Fear induced explosive transitions in the dynamics of corruption

F. Bauzá,[1,2] [ID] D. Soriano-Paños,[1,3] [ID] J. Gómez-Gardeñes,[1,3,4,a)] [ID] and L. M. Floría[1,3,b)]

AFFILIATIONS

[1]GOTHAM Laboratory, Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza,
50018 Zaragoza, Spain
[2]Department of Theoretical Physics, University of Zaragoza, 50009 Zaragoza, Spain
[3]Department of Condensed Matter Physics, University of Zaragoza, 50009 Zaragoza, Spain
[4]Center for Computational Social Science, University of Kobe, 657-8501 Kobe, Japan

Note: This article is part of the Focus Issue, Dynamics of Social Systems.
[a)]**Author to whom correspondence should be addressed:** gardenes@unizar.es
[b)]**Electronic mail:** floria@unizar.es

ABSTRACT

In this article, we analyze a compartmental model aimed at mimicking the role of imitation and delation of corruption in social systems. In particular, the model relies on a compartmental dynamics in which individuals can transit between three states: honesty, corruption, and ostracism. We model the transitions from honesty to corruption and from corruption to ostracism as pairwise interactions. In particular, honest agents imitate corrupt peers while corrupt individuals pass to ostracism due to the delation of honest acquaintances. Under this framework, we explore the effects of introducing *social intimidation* in the delation of corrupt people. To this aim, we model the probability that an honest delates a corrupt agent as a decreasing function of the number of corrupt agents, thus mimicking the fear of honest individuals to reprisals by those corrupt ones. When this mechanism is absent or weak, the phase diagram of the model shows three equilibria [(i) full honesty, (ii) full corruption, and (iii) a mixed state] that are connected via smooth transitions. However, when *social intimidation* is strong, the transitions connecting these states turn explosive leading to a bistable phase in which a stable full corruption phase coexists with either mixed or full honesty stable equilibria. To shed light on the generality of these transitions, we analyze the model in different network substrates by means of Monte Carlo simulations and deterministic microscopic Markov chain equations. This latter formulation allows us to derive analytically the different bifurcation points that separate the different phases of the system.

Published under license by AIP Publishing. https://doi.org/10.1063/5.0004826

The diffusion of behaviors, ideas, and social norms can be modeled as contagion processes. However, unlike usual epidemic models, the diffusion of behaviors generally implies more complex routes of infection than those observed for pathogens such as viruses or bacteria. Therefore, depending on the particular aspects of the human behavior whose dissemination is intended to be modeled, one must incorporate different transmission mechanisms from the usual ones in epidemic models. Here, we consider a compartmental model with three states: honesty, corruption, and ostracism, to address the imitation and delation of corrupt behaviors. Both processes involve interactions with peers, so honest agents can get corrupted after interacting with corrupt agents, while corrupt agents can be betrayed by interacting with honest individuals. However, to model these pairwise interactions, we incorporate an essential ingredient to this particular social dynamics: fear. On one hand, the fear of corrupt individuals to be betrayed is a mechanism that favors their reinsertion into the honest compartment. This fear is a growing function of the fraction of agents already betrayed and, therefore, in the state of ostracism. On the other hand, fear also plays a central role in the probability that a honest agent betrays a corrupt peer, by weakening this probability when the number of corrupt agents increases, due to the fear of possible reprisals. Both mechanisms are essential for explaining the different transitions and phases of the system. In particular, we show how the inclusion of the fear to delation by the honest leads to abrupt (explosive) transitions between honest and corrupt societies.

I. INTRODUCTION

From the seminal work of Schelling[1] in 1978, it is widely accepted that local interactions between individuals are the basic components for the development of macroscopic social behaviors. On the other hand, the connection between microscopic interactions and collective phenomena is the starting point for all applications of statistical physics. Therefore, as Anderson suggested in a work[2] a few years before that of Schelling, although the underlying mechanisms that drive the interactions between agents in society seem to be very different from fundamental physical laws, the application of the vision, techniques, and tools of statistical mechanics seems the best avenue to address many social problems in which the emergence of collective behavior is at work.[3] Examples of these problems include social segregation,[4] the spread of rumors,[5] cultural dissemination,[6] the adoption of technology and ideas,[7] or the emergence of cooperative traits,[8] among others.

The original social models were often criticized for their simplicity and the impossibility of their validation. Although the first criticism can also be made to any model of statistical physics that aims at capturing the essential ingredients that give rise to the observed behavior (that is, neglecting particular details that are irrelevant at the macroscopic level), the second criticism shows them as simple *toy models* with which to capture the qualitative behavior observed in real scenarios. This second criticism, however, is being overcome today thanks to advances in data gathering and the possibility of capturing social interactions through online networks and mobile devices.

The incorporation of interaction patterns between individuals through complex networks was the first step toward the use of these models in real contexts.[9–11] In this direction, and always aligned with the capture of data in social systems, the realism of these models has been increasing with the addition of ingredients such as the volatility of contacts,[12,13] the multilayer character of many social systems,[14–16] or the consideration of higher order interactions in different dynamics.[17]

One of the paradigmatic examples of the former view and evolution of social models is that of social contagion. Examples of these are the rumor spreading models (such as those of Daley and Kendall[5] and Maki and Thompson[18]) or those capturing complex contagions for the adoption of norms/innovations/ideas.[7,19–22] Social contagion models usually take the compartmental models of theoretical epidemiology[23,24,41] as the benchmark for capturing the interactions between individuals, the state of each agent being given by a discrete variable while the transitions among states are given by certain probabilities depending of the state of the rest of the population.

In this work, we make use of social contagion dynamics to tackle the modeling of corruption, a kind of social norm violation that is ubiquitous across the history of human cultures and societies. The social dynamics of corrupt behaviors, explicitly realized as bribery practices,[25–32] has been analyzed in the frame of (either classical or evolutionary) game theory,[33] where behaviors are formally represented by game's strategies whose spreading over the population relies on the obtained benefits.

However, when modeling the social dynamics of corruption in game theoretical terms, the existence of a large parameter space and/or strategic space often prevents us from performing a detailed analysis of model computations. In this sense, the need of a clear-cut analysis on which ingredients are most relevant and which others are not so much recommends simplifying the modeling assumptions while trying to keep at least some of those that are essential.

A recent attempt along this line of thought has been made in Ref. 34, where a simple compartmental model of corruption dynamics is introduced. There, the individuals can transit between three states, say, honesty, corruption, and ostracism, representing, respectively, the behaviors that conform to the social norm (H), violate it (C), and have been found corrupt and punished through its expulsion from the society (O). The microscopic dynamics of this HCO model is postulated through four hypothesis on the population flows among compartments, where general structured (in terms of social contacts) populations are considered:

- The first assumption is that corruption is a socially infective event: honest individuals become corrupt *only* by interaction with their corrupt acquaintances, at an infection rate f_α^{HC}, that is, a function of their local configuration.

This assumption for the *corruption* flow, $H \to C$, conceptualizes the corrupt behavior as an infective state and not as a game strategy. However, this does not prevent a game theoretic perspective because the (social) infectious power of a strategy is what allows its dissemination.

- The flow $C \to O$ is *exclusively* the outcome of the delation of corrupt individuals by their honest neighbors, at a delation rate f_β^{CO}, also a function of their local configuration. This way, we consider that only pairwise C–H interactions are the origin of this flow, thus excluding other possible sources such as administrative inspection, or police investigation.
- The third assumption is that, at a given constant rate r, the O individuals are reintroduced into the society as H individuals. This flow is called *reinsertion* flow.
- Finally, a fourth flow, the *conversion* transition $C \to H$, incorporates the *warning-to-wrongdoers* effect of social punishment. The rate governing this flow is strongly linked to the social perception of risk to be delated, here quantified as the fraction, $\langle O \rangle$, of population in the O compartment.

In Fig. 1, we represent a *chart flow* of the HCO model, where the assumptions made for the flow between any two compartments are visualized. From a non-equilibrium statistical mechanics perspective,[35] the model is a (non-Hamiltonian) version of the kinetic spin-1 Ising model, also referred to as the Blume–Emery–Griffiths model[36] in condensed matter physics, introduced by these authors as a model for the superfluid transition in He3–He4 liquid mixtures. The spin-1 Ising model has already a tradition in social dynamics, as emphasized in Ref. 37 regarding the Schelling model of urban segregation. While in the Blume–Emery–Griffiths model, the spin state of null z-component represents the helium 3 isotope and in the Schelling model, it represents an empty flat in the urban neighborhood, in the HCO model, it denotes the state of strict social isolation (ostracism).

In Ref. 34, the authors analyzed in full detail the mean field approximation to the dynamics of the HCO model, and its
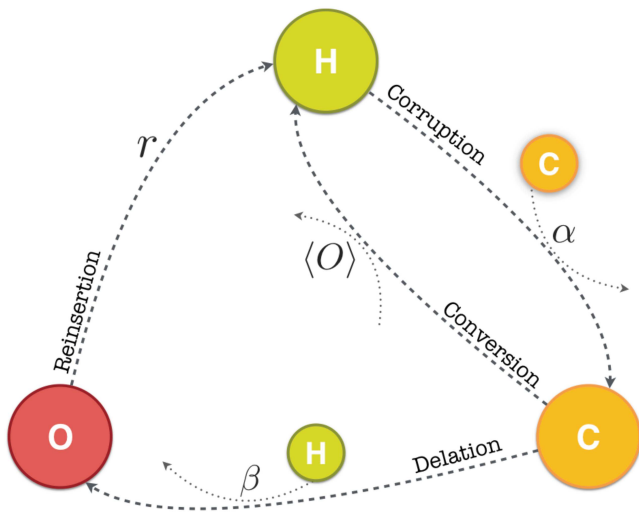
FIG. 1. *Chart flow* representation of the model. Four different flows among population compartments are considered. First, the flow $O \rightarrow H$ (reinsertion) occurs at a constant rate r. The flow $C \rightarrow H$ (conversion) is driven by the fear of being delated that we simply quantify by $\langle O \rangle$. The other two flows, namely, corruption and delation flows, are determined by pairwise (C–H) interactions and occur with rate f_α and f_β, respectively.

comparison with the numerical results of Monte Carlo (MC) simulations on random and non-random regular networks, for a type of simple one-variable contact interaction functions f_α^{HC} and f_β^{CO}. The three-dimensional phase diagram of the model shows the existence of three generic absorbing states, namely, (i) full honesty, (ii) full corruption, and (iii) a mixed state with non-zero flow through all the flow channels. The model shows no multistability and two continuous transitions "full honesty–mixed state" (corruption transition) and "full corruption–mixed state" (honesty transition), as the model parameters are tuned. The warning-to-wrongdoers has no influence on the corruption transition, but it plays an important role in the honesty transition, favoring the spreading of honest instabilities in the fully corrupt state.

In this contribution, we explore the effects of introducing *social intimidation* in the delation of corrupt people. Furthermore, we aim at unveiling the effects of introducing non-regular topologies governing the evolution of corruption. To this aim, we leverage the original equations introduced in Ref. 34 based on the Microscopic Markov Chain Approach (MMCA)[38–40] and modify the transition rate f_β^{CO} to account for the new ingredient here introduced. The resulting equations of the model dynamics for the MC simulations and both deterministic MMCA and mean field theories are formulated in Sec. II. In Sec. III, we derive closed expressions for the boundaries of the aforementioned equilibrium phases. In Sec. IV, the formalism is validated and the phase diagrams of the model for different contact networks are characterized. Finally, we discuss the relevance of our findings and their implications for future research in Sec. V.[41]

II. THE MODEL EQUATIONS

The system is composed of N agents formally represented by an unweighted undirected network given by its adjacency matrix **A**. From the assumptions that define the *HCO* model in Sec. I, the stochastic (MC) simulations of the dynamics are implemented in the following way:

Assuming that the state of each agent i at time step t is denoted as $\sigma_i(t)$:

(i) If $\sigma_i(t) = H$, then $\sigma_i(t+1) = C$ with transition probability f_α^{HC}, a (yet unspecified) function of the local network configuration around i. The agent remains honest with probability $1 - f_\alpha^{HC}$.

(ii) If $\sigma_i(t) = C$, then $\sigma_i(t+1) = H$ (warning-to-wrongdoers effect) with probability $\langle O \rangle$, the fraction of population in O state. Then, if not converted (probability $1 - \langle O \rangle$), the corrupt agents will be delated to O state with transition probability f_β^{CO}, a (yet unspecified also) function of the local configuration around i. Thus, agent i keeps corrupt at $t+1$ with probability $(1 - \langle O \rangle)(1 - f_\beta^{CO})$.

(iii) If $\sigma_i(t) = O$, then $\sigma_i(t+1) = H$ with probability r while remaining O ($\sigma_i(t+1) = O$) with probability $1 - r$.

The choice for the specific functions $f_\alpha^{HC}(i, \{\sigma_j\})$ and $f_\beta^{CO}(i, \{\sigma_j\})$ should reflect the assumptions about the pairwise nature of corruption/delation processes. In particular, as these probabilities originate exclusively from interaction among individuals in different (C, H) states, they must satisfy

$$f_\alpha^{HC}(i, \{\sigma_j\}) = 0 \text{ if } \sigma_j \neq C \text{ for all } j \text{ such that } A_{ij} = 1, \quad (1)$$

$$f_\beta^{CO}(i, \{\sigma_j\}) = 0 \text{ if } \sigma_j \neq H \text{ for all } j \text{ such that } A_{ij} = 1. \quad (2)$$

In Ref. 34, the choice made for these functions mimics the familiar implementation of infective interactions in MC simulations on compartmental epidemic models as susceptible-infected-susceptible (SIS), susceptible-infected-recovered (SIR), etc.

$$f_\alpha^{HC}(i, \{\sigma_j\}) = 1 - \prod_{j=1}^{N} (1 - \alpha A_{ij} \delta_{\sigma_j, C}), \quad (3)$$

$$f_\beta^{CO}(i, \{\sigma_j\}) = 1 - \prod_{j=1}^{N} (1 - \beta A_{ij} \delta_{\sigma_j, H}), \quad (4)$$

where $\delta_{x,y}$ is Kronecker's delta. The rationale for (3) is that a honest focal agent interact with all its corrupt neighbors, and in each one of these contacts, the probability of becoming corrupt is α. Analogously, for (4), a corrupt focal agent contacts all its honest neighbors and in each contact is delated with probability β.

At variance with Ref. 34, here we additionally want to incorporate the effect of the social intimidation (coercion) that honest agents might experience when delating corrupt neighbors. To do it in a simple way, we modify the functional form (4) for the delation rate and use instead the following:

$$f_\beta(i, \{\sigma_j\}) = 1 - \prod_{j=1}^{N} \left[1 - \beta(1 - \gamma \langle C \rangle) A_{ij} \delta_{\sigma_j, H} \right], \quad (5)$$

where $\langle C \rangle$ is the fraction of agents in the C compartment, and γ is a parameter ($0 \le \gamma \le 1$) measuring the strength of the social intimidation felt by honest people. With this choice, if $\gamma = 0$, we recover the *HCO* model without intimidation analyzed in Ref. 34. On the other hand, the effect of intimidation is determined by the level $\langle C \rangle$ of existing corruption in the system. We keep the function (3) for the corruption rate,

$$f_\alpha(i, \{\sigma_j\}) = 1 - \prod_{j=1}^{N}(1 - \alpha A_{ij}\delta_{\sigma_j, C}). \qquad (6)$$

Following Refs. 38–40, we introduce a non-linear Markov process governing the dynamics of the agents' state configurations. This so-called Microscopic Markov Chain Approach (MMCA), widely used in spreading dynamics, assigns to each agent i, at time t, a real vector $\vec{\rho}(i; t)$ whose components are the probabilities that the agent is in each of the possible states, namely,

$$\vec{\rho}(i; t) \equiv (\rho_h(i; t), \rho_c(i; t), \rho_o(i; t)).$$

The transition probabilities (i)–(iii) introduced above define a non-linear Markov process for the time evolution of these probabilities $\vec{\rho}(i; t + 1) = \mathbf{Q} \, \vec{\rho}(i; t)$, where

$$\begin{pmatrix} 1 - f_\alpha & \langle \rho_o \rangle & r \\ f_\alpha & (1 - f_\beta)(1 - \langle \rho_o \rangle) & 0 \\ 0 & f_\beta(1 - \langle \rho_o \rangle) & 1 - r \end{pmatrix}$$

is the matrix representation of $\mathbf{Q}$, and $\langle \rho_o \rangle$ is the fraction of population in O state at time t, i.e.,

$$\langle \rho_o \rangle = N^{-1} \sum_i \rho_o(i; t). \qquad (7)$$

The rates (6) and (5) translate into

$$f_\alpha(i, \{\vec{\rho}(j; t)\}) = 1 - \prod_{j=1}^{N} \left[1 - \alpha A_{ij}\rho_c(j; t) \right], \qquad (8)$$

$$f_\beta(i, \{\vec{\rho}(j; t)\}) = 1 - \prod_{j=1}^{N} \left[1 - \beta(1 - \gamma\langle\rho_c\rangle) A_{ij}\rho_h(j; t) \right], \qquad (9)$$

where $\langle \rho_c \rangle = N^{-1} \sum_i \rho_c(i; t)$ is the population average of ρ_c at time t.

The probabilities $\rho_m(i; t)$ ($m = h, c, o$) should meet the normalization condition

$$\sum_m \rho_m(i; t) = 1 \quad \forall (i, t). \qquad (10)$$

Then, their temporal evolution is given by $2N$ coupled difference equations, and one has the freedom to choose any two components (e.g., ρ_h and ρ_c, or ρ_c and ρ_o) as independent variables. For example,

choosing ρ_c and ρ_o, the evolution equations read

$$\rho_c(i; t + 1) = \rho_c(i; t) \left[\left(1 - f_\beta(i) \right) \left(1 - \langle \rho_o(t) \rangle \right) \right] + [1 - \rho_c(i; t) - \rho_o(i; t)] f_\alpha(i), \qquad (11)$$

$$\rho_o(i; t + 1) = \rho_o(i; t)(1 - r) + \rho_c(i; t) f_\beta(i) \left[1 - \langle \rho_o(t) \rangle \right]. \qquad (12)$$

The mean field approximation relies on the assumption of homogeneity of both field (agent state; *H*, *C*, or *O*) and structure of contacts (adjacency matrix *A*). Under these circumstances, it seems plausible considering average behavior as a good (least biased) estimation of agent's behavior, i.e., $\vec{\rho}(i) = \langle \vec{\rho} \rangle$ (for all i) for the associated Markov process, and the neighborhood of size k can be selected at random among the population at each time step (well-mixed population assumption). This way, the mean field discrete time dynamics is a non-linear two-dimensional map of the simplex (triangle) $\mathcal{S}_2$ (i.e., $0 \le \rho_h, \rho_c \le 1$, $\rho_h + \rho_c \le 1$) onto itself, if one chooses ρ_h and ρ_c as independent variables.

One associates to this map a two-dimensional flow (continuous time dynamics) defined by the velocity (2D vector) field $\vec{F}(\vec{\rho})$, on the simplex $\vec{F}(\vec{\rho}) = \dot{\vec{\rho}} \equiv (\dot{\rho}_h, \dot{\rho}_c)$ whose components are

$$F_h(\vec{\rho}) = -(f_\alpha + r + \rho_c)\rho_h + (r + \rho_c)(1 - \rho_c),$$
$$F_c(\vec{\rho}) = [f_\alpha + (1 - f_\beta)\rho_c]\rho_h + [(1 - f_\beta)\rho_c - 1]\rho_c, \qquad (13)$$

where f_α and f_β are now the functions

$$f_\alpha(\rho_c) = 1 - (1 - \alpha\rho_c)^k, \qquad (14)$$

$$f_\beta(\rho_h, \rho_c) = 1 - [1 - \beta(1 - \gamma\rho_c)\rho_h]^k. \qquad (15)$$

One can easily check, from the equations of motion of both MMCA [Eqs. (11) and (12)] and mean field [Eq. (13)] approaches, that the states of full honesty ($\rho_h = 1$) and full corruption ($\rho_c = 1$) are fixed points for all values of the model parameters.

In what follows, we focus our analysis on the MMCA equations, for they allow us to capture the effects of heterogeneous contact networks on the unfolding of corruption. In this sense, the homogeneous assumption of the mean-field theory limits its applicability in this scenario. However, we include a thorough analysis of the mean field (MF) equations in the Appendix, which is helpful in understanding the crucial role played by the new social intimidation mechanism. In particular, we find that, as a result of this mechanism, full honesty and full corruption can coexist as stable solutions in some regions of the parameters space.

III. LINEAR STABILITY ANALYSIS

The study of the equilibrium points for the original HCO model revealed the existence of two different transitions: one above which the fully honest state is no longer a stable configuration and the other one destabilizing a fully corrupt society. Here, we aim at deriving an analytical expression for both thresholds to shed light on the effects of the ingredients here introduced on the onset of corruption. For this purpose, let us first assume that the system has reached its steady state implying that $\rho_c(i; t + 1) = \rho_c(i; t) = \rho_c^*(i)$, $\rho_o(i; t + 1) = \rho_o(i; t) = \rho_o^*(i)$ $\forall i$. Under these conditions, Eqs. (11) and (12)

turn into

$$f_\beta(i)\rho_c^*(i) = -\langle\rho_o^*\rangle[1-f_\beta(i)]\rho_c^*(i) + [1-\rho_c^*(i)-\rho_o^*(i)]f_\alpha(i),$$ (16)

$$r\rho_o(i) = \rho_c^*(i)f_\beta(i)\left[1-\langle\rho_o^*\rangle\right].$$ (17)

A. Corruption threshold

We now address the threshold triggering the emergence of corrupt agents. Close to the full honesty state, we can assume that $\rho_c^*(i) = \varepsilon_c^i \ll 1$ and $\rho_o^*(i) = \varepsilon_o^i \ll 1 \; \forall i$. The latter allows us to linearize the equations by neglecting those terms $\mathcal{O}(\varepsilon^2)$. This way, Eqs. (16) and (17) now read

$$f_\beta(i)\varepsilon_c^i = (1-\varepsilon_c^i-\varepsilon_o^i)f_\alpha(i),$$ (18)

$$r\varepsilon_o^i = \varepsilon_c^i f_\beta(i).$$ (19)

We now must compute the probabilities of being corrupted or delated, $f_\alpha(i)$ and $f_\beta(i)$, respectively, in this scenario. Both can be expressed as

$$f_\alpha(i) = 1-\prod_{j=1}^N(1-\alpha A_{ij}\varepsilon_c^j) \simeq \alpha\sum_{j=1}^N A_{ij}\varepsilon_c^j,$$ (20)

$$f_\beta(i) = 1-\prod_{j=1}^N[1-\beta(1-\gamma\langle\varepsilon_c\rangle)A_{ij}\rho_h^j] \simeq 1-(1-\beta)^{k_i},$$ (21)

where k_i denotes the degree of node i. Plugging these values into Eq. (18) and keeping only linear terms yields

$$\varepsilon_c^i = \alpha\sum_{j=1}^N \underbrace{\frac{A_{ij}}{1-(1-\beta)^{k_i}}}_{M_{ij}} \varepsilon_c^j.$$ (22)

Therefore, obtaining the expression for the corruption threshold involves solving an eigenvalue problem. As we are interested in the minimum value of α triggering the emergence of corruption, this threshold is given by

$$\alpha_c = \frac{1}{\Lambda_{\max}(\mathbf{M})},$$ (23)

where $\Lambda_{\max}(\mathbf{M})$ represents the spectral radius of matrix $\mathbf{M}$.

B. Intimidation threshold

Let us now address the destabilization of a full corrupt population. For this purpose, we aim at obtaining the value of the intimidation γ_c below which a full corrupt state is no longer a stable solution of the dynamics. For the sake of simplicity, we explicitly write the stationary equation for the honest population, which can be easily obtained by introducing Eqs. (5) and (6) into the normalization condition. This way, the stationary probability that an agent i belongs to the honest population is given by

$$f_\alpha(i)\rho_h^*(i) = r\rho_o(i) + \rho_c^*(i)\langle\rho_o\rangle.$$ (24)

We now consider that, close to the intimidation threshold, both honest people and agents in the ostracism are negligible. In mathematical terms, this is reflected by considering $\rho_h^*(i) \simeq \varepsilon_h^i \ll 1$ and

$\rho_o^*(i) = \varepsilon_o^i \ll 1$ for all i. In this case, the probabilities $f_\alpha(i)$ and $f_\beta(i)$ are expressed as

$$f_\alpha(i) = 1-\prod_{j=1}^N[1-\alpha A_{ij}\rho_c(j)] \simeq 1-(1-\alpha)^{k_i},$$ (25)

$$f_\beta(i) = 1-\prod_{j=1}^N[1-\beta(1-\gamma\langle\rho_c\rangle)A_{ij}\varepsilon_h^j] \simeq \beta(1-\gamma)\sum_{j=1}^N A_{ij}\varepsilon_h^j.$$ (26)

After using these values in Eqs. (17) and (24) and keeping linear terms, we obtain

$$r\varepsilon_o^i = \beta(1-\gamma)\sum_{j=1}^N A_{ij}\varepsilon_h^j,$$ (27)

$$\left[1-(1-\alpha)^{k_i}\right]\varepsilon_h^i = r\varepsilon_o^i + \frac{1}{N}\sum_{j=1}^N \varepsilon_o^j.$$ (28)

Finally, we introduce Eq. (27) into Eq. (28) which, after a little bit of algebra, yields

$$\varepsilon_h^i = \beta(1-\gamma)\sum_{j=1}^N \underbrace{\frac{\frac{k_j}{Nr}+A_{ij}}{1-(1-\alpha)^{k_i}}}_{\tilde{M}_{ij}} \varepsilon_h^j.$$ (29)

Again, obtaining the intimidation threshold implies solving the spectra of a given matrix $\tilde{\mathbf{M}}$. Unlike the corruption threshold, the matrix elements do not only contain the pairwise interactions encoded by A_{ij} but also reflect the influence of the global information responsible for the warning-to-wrongdoers effect. Following the same rationale applied before, we can obtain the intimidation threshold γ_c as

$$\gamma_c = 1-\frac{1}{\beta\Lambda_{\max}(\tilde{\mathbf{M}})}.$$ (30)

IV. RESULTS

Here, we want to characterize the influence of the interplay between the social intimidation γ and the network structure for the diffusion of corruption. Let us first check the validity of the formalism introduced in Sec. II by comparing its theoretical predictions with results yielded by extensive MC simulations. For this purpose, we study the level of corruption ρ_c^* defined as the fraction of the population remaining corrupt once the dynamics has reached its stationary state. In particular, for the equations, we have adiabatically studied the dependence of ρ_c^* on the corruption rate α by following both forward and backward continuation schemes. In particular, for the forward (backward) scheme, the initial configuration for $\alpha + \delta\alpha$ ($\alpha - \delta\alpha$) is given by the steady state obtained for α with a small perturbation introduced to avoid absorbing states. Both schemes start from a value of the corruption parameter α_0 ensuring that a unique fixed point is always reached regardless of the initial configuration of agents in (H, C, O) states. Regarding MC simulations, we have carried out a similar process in which the initial microscopic configuration for $\alpha \pm \delta\alpha$ is constructed according to the macroscopic

order parameters characterizing the steady state for α. In addition, all results obtained from MC simulations are averaged over 50 realizations to remove the stochastic nature of the processes involved in them.

We represent in Fig. 2 the evolution of the level of corruption ρ_c^* as a function of the corruption rate α by considering Scale Free (SF) and Erdös–Rényi (ER) with $N = 1000$ nodes and $\langle k \rangle = 4$ as the underlying contact networks dictating the interaction among agents. There, we check that our formalism reproduces very accurately the results from MC simulations, as shown by the fair agreement between both theory and simulations, even close to both thresholds. In addition, it becomes clear that introducing the social intimidation mechanism encoded by γ changes the nature of the transition observed toward the fully corrupt society. In particular, we find that, for $\gamma = 0.4$, the system undergoes a second order transition from all-honest to all-corrupt as in the original HCO model, whereas for $\gamma = 0.8$, the transition turns into a first-order one with an hysteresis loop in which both fully honest and fully corrupt societies are stable solutions.

To further explore this phenomenon, we study in Fig. 3 the phase diagrams $\rho_c^*(\alpha, \gamma)$ for SF (left panel) and ER (right panel) networks. Here we confirm that, as explained above, the introduction of the intimidation parameter γ changes dramatically the system behavior. In particular, for both topologies, there exists a critical point $(\gamma_c'(\beta, r), \alpha_c'(\beta, r))$ above which the transition taking place toward the fully corrupt society is no longer smooth but turns into a first-order discontinuous (explosive) one. Therefore, the existence of coercion or intimidation clearly promotes the stability of widespread resilient organizations of corrupt agents.

Let us now address the relevance of the contact network on the unfolding of corruption. For this purpose, we analyze the values of both corruption and intimidation thresholds, α_c and γ_c, respectively, for both topologies. Regarding the corruption threshold α_c,

we notice that, despite the existence of *hubs* in the SF network, both configurations display similar values. This constitutes a counterintuitive result, for in many spreading processes such as epidemics or the diffusion of ideas, the existence of highly connected nodes boosts their diffusion across the population. To explain the origin of this striking phenomenon, we must realize that highly connected nodes in this model can corrupt a large number of acquaintances but are also more exposed to being delated by them. Simultaneously, the exposure of highly connected nodes to delation processes gains even more relevance close to the intimidation threshold, since a fully corrupt society is much less resilient in SF networks than in ER ones. Therefore, the observed value of both thresholds α_c and γ_c is the outcome of the competition between these two opposite effects.

Interestingly, this outcome is governed by the delation rate β as shown in Fig. 4. Let us first study the impact of the delation rate β on the corruption threshold α_c.[42] In particular, for low values of the delation rate β, the corruption threshold is the same for both networks. Indeed, by assuming that $\beta \ll 1$, it can be easily shown that Eq. (23) yields $\alpha_c \simeq \beta$, which does not depend on the underlying network and reveals the antagonistic nature of both corruption and delation processes. On the other hand, for high β values, one recovers the usual picture of spreading dynamics in which SF networks are more vulnerable than ER ones, i.e., $\alpha_c^{SF} < \alpha_c^{ER}$. This result is rooted in the fact that, if β is large enough, corrupt agents are always delated regardless of the number of honest individuals surrounding them. This way, the more neighbors a corrupt agent has, the larger is the number of agents that can be corrupted in the single step she remains corrupt, thus explaining the higher vulnerability of SF networks in this scenario.

On the other hand, the intimidation threshold is always higher for the SF network than for the ER one regardless of the β value, which implies that large corrupt organizations are more stable in homogeneous topologies than in heterogeneous ones. Remarkably,
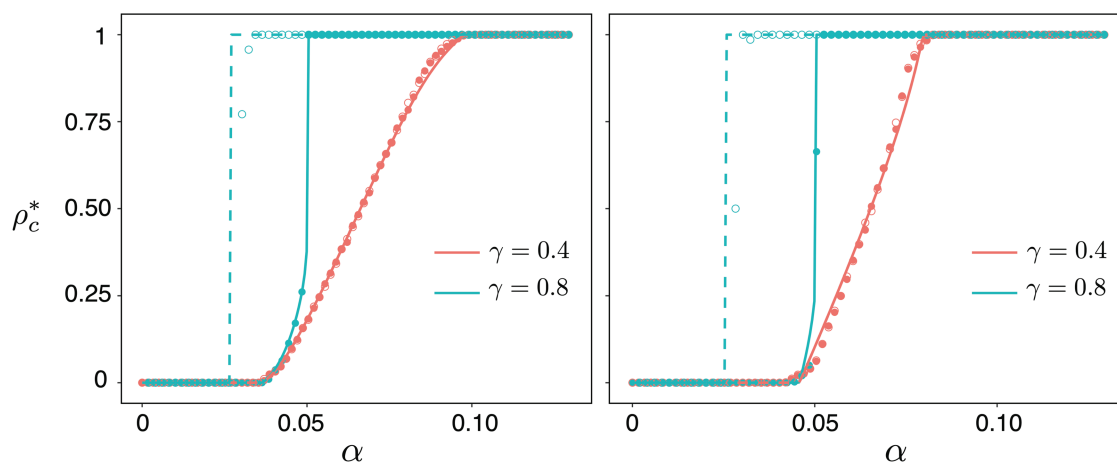

FIG. 2. Level of corruption ρ_c^* as a function of the corruption rate α and the intimidation parameter γ (color code) for scale free (left panel) and Erdös–Rényi (right panel) networks with $N = 1000$ nodes and $\langle k \rangle = 4$. Lines depict theoretical predictions computed by iterating equations (11) and (12), whereas points show results from MC simulations averaged over 50 realizations. In particular, solid lines and filled points are obtained following a forward continuation scheme (see text for details), while dashed lines and empty points correspond to the backward one. In both panels, the rest of parameters have been fixed to $(\beta, r) = (0.05, 0.75)$.
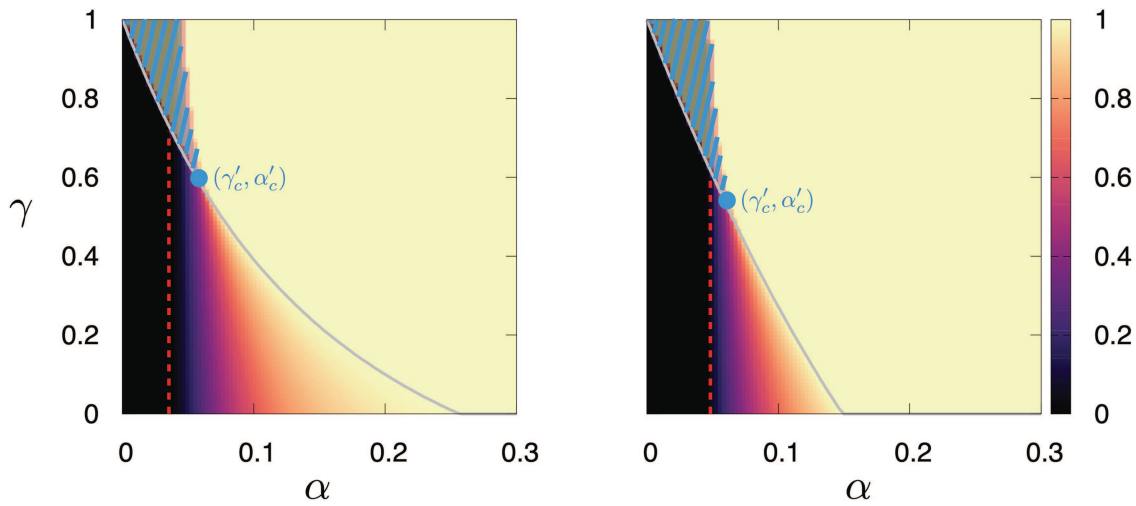
FIG. 3. Level of corruption ρ_c^* (color code) obtained from iterating equations (2)–(4) as a function of the corruption parameter α and the intimidation factor γ. The underlying structures governing both corruption and delation processes are SF (left panel) and ER (right panel) networks of $N = 1000$ nodes with $\langle k \rangle = 4$. The red dashed line indicates the value of the corruption threshold α_c triggering the onset of corruption. The stripped zone highlights the bistability regions, where the fully corrupted society and the existence of honest agents are both stable configurations, whereas the blue dot pinpoints the existence of a critical point $(\alpha_c'(\beta, r), \gamma_c'(\beta, r))$ changing the order of the transition toward the fully corrupted state. The rest of model parameters have been fixed to $(\beta, r) = (0.05, 0.75)$.

unlike the corruption threshold α_c, the most noticeable differences between both configurations appear for intermediate β values. To explain this fact, we must realize that, for low β values, the corrupt state is very stable for both networks, so no intimidation is needed for holding it. In contrast, when the delation rate is high enough, any corrupt agent is rather prone to be delated, so high intimidation values are needed for both configurations to keep a fully corrupt society.
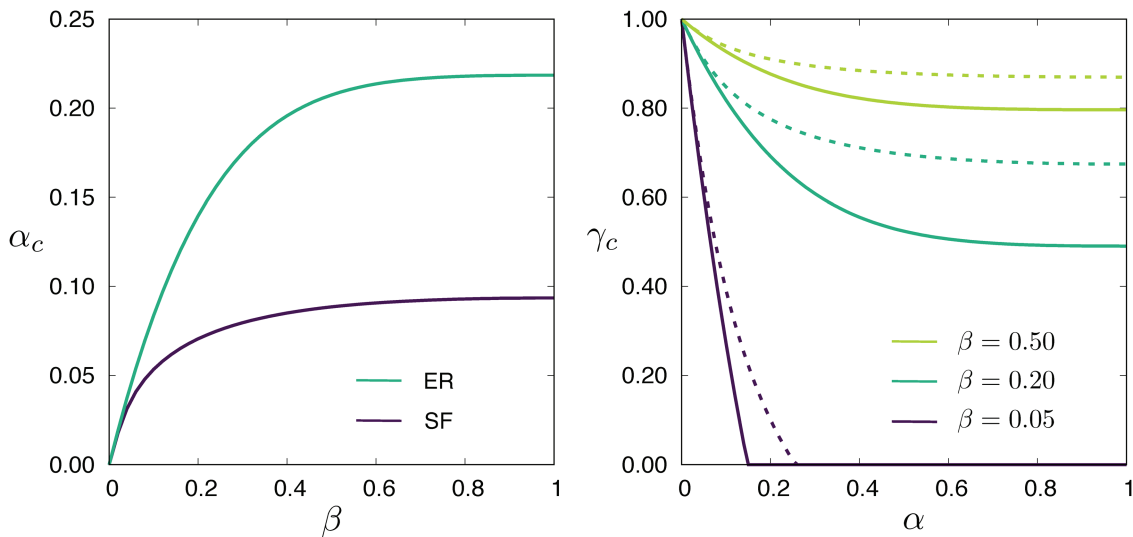

FIG. 4. Left panel: Corruption threshold α_c as a function of the delation rate β. The underlying structures governing both corruption and delation processes are SF (dark blue line) and ER (light blue line) networks of $N = 1000$ nodes with $\langle k \rangle = 4$. Right panel: Intimidation threshold as a function of the corruption rate α for several values of the delation rate (color code). Solid lines depict the estimation for the ER network, whereas dashed ones correspond to the SF network. The reinsertion rate is fixed to $r = 0.75$.

V. CONCLUSIONS

Game theory has been traditionally the most used benchmark to study the evolution of corruption. In this sense, game theoretical frameworks allow for incorporating many of the features associated with this phenomenon such as the existence of social norm violations and its consequent punishment. Nonetheless, the large number of parameters involved in these models hinders getting any general insights into the relevance of these processes. For this reason, we proposed an alternative approach in Ref. 34, the HCO model that consists in a minimal compartmental model which, apart from incorporating these ingredients, enables us to understand their impact on the diffusion of corruption.

The original HCO model, though novel with respect to other frameworks, has some limitations such as considering that the interaction among agents is governed by regular topologies or assuming the behavior of honest agents not to depend on the presence of corrupt counterparts. To increase the degree of realism of the model, in this manuscript we have extended the HCO model to account for the heterogeneity observed in most of the real networks. In addition, we also introduce a new mechanism, the social intimidation, which captures coercion practices made by corrupt agents aimed at preventing honest agents from delating them.

Formally, to monitor the evolution of corruption, we have proposed a set of equations based on the MMCA, which have been validated by comparing their predictions with results obtained from extensive individual MC simulations. The analysis of the equations has revealed the existence of two different transitions: one destabilizing a fully honest population and the other one giving rise to a totally corrupt society. In this sense, the linearization of the model has allowed us to accurately determine the thresholds triggering both transitions.

Remarkably, we have proved that the social intimidation mechanism changes the nature of the transition toward the fully corrupt society. Thus, enforcing this mechanism turns the smooth continuous transition reported in the original HCO model into a first-order discontinuous one,[43] reminding other types of abrupt transitions that occur on social norms adoption models.[7,19–22] In particular, we have found some regions of the parameters space, where both a fully honest population and a fully corrupt one can coexist as stable solutions, thus highlighting the resilience of the latter state due to the new mechanism. Moreover, we have characterized the role that the contact network plays in the evolution of corruption. In particular, we have found that the antagonistic nature of corruption and delation processes may hinder the characteristic spreading enhancement associated with heterogeneous topologies. This is explained by noticing that the most connected nodes in heterogeneous networks, responsible for the onset of corruption, are also the most exposed ones to the delation of their acquaintances. Despite being a counterintuitive result, this spreading detriment in heterogeneous networks has been already reported for other antagonistic coupled dynamics such as the spread of diseases and vaccination[44] or the spread of ideas and lack of consensus.[45]

In a nutshell, the main contribution of this work is to provide a theoretical framework that allows one to incorporate and understand the relevance of different features inherent to corruption dynamics that cannot be characterized neither by following game theoretical approaches nor by considering traditional compartmental models widely used for other spreading processes such as epidemics or ideas diffusion. We believe that this formalism paves the way to construct more refined models including more realistic ingredients such as the multiplex nature of human interactions[14–16] or the existence of gangs formed by different types of agents.[46]

APPENDIX: LINEAR STABILITY ANALYSIS OF FIXED POINTS. MEAN FIELD APPROXIMATION

In Sec. II, the mean field dynamics of the model was derived as the flow in the simplex $\mathcal{S}_2$ defined by the velocity field $\vec{F}(\vec{\rho})$ of components,

$$F_h(\vec{\rho}) = -(f_\alpha + r + \rho_c)\rho_h + (r + \rho_c)(1 - \rho_c), \qquad (A1)$$

$$F_c(\vec{\rho}) = (f_\alpha + (1 - f_\beta)\rho_c)\rho_h + ((1 - f_\beta)\rho_c - 1)\rho_c, \qquad (A2)$$

where the functions f_α and f_β are given, respectively, by Eqs. (14) and (15).

The Jacobian matrix, $J_{ml} = \frac{\partial F_m}{\partial \rho_l}$, at the full honesty fixed point ($\rho_h = 1$) is

$$\begin{pmatrix} -r & -r - f_\alpha(0) \\ 0 & f_\alpha(0) - f_\beta(1,0) \end{pmatrix}, \qquad (A3)$$

from where the stability condition of the full H state is obtained as $f_\alpha(0) < f_\beta(1,0)$, i.e.,

$$\alpha < \alpha_c(\beta, \gamma, r) = \frac{1 - (1 - \beta)^k}{k}. \qquad (A4)$$

The corruption threshold α_c is independent of the values of both reinsertion and intimidation parameters, r and γ. Thus, we see that social intimidation has no effects on the onset of corruptive fluctuations in the full H macrostate. As analyzed in Ref. 34, the warning-to-wrongdoers does not influence this transition either so that only the balance among α and β determines the corruption transition.

To analyze the stability of the full corruption state ($\rho_c = 1$), we compute the Jacobian matrix at this fixed point as

$$\begin{pmatrix} -f_\alpha(1) - r - 1 & -r - 1 \\ f_\alpha(1) + 1 - \frac{\partial f_\beta}{\partial \rho_h}(0,1) & 1 \end{pmatrix}, \qquad (A5)$$

from which the stability condition of the full C state is given by

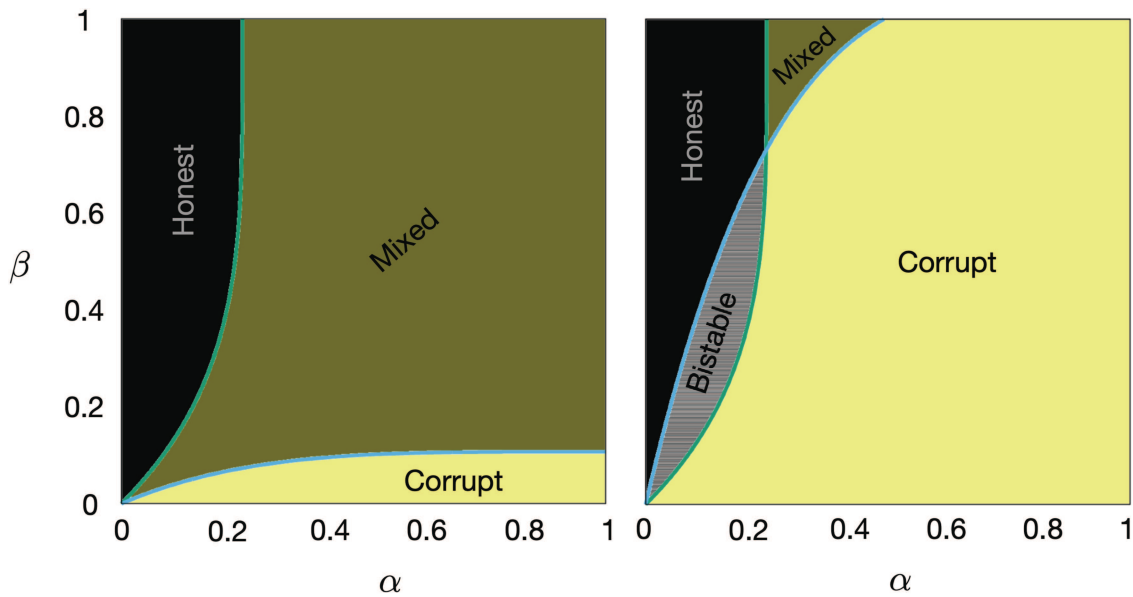$$\left(1 + \frac{1}{r}\right)\frac{\partial f_\beta}{\partial \rho_h}(0,1) < f_\alpha(1). \qquad (A6)$$

FIG. 5. Phase diagram in the (α, β) plane. The intimidation parameter has been fixed to $\gamma = 0$ (left panel) and $\gamma = 0.9$ (right panel). The reinsertion rate has been fixed in both panels to $r = 0.75$. The green line shows the value of $\alpha_c(\beta, \gamma, r)$ computed according to Eq. (A4), whereas the blue one represents $\beta_c(\alpha, \gamma, r)$ obtained from Eq. (A7). The bistable area encodes a region of the parameters space in which both fully honest and fully corrupt configurations are stable states.

For fixed values of α, γ, and r, the stability condition (A6) tells us that the delation threshold is given by

$$\beta_c(\alpha, \gamma, r) = \frac{r}{k(1-\gamma)(1+r)}\left(1 - (1-\alpha)^k\right). \quad (A7)$$

Alternatively, for fixed values of α, β, and r, the full corruption macrostate is stable provided the intimidation parameter satisfies

$$\gamma > \gamma_c(\alpha, \beta, r) = 1 - \frac{r}{k\beta(1+r)}\left(1 - (1-\alpha)^k\right). \quad (A8)$$

Let us fix the values of γ and r, and consider (see Fig. 5) the unit square $0 \le \alpha \le 1, 0 \le \beta \le 1$. The region to the left of the line $\alpha_c(\beta)$ [Eq. (A4)] connecting the points $(0,0)$ and $(1/k, 1)$ is the stability region of the full H macrostate, while the region below the line $\beta_c(\alpha)$ [Eq. (A7)] corresponds to the stability region of the full C macrostate. These two regions do not overlap for low enough values of the intimidation parameter γ (see the left panel of Fig. 5). However, they do overlap for higher values of γ, as shown in the right panel of Fig. 5. The value γ^* of the intimidation parameter above which both macrostates, full H and full C, coexist as stable fixed points is the one for which the slopes at $(0,0)$ of both lines coincides, i.e.,

$$\gamma^* = 1/(1+r). \quad (A9)$$

When the full H and full C fixed points are stable, an invariant manifold (containing an unstable fixed point) must exist in the interior of the simplex, that is, the basin boundary separating the basins of attraction of these stable fixed points. Therefore, the incorporation of social intimidation in the *HCO* model introduces new scenarios regarding the evolution of asymptotic states when

parameters are tuned: multistability, hysteresis, and discontinuous (explosive) transitions.

Let us think of a cyclic process at fixed values of β, γ, and r in which the value of the parameter α first increases from 0 to 1 and then decreases from 1 to 0. For very low values of γ, the possible sequences of absorbing states along the cycle are those of the *HCO* model without intimidation,[34] i.e.,

(a) If $\beta < \beta_c(1)$, "full H → mixed state → full C" for the forward semi-cycle, and "full C → mixed state → full H" for the backward semi-cycle. The arrows represent continuous transitions (at $\alpha = \alpha_c(\beta)$ and at $\alpha = \bar{\alpha}$ such that $\beta_c(\bar{\alpha}) = \beta$, respectively) so that the cycle shows no hysteresis.

(b) If $\beta > \beta_c(1)$, "full H → mixed state," followed by "mixed state → full H." However, for higher values of γ, besides the previous ones, the following sequences of asymptotic states during the cycle are also possible:

(c) For the forward semi-cycle, "full H → mixed state $\rightsquigarrow$ full C," where the curly arrow represents a discontinuous transition, followed by "full C $\rightsquigarrow$ mixed state → full H," with hysteresis, so that the discontinuous transition during the backward semi-cycle occurs at a value of α lower than that of the discontinuous transition in the forward semi-cycle.

(d) "Full H → mixed state $\rightsquigarrow$ full C" for the forward semi-cycle, followed by "full C $\rightsquigarrow$ full H."

(e) "Full H $\rightsquigarrow$ full C," followed by "full C $\rightsquigarrow$ full H."

Note that in the scenario described by item (c), there is no coexistence of full H and full C states as stable fixed points along the cycle, while its coexistence is required in the scenarios (d) and (e). On
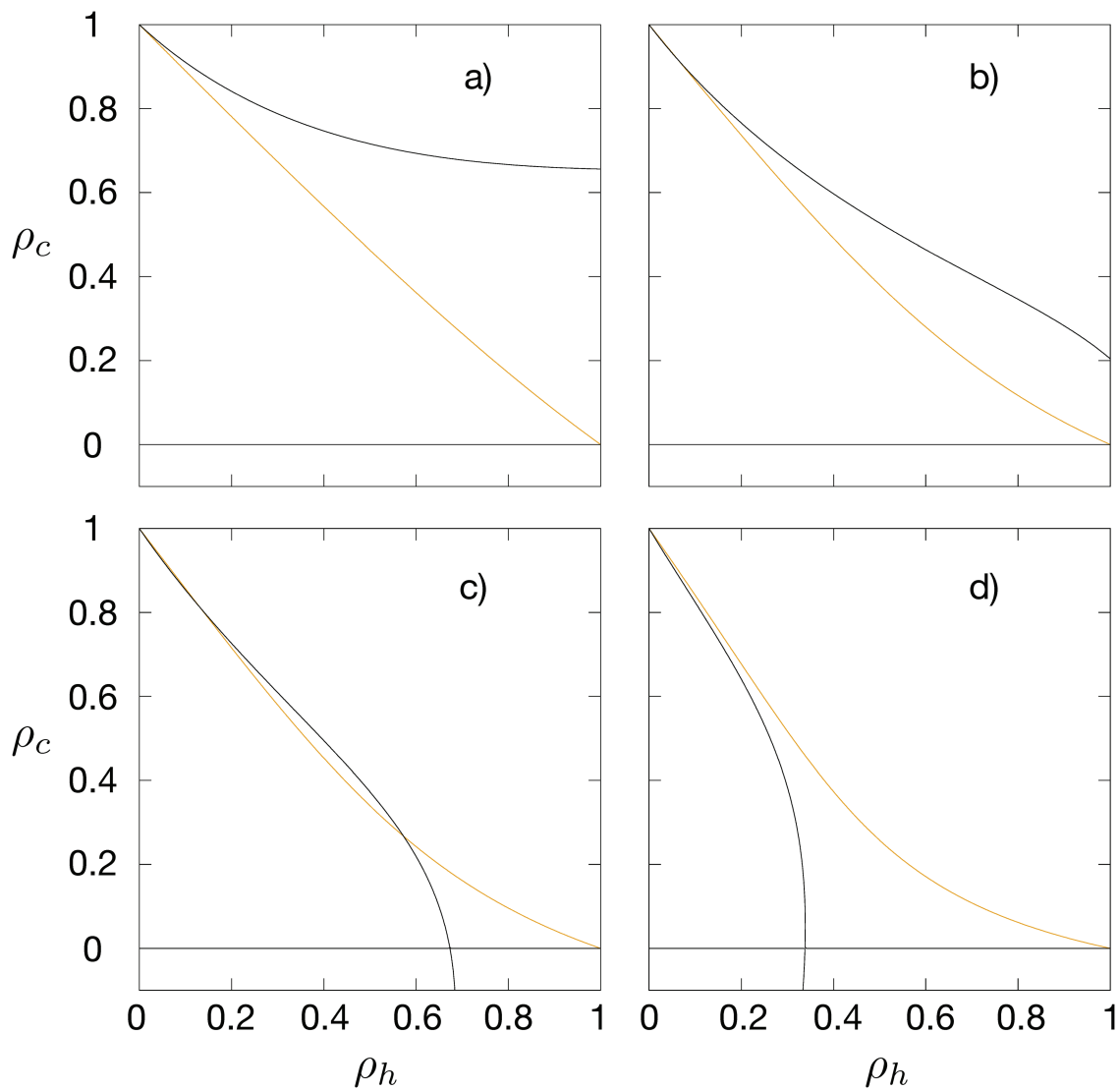
FIG. 6. Phase portraits according to Eqs. (A1) and (A2) by setting $(k, \beta, r, \gamma) = (4, 0.4, 0.75, 0.9)$. The F_h nullcline is represented by orange lines, whereas the F_c one is depicted by black lines. The chosen values for the corruption rate are: (a) $\alpha = 0.05$, (b) $\alpha = 0.20$, (c) $\alpha = 0.30$, and (d) $\alpha = 0.60$.

the other hand, both scenarios (c) and (d) require the simultaneous stability of the full C state and an interior fixed point.

Our numerical exploration in the 4D space of parameters reveals that these new scenarios are indeed observed. In Fig. 6, we show some examples of phase portraits along a path where scenario (d) is observed.

DATA AVAILABILITY

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

[1] T. C. Schelling, *Micromotives and Macrobehaviors* (Norton, 1978).
[2] P. W. Anderson, "More is different," Science **177**, 393–396 (1972).
[3] C. Castellano, S. Fortunato, and V. Loreto, Rev. Mod. Phys. **81**, 591 (2009).
[4] T. C. Schelling, J. Math. Sociology **1**, 143–186 (1971).
[5] D. J. Daley and D. G. Kendall, J. Inst. Math. Appl. **1**, 42 (1965).
[6] R. Axelrod, J. Conflict Resolut. **41**, 203–226 (1997).
[7] M. Granovetter, Am. J. Sociol. **83**, 1420 (1978).
[8] R. Axelrod, *The Evolution of Cooperation* (Basic Books, 1984).
[9] M. E. J. Newman, *Networks: An Introduction* (Oxford University Press, 2018).
[10] V. Latora, V. Nicosia, and G. Russo, *Complex Networks: Principles, Methods and Applications* (Cambridge University Press, 2017).

[11]E. Estrada, *The Structure of Complex Networks: Theory and Applications* (Oxford University Press, 2016).

[12]P. Holme and J. Saramaki, Phys. Rep. **519**, 102 (2012).

[13]P. Holme and J. Saramaki, *Temporal Network Theory* (Springer, 2019).

[14]S. Boccaletti, G. Bianconi, R. Criado, C. I. Del Genio, J. Gómez-Gardeñes, M. Romance, I. Sendiña-Nadal, Z. Wang, and M. Zanin, Phys. Rep. **544**, 1 (2014).

[15]M. Kivela, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter, J. Complex Netw. **2**, 203–271 (2014).

[16]M. De Domenico, C. Granell, M. A. Porter, and A. Arenas, Nat. Phys. **12**, 901 (2016).

[17]I. Iacopini, G. Petri, A. Barrat, and V. Latora, Nat. Commun. **10**, 2485 (2019).

[18]D. P. Maki and M. Thompson, *Mathematical Models and Applications, With Emphasis on the Social, Life, and Management Sciences* (Prentice Hall, Englewood Cliffs, NJ, 1973).

[19]D. J. Watts, Proc. Natl. Acad. Sci. U.S.A. **99**, 5766 (2002).

[20]J. Gómez-Gardeñes, L. Lotero, S. N. Taraskin, and F. J. Pérez-Reche, Sci. Rep. **6**, 19767 (2016).

[21]R. Amato, L. Lacasa, A. Díaz-Guilera, and A. Baronchelli, Proc. Natl. Acad. Sci. U.S.A. **115**, 33 (2018).

[22]J. Gómez-Gardeñes, A. S. de Barros, S. T. R. Pinho, and R. F. S. Andrade, Europhys. Lett. **110**, 58006 (2015).

[23]W. O. Kermack and A. G. McKendrick, Proc. Roy. Soc. Lond. Mat. A **115**, 700–721 (1927).

[24]R. M. Anderson, R. M. May, and B. Anderson, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1992).

[25]J.-H. Lee, Y. Iwasa, U. Diekmann, and K. Sigmund, Proc. Natl. Acad. Sci. U.S.A. **116**, 13276–13281 (2019).

[26]V. N. Kolokoltsov, Int. J. Stat. Probab. **1**, 77 (2012).

[27]V. N. Kolokoltsov and O. A. Malafeyev, Dyn. Games Appl. **7**, 34–47 (2017).

[28]J.-H. Lee, K. Sigmund, U. Dieckmann, and Y. Iwasa, J. Theor. Biol. **367**, 1–13 (2015).

[29]J.-H. Lee, M. Jusup, and Y. Iwasa, J. Theor. Biol. **428**, 76–86 (2017).

[30]P. Verma and S. Sengupta, PLoS One **10**, e0133441 (2015).

[31]P. Verma, A. K. Nandi, and S. Sengupta, Sci. Rep. **7**, 42735 (2017).

[32]P. Verma, A. K. Nandi, and S. Sengupta, J. Theor. Biol. **450**, 43–52 (2018).

[33]J. von Neumann and O. Morgenstern, *Theory of Games and Economic Behavior* (Princeton University Press, 1944).

[34]D. Lu, F. Bauza, D. Soriano-Paños, J. Gómez-Gardeñes, and L. M. Floría, Phys. Rev. E **101**, 022306 (2020).

[35]J. Marro and R. Dickman, *Nonequilibrium Phase Transitions in Lattices* (Cambridge University Press, 2005).

[36]M. Blume, V. J. Emery, and R. B. Griffiths, Phys. Rev. A **4**, 1071 (1971).

[37]L. Gauvin, J.-P. Nadal, and J. Vannimenus, Phys. Rev. E **81**, 066120 (2010).

[38]S. Gómez, A. Arenas, J. Borge-Holthoefer, S. Meloni, and Y. Moreno, Europhys. Lett. **89**, 38009 (2010).

[39]B. Guerra and J. Gómez-Gardeñes, Phys. Rev. E **82**, 035101 (2010).

[40]S. Gómez, J. Gómez-Gardeñes, Y. Moreno, and A. Arenas, Phys. Rev. E **84**, 036105 (2011).

[41]X. Chen *et al.*, New J. Phys. **20**, 013007 (2018).

[42]We do not include the impact of the intimidation γ since the corruption threshold α_c does not depend on this parameter.

[43]R. M. D'Souza, J. Gómez-Gardeñes, J. Nagler, and A. Arenas, Adv. Phys. **68**, 123 (2019).

[44]H. Zhang, J. Zhang, C. Zou, M. Small, and B. Wang, New J. Phys. **12**, 023015 (2010).

[45]D. Soriano-Paños, Q. Guo, V. Latora, and J. Gómez-Gardeñes, Phys. Rev. E **99**, 062311 (2019).

[46]H. V. Ribeiro, L. G. A. Alves, A. F. Martins, E. K. Lenzi, and M. Perc, J. Complex Netw. **6**, 989 (2018).