  sorted by their degree:

$$S_j \subseteq V_j, \quad S_j = \{v_1, v_2, \dots, v_N\},$$

$$\deg(v_i) \leq \deg(v_l), \quad i < l$$

Finally we assign to each agent in the  $j$ th network a microscopic state according to the following relation:

$$\forall v_i \in S_j : U_{v_i}(x) = \text{obs}_{\text{floor}[(i-1)/(N/m)]+1} \quad (7)$$

where  $\text{floor}(h/b)$  is the integer part of the division  $h/b$ .

Alternative ways of constructing the lifting and restricting operators based on the cumulative distribution function of the degree distribution [Newman, 2003] can also be considered.

### 3. The Agent-Based Model

For our illustrations we used a simple dynamic agent-based model inspired by the work presented in [Omurtag & Sirovich, 2006]; we altered their basic model to include a social network structure, whose construction is presented in the following section, and to implement a kind of information propagation through agent interactions. Our model consists of  $N$  agents that “live” in our artificial social system caricature and a graph that describes the connections between these individuals.

The agents are characterized by a single real variable which quantifies their emotional state, denoted by  $x$ , where  $x \in [-1, 1]$ . Agents receive information at discrete time instances (Poisson-distributed in time) from the external environment (public information) as well as from their links in their social environment (private information) changing their state according to the following rules:

- (1) In the absence of any public or private information arrival during a time interval  $\Delta t$ , an agent’s state decays exponentially towards zero in time as  $x(t) = x(0) \exp[-\gamma \delta t]$ , where  $\gamma$  is the decay rate (here constant, a model parameter) and  $x(0)$  is the state at the beginning of the time interval. Time intervals in our model are defined in terms of the time unit of our system, *timeUnit*.
- (2) *Public* information arrives to each agent as “good” or “bad” news, each type according

to a Poisson distribution with arrival rates  $v^+$  and  $v^-$ , respectively. When an agent receives good (bad) external news then its emotional state instantaneously jumps by a finite (typically small and constant, a parameter of the model) positive (negative) amount denoted by  $\varepsilon^+$  ( $\varepsilon^-$ ).

- (3) *Private* information is defined as information obtained through the agent's social environment. The arrivals of such information also follow a Poisson distribution but with rate equal to  $k = \alpha * \deg(i)$ , where  $\alpha$  is a scalar (constant, a parameter of our model) denoting the frequency of "encounters" that an agent initiates with another one of its "environment". When an agent initiates an "encounter" with another one the information received (the change in its emotional state) is considered "good" or "bad" (results in a positive change  $e^+$  or a negative change  $e^-$ , respectively) based on a comparison of its current emotional state with the emotional state of the agent encountered *as recorded at the previous observation time step* (see below). If the initiating agent is at a worse (algebraically smaller) emotional state than the encountered one, then the incoming information to him/her is classified as "good" news; if he/she has a better (algebraically bigger) emotional state, it is classified as "bad" news.
- (4) If during the simulations the emotional state of an agent reaches or exceeds one of the boundary values ( $-1$  and  $1$ ) then its state "locks" there for the remaining fragment of the time interval  $\delta t$ , independently of the arrival of new information. From this, we intend to simulate situations where the emotional state "saturates" or where the agent has "made up his/her mind".

### 3.1. The construction of the social network

Consider a graph  $G = (V, E)$ , where  $V$  is the set of vertices — which in our model correspond to the  $N$  individuals  $\{v_i | i = 1, \dots, N\}$  — and  $E$  is the set of edges, the links between the individuals. An edge  $e$  is defined by  $\{v_i, v_j\}$  where  $v_i, v_j \in V$  are the nodes connected by the edge, and will be referred to as  $e_{v_i v_j}$ . The graph is undirected so that  $e_{v_i v_j} \equiv e_{v_j v_i}$ ; multiple edges and self-loops are precluded, i.e.  $\nexists e_{v_i v_j} \in E$  where  $v_i = v_j$ . We considered

random connected graphs whose degree sequences are sampled from a slightly altered (in terms of the range of  $k$ ) discrete geometric distribution<sup>1</sup> with parameters  $p$  and maximum degree  $\max \text{Deg}$ :

$$y = y(k; p) = \frac{p(1-p)^k}{\sum_{i=1}^{\max \text{Deg}} p(1-p)^i} \quad (8)$$

where  $y$  is the probability of a vertex with degree  $k \geq 1$ .

Here  $\max \text{Deg}$  is set to 140; this choice was based on the so-called *Dunbar's number* (see [Edney, 1981; Dunbar, 1993; Gladwell, 2000]);  $1 \leq k \leq 140$  so that the degree  $\deg(v_i)$  of each node  $v_i$  lies between  $1 \leq \deg(v_i) \leq 140$ . At this point, we should note that for our illustrations we used a geometric distribution in order to create a right-hand skewed degree of vertices as this characteristic pertains to most real-world networks [Newman, 2003]. In fact, we could use any other distribution, skewed or not.

Our main goal is to examine the dynamic behavior of the system described above as the graph topology (in terms of the parameter  $p$ ) varies.

More details on the construction of a connected random graph with a given degree probability distribution are given in the appendix.

## 4. Simulation Results and Discussion

### 4.1. Model dynamics

Simulation results were obtained with the following parameter values:  $N = 20\,000$ ,  $\delta t = 0.25$ ,  $\gamma = 1$ ,  $v^+ = 20$ ,  $v^- = 20$ ,  $a = 2$ ,  $\varepsilon^+ = 0.075$ ,  $\varepsilon^- = 0.072$ ,  $e^- = 0.33$ ,  $e^+ = 0.35$ .

Figure 2 depicts the temporal evolution of the averaged (over  $l = 1920$  different network realizations over all nodes with the same degree, and over all degrees) emotional state at each reporting time-instant. Depending on the value of the probability  $p$  of the geometric discrete distribution, our network is capable of exhibiting interesting nonlinear behavior. For large values of  $p$  the average emotional state of the social network converges to a single stationary point, independently of the initial conditions of the emotional state of our agents [see Figs. 2(a) and 2(b)]. As the value of  $p$  decreases

<sup>1</sup>From now on we will refer to it as geometric discrete distribution.

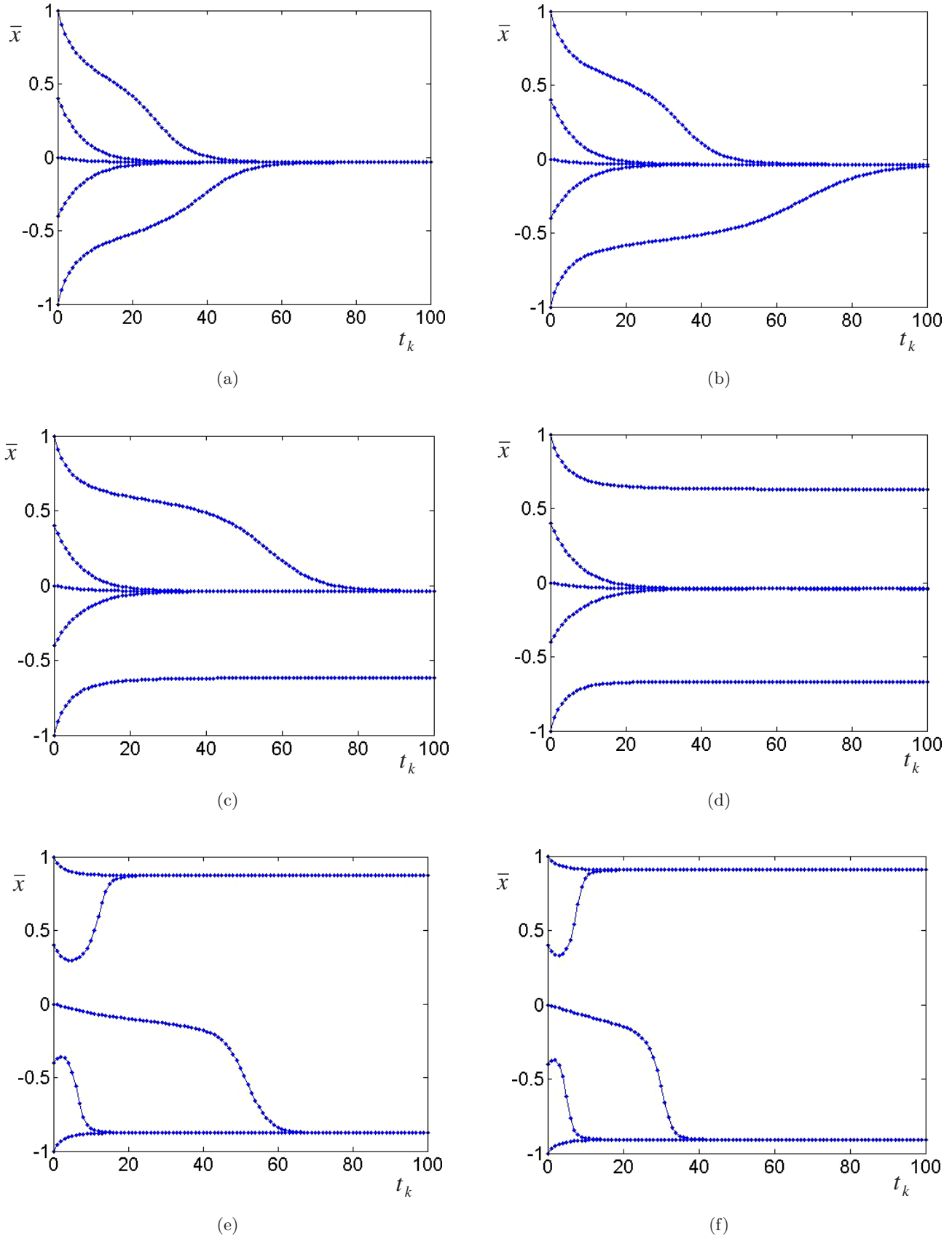


Fig. 2. Temporal simulations depicting the evolution of the network average (over 1920 independent realizations) emotional state starting from different initial conditions as the value of the parameter  $p$  decreases: (a)  $p = 0.094$ , (b)  $p = 0.089$ , (c)  $p = 0.084$ , (d)  $p = 0.079$ , (e)  $p = 0.03$ , (f)  $p = 0.02$ . A maximum of three stable stationary solutions is possible.

(making our graph denser, as far as the number of edges is concerned) — and now depending on the initial conditions — a second stable stationary point is revealed [Fig. 2(c)]. Further decrement of the value of  $p$ , depending again on the initial conditions, brings up a new stationary point [Fig. 2(d)] making the number of stable stationary average states of the network equal to three. Eventually, when  $p$  decreases beyond a critical value, the middle stable stationary point disappears [Figs. 2(e) and 2(f)]. The above results are indicative of the existence of several critical parameter values (values of the discrete geometric distribution parameter  $p$ ) marking qualitative transitions in the dynamic behavior of our system — and as the bifurcation theory dictates — the onset of unstable coarse-grained states. Based on the results of our temporal simulations, these three critical values lie in the ranges:

$$0.084 < p_{\text{critical}}^1 < 0.089, \quad 0.079 < p_{\text{critical}}^2 < 0.084, \\ 0.03 < p_{\text{critical}}^3 < 0.04$$

The temporal simulations shown in Fig. 2 indicate that networks with the same degree distribution (and even further, networks with degrees sampled from the discrete geometric distribution with the same probability  $p$ ) can have one, two or even three stable stationary states and unstable ones (which are unreachable through long-temporal simulations). To systematically study the nonlinear behavior of the system we employed the Equation-Free approach to construct the coarse-grained bifurcation diagram.

As we noted in the previous sections the key idea underlying the Equation-Free approach is that the higher-order moments of the evolving distribution evolve very fast, compared to the macroscopic time-horizon  $T$ , becoming quickly functionals of a few lower-order moments. In other words, the system evolves quickly to a low-dimensional, attracting slow manifold. In order to provide computational support for this slaving we performed various temporal runs near the stationary solutions of all three apparently stable solution branches involving the first moments of the underlying microscopic distribution, namely the overall network average emotional state  $\bar{x}$  and the norm  $\|\bar{s}\|$  of the vector of variances of the average emotional state for each degree.

These variances are computed as follows: assume that we have created  $l$  consistent networks:

$G_j = (V_j, E_j)$ , where  $1 \leq j \leq l$  with  $N$  agents per network and a maximum degree equal to  $\max \text{Deg}$ . Based on the sets  $V_j$  for each social network  $j$  we create the following new sets:

$$D_{ij} \subseteq V_j \quad \text{where } \forall v \in V_j \\ \text{if } \deg(v) = i \Leftrightarrow v \in D_{ij}$$

Each agent  $v$  has an emotional state denoted by  $x(v)$ . For each network  $j$  and each degree  $i$  the average emotional state and the corresponding variance are  $\bar{x}_{ij}$  and  $s_{ij}$ , respectively. These are computed by the following relations:

$$\bar{x}_{ij} = \sum_{v \in D_{ij}} \frac{x(v)}{d_{ij}}, \quad \text{where } d_{ij} = |D_{ij}| \quad (9)$$

and

$$s_{ij} = \frac{\left( \sum_{v \in D_{ij}} (x(v) - \bar{x}_{ij})^2 \right)}{d_{ij}} \quad (10)$$

For each  $j$  we compute the network average emotional state  $\bar{x}_j$ , reading

$$\bar{x}_j = \sum_{v \in V_j} \frac{x(v)}{|V_j|} \quad (11)$$

Over the set of networks the resulting overall network average emotional state is

$$\bar{x} = \frac{\left( \sum_{j=1}^l \bar{x}_j \right)}{l} \quad (12)$$

We construct a vector  $\bar{s} = \{\bar{s}_1, \bar{s}_2, \dots, \bar{s}_{\max \text{Deg}}\}$  where

$$\bar{s}_i = \frac{\left( \sum_{j=1}^l s_{ij} \right)}{l}, \quad \text{where } 1 \leq i \leq \max \text{Deg} \quad (13)$$

which is the average variance of the emotional state for each degree over the entire set of social networks that we have created and simulated.

All the phase portraits [see Figs. 3(a)–3(f)] depict four transients, each initialized with all individuals with the same initial emotional state (thus having a zero variance for each degree). A fifth transient simulation, initialized based on the



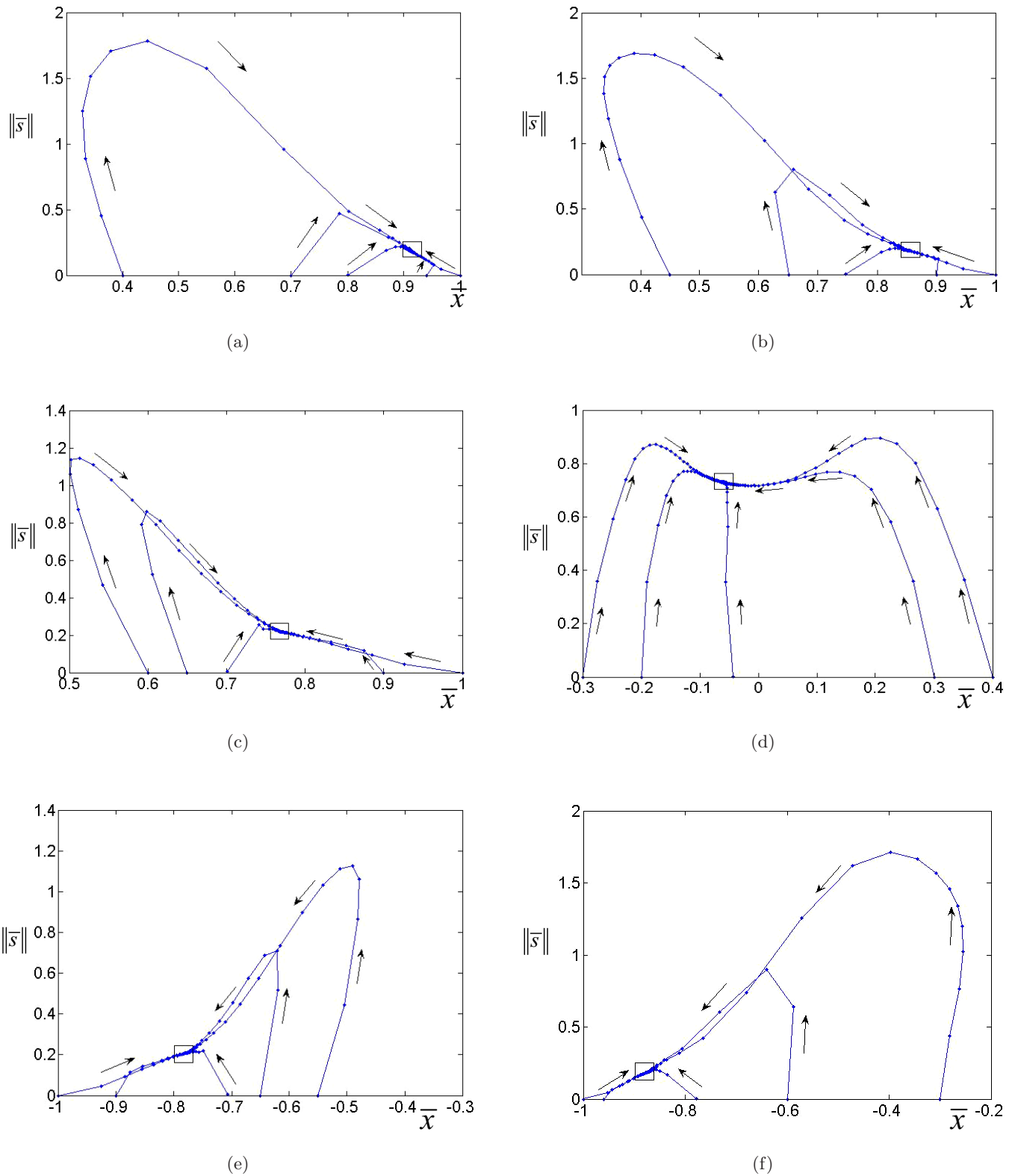


Fig. 3. Phase portraits in terms of  $\bar{x}$  (the overall average network emotional state, the first moment) and the norm  $\|\bar{s}\|$  of the vector of variances of the average emotional state for each degree (a higher-order feature) for (a)  $p = 0.0197$ , (b)  $p = 0.0384$ , (c)  $p = 0.0589$ , (d)  $p = 0.0572$ , (e)  $p = 0.0585$ , (f)  $p = 0.0316$ . Notice the fast evolution of all plotted transients to a one-dimensional slow manifold, apparently well-parameterized by  $\bar{x}$ .

statistics of a stable stationary point of a different set of networks, sampled with different  $p$  from the discrete geometric distribution, is also depicted. Initialization of this transient involves application of the restriction and then of the lifting operations. Clearly, as is shown in all phase portraits, after a relatively short transient, the norm  $\|\bar{s}\|$  of the vector of variances becomes a function of  $\bar{x}$ : it evolves quickly towards, and subsequently on a *slow manifold* parameterized by the first-order moment  $\bar{x}$  of the underlying distribution. The alternatively initialized fifth transient is the middle one in each panel; it evolves (even faster) to the same slow manifold.

#### 4.2. The coarse-grained bifurcation diagram and the associated stability analysis

Having described the choice of coarse-grained variables as well as our restriction and lifting operations, we are ready to proceed to the construction

of the coarse-grained bifurcation diagram for our system. The stationary states on the coarse-grained bifurcation diagram have been obtained as fixed points of the agent-based timestepper averaged over 1920 realizations (this particular number,  $6 \times 320$ , was dictated by our six-node cluster), by wrapping Newton-GMRES [Kelley, 1995], as a shell, around the coarse timestepper. Continuation around the coarse-grained turning points is accomplished by coupling the fixed point algorithm with pseudo-arc-length continuation. The critical eigenvalues are computed by implementing an Arnoldi iterative eigensolver which was again wrapped around our coarse timestepper. In our computations we used eighty coarse-grained observables ( $m = 80$  in restriction and lifting operators) and the time horizon  $T$  was set to seven  $\delta t$ s. The coarse-grained bifurcation diagram constructed through this approach is shown in Fig. 4. Two sets of branches (which we will refer to as the “upper” and the “lower” ones) are found. The five insets show the average emotional state for each degree at the five stationary

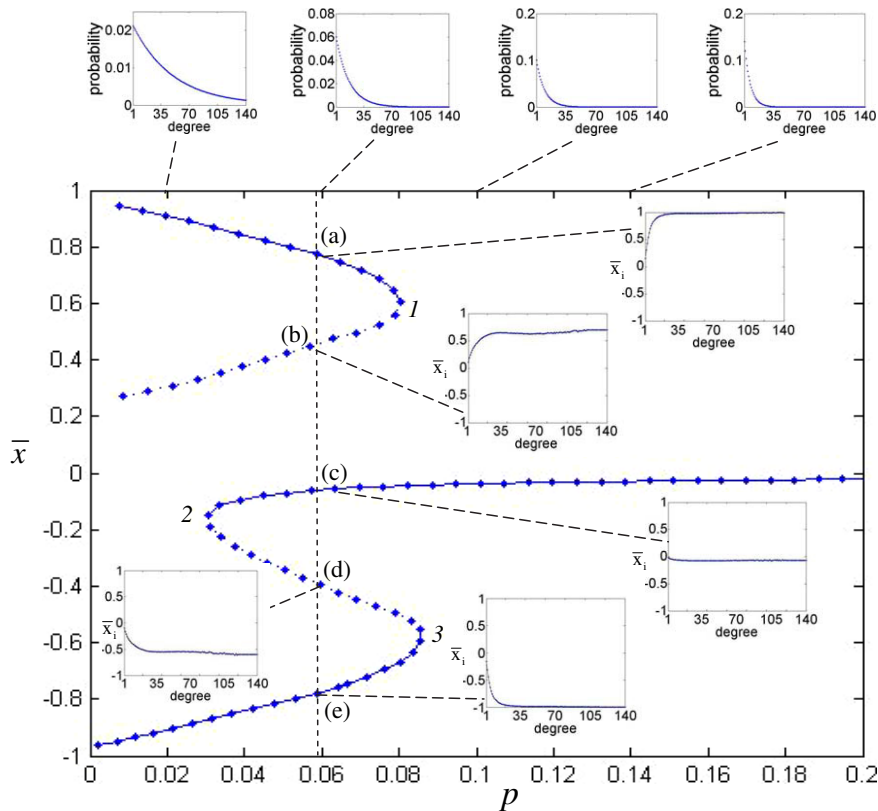


Fig. 4. Coarse-grained bifurcation diagram, plotted in terms of the overall network average emotional state  $\bar{x}$  versus  $p$ . It has been obtained by wrapping a Newton-GMRES iterative solver around the coarse timestepper; five coexisting stationary states, for approximately same  $p$  value, are marked (a)–(e). The insets show the average emotional state ( $\bar{x}_i$ , ordinate) for each degree (abscissa) at each of these stationary points, whether stable or unstable. Representative degree probability distributions at four distinct values of  $p$  are included above the bifurcation diagram.

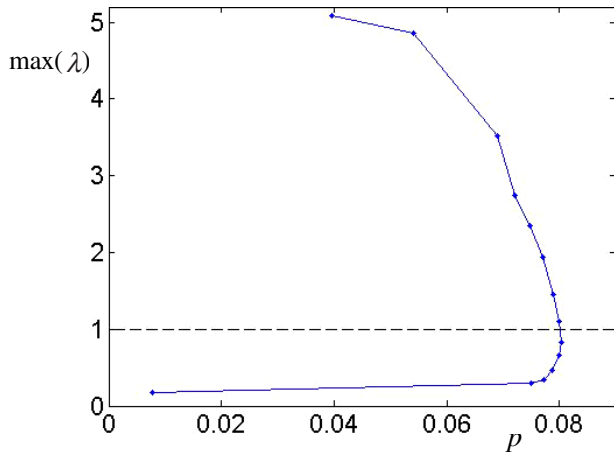


Fig. 5. The leading eigenvalue of the linearized coarse-grained dynamics versus the bifurcation parameter  $p$ ; it was computed by implementing Arnoldi's algorithm around the coarse timestepper along the "upper" solution branches of Fig. 3.

points (marked a–e) coexisting at approximately the same  $p$  value. The insets above the bifurcation diagram depict representative connectivity degree probability distributions at four different values

of  $p$ . As it is shown, steeper distribution curves (indicating that the majority of the nodes in the network have a rather small number of links) drive the system to unique stable stationary states. On the other hand, looser distribution curves (indicating that a large percentage of nodes have a large number of links) result in the emergence of multiple stationary states; depending if the system reaches one of the initial conditions. At this point, we should note that the coarse-grained bifurcation diagram has been obtained using values of  $e^+$ ,  $e^-$  which correspond to "severe" overreaction. Qualitatively similar behavior is observed for smaller values of these parameters but for bigger rates of contacts between individuals.

To perform linearized stability analysis of the obtained coarse-grained solutions we numerically approximated the dominant/critical eigenvalues of the macroscopic dynamics using the Arnoldi iterative eigensolver. Figure 5 shows the variation of the largest eigenvalue as the bifurcation parameter  $p$  varies along the upper solution branches in the coarse-grained bifurcation diagram. Figure 6 shows

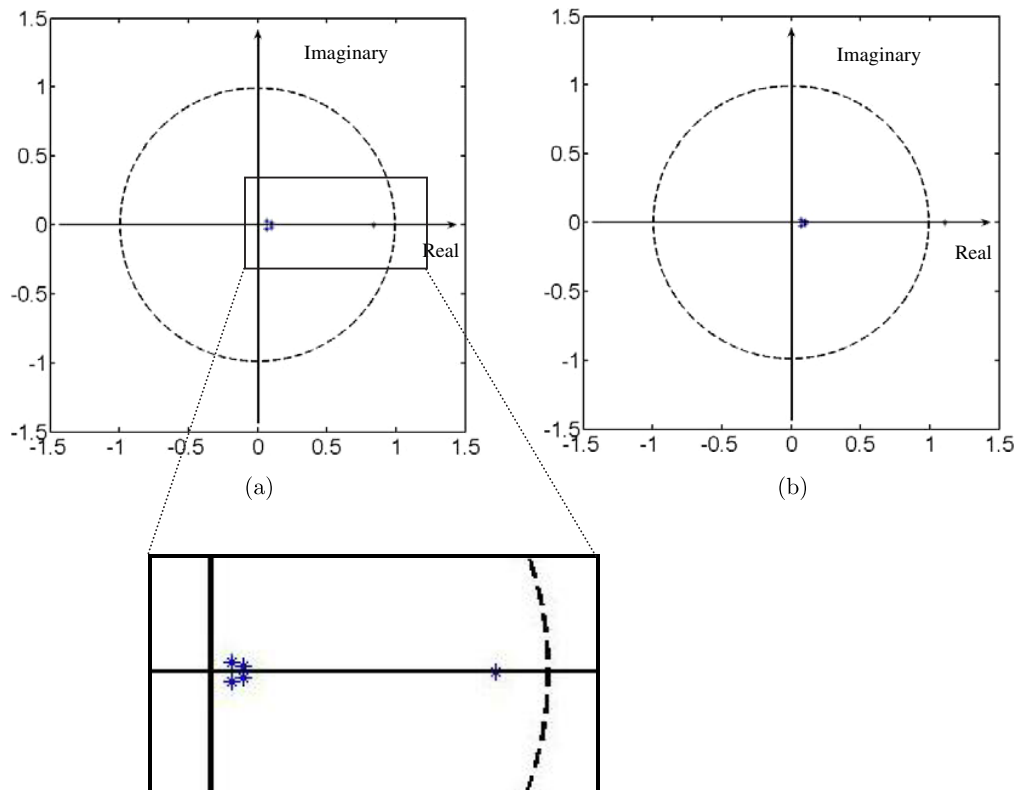


Fig. 6. Five largest eigenvalues of the linearized coarse-grained dynamics along the upper solution branches of the coarse-grained bifurcation diagram (Fig. 3). (a) Just before the "1" saddle-node bifurcation, on the stable branch of coarse-grained equilibria for  $p = 0.0804$  and (b) just after the "1" saddle-node bifurcation on the lower-lying, unstable branch for  $p = 0.0801$ . The blowup in (a) is included to clearly show the four eigenvalues forming an inner "cluster".

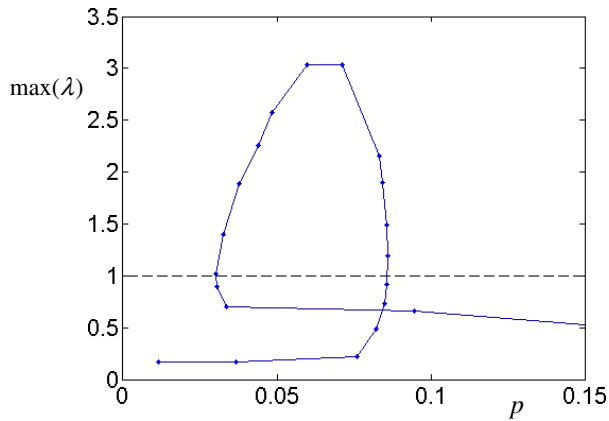


Fig. 7. The leading eigenvalue of the linearized coarse-grained dynamics versus the bifurcation parameter along the lower solution branches of the coarse-grained bifurcation diagram (Fig. 3).

the five largest coarse eigenvalues at two points, just before [Fig. 6(a)] and just after [Fig. 6(b)] the “1” saddle-node bifurcation, in the complex plane. Figure 7 shows the largest eigenvalue of the coarse-grained linearization as the bifurcation parameter  $p$  varies along the lower part of branches of the coarse-grained bifurcation diagram. Figure 8 shows the five largest eigenvalues at four points, just before [Figs. 8(a) and 8(c)] and just after [Figs. 8(b) and 8(d)] the two saddle-node bifurcations marked as “2” and “3”, respectively.

The obtained state distributions and coarse-grained stability results are consistent with temporal agent-based simulations.

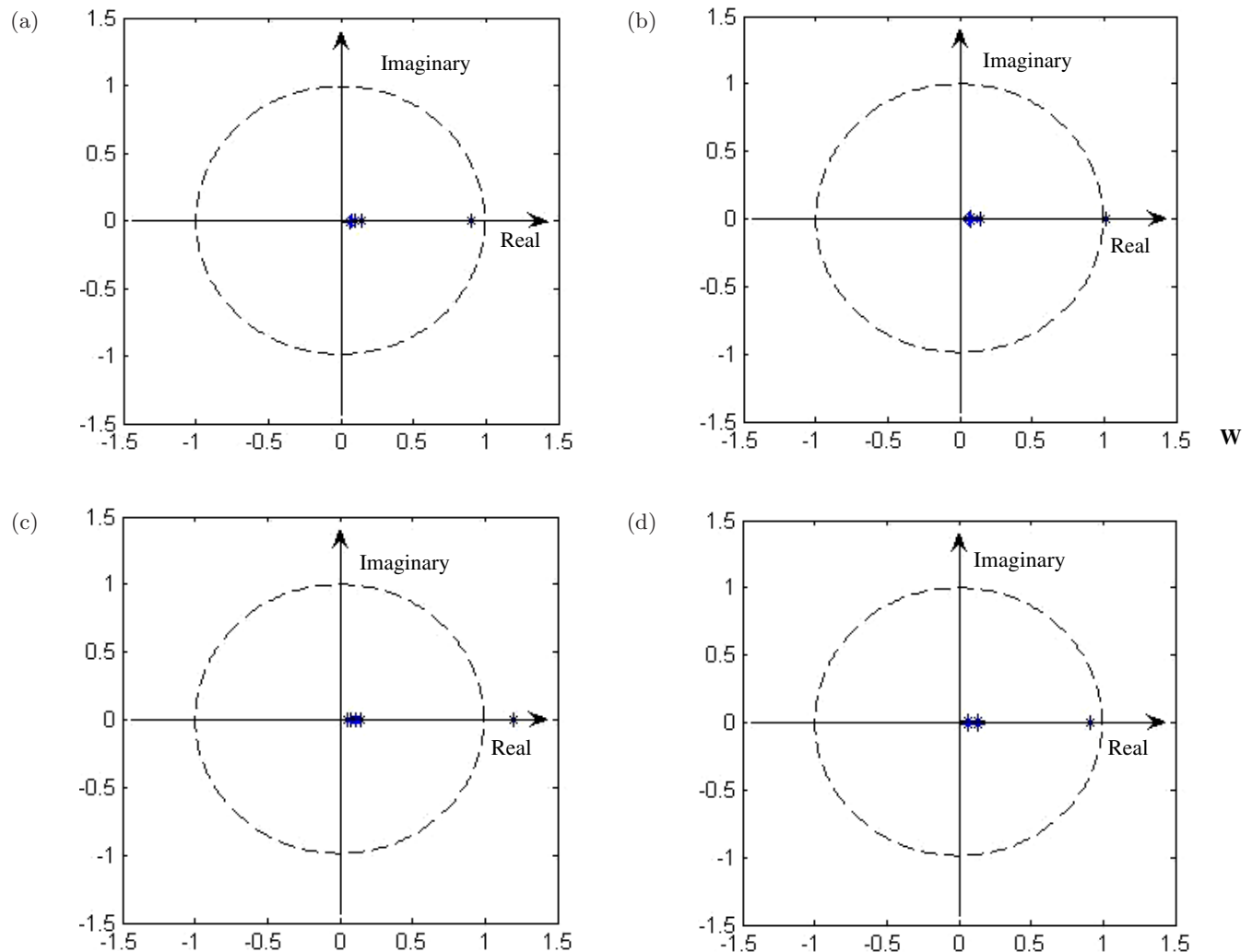


Fig. 8. Five largest eigenvalues of the linearized coarse-grained dynamics along the lower solution branches of the coarse-grained bifurcation diagram (Fig. 3): (a) just before the “2” saddle-node bifurcation on the branch of coarse-grained stable equilibria for  $p = 0.0305$ , (b) just after the “2” saddle-node bifurcation on the branch of coarse-grained unstable equilibria for  $p = 0.0303$ , (c) just before the “3” saddle-node bifurcation on the branch of coarse-grained unstable equilibria for  $p = 0.0861$ , (d) just after the “3” saddle-node bifurcation on the branch of coarse-grained stable equilibria for  $p = 0.0857$ .

## 5. Discussion and Conclusions

We have shown how the so-called “Equation-Free” approach for complex/multiscale computations can be exploited to systematically analyze the coarse-grained dynamics of a certain class of social networks. This was accomplished through the design of computational experiments using the detailed, agent-based simulation as a black-box coarse timestepper. For our illustrations, we used a computational caricature of a mimesis-like propagation process in a simple social network of individuals. The agent-based simulations show some interesting nonlinear behavior including apparent phase transitions and multiplicity of stationary states. We constructed coarse-grained bifurcation diagrams with respect to a parameter dictating the degree distribution of the networks and analyzed the (coarse) stability of the stationary solutions. At this point we should note that the identification of multiple stable equilibria is in agreement with empirical evidence and studies of behavioral and financial processes as well as other instances of complex systems (such as information propagation in networks and neuroscience) [Topol, 1991; Callaway *et al.*, 2001; Jones, 2008].

Further research could proceed along several directions including the study of “rare event” transitions at the neighborhood of coarse-grained bifurcation points, and the use of Diffusion Map techniques [Coifman *et al.*, 2005; Kolpas *et al.*, 2007] that can be used to both extract the right variables and to efficiently extract information from more complicated and thus more realistic agent-based social network simulations.

One can also envision using such coarse-grained techniques to enhance the study of the dynamics of complex network/population models which incorporate decision making based on agent-based sociological models (econophysics [Mantegna & Stanley, 1999; Chakrabarti *et al.*, 2006]), to explore strategy formation, formulate optimal control policies, and more generally, in order to prescribe/control the system (network) dynamics. Traditional numerical tools for performing such tasks in continuum models are, in principle, not applicable when the model is a detailed agent-based one. Equation-Free techniques hold the promise of linking agent-based codes with traditional continuum numerical analysis; the compactness of the coarse descriptions and the fast and quantitative extraction of macroscopic information from realistic fine scale codes may well

enhance the use of agent-based models for understanding, analyzing and even controlling certain social phenomena.

## Acknowledgments

This work was partially supported by the State Scholarships Foundation of Greece (IKY) (ACT, GVB) and the National Technical University of Athens through the Basic Research Program “Constantin Caratheodory” (ACT, CIS). IKG acknowledges partial support by the US NSF and DTRA.

## References

- Albert, R. & Barabasi, A. L. [2002] “Statistical mechanics of complex networks,” *Rev. Mod. Phys.* **74**, 47–97.
- Axelrod, R. [1997] *The Complexity of Cooperation* (Princeton University Press, Princeton).
- Balmer, M., Nagel, K. & Raney, B. [2004] “Large scale multi-agent simulations for transportation applications,” *J. Intell. Transp. Syst.* **8**, 205–221.
- Barberis, N. [2003] *Elsevier Handbook of the Economics of Finance*, Vol. 1B, Chapter 18 (Elsevier North-Holland), p. 1052.
- Boginski, V., Butenko, S. & Pardalos, P. M. [2006] “Mining market data: A network approach,” *Comput. Operat. Res.* **33**, 3171–3184.
- Bogunà, M. & Pastor-Satorras, R. [2002] “Epidemic spreading in correlated complex networks,” *Phys. Rev. E* **66**, 047104.
- Bollobás, B. [1979] *Graph Theory* (Springer, NY).
- Bonabeau, E., Dorigo, M. & Theraulaz, G. [2000] “Inspiration for optimization from social insect behavior,” *Nature* **406**, 39–42.
- Bonabeau, E. [2002] “Agent-based modelling: Methods and techniques for simulating human systems,” *PNAS* **99**, 7280–7287.
- Boots, M. & Sasaki, A. [1999] “‘Small worlds’ and the evolution of virulence: Infection occurs locally and at a distance,” *Proc. R. Soc. Lond. B* **266**, 1933–1938.
- Callaway, D. S., Hopcroft, J. E., Kleinberg, J. M., Newman, M. E. J. & Strogatz, S. H. [2001] “Are randomly grown graphs really random?” *Phys. Rev. E* **64**, 041902.
- Castellano, C., Fortunato, S. & Loreto, V. [2009] “Statistical physics of social dynamics,” *Rev. Mod. Phys.* **81**, 591–646.
- Castellano, C., Marsili, M. & Vespignani, A. [2000] “Nonequilibrium phase transition in a model for social influence,” *Phys. Rev. Lett.* **85**, 3536–3539.
- Chakrabarti, B. K., Chakraborti, A. & Chatterjee, A. [2006] *Econophysics and Sociophysics: Trends and Perspectives* (Wiley-VCH, Berlin).



- Christodoulou, K. N. & Scriven, L. E. [1998] "Finding leading modes of a viscous free surface flow: An asymmetric generalized eigenproblem," *SIAM J. Sci. Comput.* **3**, 355–406.
- Clifford, P. & Sudbury, A. [1973] "A model for spatial conflict," *Biometrika* **60**, 581–588.
- Coifman, R. R., Lafon, S., Lee, A. B., Maggioni, M., Nadler, B., Warner, F. & Zucker, S. W. [2005] "Geometric diffusions as a tool for harmonic analysis and structure definition of data: Diffusion maps," *PNAS* **102**, 7426–7431.
- Cooley, W. W. & Lohnes, P. R. [1971] *Multivariate Data Analysis* (John Wiley & Sons, Inc., NY).
- Daniel, K., Hirshleifer, D. & Teoh, S. H. [2002] "Investor psychology in capital markets: Evidence and policy implications," *J. Monetary Econ.* **49**, 139–209.
- DeAngelis, D. L., Rose, K. A. & Huston, M. A. [1994] *Frontiers in Mathematical Biology* (Springer, Berlin).
- Dijkstra, J., Jessurun, A. J. & Timmermans, H. J. P. [2001] *Pedestrian and Evacuation Dynamics* (Springer-Verlag, Berlin).
- Dorogovtsev, S. N. & Mendes, J. F. F. [2001] "Language as an evolving word web," *Proc. Roy. Soc. London Ser. B* **268**, 2603–2606.
- Dunbar, R. I. M. [1993] "Coevolution of neocortical size, group size and language in humans," *Behav. Brain Sci.* **16**, 681–735.
- Edney, J. J. [1981] "Paradoxes on the commons: Scarcity and the problem of equality," *J. Comm. Psych.* **9**, 3–34.
- Eubank, S. H., Guclu, V. S. A., Kumar, M., Marathe, M. V., Srinivasan, A., Toroczkai, Z. & Wang, N. [2004] "Modelling disease outbreaks in realistic urban social networks," *Nature* **429**, 180–184.
- Ferguson, N. M., Cummings, D. A. T., Cauchemez, S., Fraser, C., Riley, S., Meeyai, A., Iamsirithaworn, S. & Burke, D. S. [2005] "Strategies for containing an emerging influenza pandemic in Southeast Asia," *Nature* **437**, 209–214.
- Gear, C. W., Kevrekidis, I. G. & Theodoropoulos, C. [2002] "Coarse integration/bifurcation analysis via microscopic simulators: Micro-Galerkin methods," *Comput. Chem. Engin.* **26**, 941–963.
- Gladwell, M. [2000] *The Tipping Point — How Little Things Make a Big Difference* (Little, Brown and Company).
- Grimm, V. [1999] "Ten years of individual-based modelling in ecology: What we have learned, and what could we learn in the future?" *Ecol. Model.* **115**, 129–148.
- Guardiola, X., Diaz-Guilera, A., Perez, C. J., Arenas, A. & Llas, M. [2002] "Modelling diffusion of innovations in a social network," *Phys. Rev. E* **66**, 026121.
- Hein, O., Schwind, M. & Spiwoks, M. [2008] "Frankfurt artificial stock market: A microscopic stock market model with heterogeneous interacting agents in small-world communication networks," *J. Econ. Inter. Coord.* **1**, 59–71.
- Helbing, D. & Molnár, P. [1995] "Social force model for pedestrian dynamics," *Phys. Rev. E* **51**, 4282–4286.
- Helbing, D., Farkas, I. & Vicsek, T. [2000] "Simulating dynamical features of escape panic," *Nature* **407**, 487–490.
- Helbing, D. [2001] "Traffic and related self-driven many-particle systems," *Rev. Mod. Phys.* **73**, 1067–1141.
- Hirshleifer, D., Subrahmanyam, A. & Titman, S. [2006] "Feedback and the success of the irrational investors," *J. Finan. Econ.* **81**, 311–338.
- Holley, R. & Liggett, T. [1975] "Ergodic theorems for weakly interacting infinite systems and the voter model," *Ann. Probab.* **3**, 643–663.
- Hughes, R. L. [2000] "The flow of large crowds of pedestrians," *Math. Comp. Simul.* **53**, 367–370.
- Iori, G. [2002] "A microsimulation of traders activity in the stock market: The role of heterogeneity, agents' interactions and trade frictions," *J. Econ. Behav. Organiz.* **49**, 269–285.
- Jones, C. A. [2008] "Connecting blazons and neurons," *Science* **319**, 35–36.
- Julka, N., Srinivasan, R. & Karimi, I. [2002] "Agent-based supply chain management 1: Framework," *Comput. Chem. Engin.* **26**, 1755–1769.
- Kavousanakis, M., Russo, L., Siettos, C. I., Boudouvis, A. G. & Georgiou, G. C. [2008] "A timestepper approach for the systematic bifurcation and stability analysis of polymer extrusion dynamics," *J. Non-Newtonian Fluid Mech.* **151**, 59–68.
- Keeling, M. J. [2005] "Implications of network structure for epidemic dynamics," *Theor. Popul. Biol.* **67**, 1–8.
- Keeling, M. J. & Eames, K. T. D. [2005] "Networks and epidemic models, networks and epidemic models," *J. R. Soc. Interf.* **2**, 295–307.
- Kelley, C. T. [1995] *Iterative Methods for Linear and Nonlinear Equations* (SIAM, Philadelphia).
- Kevrekidis, I. G., Gear, C. W., Hyman, J. M., Kevrekidis, P. G., Runborg, O. & Theodoropoulos, C. [2003] "Equation-free coarse-grained multiscale computation: Enabling microscopic simulators to perform system-level tasks," *Comm. Math. Sci.* **1**, 715–762.
- Kevrekidis, I. G., Gear, C. W. & Hummer, G. [2004] "Equation-free: The computer-assisted analysis of complex, multiscale systems," *AI.Ch.E. J.* **50**, 1346–1254.
- Khulood, A. M. & Raed, A. Z. [2007] "Emotional agents: A modeling and an application," *Inform. Softw. Technol.* **49**, 695–716.
- Kim, K., Kim, S. Y. & Ha, D.-H. [2007] "Characteristics of networks in financial markets," *Comput. Phys. Commun.* **177**, 184–185.
- Kolpas, A., Moehlis, J. & Kevrekidis, I. G. [2007] "Coarse analysis of stochasticity-induced switching

- between collective motion states,” *PNAS* **104**, 5931–5936.
- Lambiotte, R., Ausloos, M. & Holyst, J. A. [2007a] “Majority model on a network with communities,” *Phys. Rev. E* **75**, 030101(R).
- Lambiotte, R., Thurner, S. & Hanel, R. [2007b] “Unanimity rule on networks,” *Phys. Rev. E* **76**, 046101.
- Levine, H., Rappel, W. J. & Cohen, I. [2001] “Self-organization in systems of self-propelled particles,” *Phys. Rev. E* **63**, 017101 1–4.
- Liu, Y. & Passino, K. M. [2004] “Stable social foraging swarms in a noisy environment,” *IEEE Trans. Autom. Contr.* **49**, 30–44.
- Liu, X., Gregor, S. & Yang, J. [2008] “The effects of behavioral and structural assumptions in artificial stock market,” *Physica A* **387**, 2535–2546.
- Llas, M., Gleiser, P. M., Lopez, J. M. & Diaz-Guilera, A. [2003] “Nonequilibrium phase transition in a model for the propagation of innovations among economic agents,” *Phys. Rev. E* **68**, 066101.
- Longini, I. M., Fine, P. E. & Thacker, S. B. [1986] “Predicting the global spread of new infectious agents,” *Am. J. Epidemiol.* **123**, 383–391.
- Loreto, V. & Steels, L. [2007] “Social dynamics: Emergence of language,” *Nature Phys.* **3**, 758–760.
- Makeev, A., Maroudas, D. & Kevrekidis, I. G. [2002] “Coarse stability and bifurcation analysis using stochastic simulators: Kinetic Monte Carlo examples,” *J. Chem. Phys.* **116**, 10083–10091.
- Mantegna, R. N. & Stanley, H. E. [1999] *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge).
- Molloy, M. & Reed, B. [1995] “A critical point for random graphs with a given degree sequence,” *Rand. Struct. Algorith.* **6**, 161–180.
- Newman, M. E. J., Watts, D. J. & Strogatz, S. H. [2002] “Random graph models of social networks,” *PNAS* **99**, 2566–2572.
- Newman, M. E. J. [2003] “The structure and function of complex networks,” *SIAM Rev.* **45**, 167–256.
- Omurtag, A. & Sirovich, L. [2006] “Modelling a large population of traders: Mimesis and stability,” *J. Econ. Behav. Organiz.* **61**, 562–576.
- Palatella, L., Perellò, J., Montero, M. & Masoliver, J. [2005] “Diffusion entropy technique applied to the study of the market activity,” *Physica A* **355**, 131–137.
- Parisi, D. R. & Dorso, C. O. [2005] “Microscopic dynamics of pedestrian evacuation,” *Physica A* **354**, 606–618.
- Raberto, M., Cincotti, S., Focardi, S. M. & Marchesi, M. [2001] “Agent-based simulation of a financial market,” *Physica A* **299**, 319–327.
- Raney, B., Cetin, N., Völlmy, A., Vrtic, M., Axhausen, K. & Nagel, K. [2003] “An agent-based microsimulation model of Swiss travel: First results,” *Networks and Spatial Econ.* **3**, 23–42.
- Reichardt, J. & White, D. R. [2007] “Role models for complex networks,” *Eur. Phys. J. B* **60**, 217–224.
- Ritter, J. R. [2003] “Behavioral finance,” *Pacific-Basin Fin. J.* **11**, 429–437.
- Runborg, O., Theodoropoulos, C. & Kevrekidis, I. G. [2002] “Effective bifurcation analysis: A timestepper-based approach,” *Nonlinearity* **15**, 491–511.
- Saad, Y. [1992] *Numerical Methods for Large Eigenvalue Problems* (Manchester University Press: Oxford-Manchester).
- Samanidou, E., Zschischang, E., Stauffer, D. & Lux, T. [2007] “Agent-based models of financial markets,” *Rep. Prog. Phys.* **70**, 409–450.
- Siettos, C. I., Graham, M. & Kevrekidis, I. G. [2003] “Coarse Brownian dynamics for nematic liquid crystals: Bifurcation diagrams via stochastic simulation,” *J. Chem. Phys.* **118**, 10149–10156.
- Steels, L. & Baillie, J. C. [2000] “Evolving grounded communication for robots,” *Trends Cogn. Sci.* **7**, 308–312.
- Strogatz, S. H. [2001] “Exploring complex networks,” *Nature* **410**, 268–276.
- Szabó, G. & Fàth, G. [2007] “Evolutionary games on graphs,” *Phys. Rep.* **446**, 97–216.
- Tessone, C. J. & Toral, R. [2005] “System size stochastic resonance in a model for opinion formation,” *Physica A* **351**, 106–116.
- Theodoropoulos, C., Qian, Y. H. & Kevrekidis, I. G. [2000] “Coarse stability and bifurcation analysis using time-steppers: A reaction-diffusion example,” *PNAS* **97**, 9840–9843.
- Topol, R. [1991] “Bubbles and volatility of stock prices: Effect of mimetic contagion,” *Econ. J.* **101**, 786–800.
- Wang, S. & Zhang, C. [2005] “Microscopic model of financial markets based on belief propagation,” *Physica A* **354**, 496–504.
- Weidlich, W. [2002] *Sociodynamics: A Systematic Approach to Mathematical Modelling in Social Sciences* (Taylor and Francis, London, UK).
- Wu, F. & Huberman, B. A. [2007] “Novelty and collective attention,” *PNAS* **104**, 17599–17601.

## Appendix

### The Construction of a Connected Random Graph with a Given Degree Probability Distribution Without Self-Loops and Multiple Edges

We constructed the graph using a modified version of the algorithm described in [Molloy & Reed, 1995] in order to create a network which precludes the presence of self-loops, multiple edges and ensures that there is at least one path connecting any two

nodes of the network. We should note that the algorithm as described in [Molloy & Reed, 1995] does not guarantee a connected graph and the occurrence of self-loops and multiple edges is not ruled out. The modified algorithm reads as follows:

The degree  $\deg(v_i)$  of a vertex  $v_i$  of a simple graph  $G(V, E)$  is defined as

$$\deg(v_i) = |e_{v_l v_j}| \quad \forall v_i \in V, \quad \text{where } e_{v_l v_j} \in E$$

$$\text{and } v_j = v_i \quad \text{or} \quad v_l = v_i.$$

(a) For each node choose randomly a degree according to the given discrete geometric distribution. In an undirected graph the sum of the node degrees should be even, since the number of edges is an integer equal to  $(\sum_{i=1}^N \deg(v_i))/2$ . To satisfy this constraint, we may need to repeatedly select the degree of the last node randomly so that the sum becomes even. An undirected connected graph has a minimum of  $N - 1$  edges, where  $N$  is the number of vertices. If by sampling there are less than the minimum needed edges we resample the probability distribution.

(b) We then randomly connect the vertices of our graph in such a way that there are no two vertices connected by more than one edge and no self-loops.

(c) After that, as we want to construct a one component graph, we identify the components of our graph, count them, we identify the edges that participate in circuits<sup>2</sup> of the components, and we finally link these components. The first two steps are achieved using a modified breadth-first search algorithm with “Last In First Out” buckets, for efficiency, in order to create the spanning trees, thus identifying the components and registering those edges that are not part of the spanning trees (these edges belong to circuits). We list the graph components in descending order of the number of registered edges of each component. Knowing the edges that are part of circuits, we can use them in order to progressively link components without causing any subcomponent to become disconnected (the structure of the spanning tree created from each component that has at least a circuit remains unaltered). The steps of the linking algorithm are then:

(a) Register all components in descending order of the number of circuits in a components vector,

*component*[], set *componentCounter* := 1 and set *numOfComponents* equal to the number of the components of the graph;

- (b) Insert all edges that create a circuit (the edges not used in the spanning tree) of the *component*[*componentCounter*] into the *circuit\_vector*;
- (c) Select one of the edges randomly and remove it from the *circuit\_vector*;
- (d) If the number of the circuits of the *component*[*componentCounter* + 1] are greater than zero then

- (i) select one of the edges that create a circuit and insert all the other edges to the *circuit\_vector*;
- (ii) delete the two selected edges and using the four vertices involved create (randomly among the two options available) two new edges, different from the deleted ones;
- (iii) insert one of the new edges (chosen randomly) to the *circuit\_vector*;

else

- (i) select one of the edges of the component randomly;
- (ii) delete the two selected edges, randomly create two other between the four vertices (different than the deleted ones)

endif;

- (e) Set *componentCounter* := *componentCounter* + 1;
- (f) If *componentCounter* < *numOfComponents* then  
     goto c,  
   endif;
- (g) Terminate.

The nature of the problem makes the linking-part of the algorithm simple, in the sense that linking components do not violate, by construction, certain constraints (no self-loops, no multiple edges between nodes) of our graphs. The two largest components of the graph have been constructed randomly and thus a direct connection to them does not insert any bias into the algorithm. Moreover, all rewiring parts of the algorithm maintain the degree chosen for each vertex at the first step of the algorithm.

<sup>2</sup>A circuit is a walk starting from a node  $x$  and ending at the same node without having visited any other node more than once or used any edge more than once.