A 1:1000 scale model of the digital world: Global connectivity can lead to the extinction of local networks

Kaj-Kolja Kleineberg[1,*] and Marián Boguñá[1]

[1]*Departament de Física Fonamental, Universitat de Barcelona, Martí i Franquès 1, 08028 Barcelona, Spain*

(Dated: April 7, 2015)

The overwhelming success of online social networks, the key actors in the cosmos of the Web 2.0, has reshaped human interactions on a worldwide scale. To understand the fundamental mechanisms which determine the fate of online social networks at the system level, we recently introduced a general ecological theory of the digital world. In this paper, we discuss the impact of heterogeneity in the network intrinsic fitness and present how the general theory can be applied to understand the competition between an international network, like Facebook, and local services. To this end, we construct a 1:1000 scale model of the digital world enclosing the 80 countries with most Internet users. We find that above a certain threshold the level of global connectivity can lead to the extinction of local networks. In addition, we reveal the complex role the tendency of individuals to engage in more active networks plays for the probability of local networks to become extinct and provide insights into the conditions under which they can prevail.

Keywords: complex systems — complex networks — online social networks — digital ecology — digital world — network of networks — double mean field approximation

I. INTRODUCTION

The rapid growth of online social networks (OSNs), such as Twitter or Facebook, has led to over two billion active accounts in 2014 [1], connecting over one quarter of the world population and 72% of online U.S. adults [2]. Bridging the gap between social sciences and information and communication technologies, OSNs constitute a crucial building block in the development of innovative approaches to the challenges society faces nowadays. However, technological progress in the last decade has dramatically outpaced our understanding of the new systems and their impact for society.

This lack of understanding of the complex dynamics in the digital world urges the need of a comprehensive and concise theory in which the Web 2.0 is described by a set of interacting networks. In this context, the activity of users has become a scarce resource driving the competition in the digital world. Digital services only persist if they can attract and maintain users' attention. In recent studies [3, 4], we revealed the fundamental mechanisms for the evolution of online social networks, the key players in the cosmos of the Web 2.0, and developed a general and concise ecological theory of interacting networks. Interestingly, this theory predicts the possibility of coexistence of multiple *a priori* identical networks, in contrast to the principle of competitive exclusion [5].

In reality not all networks are initially identical. They can differ in functionality, features, and –most importantly– they can address different peer groups. In this paper we discuss how the effect of different overlapping peer groups can be described in terms of different network intrinsic fitnesses and show how heterogeneity in these fitnesses can impede the coexistence of networks under certain conditions.

This effect is particularly important for the competition between local networks and globally operating ones. We enrich our theory with empirical data to describe the competition between an internationally operating network like Facebook and a set of local networks which operate exclusively in their countries of origin. Empirical observations have shown that Facebook expanded in the mid 2000s starting in the US when local networks were the most popular services in most countries worldwide. Several years later, Facebook had become the most popular network in most countries. Is then the final fate of the digital world one with a single predominant "big brother" taking over all our digital interactions? Is, instead, digital diversity possible from a system's level perspective? In this paper, we show that due to the non-linear character of the underlying laws at play, the answer to both questions can be positive or negative depending on the range of parameters and, quite surprisingly, depending on chance.

II. RESULTS

To shed light on these questions, we extend our general theory by incorporating the effect of different underlying communities given by the countries in which networks operate. In section II B we show how this effect can be taken into account by an effective activity which leads to an increased intrinsic fitness of the international network. In section II C we show how heterogeneity in network intrinsic fitnesses can impede coexistence which can lead to the extinction of local networks as well as their prevalence under certain conditions. In section II D we present a 1:1000 scale model of the digital world by creating synthetic networks for the underlying communities and per-

* kkl@ffn.ub.edu

form numerical simulations to provide insights about the conditions which allow local networks to persist.

A. Evolution and ecology of online social networks

In this section, we provide a brief summary of the evolution [3] and ecology [4] of online social networks. The evolution of an OSN is coupled to the pre-existing offline social structure. The following four dynamical processes drive the evolution of the system (see Fig. 1d):

1. *Viral activation:* a susceptible node can be virally activated and added to the OSN by contact with an active neighbor in the traditional off-line network. Such events happen at rate λ for each active link.

2. *Mass media effect:* each susceptible individual becomes active spontaneously at rate μ and may thus be added to the OSN in response to the visibility of the OSN.

3. *Deactivation:* active users become spontaneously passive at rate δ and no longer trigger viral activations or reactivate other passive nodes.

4. *Viral reactivation:* at rate λ, active users can reactivate their passive neighbors. The neighbor then becomes active and can trigger both viral activations and viral reactivations.

The balance between the mass media influence, μ, and the viral effect, λ, can be estimated from the topological evolution of the corresponding empirical network. This estimation can be performed by making use of the network exhibiting a dynamical percolation transition. The critical point of the transition depends on the ratio between λ and μ. This is due to the complementary roles the respective effects play in the topological evolution. Matching the system size at the critical point then yields a linear relationship between λ and μ (see [3] for further details).

The simultaneous existence of multiple digital services in competition for the attention of users suggests an ecological perspective to explain the prevalence of one network or the coexistence of multiple networks. In ecology theory, the principle of competitive exclusion [5] states that multiple species in competition for the same only resource cannot coexist as even the slightest advantage of one species is amplified successively, a mechanism referred to as rich-get-richer or preferential attachment [6–14]. This eventually leads to the extinction of the inferior species. Analogously, we assume such a mechanism for the interaction of multiple networks where the activity plays the role of the networks fitness. The total amount of the viral parameter, λ, constitutes a conserved quantity related to the physical and cognitive limitations of users. Each network obtains a share $\lambda_i = \omega_i(\boldsymbol{\rho}^{\mathrm{a}})\lambda$, where $\omega_i(\boldsymbol{\rho}^{\mathrm{a}})$ represents a normalized set of weights. The weights depend on the activities of the networks, which are given by the fraction of active nodes in the network divided by the total number of users in the traditional off-line network. We denote the activities as a vector $\boldsymbol{\rho}^{\mathrm{a}} = (\rho_1^{\mathrm{a}}, \rho_2^{\mathrm{a}}, \ldots, \rho_{n_l}^{\mathrm{a}})^T$ of length n_l. n_l corresponds to the number of OSNs competing for the same set of users. Users are more likely to subscribe to or engage in more active networks, hence $\partial\omega_i(\boldsymbol{\rho}^{\mathrm{a}})/\partial\rho_i^{\mathrm{a}} > 0$. In [4], we proposed

$$\omega_i(\boldsymbol{\rho}^{\mathrm{a}}) = \frac{\psi(\rho_i^{\mathrm{a}})}{\sum_{j=1}^{n_l} \psi(\rho_j^{\mathrm{a}})} \tag{1}$$

which governs the competitive interaction between digital services. For the treatment of intrinsically equal networks, the choice

$$\psi(\rho_i^{\mathrm{a}}) = [\rho_i^{\mathrm{a}}]^\sigma \tag{2}$$

allows us to to interpolate between a set of independent networks and highly coupled ones as a function of the activity affinity σ, which denotes the tendency of users to subscribe to or engage in more active networks. Interestingly, in contrast to the principle of competitive exclusion, multiple networks can coexist because the rich-get-richer mechanism is damped by the diminishing returns of the dynamics of the network evolution. Although any number of networks can coexist in a certain range of parameters, it is highly probable to observe only a moderate number of coexisting services. For details we refer the reader to [4].

In the following, we introduce additional mechanisms and enrich the theory with empirical data to account for the situation of competition between a large internationally operating online social network and locally operating services.

B. Effective activity

As mentioned before, since its official launch in 2004, Facebook has become the most popular online social network in most countries, even in countries where there was already a popular one before Facebook was launched.

To mimic the real evolution, we assume that one local network exists in each country in addition to a globally operating, international network (see Fig. 1a). In the US, both networks are launched at the same time whereas the international network is launched with a delay in the remaining countries to take into account the initial prevalence of the local networks.

Consistent with our findings in [4], we assume that in general users are more prone to subscribe to or engage in more active networks. However, once launched, the international network provides the user with the possibility to connect to individuals from different countries, in contrast to local networks. This makes the international network more attractive. The attractiveness difference is the result of inter-country social ties, thus its amount is proportional to the abundance of those ties.

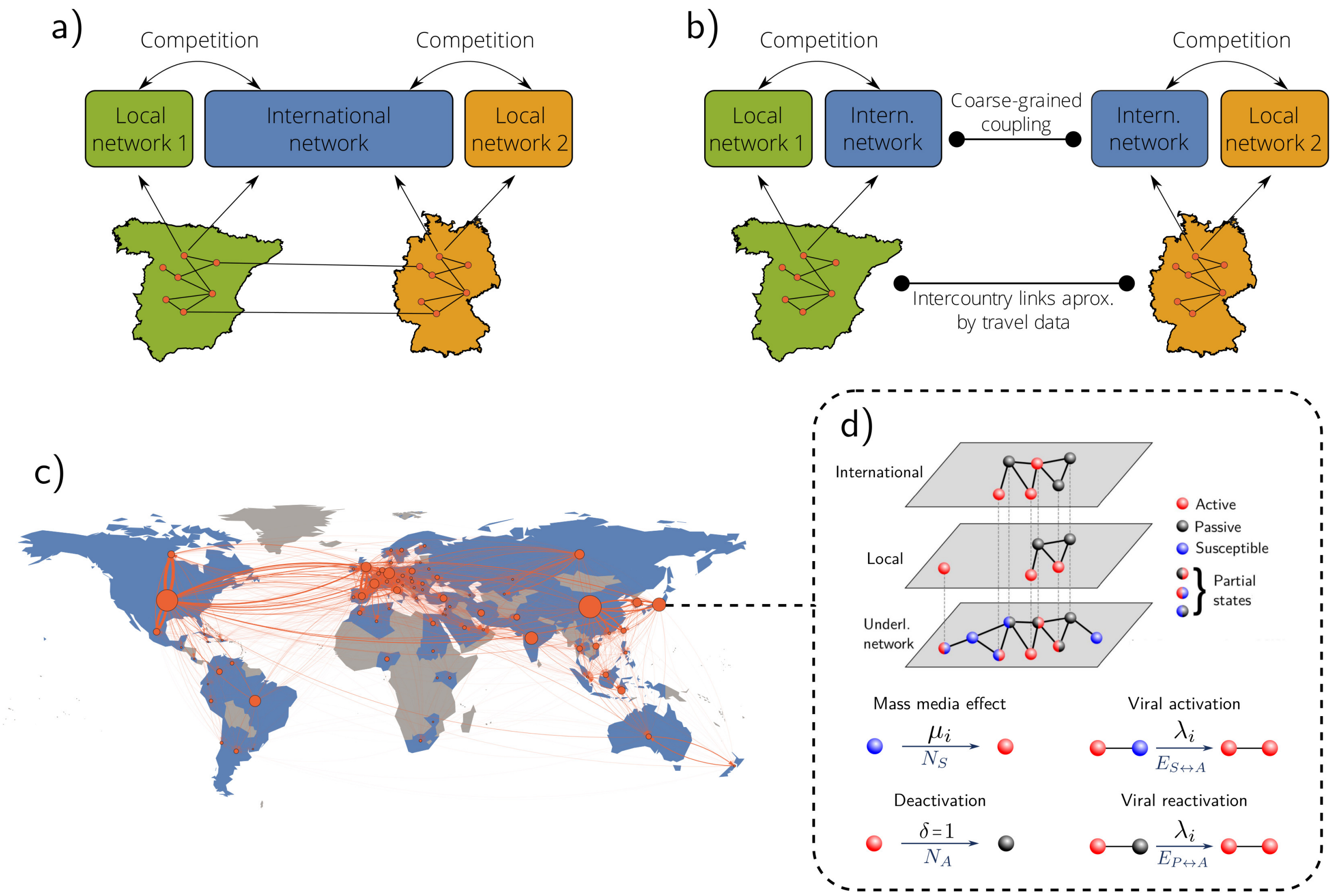

Figure 1. Constituents of our model. **a)** Design of the international network and local networks. **b)** Sketch of our model using a coarse-grained coupling. **c)** Visualization of the flight network. The area of the nodes is proportional to the number of Internet users in the respective country with Internet access. The transparency and thickness of the links represents the density of passengers between the involved countries. **d)** Illustration of the competition of the international network and the local network within one country.

We use data from air travel passengers as a proxy for the frequency of inter-country social ties. Note that air travel data has proven to be a good proxy for large-scale contagion processes [15].

When a user in country i evaluates the attractiveness of the international network she considers its activity with respect to the population of her own country but also might have contacts in other countries. To account for the latter, on a coarse grained level, the attractiveness of the international network will be augmented by its activity in all the other countries discounted by some factor which depends on the frequency of the respective inter-country social ties. Such evaluation of the attractiveness allows us to define an effective activity which takes into account the coarse-grained coupling induced by the previously mentioned mechanism (sketched in Fig. 1b). In detail, we define the effective activity of the international network perceived in country i

$$\tilde{\rho}^{\mathrm{a}}_{i,\mathrm{int}} = \rho^{\mathrm{a}}_{i,\mathrm{int}} + \alpha \sum_j \Omega_{ij} \rho^{\mathrm{a}}_{j,\mathrm{int}} \qquad (3)$$

where

$$\Omega_{ij} = \frac{W_{ij}/N_i}{\max[W_{ij}/N_i]}(1-\delta_{ij}) \qquad (4)$$

denotes the fraction of the number of air travel passengers between countries i and j, W_{ij}, and the size of country i, N_i. In the following, we refer to the network given by the adjacency matrix Ω_{ij} as the generalized air travel network. α is a constant which represents the proportionality between the number of passengers and the number of contacts in the respective country, namely $N_{ij} \propto \alpha W_{ij}$. The parameter α controls the overall frequency of inter-country social ties. The normalization in Eq. (4) leads to reasonable values for the parameter α in the order of unity.

The definition of the effective activity allows us to treat the international network as a set of disjunct coupled networks operating in each country. A user in country i then perceives the activity of the local network, $\rho^{\mathrm{a}}_{i,\mathrm{loc}}$, and the effective activity of the international network with respect to country i, $\tilde{\rho}^{\mathrm{a}}_{i,\mathrm{int}}$. To account for this effect, we replace the activity of the international network in the argument of the weight function by the effective activity. Then, the competition in this country is equivalent to the one of two local networks if we replace the activity of one of them with the effective activity of the international network (see Fig. 1b). Note that the latter depends on the state of the whole system.

The effective activity dynamically couples the evolution of the networks in different countries via the generalized air travel network. Hence, our model forms a network of networks [16–19], where each node in Fig. 1c represents a three layer multiplex network [20, 21] in which the bottom layer corresponds to the underlying social structure of the respective country and the two upper layers denote the local and international network operating in the respective country (see Fig. 1d).

C. Double meanfield approximation

We assume that the physical and cognitive limitations of users are country independent, hence each country contains the same total amount of virality λ. In each country the virality is distributed between the international and local network via the weight function $\omega_l(\boldsymbol{\rho}_i^{\mathrm{a}})$ as introduced in Eq. 1, where $l = (\mathrm{loc}, \mathrm{int})$ denotes the local or international network and $i = (1, \dots, n_c)$ indicates the country. The normalization of the weight function then reads $\omega_{\mathrm{loc}}(\boldsymbol{\rho}_i^{\mathrm{a}}) + \omega_{\mathrm{int}}(\boldsymbol{\rho}_i^{\mathrm{a}}) = 1$. Here, $\boldsymbol{\rho}_i^{\mathrm{a}} = \left(\rho_{i,\mathrm{loc}}^{\mathrm{a}}, \tilde{\rho}_{i,\mathrm{int}}^{\mathrm{a}}\right)^{\mathrm{T}}$ denotes the vector which contains the activity of the local network and the effective activity of the international network in country i. Recall that the effective activity of the international network $\tilde{\rho}_{i,\mathrm{int}}^{\mathrm{a}}$ defined in Eq. (3) depends on the state of the whole system.

To understand the qualitative behavior of the system, in this section we present a double meanfield approximation of the system. This reduces the system given by a network of networks to a set of evolution equations of the average activity in the international network and in local networks. As we show in section II D, the results of the full model with heterogeneous topologies exhibits a similar behavior as found by the double meanfield approximation.

The first meanfield approximation consists in assuming a fully mixed homogeneous population in each country. Let $\rho_{i,l}^{\mathrm{a}}$ denote the fraction of active users in network $l \in (\mathrm{loc}, \mathrm{int})$ in country i and $\rho_{i,l}^{\mathrm{s}}$ the fraction of nodes susceptible to joining this network. Then, the fraction of passive users is given by $1 - \rho_{i,l}^{\mathrm{s}} - \rho_{i,l}^{\mathrm{a}}$. The evolution equations of the resulting system is a generalization of the evolution equations for identical networks which we derived in [4] where one replaces the activity of the international network with the effective activity. This procedure yields

$$\begin{aligned} \dot{\rho}_{i,l}^{\mathrm{a}} &= \rho_{i,l}^{\mathrm{a}} \left\{ \lambda \langle k \rangle \, \omega_l(\boldsymbol{\rho}_i^{\mathrm{a}}) \left[1 - \rho_{i,l}^{\mathrm{a}}\right] - 1 \right\} + \frac{\lambda}{\nu} \omega_l(\boldsymbol{\rho}_i^{\mathrm{a}}) \rho_{i,l}^{\mathrm{s}} \\ \dot{\rho}_{i,l}^{\mathrm{s}} &= -\frac{\lambda}{\nu} \omega_l(\boldsymbol{\rho}_i^{\mathrm{a}}) \rho_{i,l}^{\mathrm{s}} \left\{ 1 + \nu \langle k \rangle \rho_{i,l}^{\mathrm{a}} \right\}, \end{aligned} \tag{5}$$

where we assume the same linear relationship between virality and media influence in each country, thus we have in country i

$$\mu_{i,\mathrm{loc}} = \lambda \omega_{\mathrm{loc}}(\boldsymbol{\rho}_i^{\mathrm{a}})/\nu \tag{6}$$

$$\mu_{i,\mathrm{int}} = \lambda \omega_{\mathrm{int}}(\boldsymbol{\rho}_i^{\mathrm{a}})/\nu \tag{7}$$

(see [3, 4]). As shown in [4] the value of ν does not affect the stability of the system. In the following we perform the stability analysis for $\nu \to \infty$. This decouples the evolution of $\rho_{i,l}^{\mathrm{a}}$ from $\rho_{i,l}^{\mathrm{s}}$, so that in the following we only have to consider $\rho_{i,l}^{\mathrm{a}}$. Plugging in the weights function defined in Eq. (1) and the effective activity from Eq. (3) yields the evolution equations for the activities of the local and international network in country i

$$\begin{aligned} \dot{\rho}_{i,\mathrm{loc}}^{\mathrm{a}} &= \rho_{i,\mathrm{loc}}^{\mathrm{a}} \left[\frac{\lambda \langle k \rangle \, [\rho_{i,\mathrm{loc}}^{\mathrm{a}}]^\sigma}{[\rho_{i,\mathrm{loc}}^{\mathrm{a}}]^\sigma + (\rho_{i,\mathrm{int}}^{\mathrm{a}} + \delta_i)^\sigma} [1 - \rho_{i,\mathrm{loc}}^{\mathrm{a}}] - 1 \right] \\ \dot{\rho}_{i,\mathrm{int}}^{\mathrm{a}} &= \rho_{i,\mathrm{int}}^{\mathrm{a}} \left[\frac{\lambda \langle k \rangle \, (\rho_{i,\mathrm{int}}^{\mathrm{a}} + \delta_i)^\sigma}{[\rho_{i,\mathrm{loc}}^{\mathrm{a}}]^\sigma + (\rho_{i,\mathrm{int}}^{\mathrm{a}} + \delta_i)^\sigma} [1 - \rho_{i,\mathrm{int}}^{\mathrm{a}}] - 1 \right], \end{aligned} \tag{8}$$

where $\delta_i = \alpha \sum_j \Omega_{ij} \rho_{j,\mathrm{int}}^{\mathrm{a}}$.

The second meanfield approximation consists in applying the hypothesis of a fully mixed homogeneous network to the generalized air travel network given by the adjacency matrix Ω_{ij}. We denote $\bar{\Omega} = \alpha \langle \Omega_{ij} \rangle$ and define the mean activity of the local networks as $x \equiv \left\langle \rho_{i,\mathrm{loc}}^{\mathrm{a}} \right\rangle$ and the mean activity of the international network as $y \equiv \langle \rho_{i,\mathrm{int}}^{\mathrm{a}} \rangle$. In double meanfield approximation, one obtains

$$\begin{aligned} \dot{x} &= x \left[\lambda \langle k \rangle \frac{x^\sigma}{x^\sigma + (y(1+\bar{\Omega}))^\sigma} [1 - x] - 1 \right] \\ \dot{y} &= y \left[\lambda \langle k \rangle \frac{(y(1+\bar{\Omega}))^\sigma}{x^\sigma + (y(1+\bar{\Omega}))^\sigma} [1 - y] - 1 \right]. \end{aligned} \tag{9}$$

Note that Eq. (9) incorporates an increased intrinsic fitness of the international network as a result of inter-country social ties. However, different classes of increased intrinsic fitnesses –for example superior features or functionality– can be described by the same system. As a consequence, the following results are also valid in these cases and hence can be applied to a broad spectrum of empirical situations.

For constant σ, the system of Eq. (9) exhibits a saddlenode bifurcation at a critical value of the global connectivity $\bar{\Omega}_c$ (see Fig. 2). Above this point, coexistence is not possible and the only stable solutions correspond

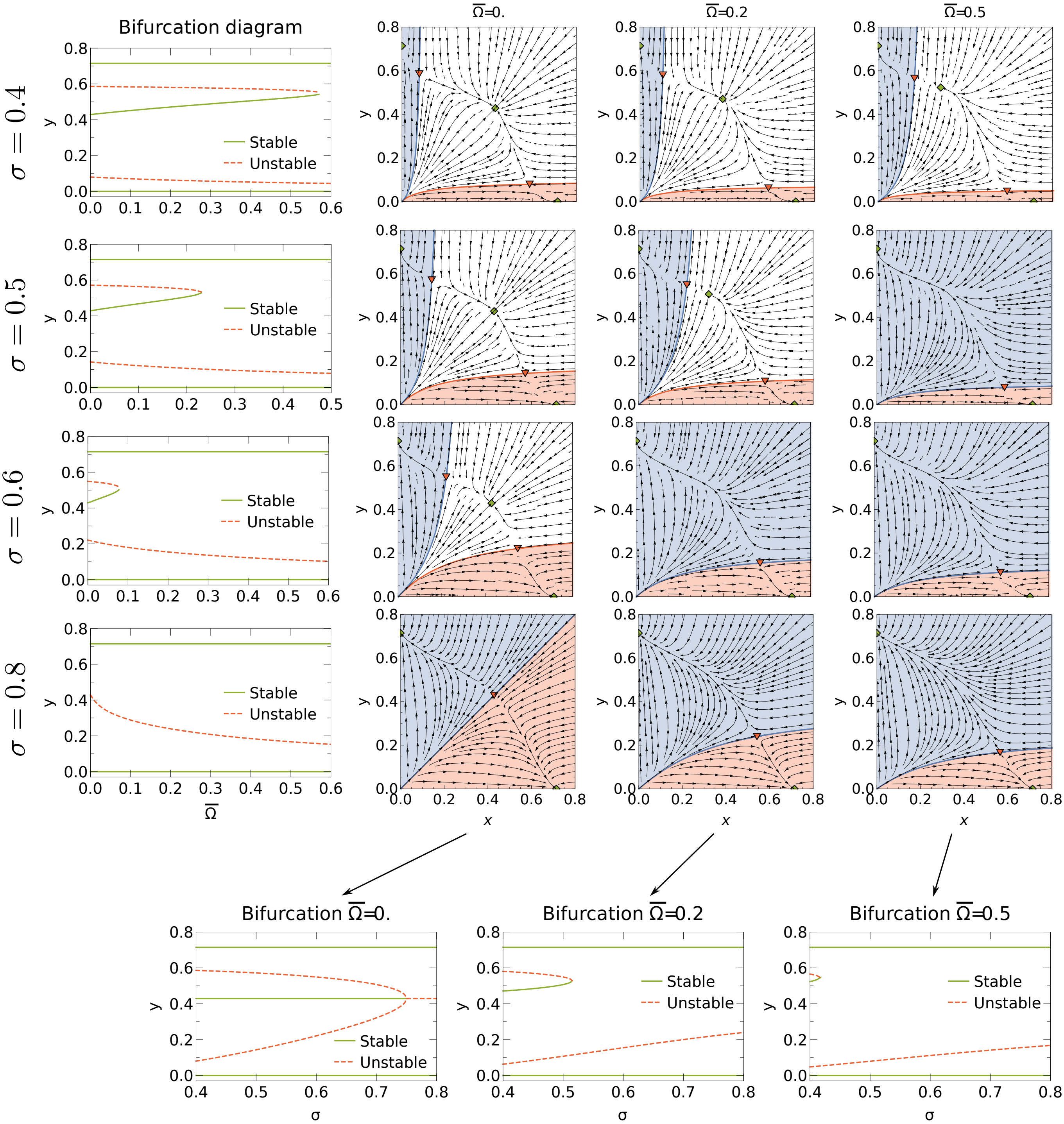

Figure 2. Bifurcation diagram and stream plots for the double meanfield approximation (9) for $\lambda \langle k \rangle = 3.5$ and $\nu \to \infty$. The basins of attraction for the domination of the international network is marked blue, the basin of attraction for the domination of local networks is marked red, and the white area corresponds to the basin of attraction of the coexistence solution (if it exists).

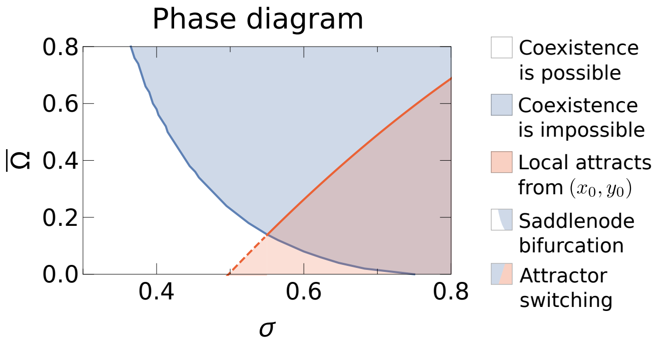

Figure 3. Phase diagram of the double meanfield approximation for $\lambda \langle k \rangle = 3.5$. The white area denotes the parameters for which a coexistence is possible. In the blue area the domination of the international network is reached and in the red region local networks dominate. At the blue line, the system undergoes a saddlenode bifurcation in which the stable coexistence solution disappears. The red dashed line denotes the combination of parameters for which the system switches attractors for the initial conditions given by Eq. (10).

either to the domination of local networks or the international network. Above and below the critical point, the basin of attraction for the solution corresponding to the domination of local networks decreases with $\bar{\Omega}$ whereas the one for the international network increases (see the rows of Fig. 2).

For constant $\bar{\Omega} > 0$, the system also exhibits a saddlenode bifurcation at a critical value of the activity affinity σ_c. The evolution of the basins of attraction is more complex compared to the case of constant σ. Below the critical point, both basins of attraction increase with σ. Above the critical point, the basin of attraction for the local network increases but the basin of attraction for the international network decreases with σ (see the columns of Fig. 2). This is particularly interesting as it implies that an intermediate value of the activity affinity just slightly above the critical point represents the worst scenario for the survival of local networks.

In [4] we showed that the system undergoes a subcritical pitchfork bifurcation with respect to the control parameter σ above which no stable coexistence is possible. $\bar{\Omega} > 0$ breaks the symmetry of the pitchfork bifurcation and in this case the system undergoes a saddlenode bifurcation with respect to σ instead (see bottom of Fig. 2). This behavior is well known in bifurcation theory by adding a small error term to the normal form of the pitchfork bifurcation (see [22] and Appendix C).

In Fig. 3 we show the critical line $\bar{\Omega}_c(\sigma)$ which separates a phase in which coexistence is possible and one in which only domination can occur.

The increasing size of the basin of attraction for the domination of local networks above the critical point with respect to σ can dramatically alter the fate of the system for a given initial condition. Assume that the international network dominates in the US and starts with a significant delay in each other country which causes the local networks to dominate in those countries. The US constitutes about 20% to the total population taken into account here (see table I). To illustrate the effect of the change of the basin of attraction, consider the initial conditions

$$
\begin{aligned}
x_0 &= 0.8 \left[1 - \frac{1}{\lambda \langle k \rangle}\right] \\
y_0 &= 0.2 \left[1 - \frac{1}{\lambda \langle k \rangle}\right] .
\end{aligned}
\tag{10}
$$

If one network dominates in country i, its activity is given by $\rho^{a}_{i,l} = 1 - \frac{1}{\lambda \langle k \rangle}$. Hence, the initial conditions given by Eq. (10) reflect that local networks dominate in 80% of the system and the international one dominates in the remaining 20%. The evolution of the basins of attraction makes the system approach different stationary solutions from this initial condition for different parameters. As shown by the red area in Fig. 3, for low values of the global connectivity and high activity affinity the domination of local networks is approached from the initial conditions x_0, y_0. This means that in this parameter region, when the international network is launched globally, it is not able to overcome the initial advantage of local networks due to its earlier launch.

To conclude, the double meanfield approximation predicts that intermediate values of the activity affinity most favor the international network whereas the local networks can dominate for a high activity affinity and low global connectivity. We confirm these findings by numerical simulations in the following section.

D. A 1:1000 scale model of the digital world

To investigate the full dynamics of the system, we now present a miniature model of the digital world. To this end, we construct 1:1000 scaled synthetic networks for the 80 countries with the largest number of Internet users (see Tab. I). To generate these networks, we make use of a model introduced in [23–25], which produces realistic topologies of the traditional off-line social networks, including heterogeneous node degrees and a high level of clustering (see Methods and Material A). To complete the network of networks description, we use the empirical network of air travel passengers (see Fig. 1c and Methods and Material) which couples the evolution in different countries by a dynamic heterogeneity in the network intrinsic fitnesses as a result of the previously introduced effective activity.

The value ν resembles the relative importance of the influence of mass media compared to the viral spreading mechanism. Due to the existence of multiple stable fixed points, a higher influence of mass media initially drives the system closer to the coexistence point, as a consequence the probability to reach the coexistence solution increases [4]. In the following we set $\nu = 4$, the value found empirically in [3].

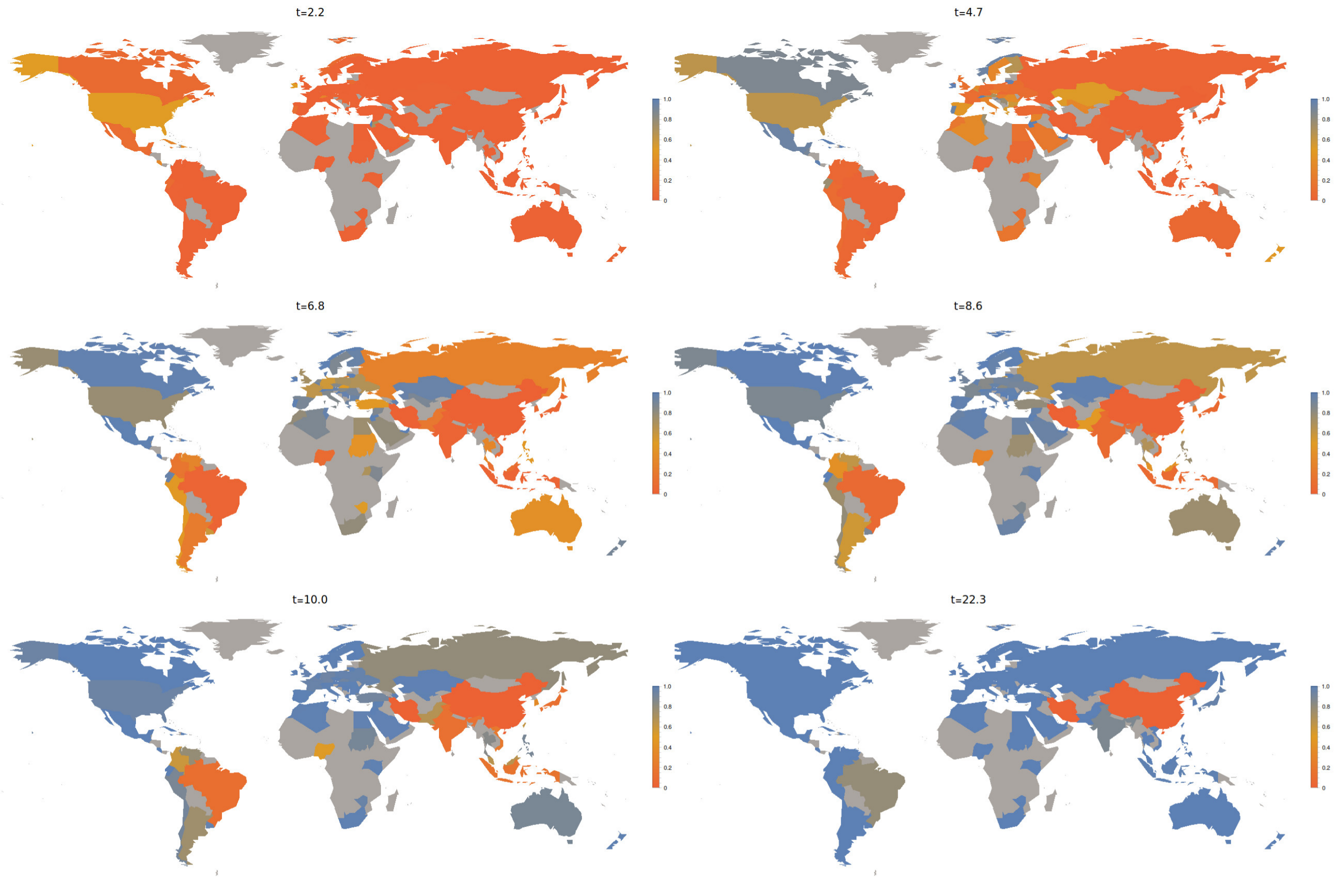

Figure 4. Typical realization. Color coded is the relative prevalence of the international network, given by $\rho^{\mathrm{a}}_{i,\mathrm{int}}/(\rho^{\mathrm{a}}_{i,\mathrm{int}}+\rho^{\mathrm{a}}_{i,\mathrm{loc}})$. Here, $\lambda = 0.2$ per country, $\sigma = 0.75$, $\Delta t = 2$ and $\alpha = 2$.

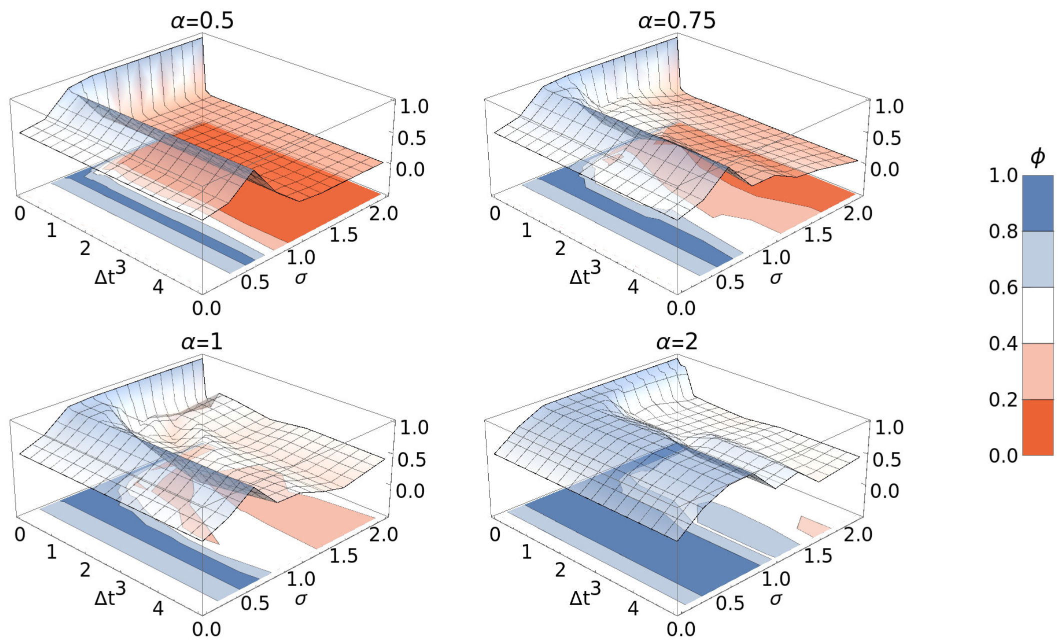

Figure 5. Mean prevalence ϕ of the international network is shown on the z axis as a function of the launch time delay Δt and the coupling strength σ. Here, $\lambda = 0.2$ per country. Averaged over 30 realizations.

We consider the international network to be banned in China and Iran. To model this, we set the values of $\Omega_{ij} = 0$ for each entry which involves one of these countries. This is equivalent to assuming that in these countries two local networks compete without any coupling to the rest of the world.

In Fig. 4 we show a typical realization of our model for parameters that allow the international network to dominate eventually. The international network starts delayed in all countries except the US, so that at the beginning in these countries the respective local network dominates. After some time, the international network obtains a significant advantage and quickly takes over North America and Europe, followed by Africa, South America, and Asia. The main patterns are in agreement with empirical observations, as reflected by the "world map of social networks" [26]. However, on a country by country level, there are deviations regarding our model and the observations. To improve the performance, it is essential to account for the specific situation of each country, for instance by adjusting λ and ν for each country separately. In addition, some networks are not really local but are present in several countries, like the network "VKontakte" (see [26]). The inclusion of further, more detailed, small-grained mechanisms and the enrichment with such data is left for future research.

To further investigate the properties of the model presented here, we define the global prevalence of the international network as

$$\Phi = \frac{1}{n_c} \sum_{i=1}^{n_c} \frac{\rho^{\mathrm{a}}_{i,\mathrm{int}}\big|_{\mathrm{st}}}{\rho^{\mathrm{a}}_{i,\mathrm{int}}\Big|_{\mathrm{st}} + \rho^{\mathrm{a}}_{i,\mathrm{loc}}\Big|_{\mathrm{st}}} \tag{11}$$

where n_c denotes the number of countries, $\rho^{\mathrm{a}}_{i,\mathrm{int}}\big|_{\mathrm{st}}$ the activity of the international network in country i at the stationary state, and $\rho^{\mathrm{a}}_{i,\mathrm{loc}}\Big|_{\mathrm{st}}$ the activity of the local network in country i at the steady state.

The results are shown in Fig. 5 for different levels of global connectivity, α, which is proportional to the control parameter $\bar{\Omega}$ used in the double meanfield approximation. We observe that for large enough launch time delays, the actual amount of time delay becomes irrelevant. We can understand this in analogy to the basin of attraction for the domination of local networks in Fig. 2, where the line separating this attractor saturates for higher activity values of the local network. Note, that a launch time delay means a shift to the right in the stream plots. In addition, note that noise in the full stochastic model can make the system jump to another attractor as it is near its border.

The independence of the actual amount of launch time delay in the limit $\Delta t \gg 0$ allows us to average over launch time delays $\Delta t \geq 2$. The result is shown in Fig. 6. Indeed, numerical simulations of the full model confirm the results from the meanfield analysis, in particular the complex role of the activity affinity σ. We can distinguish different regions depending on the global connectivity α and the activity affinity σ as denoted in Fig. 6. For small α and σ, local networks and the international network can coexist. Increasing σ and/or α favors the international network which gives rise to the blue "V"-shaped region around $\sigma = 0.5$ corroborating the saddlenode bifurcation in the double meanfield approximation. In this region, the international network dominates. See supplementary video [27] for an explicit realization of this case. For high values of σ and small values of α (red region in the right bottom corner of Fig. 6), local networks tend to dominate. Note that partial states are also possible in which the international networks dominates in some countries and local networks dominate in the remaining countries. See supplementary video [28] for an explicit realization of this case. Between the region of domination of the international network and local networks, a region is found in which the final fate of the system varies significantly between different realizations of the model ("coinflip region"). In this region, the network which wins in the US will become dominant globally. Although in this region the prevalence of the international network averaged over many realizations is about 0.5 as in the coexistence region in the left bottom corner of Fig. 6, the behavior of the system differs dramatically between them. In the coexistence region, each realization of the model leads to the same final state, namely the coexistence of local networks and the international one. In contrast, in the coinflip region, coexistence is not possible, as this region of the parameter space corresponds to the supercritical regime (the blue area in Fig. 3). In the coinflip region, about 50% of the realizations show the domination of the international network whereas the remaining 50% lead to the domination of local networks. As a consequence, even knowing the exact parameters, it is impossible to predict the fate of the system beforehand.

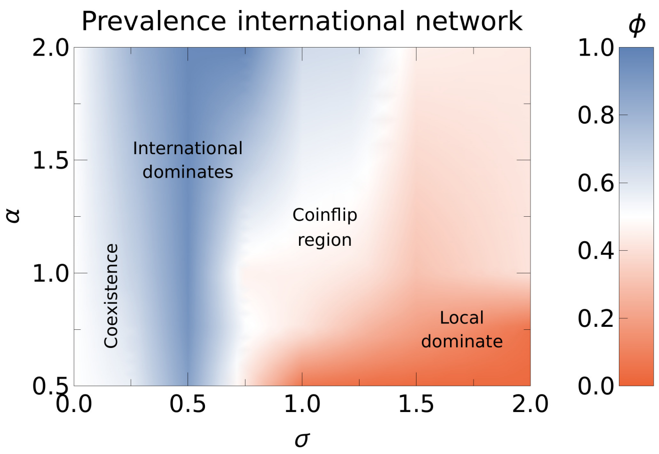

Figure 6. The prevalence of the international network averaged over time delays for $\Delta t \geq 2$ as a function of the activity affinity (σ) and the global connectivity α. Averaged over 30 realizations.

We can summarize our findings as follows. A higher value of α, which is a measure of the global connectivity of society, favors the prevalence of the international network and hinders the survival of the local ones. The role of the tendency of individuals to participate in more active networks (activity affinity), σ, is particularly interesting. Low values allow the networks to coexist whereas intermediate values always lead to the prevalence of the international network and the extinction of local ones. A high activity affinity, however, enables the prevalence of local networks and thus can even lead to the extinction of the international network.

III. DISCUSSION

The understanding of the digital world constitutes an important challenge for interdisciplinary science nowadays. In recent studies we presented a general theory for the evolution and ecology of online social networks which explains under which conditions networks can coexist and why we observe a moderate number of coexisting digital services. However, the heterogeneity found in empirical networks demands a more detailed description.

In this paper, we showed how inter-country social ties can lead to an increased intrinsic fitness of an internationally operating network compared to local networks. Interestingly, heterogeneity in network intrinsic fitnesses can impede the coexistence of networks as the system undergoes a saddlenode bifurcation. However, under certain conditions local networks can persist even if coexistence is impossible, a scenario which leads to the extinction of the international network. We confirm our analytical analysis by constructing a 1:1000 scale model of the digital world incorporating synthetic networks for the underlying societies and use data from the air transportation network as a proxy for inter-country social ties. We find that in addition to the previously mentioned scenarios, a partial state in which the international network dominates in some countries and local networks dominate in the remaining is possible as well. Finally, we note that depending on the parameters the final state of the system –whether local networks dominate or the international one wins– can be completely unpredictable as it varies randomly between different realizations of the model.

Our findings here suggest interesting future research lines. On the one hand, even without adjusting parameters on a country-by-country level, our model reproduces the main features empirically observed in the overtake of Facebook and the extinction of local networks in most countries for a certain parameter region. It remains an interesting task for future research to further increase the precision of the model. This can be done by improving the proxy for the similarity between countries and by adjusting parameters on a country-by-country level. On the other hand, the model can be extended to account for several international networks to investigate their global competition. Different properties of the networks like features or functionalities can lead to different intrinsic fitnesses. For a second international network to overcome the first one a certain minimal difference of intrinsic fitness is needed. Finally, random fluctuations of the intrinsic fitness can be incorporated to describe Darwinian selection in the digital ecosystem.

ACKNOWLEDGMENTS

This work was supported by: the European Commission within the Marie Curie ITN "iSocial" grant no. PITN-GA-2012-316808; a James S. McDonnell Foundation Scholar Award in Complex Systems; the ICREA Academia prize, funded by the *Generalitat de Catalunya*; the MINECO projects nos. FIS2010-21781-C02-02 and FIS2013-47282-C2-1-P; and the *Generalitat de Catalunya* grant no. 2014SGR608. Furthermore, M. B. acknowledges support from the European Commission LASAGNE project no. 318132 (STREP).

Appendix A: S1 model

We use the $\mathbb{S}1$ model [23–25] to generate the synthetic networks for the underlying societies of each country. The model allows to specify the degree distribution and the level of clustering. The model is based on a circle as a hidden metric space and works as follows:

1. All nodes are placed on the circle with randomly assigned variable θ which represent the polar coordinate. θ is uniformly distributed in $[0, 2\pi)$. To keep the average node density on the circle constant, its radius grows linearly with the number of nodes to satisfy $N = 2\pi R$.

2. We assign each node a second hidden variable, κ, which represent its expected degree. κ is drawn from an arbitrary distribution $\rho(\kappa)$.

3. A pair of nodes are connected with a probability r that depends on their hidden variables (θ, κ) and

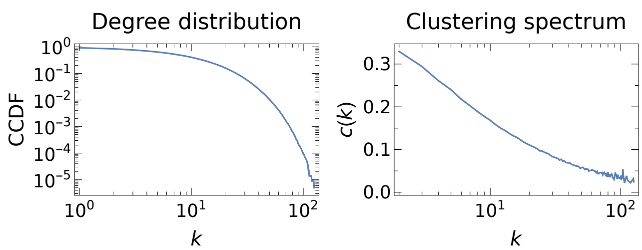

Figure 7. Degree distribution and clustering spectrum for generated networks for the example of the US (≈ 230.000 nodes).

(θ', κ')

$$r(\theta,\kappa;\theta',\kappa') = \left(1 + \frac{d(\theta,\theta')}{\mu\kappa\kappa'}\right)^{-\alpha}, \quad \text{(A1)}$$

with $\mu = \frac{\alpha-1}{2\langle k\rangle}$. Here, $d(\theta,\theta')$ denotes the geodesic distance between the two nodes on the circle and $\langle k\rangle$ the mean degree. Then, a the expected degree $\bar{k}(\kappa)$ of a node with hidden variable κ can be shown to be proportional to the latter [25]. As a consequence, the degree distribution $p(k)$ of the network follows the shape of the distribution $\rho(\kappa)$.

Here, we use an exponential distribution $\rho_\xi(\kappa) = \xi e^{-\xi\kappa}$ with $\xi = 10$. We set the parameter $\alpha = 1.5$ and have $\mu = 0.02$. After generating the networks, we remove nodes with zero degree. In Fig. 7 we show the degree distribution and the clustering spectrum for the synthetic network created for the US.

Appendix B: Air travel data

Air travel data aggregated on country basis was taken from http://www.visualizing.org/datasets/global-flights-network. Original data can be accessed at http://openflights.org/data.html.

Appendix C: Double meanfield approximation: $\bar{\Omega} > 0$ breaks symmetry of pitchfork bifurcation

The evolution equations for the double meanfield approximation (9) contain an additional control parameter $\bar{\Omega}$. For $\bar{\Omega} = 0$ we recover the case of two competing identical networks as discussed in [4]. In this case, the system undergoes a subcritical pitchfork bifurcation. Such bifurcation is symmetric locally near the critical point. However, the additional control parameter $\bar{\Omega} > 0$ breaks this symmetry. As a consequence, the system undergoes a saddlenode bifurcation instead of the former pitchfork. See Fig. 8.

[1] WeAreSocial. Global digital statshot 001 (2015, Date of access: 31/03/2015). URL `http://wearesocial.net/tag/sdmw/`.

[2] Brenner, J. & Smith, A. 72% of online adults are social networking site users. *Pew Research Internet Project* (2013, Date of access: 31/03/2015). URL `http://www.pewinternet.org/2013/08/05/72-of-online-adults-are-social-networking-site-users`.

[3] Kleineberg, K.-K. & Boguñá, M. Evolution of the digital society reveals balance between viral and mass media influence. *Phys Rev X* **4**, 031046 (2014).

[4] Kleineberg, K.-K. & Boguñá, M. Ecology 2.0: Coexistence and domination among interacting networks. *ArXiv 1410.8865* (2014).

[5] Hardin, G. The competitive exclusion principle. *Science* **131**, 1292–1297 (1960).

[6] Barabási, A.-L. & Albert, R. Emergence of scaling in random networks. *Science* **286**, 509–512 (1999).

[7] Dorogovtsev, S. N., Mendes, J. & Samukhin, A. Size-dependent degree distribution of a scale-free growing network. *Phys Rev E* **63**, 062101 (2001).

[8] Bianconi, G. & Barabási, A.-L. Bose-Einstein Condensation in Complex Networks. *Phys Rev Lett* **86**, 5632–5635 (2001).

[9] Caldarelli, G., Capocci, A., Rios, P. D. L. & and M. A. Muñoz. Scale-Free Networks from Varying Vertex Intrinsic Fitness. *Phys Rev Lett* **89**, 258702 (2002).

[10] Vázquez, A. Growing network with local rules: Preferential attachment, clustering hierarchy, and degree correlations. *Phys Rev E* **67**, 056104 (2003).

[11] Pastor-Satorras, R., Smith, E. & Sole, R. V. Evolving protein interaction networks through gene duplication. *J Theor Biol* **222**, 199–210 (2003).

[12] Fortunato, S., Flammini, A. & Menczer, F. Scale-Free Network Growth by Ranking. *Phys Rev Lett* **96**, 218701 (2006).

[13] D'Souza, R. M., Borgs, C., Chayes, J. T., Berger, N. & Kleinberg, R. D. Emergence of tempered preferential attachment from optimization. *Proc Natl Acad Sci USA* **104**, 6112–7 (2007).

[14] Papadopoulos, F., Kitsak, M., Serrano, M., Boguñá, M. & Krioukov, D. Popularity versus similarity in growing networks. *Nature* **489**, 537–540 (2012).

[15] Brockmann, D. & Helbing, D. The hidden geometry of complex, network-driven contagion phenomena. *Science* **342**, 1337–42 (2013).

[16] Reis, S. D. S. *et al.* Avoiding catastrophic failure in correlated networks of networks. *Nature Physics* (2014).

[17] Radicchi, F. Driving Interconnected Networks to Supercriticality. *Phys Rev X* **4**, 021014 (2014).

[18] Bianconi, G. & Dorogovtsev, S. N. Multiple percolation transitions in a configuration model of a network of networks. *Phys Rev E* **89**, 062814 (2014).

[19] *Networks of Networks: The Last Frontier of Complexity*, vol. XII of *Understanding Complex Systems* (Springer, 2014).

[20] De Domenico, M. *et al.* Mathematical formulation of multilayer networks. *Phys Rev X* **3**, 041022 (2013).

[21] Kivelä, M. *et al.* Multilayer networks. *Journal of Complex Networks* (2014).

[22] Argyris, J. H., Faust, G., Haase, M. & Friedrich, R. *Die Erforschung des Chaos: eine Einführung in die Theorie nichtlinearer Systeme* (Springer, Heidelberg, 2010), zweite edn.

China	253.	UnitedStates	231.	Japan	90.9	India	81.	Brazil	64.9
Germany	62.	UnitedKingdom	48.8	Russia	45.2	France	42.9	SouthKorea	37.5
Indonesia	30.	Spain	25.2	Canada	25.1	Italy	25.	Turkey	24.5
Mexico	23.3	Iran	23.	Vietnam	20.8	Poland	18.7	Pakistan	18.5
Colombia	17.1	Malaysia	16.9	Thailand	16.1	Australia	15.2	Taiwan	15.1
Netherlands	14.3	Egypt	11.4	Argentina	11.2	Nigeria	11.	Ukraine	10.4
Morocco	10.3	Sweden	8.1	SaudiArabia	7.7	Belgium	7.3	Venezuela	7.2
Peru	7.1	Romania	6.1	CzechRepublic	6.	Austria	5.9	Hungary	5.9
Switzerland	5.7	Philippines	5.6	Chile	5.5	Denmark	4.6	Portugal	4.5
Finland	4.4	Greece	4.3	Sudan	4.2	SouthAfrica	4.2	HongKong	4.1
Algeria	4.1	Norway	3.9	Slovakia	3.6	Syria	3.6	Singapore	3.4
Kenya	3.4	Belarus	3.1	NewZealand	3.	Serbia	2.9	UnitedArabEmirates	2.9
Ireland	2.8	Tunisia	2.8	Bulgaria	2.6	Uganda	2.5	Uzbekistan	2.5
Kazakhstan	2.3	Lebanon	2.2	DominicanRepublic	2.1	Israel	2.1	Guatemala	2.
Croatia	1.9	Lithuania	1.8	Jamaica	1.5	Jordan	1.5	Azerbaijan	1.5
CostaRica	1.5	Cuba	1.4	Zimbabwe	1.4	Uruguay	1.3	Ecuador	1.3

Table I. List of countries and estimated number of Internet users ($\times 10^6$) according to Wolfram Alpha database.

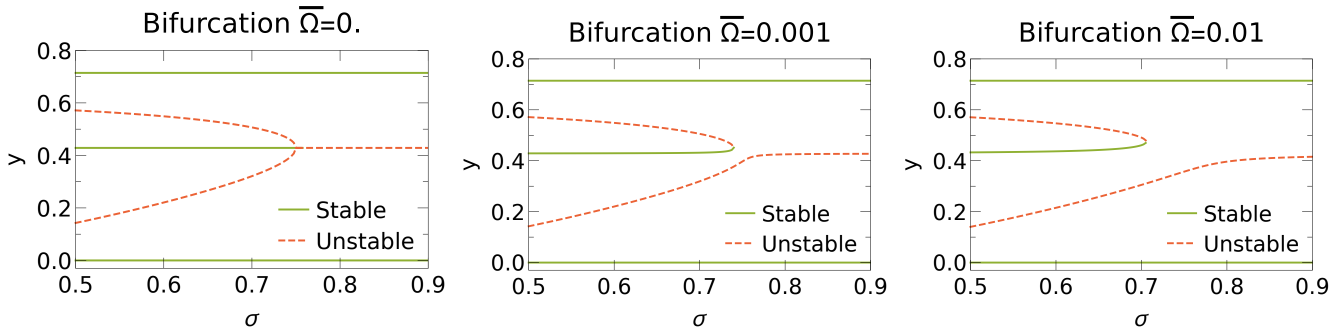

Figure 8. Bifurcation diagram as a function of the control parameter σ for different values of $\bar{\Omega}$. Here, $\lambda \langle k \rangle = 3.5$.

[23] Serrano, M., Krioukov, D. & Boguñá, M. Self-Similarity of Complex Networks and Hidden Metric Spaces. *Phys Rev Lett* **100**, 078701 (2008).

[24] Boguñá, M., Krioukov, D. & Claffy, K. C. Navigability of complex networks. *Nature Physics* **5**, 74–80 (2008).

[25] Boguñá, M. & Pastor-Satorras, R. Class of correlated random networks with hidden variables. *Phys Rev E* **68**, 036112 (2003).

[26] Cosenza, V. World map of social networks (2014, Date of access: 31/03/2015). URL `http://vincos.it/world-map-of-social-networks`.

[27] Kleineberg, K.-K. 1:000 scale model of the digital world: Extinction of local networks (2015, Date of access: 31/03/2015). URL `https://www.youtube.com/watch?v=z3dP3PD7ueA`.

[28] Kleineberg, K.-K. 1:000 scale model of the digital world: Partial persistence of local networks (2015, Date of access: 31/03/2015). URL `https://www.youtube.com/watch?v=XkZTxnJd-eI`.