Agreement dynamics on interaction networks with divese topologies

Alain Barrat,[1,2,3] Andrea Baronchelli,[4] Luca Dall'Asta,[1,2] and Vittorio Loreto[4]

[1]*LPT, CNRS, UMR 8627, Orsay, F-91405 (France)*
[2]*Univ Paris-Sud, Orsay, F-91405 (France)*
[3]*Complex Networks Lagrange Laboratory, ISI Foundation, Turin, Italy*
[4]*Dipartimento di Fisica, Università "La Sapienza" and SMC-INFM, P.le A. Moro 2, 00185 ROMA, (Italy)*

We review the behaviour of a recently introduced model of agreement dynamics, namely the Naming Game. This model describes the self-organized emergence of a linguistic convention or a communication system in a population of agents with pairwise local interactions. The mechanisms of convergence towards agreement strongly depend on the network interaction between the agents. In particular, the mean-field case in which all agents communicate with all the others is not efficient in the sense that a large temporary memory is needed. On the other hand, regular lattice topologies lead to a fast local convergence but to slow global dynamics similar to coarsening phenomena. The embedding of the agents in a small-world network represents an interesting trade-off: a local consensus is easily reached, while the long-range links allow to bypass coarsening-like convergence. We also consider alternative adaptive strategies which can lead to faster global convergence.

I. INTRODUCTION

The recent past has witnessed an important development of the activities of statistical physicists in the area of social sciences (for a recent collection of papers see [1]). Indeed, the standard methods of statistical physics are very appropriate to study collective behaviors, neglecting details and retaining only few general ingredients observed in real social interactions. For this reason, physicists have put forward a large number of theoretical models of *social dynamics*, borrowing a suite of statistical methods from the theory of interacting particles systems [2–5]. In particular, many models have been proposed for the study of opinion formation, such as the Voter model [2, 6–10], the Sznajd-Weron model [11], the Axelrod model for the dissemination of culture [12], and their variants (e.g. the models proposed by Deffuant et al. [13] and by Krause and Hegselmann [14]).

The behavior of these models has been much studied on regular topologies or in situations where each agent can interact with all the others. Recently however, network science [15–17] has led to a better knowledge of the topological properties of real social groups [18], and in particular to show that the topology of the network on which agents interact is not regular. Models of social dynamics have thus been reconsidered, in order to integrate the new framework of complex networks, and to study the influence of various complex topologies on the corresponding dynamical behavior.

In this article, we review the behaviour of a recently proposed model for the emergence of a communication system, called Naming Game, investigating its dynamical behavior on both regular topologies and complex networks [19–22]. Social interactions are indeed based on the existence of a communication system among the agents, who are able to understand each other by means of common linguistic patterns or, more generally, by means of a common vocabulary of symbols.

Such a communication system is the result of a self-organized process in which individuals select specific symbols (*words*) and associate them to concepts and ideas (*objects*). The emergence of a shared lexicon inside social groups and communities of people is very likely to be driven by simple criteria, like popularity, imitation, negotiation, and agreement. When a new concept is introduced, people refer to it using several different names or words. These words start spreading among the population, competing one against the other, until the choice of one of them is taken (with a sudden transition or with a long process) and everyone uses the same word (or symbol, etc) [23–25]. This kind of dynamics has recently become of broad interest after the diffusion of a new generation of web-tools which enable human users to self-organize a system of tags in such a way to ensure a shared classification of information about different arguments (see, for instance, del.icio.us or www.flickr.com and Refs. [26, 27]). Another application concerns global coordination problems in artificial intelligence, where a group of artificial embodied agents moving in an unknown environment have to exchange informations about the objects they gradually discover. The emergence of consensus about the objects names allows to establish a communication system. A practical example of this type of dynamics is provided by the well-known Talking Heads experiment [28, 29], in which embodied software agents develop their vocabulary observing objects through digital cameras, assigning them randomly chosen names and negotiating these names with other agents.

The paper is organized as follows..........

II. MODEL DEFINITION

A minimal model of Naming Game has been put forward by Baronchelli et al. in Ref. [30] to reproduce the main features of Semiotic Dynamics and the fundamental results of adaptive coordination observed in the Talking Heads experiment. The minimal Naming Game model

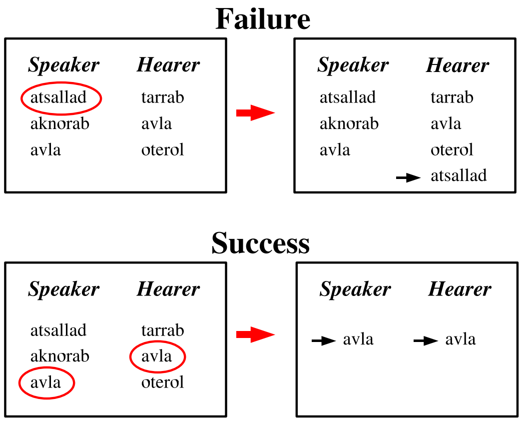

FIG. 1: Agents interaction rules. Each agent is described by its inventory, i.e. the repertoire of known words. The speaker picks up at random a name in its inventory and transmits it to the hearer. If the hearer does not know the selected word the interaction is a failure (top), and it adds the new name to its inventory. Otherwise (bottom), the interaction is a success and both agents delete all their words but the winning one. Note that if the speaker has an empty inventory (as it happens at the beginning of the game), it invents a new name and the interaction is a failure.

consists of a population of N agents observing a single object, for which they invent names that they try to communicate to one another through pairwise interactions, in order to reach a global agreement. The agents are identical and dispose of an internal inventory, in which they can store an a priori unlimited number of names (or opinions). All agents start with empty inventories. At each time step, a pair of neighboring agents is chosen randomly, one playing as “speaker”, the other as “hearer”, and negotiate according to the following rules (see also Fig. 1):

- the speaker selects randomly one of its words and conveys it to the hearer;
- if the hearer's inventory contains such a word, the two agents update their inventories in order to keep only the word involved in the interaction (*success*);
- if the hearer does not possess the uttered word, the latter is added to those already stored in the hearer's inventory (*failure*), i.e. it learns the word.

Before entering in the detailed description of the dynamics, it is worth noting some visible differences of the Naming Game with other commonly studied models of social dynamics and, in particular, of opinion formation [2, 11–13]. First of all, each agent can potentially be in an infinite number of possible discrete states (words, names, opinions), and *the maximum number of states depends on the dynamical evolution* itself. This is in contrast with traditional models (Voter, Potts, etc) in which the number of states is a fixed external parameter taking finite (and usually small) values. The *two-steps decision process* is moreover rather realistic: an agent can accumulate in its memory different possible names for the object, waiting before reaching a decision. Each dynamical step can be seen as a *negotiation* between speaker and hearer, with a certain degree of stochasticity, that is absent in deterministic models such as the Voter model. The stochastic component is however of a different nature compared to that of standard Glauber dynamics used in majority rule models [31], since here it comes from an internal selection criterion, and involves only the speaker, without affecting the (deterministic) decision process of the hearer.

An important remark also concerns the random extraction of the word in the speaker's inventory. Most previously proposed models of semiotic dynamics attempted to give a more detailed representation of the negotiation interaction assigning weights to the words in the inventories. In such models, the word with largest weight is automatically chosen by the speaker and communicated to the hearer. Success and failures are translated into updates of the weights: the weight of a word involved in a successful interaction is increased to the detriment of those of the others (with no deletion of words); a failure leads to the decrease of the weight of the word not understood by the hearer. An example of a model including weights dynamics can be found in Ref. [32] (and references therein). For the sake of simplicity the minimal Naming Game avoids the use of weights, that are apparently more realistic, but their presence is not essential for the emergence of a global collective behavior of the system.

Finally, we stress that, in the minimal Naming Game, *all agents refer to the same single object*, while in the original experiments the embodied agents could observe a set of different objects. This is actually possible only if we assume that homonymy is excluded, i.e. two distinct objects cannot have the same name. Consequently, in this model, all objects are independent and the general problem reduces to a set of independently evolving systems, each one described by the minimal model. In more realistic situations, however, the occurrence of homonymy cannot be neglected.

III. MEAN-FIELD

Many studies of social dynamics have focused on populations of agents in which all pairwise interactions are allowed, i.e. the agents are placed on the vertices of a fully-connected graph. In statistical mechanics, this topological structure is commonly referred as "mean-field" topology. The main quantities of interest which describe the system's evolution are [30]

- the total number $N_w(t)$ of words in the system at the time t (i.e. the total size of the memory);
- the number of different words $N_d(t)$ in the system

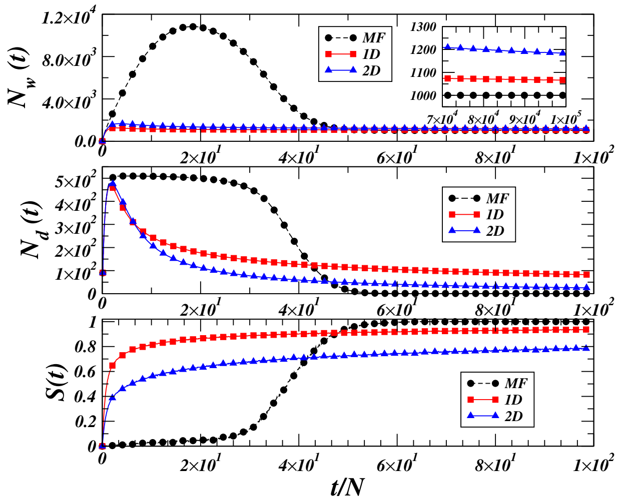

FIG. 2: Evolution of the total number of words N_w (top), of the number of different words N_d (center), and of the average success rate $S(t)$ (bottom), for a mean-field system (black circles) and low-dimensional lattices ($1D$, red squares and $2D$, blue triangles) with $N = 1024$ agents, averaged over 10^3 realizations. The inset in the top graph shows the very slow convergence in low-dimensional lattices.

at the time t;

- the average success rate $S(t)$, i.e. the probability, computed averaging over many simulation runs, that the chosen agent gets involved in a successful interaction at a given time t.

The consensus state is obtained when $N_d = 1$ and $N_w = N$ (moreover it implies $S = 1$).

The temporal evolution of the three main quantities is depicted in Fig. 2 (circles). At the beginning, many disjoint pairs of agents interact, with empty initial inventories: they invent a large number of different words, that start spreading throughout the system, through failure events. Indeed, the number of words decreases only by means of successful interactions. In the early stages of the dynamics, the overlap between the inventories is very low, and successful interactions are limited to those pairs which have been chosen at least twice. Since the number of possible partners of an agent is of order N, it rarely interacts twice with the same partner, the probability of such an event growing as t/N^2. The initial trend of $S(t)$ (black circles) is indeed linear with a slope of order $1/N^2$. In this phase of uncorrelated proliferation of words, the number of different words N_d invented by the agents grows, rapidly reaching a maximum that scales as $\mathcal{O}(N)$. Then N_d saturates, displaying a plateau, during which no new word is invented anymore (since every inventory contains at least one word). The total number N_w of words stored in the system has a similar behavior, but it keeps growing after N_d has saturated, since the words continue to propagate throughout the system even if no new one is introduced. The peak of N_w has been shown to scale as $\mathcal{O}(N^{1.5})$ [30], meaning that each agent stores $\mathcal{O}(N^{0.5})$ words. This peak occurs after the system has evolved for a time $t_{max} \sim \mathcal{O}(N^{1.5})$. In the subsequent dynamics, strong correlations between words and agents develop, driving the system to a final rather fast convergence to the absorbing state in a time $t_{conv} \sim \mathcal{O}(N^{1.5})$. The S-shaped curve of the success rate in Fig. 2 summarizes the dynamics: initially, agents hardly understand each others ($S(t)$ is very low); then the inventories start to present significant overlaps, so that $S(t)$ increases until it reaches 1, and the communication system is completely set in.

IV. COARSENING PHENOMENON ON LATTICES

A first study of the effects of topological embedding on the Naming Game dynamics is reported in Ref. [19]. When the interacting agents sit on the nodes of low-dimensional lattices, the long-time behavior is still characterized by the convergence to a homogeneous consensus state, but the evolution of the system changes considerably. In particular, the time required by the system to reach the global consensus displays a different scaling with the size N, and the effective size of the inventories is considerably diminished. Actually, the existence of different dynamical patterns are clearly visible in Fig. 2. Since each agent can interact only with a limited number of neighbors ($2d$ in a d-dimensional lattice), at the local scale the dynamics is very fast: agents can rapidly interact two or more times with their neighbors, favoring the establishment of a local consensus with a high success rate, i.e. of small sets of neighboring agents sharing a common unique word. These "clusters" of neighboring agents with a common unique word are separated by individuals having a larger inventory with two or more words, playing the role of "interfaces". These interfaces then start a diffusion process, and the clusters of unique words grow in time with a law that is typical of coarsening phenomena [19]: the competition among the clusters is driven by the fluctuations of the interfaces. The easily reached local consensus thus leads to a slow dynamics, and the global consensus takes much longer to be reached than in mean-field: for example, $\mathcal{O}(N^3)$ in dimension 1 vs. $\mathcal{O}(N^{1.5})$ in mean-field. However, another important aspect of the problem concerns the memory used by the agents. In mean-field indeed, each agent needs a memory capacity scaling as $\mathcal{O}(N^{1/2})$, i.e. diverging with the system size. In contrast, the consequence of the embedding in a finite-dimensional lattice (with finite number of neighbors), and of the subsequent coarsening like phenomena, with rapid local consensus, is that each agent uses only a finite capacity: the maximum total number of words in the system (maximal memory capacity) scales linearly with the system size N (as for the number of different words). In summary, *low-dimensional lattice systems require more time to reach the consensus compared to mean-field, but a lower use of memory.*

V. FAST CONVERGENCE IN SMALL-WORLD NETWORKS

The precise knowledge of the dynamical behavior of the Naming Game model on low-dimensional lattices, and in particular on the one-dimensional ring, makes possible to understand, by means of simple arguments and numerical simulations, the effect of the small-world property, that is a relevant feature of real complex networks.

In the following, indeed, we investigate the effect of introducing long-range connections which link agents that are far from each other on the regular lattice. In other words, we study the Naming Game on the small-world model proposed by Watts and Strogatz [33]. Starting from a quasi-one-dimensional banded network in which each node has $2m$ neighbors, the edges are rewired with probability p, i.e. p represents the density of long-range connections introduced in the network. For $p = 0$ the network retains a one-dimensional topology, while the random network structure is approached as p goes to 1. At small but finite p ($1/N \ll p \ll 1$), a small-world structure with short distances between nodes, together with a large clustering, is obtained. When $p = 0$, the system is one-dimensional and the dynamics proceeds by slow coarsening. At small p, the typical distance between shortcuts is $\mathcal{O}(1/p)$, so that the early dynamics is not affected and proceeds as in one-dimensional systems. In particular, at very short times many new words are invented since the success rate is small. The maximum number of different words scales as $\mathcal{O}(N)$, as in the other cases, while the average used memory per agent remains finite, since the number of neighbors of each site is bounded (the degree distribution decreases exponentially [34]).

The typical cluster dynamics on a small-world network is graphically represented in Fig. 3. As long as the typical cluster size is smaller than $1/p$, the clusters are typically one-dimensional, and the system evolves by means of the usual coarsening dynamics. However, as the average cluster size reaches the typical distance between two shortcuts $\sim 1/p$, a crossover phenomena toward an accelerated dynamics takes place. Since the cluster size grows as $\sqrt{t/N}$, this corresponds to a crossover time $t_{cross} = \mathcal{O}(N/p^2)$. For times much larger than this crossover, one expects that the dynamics is dominated by the existence of shortcuts, entering a mean-field like behavior. The convergence time is thus expected to scale as $N^{3/2}$ and not as N^3. The condition in order for this picture to be possible is exactly the small-world condition; indeed, the crossover time N/p^2 has to be much larger than 1, and much smaller than the consensus time for the one-dimensional case N^3, that together imply $p \gg 1/N$. In summary, *the small-world topology allows to combine advantages from both finite dimensional lattices and mean-field networks*: on the one hand, only a finite memory per node is needed, in opposition to the $\mathcal{O}(N^{1/2})$ in mean-field; on the other hand the convergence time is expected to be much shorter than in finite dimensions.

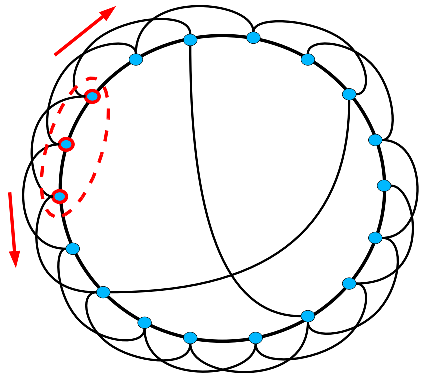

FIG. 3: A naive representation of clusters growth in the small-world model of Watts and Strogatz. A cluster (in red) starts to expand locally by coarsening dynamics like in dimension one. When the size of the cluster is of the order of the average distance between shortcuts, long-range interactions take place. The effect of these long-range interactions is that of boosting up the dynamics.

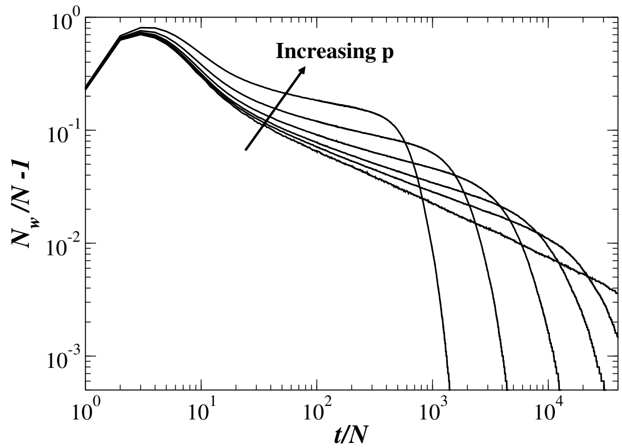

FIG. 4: Average number of words per agent in the system, N_w/N as a function of the rescaled time t/N, for small-world networks with $\langle k \rangle = 8$ and $N = 10^3$ nodes, for various values of p. The curve for $p = 0$ is shown for reference, as well as $p = 5.10^{-3}$, $p = 10^{-2}$, $p = 2.10^{-2}$, $p = 4.10^{-2}$, $p = 8.10^{-2}$, from bottom to top on the left part of the curves.

Figure 4 displays the evolution of the average number of words per agent as a function of time, for a small-world network with average degree $\langle k \rangle = 8$, and various values of the rewiring probability p. While $N_w(t)$ in all cases decays to N, after an initial peak whose height is proportional to N, the way in which this convergence is obtained depends on the parameters. At fixed N, for $p = 0$ a power-law behavior $N_w/N - 1 \propto 1/\sqrt{t}$ is observed due to the one-dimensional coarsening process. As soon as $p \gg 1/N$ however, we observe deviations getting stronger as p is increased: the decrease of N_w is first slowed down after the peak, but leads in the end to a very fast convergence.

As previously mentioned, a crossover phenomenon is expected when the one-dimensional clusters reach sizes of order $1/p$, i.e. at a time of order N/p^2. Since the agents with more than one word in memory are localized at the interfaces between clusters, their number is $\mathcal{O}(Np)$. The average excess memory per site (with respect to global consensus) is thus of order p, so that one expects $N_w/N - 1 = p\mathcal{G}(tp^2/N)$. Detailed numerical investigations confirm this picture [21], and allow to show that the convergence towards consensus is reached on a timescale of order $N^{\beta_{SW}}$, with $\beta_{SW} \approx 1.4 \pm 0.1$, close to the mean-field case $N^{3/2}$ and in strong contrast with the N^3 behavior of purely one-dimensional systems. Note that the time to converge scales as $p^{-1.4\pm.1}$, that is consistent with the fact that for p of order $1/N$ one should recover an essentially one-dimensional behavior with convergence times of order N^3.

The combination of the results concerning average used memory and convergence time shows that the small-world topology in fact combines of both finite dimensional and mean-field topologies, while corresponding to more realistic interaction networks.

VI. THE NAMING GAME ON GENERAL COMPLEX NETWORKS

A. Pair selection strategies

Before describing the behaviour of the Naming Game dynamics on general networks, it is worth noting that the definition of the model itself has in fact to be precised. Indeed, the two neighboring agents chosen to interact have different roles: one (the speaker) transmits a word and is thus more "active" than the other (the hearer). One should therefore specify whether, when choosing a pair, one chooses first a speaker and then a hearer among the speaker's neighbors, or the reverse order. If the agents sit on a fully connected graph or on a regular lattice, or even on a random graph with homogeneous degree distribution, they have an equivalent neighborhood so the order is not important. In the case of heterogeneous networks however, the degrees of the first and the second chosen nodes can have very different distributions (respectively $P(k)$ and $kP(k)/\langle k\rangle$). The asymmetry between speaker and hearer can couple to the asymmetry between a randomly chosen node and its randomly chosen neighbor, leading to different dynamical properties (this is the case for example in the Voter model, as studied by Castellano [35]). We therefore distinguish more possibilities for the definition of the Naming Game on generic networks.

- (i) A randomly chosen speaker selects (again randomly) a hearer among its neighbors. This is probably the most natural generalization of the original rule. We call this strategy *direct Naming Game*. In this case, larger degree nodes will preferentially act as hearers.

- (ii) The opposite strategy, here called *reverse Naming Game*, can also be carried out: we choose the hearer at random and one of its neighbors as speaker. In this case the hubs are preferentially selected as speakers.

- (iii) A *neutral* strategy to pick up pairs of nodes is that of considering the extremities of an edge taken uniformly at random. The role of speaker and hearer are then assigned randomly with equal probability among the two nodes.

As shown in [36], a larger memory is used for the reverse rule, although the number of different words created is smaller, and a faster convergence is obtained. This corresponds to the fact that the hubs, playing principally as speakers, can spread their words to a larger fraction of the agents, and remain more stable than when playing as hearers, enhancing the possibility of convergence. Depending on the network under study, and similarly to the Voter model case [35], the scaling laws of the convergence time can even be modified. From the point of view of a realistic interaction among individuals or computer-based agents, the direct Naming Game in which the speaker chooses a hearer among its neighbors seems somehow more natural than the other ones. In the remainder of this section therefore, we will focus on the direct Naming Game.

B. Global quantities

Figure 5 reports, for homogeneous Erdös-Renyi networks (left) and heterogeneous Barabási-Albert networks (right), the temporal evolution of the three main global quantities: the total number $N_w(t)$ of words in the system, the number of different words $N_d(t)$, and the rate of success $S(t)$. The curves for the average use of memory $N_w(t)$ show a rapid growth at short times, a peak and then a plateau whose length increases as the size of the system is increased. The time and the height of the peak, and the height of the plateau, are proportional to N. A systematic study of the scaling behavior shows that the convergence time t_{conv} scales as N^β with $\beta \simeq 1.4$ for both ER and BA. The apparent plateau of N_w does however not correspond to a steady state, as revealed by the continuous decrease of the number of different words N_d in the system: in this re-organization phase, the system keeps evolving by elimination of words, although the total used memory does not change significantly.

The observed scaling law for the convergence time is a general robust feature that is not affected by other topological details (average degree, clustering, etc), and more surprisingly it seems to be independent of the particular form of the degree distribution. We have indeed checked the value of the exponent $\beta \simeq 1.4\pm0.1$ for various

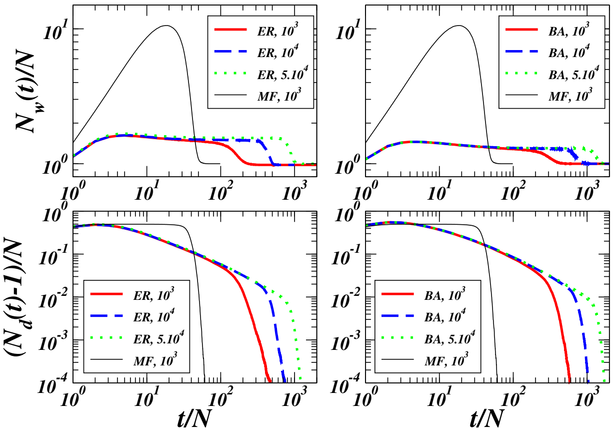

FIG. 5: ER random graph (left) and BA scale-free network (right) with $\langle k\rangle = 4$ and sizes $N = 10^3, 10^4, 5.10^4$. Top: evolution of the average memory per agent N_w/N versus rescaled time t/N. For increasing sizes a plateau develops in the reorganization phase preceding the convergence. The height of the peak and of the plateau collapse in this plot, showing that the total memory used scales with N. Bottom: evolution of the number of different words N_d in the system. $(N_d - 1)/N$ is plotted in order to emphasize the convergence to the consensus with $N_d = 1$. A steady decrease is observed even if the memory N_w displays a plateau. The mean-field (MF) case is also shown (for $N = 10^3$) for comparison.

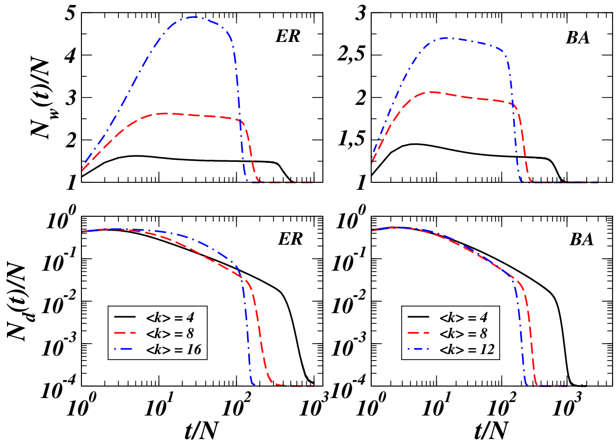

FIG. 6: ER networks (left) and BA networks (right) with $N = 10^4$ agents and average degree $\langle k\rangle = 4$, 8, 16. The increase of average degree leads to a larger memory used (N_w, top) but a faster convergence. The maximum in the number of different words is not affected by the change in the average degree (bottom).

$\langle k\rangle$, clustering, and exponents γ of the degree distribution $P(k) \sim k^{-\gamma}$ for scale-free networks constructed with the uncorrelated configuration model. These parameters have instead an effect on other quantities such as the time and the value of the maximum of memory, as shown in Fig. 6, which displays the effects of increasing the average degree on the behavior of the main global quantities. In both ER (left) and BA (right) models, increasing the average degree provokes an increase in the memory used, while the global convergence time is decreased: there is a trade-off between memory and rapidity. We have also investigated the effect of increasing the clustering at fixed average degree and degree distributions: the number of different words is not changed, but the average memory used is smaller and the convergence takes more time: it is more probable for a node to speak to 2 neighbors that share common words because they are themselves connected and have already interacted, so that it is less probable to learn new words. At fixed average degree, i.e. global number of links, less connections are available to transmit words from one part of the network to the other since many links are used in "local" triangles. The local cohesiveness is therefore in the long run an obstacle to the global convergence. This effect is similar to the observation of an increase in the percolation threshold in clustered networks, due to the fact that many links are "wasted" in redundant local connections [37].

C. Cluster statistics

Further insights into the convergence mechanisms is obtained by the investigation of the behavior of clusters of words (recall that a cluster is any set of neighboring agents sharing a common unique word). In low-dimensional lattices indeed, the dynamics of the Naming Game proceeds by formation of such clusters, that grow through a coarsening phenomenon. As shown instead in Fig. 7 for ER networks (a similar behaviour is observed in the BA case), the normalized average cluster size remains very close to zero (in fact, of order $1/N$) during the reorganization phase that follows the peak in the number of words, and converges to one with a sudden transition. The same behavior is shown also by the number of clusters $N_{cl}(t)$, that decreases to one very sharply. The emerging picture is not that of a coarsening or growth of clusters, but that of a slow process of correlations between inventories, followed by a multiplicative process of cluster growth triggered by a sort of symmetry breaking event in the success probability of the words (in favor of the word that will ultimately survive).

VII. STRATEGIES FOR FASTER CONVERGENCE

In all the investigated cases, the time to convergence grows quite fast as a function of the system size. A natural and important question is therefore whether it is possible to improve the performance of the system. More precisely, a major challenge would be to improve the population-scale performances of the process without loosing the simplicity of the microscopic rules, which is the precious ingredient that allows for in-depth investiga-

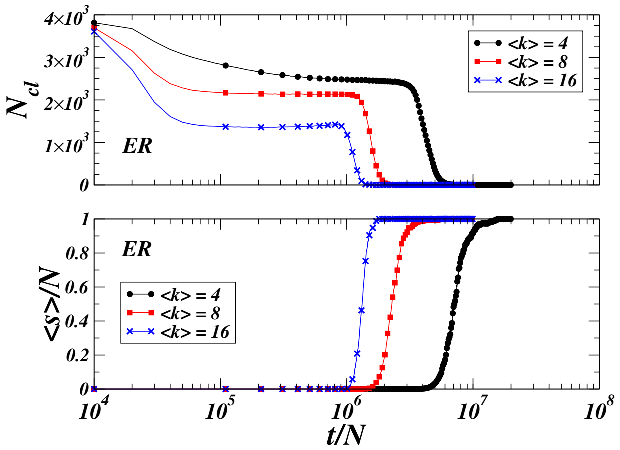

FIG. 7: Number of clusters N_{cl} and normalized average cluster size $\langle s \rangle/N$ vs. time for ER networks with $N = 10^4$, $\langle k \rangle = 4$ (circles), $\langle k \rangle = 8$ (squares), $\langle k \rangle = 16$ (crosses).

tions of global-scale dynamics. To investigate this issue, we come back for simplicity to the framework of mean-field dynamics, in which all agents can interact with all others.

Up to this point agents, when playing as speakers, extract randomly a word in their inventories. This feature, along with the drastic deletion rule that follows a successful game, is the distinctive trait of the model. Other previously proposed models use the possibility of associating a weight to each word in each inventory [32]: weights are updated at each interaction, with rewards for winning words and penalties for the others. While such sophisticated structures could in principle lead to faster convergence, they also make the models more complicated, compromising the possibility of a clear global picture of the convergence process. We follow here a different route which maintains the simplicity of the dynamical rules. Among the words of a given agent, two words can be easily distinguished: the last recorded one and the last one that gave rise to a successful game, i.e. the first that was recorded in the new inventory generated after the successful interaction. Natural strategies to investigate consist therefore in choosing systematically one of these particular words. We shall refer to these strategies as "play-last" and "play-first" respectively.

A. "Play-last" strategy

When the "play-last" strategy is adopted, the peak time and height scale respectively as $t_{max} \sim N^{\alpha}$ with $\alpha \approx 1.3$ and $N_w^{max} \sim N^{\gamma}$ with $\gamma \approx 1.3$, i.e. the used memory is reduced, while the convergence time scales as $t_{conv} \sim N^{\beta}$ with $\beta \approx 2.0$. At the beginning of the process, playing the last registered word creates a positive feedback that enhances the probability of a success. In particular a circulating word has more probabilities of being played than with the usual stochastic rule, thus creating a scenario in which less circulating words are known by more agents. On the other hand the "last in first out" approach is highly ineffective when agents start to win, i.e. after the peak. In fact, the scaling $t_{conv} \sim N^{\beta}$ can be explained through simple analytical arguments. Let us denote by N_a the number of agents having the word "a" as last recorded one. This number can increase by one unit if one of these agents is chosen as speaker, and one of the other agents is chosen as hearer, i.e. with probability $N_a/N \times (1 - N_a/N)$; the probability to decrease N_a of one unit is equal to the probability that one of these agents is a hearer and one of the others is a speaker, i.e. $(1 - N_a/N)N_a/N$. These two probabilities are perfectly balanced so that the resulting process for the density $\rho_a = N_a/N$ can be written as an unbiased random walk (with actually a diffusion coefficient $\rho_a(1 - \rho_a)/N^2$); it is then possible to show that the time necessary for one of the ρ_a to reach 1 is of order N^2. In summary, in this framework it is much more difficult to bring to convergence all the agents, since each residual competing word has a good probability of propagating to other individuals.

B. "Play-first" strategy

The "play-first" strategy, on the other hand, leads to a faster convergence. Due to a sort of arbitrariness in the strategy before the first success of the speaker, the peak related quantities keep scaling as in the usual model, so that $t_{max} \sim N^{\alpha}$ and $N_w^{max} \sim N^{\gamma}$ with $\alpha \approx \gamma \approx 1.5$. This seems natural, since playing the first recorded word is essentially the same as extracting it randomly when most agents have only few words. In fact, in both cases no virtuous correlations or feedbacks are introduced between circulating and played words. However, playing the last word which gave rise to a successful interaction strongly improves the system-scale performances once the agents start to win. In particular it turns out that for the difference between the peak and convergence time we obtain $(t_{conv} - t_{max}) \sim N^{\delta}$ with $\delta \approx 1.15$, so that the behavior of the convergence time is the result of the combination of two different power law regimes, i.e. $t_{conv} \sim aN^{\alpha} + bN^{\delta}$. On the other hand, the usual stochastic rule leads to $(t_{conv} - t_{max}) \sim N^{1.5}$. This means that the "play-first" strategy is able to reduce the time that the system has to wait before reaching the convergence, after the peak region. This seems the natural consequence of the fact that successful words increase their chances to be played while suppressing the spreading of other competitors.

C. "Play-smart", an adaptive strategy

Compared to the usual random extraction of the played word, the "play-last" strategy is more performing at the beginning of the process, while the "play-first" one allows to fasten the convergence of the process, even if it

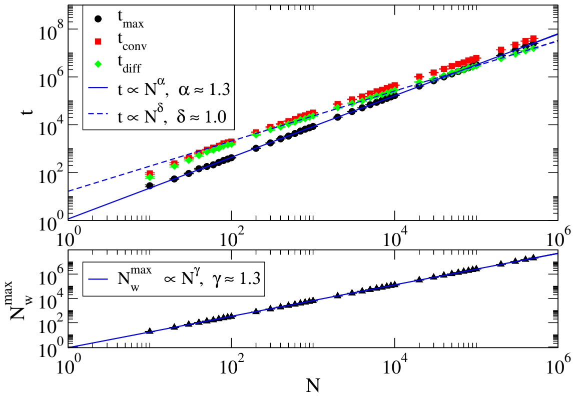

FIG. 8: Play-smart strategy - scaling with the population size N. Top - For the time of the peak $t_{max} \sim N^{\alpha}$, $\alpha \approx 1.3$, while for the convergence time we have $t_{conv} \sim aN^{\alpha} + bN^{\delta}$ with $\delta \approx 1.3$, $\delta \approx 1.0$. Bottom - the maximum number of words scales as $N_w^{max} \sim N^{\gamma}$ with $\gamma \approx 1.3$. The "play-smart" rule gives rise to a more performing process, from the point of view of both convergence time and memory needed.

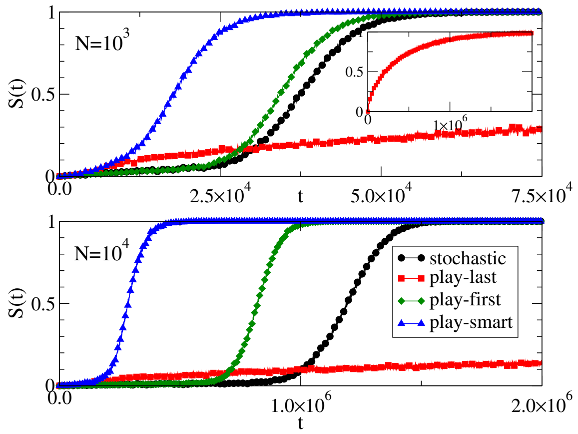

FIG. 9: Success rate curves $S(t)$ for the various strategies: stochastic, "play-last", "play-first" and "play-smart". At the beginning of the process the stochastic and "play-first" strategies yield similar succe ss rates, but then the deterministic rule speeds up convergence. On the other hand also the "play-smart" and the "play last" evolve similarly at the beginning, but the latter reaches the final state much earlier through a steep jump. It is worth noting that for three strategies the $S(t)$ curves present a characteristic $S-$*shaped* behavior, while in the "play-last" one the disorder-order transition is more continuous (see inset in the top figure). All curves, both for $N = 10^3$ and $N = 10^4$, have been generated averaging over 3×10^3 simulation runs.

is effective only after the peak of the total number of words. We therefore define a third alternative strategy which results from the combination of the two. The new prescription, called "play-smart", is the following:

- If the speaker has never took part in a successful game, it plays the last word recorded;

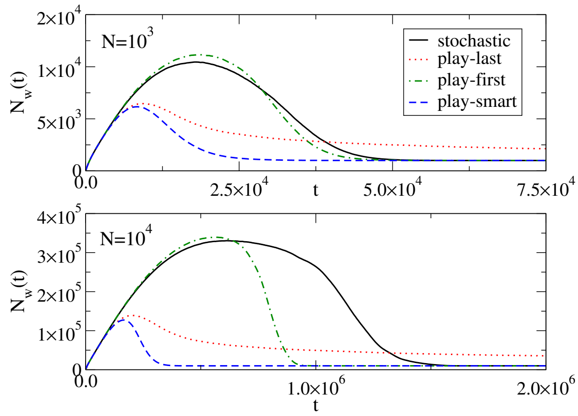

FIG. 10: Total number of words $N_w(t)$ for the various strategies: stochastic, "play-last", "play-first" and "play-smart". Due to different scaling behaviors of the process, differences become more and more relevant for larger N (top figure: $N = 10^3$; bottom figure: $N = 10^4$). The "play-smart" approach combines the advantages of "play-last" and "play-first" strategies.

- Else, if the speaker has won at least once, it plays the last word it had a communicative success with.

The first rule will thus be applied mostly at the beginning, and as the system evolves, the second rule will be progressively adopted by more and more agents. Since the change in strategy is not imposed at a given time, but takes place gradually, in a way depending of the evolution of the system, such a strategy has also the interest of being in some sense self-adapting to the system's actual state. In Figure 8, the scaling behaviors relative to the "play-smart" strategy are reported. Both the height and time of the maximum follow the scaling of the "play-last" strategy: $t_{max} \sim N^{\alpha}$ and $N_w^{max} \sim N^{\gamma}$ with $\alpha \approx \gamma \approx 1.3$. The convergence time, on the other hand, scales as a superposition of two power laws: $t_{conv} \sim aN^{\alpha} + bN^{\delta}$ with $\alpha \approx 1.3, \delta \approx 1.0$. Thus, the global behavior determined by the "play-smart" modification is indeed less demanding in terms of both memory and time. In particular, while the lowering of the peak height yields in fact a slower convergence for the "play-last" strategy, the progressive self-driven change in strategy allows to fasten the convergence further than for the "play-first" strategy.

Finally, in order to have an immediate feeling of what different playing word selection strategies imply, we report in Figures 9 and 10 the success rate $S(t)$ and the total number of words, $N_w(t)$ relative to the four strategies described previously, for two different sizes. The "play-first" and "play-smart" curves exhibit the same "S-shaped" behavior for $S(t)$ as in the case of the stochastic model, while the "play-last" rule affects qualitatively the way in which the final state is reached. Indeed, in this case the transition between the initial disordered state and the final ordered one is more continuous (see the inset in the top figure). Moreover, Figure 10 illustrates that the choice of the strategy has substantial quantitative consequences for both necessary memory and time needed to reach convergence, even if the changes in scaling behavior could at first appear rather limited (from $N^{1.5}$ to $N^{1.3}$). In particular, the "play-smart" strategy, which adapts itself to the state of the system, leads to a drastic reduction of the memory and time costs and thus to a dramatic increase in efficiency.

VIII. CONCLUSIONS

IX. ACKNOWLEDGMENTS

A. Baronchelli and V. L. are partially supported by the EU under contract IST-1940 (ECAgents). A. Barrat and L.D. are partially supported by the EU under contract 001907 "Dynamically Evolving, Large Scale Information Systems" (DELIS).

[1] P.L. Garrido, J. Marro and M.A. Munoz (eds.), Eighth Granada Lectures on *Modeling cooperative behaviour in the social sciences*, Granada, Spain, 7-11 February 2005, AIP Conference Proceedings **779** (2005).

[2] Ligget, T., *Interacting particle systems,* New York, Springer-Verlag (1985).

[3] de Oliveira, S. M., de Oliveira, P. M. C., and Stauffer, D., *Evolution, Money, War and Computers - Non-Traditional Applications of Computational Statistical Physics,* Teubner, Stuttgart (1999).

[4] Durlauf, S. N., 'Statistical Mechanics Approaches to Socioeconomic Behavior,' in *The Economy as a Complex Evolving System II,* W.B. Arthur, S.N. Durlauf and D. Lane eds., Redwood City, Addison-Wesley (1997).

[5] Blume, L., *The statistical Mechanics of Social Interaction,* Mimeo, Cornell University, Ithaca, N.Y. (1991).

[6] Krapivsky, P. L., 'Kinetics of monomer-monomer surface catalytic reactions,' *Phys. Rev. A* **45**, 1067 (1992).

[7] Krapivsky, P. L., and Redner, S., 'Dynamics of Majority Rule in Two-State Interacting Spin Systems,' *Phys. Rev. Lett.* **90**, 238701 (2003).

[8] Ben - Naim, E., Frachebourg, L., and Krapivsky, P. L., 'Coarsening and Persistence in the Voter Model,' *Phys. Rev. E* **53**, 3078 (1996).

[9] Sood, V., and Redner, S., 'Voter Model on Heterogeneous Graphs,' *Phys. Rev. Lett.* **94**, 178701 (2005).

[10] Frachebourg, L., and Krapivsky, P. L., 'Exact results for kinetics of catalytic reactions,' *Phys. Rev. E* **53**, R3009 (1996).

[11] Sznajd-Weron, K., and Sznajd, J., 'Opinion evolution in closed community,' *Int. J. Mod. Phys. C* **11**, 1157 (2000).

[12] Axelrod, R., 'The Dissemination of Culture: A Model with Local Convergence and Global Polarization,' *J. of Conflict Resolut.* **41**, 203, (1997).

[13] Deffuant, G., Neau, D., Amblard, F., and Weisbuch, G., 'Mixing beliefs among interacting agents,' *Adv. Compl. Syst.* **3**, 87 (2001).

[14] Hegselmann, R., and Krause, U., 'Opinion dynamics and bounded confidence models,. analysis, and simulation,' *J. Art. Soc. Soc. Sim.* **5(3)**, paper 2 (2002).

[15] A.-L. Barabási and R. Albert, Rev. Mod. Phys. **74**, 47-97 (2002).

[16] S.N. Dorogovtsev and J.F.F. Mendes, *Evolution of Networks: from biological nets to the Internet and WWW* (Oxford University Press, Oxford 2003).

[17] R. Pastor-Satorras and A. Vespignani, *Evolution and Structure of the Internet: A statistical physics approach* (Cambridge University Press, Cambridge, 2004).

[18] M. Granovetter, 'The strength of weak ties' American Journal of Sociology, **78**, 1360 (1973).

[19] Baronchelli, A., Dall'Asta, L., Barrat, A., and Loreto, V., 'Topology Induced Coarsening in Language Games,' *Phys. Rev. E* **73**, 015102(R) (2006).

[20] Baronchelli, A., Dall'Asta, L., Barrat, A., and Loreto, V., 'Strategies for fast convergence in semiotic dynamics,' *ALIFE X*, Bloomington Indiana (2006), (preprint arxiv:physics/0511201).

[21] Dall'Asta, L., Baronchelli, A., Barrat, A., and Loreto, V., 'Agreement dynamics on small-world networks,' *Europhys. Lett.* **73**, 969 (2006).

[22] Dall'Asta, L., Baronchelli, A., Barrat, A., and Loreto, V., 'Non-equilibrium dynamics of language games on complex networks,' *Phys. Rev. E* **73**, 036105 (2006).

[23] Lass, R., *Historical Linguistics and Language Change*, Cambridge University Press (Cambridge), 1997.

[24] Briscoe, T., *Linguistic evolution through language acquisition: formal and computational models*, Cambridge University Press, Cambridge (1999).

[25] Hurford, J., Knight, C., and Studdert-Kennedy, M., (eds.), *Approaches to the Evolution of Human Language*, Cambridge University Press (Cambraidge), 1999.

[26] Cattuto, C., Loreto, V., and Pietronero, L., 'Semiotic dynamics and collaborative tagging,' preprint (2006).

[27] Golder, S. A., and Huberman, B. A., 'The structure of collaborative tagging systems,' preprint arxiv:cs.DL/0508082 (2005).

[28] Steels, L., 'The Origins of Ontologies and Communication Conventions in Multi-Agent Systems', *Autonomous Agents and Multi-Agent Systems* **1**, 169-194 (1998).

[29] Steels, L., *The Talking Heads Experiment. Volume 1. Words and Meanings*, Laboratorium, Antwerpen (Belgium), 1999.

[30] A. Baronchelli, M. Felici, V. Loreot, E. Caglioti, and L. Steels, J. Stat. Mech. P06014 (2006).

[31] Glauber, R. J., 'Time-dependent statistics of the Ising model,' *J. Math. Phys.* **4** 294 (1963).

[32] Lenaerts, T., Jansen, B., Tuyls, K., and de Vylder, B., 'The evolutionary language game: An orthogonal approach,' *Journal of Theoretical Biology*, **235**, 566-582 (2005).

[33] Watts, D. J., and Strogatz, S. H., 'Collective dynamics of small-world networks,' *Nature* **393**, 440–442 (1998).

[34] Barrat, A., and Weigt, M., 'On the properties of small-world network models,' *Eur. Phys. J. B* **13**, 547 (2000).

[35] C. Castellano 'Effect of network topology on the ordering dynamics of voter models' *AIP Conf. Proc.* **779**, 114 (2005).

[36] A. Baronchelli, L. Dall'Asta, A. Barrat, and V. Loreto, 'Bootstrapping communication in language games. Strategy, topology and all that' *Proceedings of Evolang6*, 2006.

[37] Serrano, M. Á., Boguna, M., 'Percolation and epidemic thresholds in clustered networks,' preprint arxiv:cond-mat/0603353 (2006).