Reconciling long-term cultural diversity and short-term collective social behavior

Luca Valori[a], Francesco Picciolo[a], Agnes Allansdottir[b], and Diego Garlaschelli[c,d,1]

[a]Department of Chemistry, University of Siena, Via A. De Gasperi 2, 53100 Siena, Italy; [b]University of Siena, 53100 Siena, Italy; [c]Lorentz Institute for Theoretical Physics, University of Leiden, Niels Bohrweg 2, NL-2333 CA Leiden, The Netherlands; and [d]CABDyN Complexity Centre, University of Oxford, Park End Street OX1 1HP, Oxford, United Kingdom

Edited* by Robert Axelrod, University of Michigan, Ann Arbor, MI, and approved November 22, 2011 (received for review June 15, 2011)

An outstanding open problem is whether collective social phenomena occurring over short timescales can systematically reduce cultural heterogeneity in the long run, and whether offline and online human interactions contribute differently to the process. Theoretical models suggest that short-term collective behavior and long-term cultural diversity are mutually excluding, since they require very different levels of social influence. The latter jointly depends on two factors: the topology of the underlying social network and the overlap between individuals in multidimensional cultural space. However, while the empirical properties of social networks are intensively studied, little is known about the large-scale organization of real societies in cultural space, so that random input specifications are necessarily used in models. Here we use a large dataset to perform a high-dimensional analysis of the scientific beliefs of thousands of Europeans. We find that interopinion correlations determine a nontrivial ultrametric hierarchy of individuals in cultural space. When empirical data are used as inputs in models, ultrametricity has strong and counterintuitive effects. On short timescales, it facilitates a symmetry-breaking phase transition triggering coordinated social behavior. On long timescales, it suppresses cultural convergence by restricting it within disjoint groups. Moreover, ultrametricity implies that these results are surprisingly robust to modifications of the dynamical rules considered. Thus the empirical distribution of individuals in cultural space appears to systematically optimize the coexistence of short-term collective behavior and long-term cultural diversity, which can be realized simultaneously for the same moderate level of mutual influence in a diverse range of online and offline settings.

collective social phenomena | homophily | social influence | sociophysics | ultrametricity

How a society spontaneously organizes macroscopically from the microscopic, uncoordinated behavior of individuals is one of the most studied and exciting problems of modern science (1–3). Collective social phenomena are systematically observed in several aspects of everyday life, including the onset of large-scale popularity and fashion [both offline (4) and online (5–7)], the existence of large speculative bubbles and herding behavior in financial markets (8–10), the spontaneous emergence of order in traffic and crowd dynamics (11), the properties of voting dynamics (4, 12), the structure of countrywide communication networks (3, 13, 14), individual specialization and the spreading of cooperation, habits, fear, gossip, rumors, etc. (2, 11, 15). When collective phenomena occur, large parts of a population turn out to be globally correlated or synchronized as a result of the combination of many local interactions, even if no centralized mechanism takes place. Importantly, the collective outcome is different from a mere superposition of noninteracting individual behaviors, and contrasts the "representative agent" scenario often postulated in economic theories (16). The different characteristics, choices and behaviors of individuals, rather than being 'averaged out' in the long run and at a large scale, may in some circumstances become amplified at the societal level (1).

The substrate for collective phenomena is being expanded by the increasing connectedness and integration of people across the world (17, 18), made possible on one hand by the steady enhancement of communication technologies (18–21) and on the other hand by the intensification of global trade, worldwide travel, and international education (22–24). All these factors virtually reduce physical and communication distances among people, irrespective of their cultural traits. In particular, large-scale electronic platforms where people can exchange information with unprecedented speed and breadth are triggering a major shift toward effectively infinite-ranged social interactions (see *SI Text*, *Discussion*). This naturally leads to the question of whether the diversity of behaviors, attitudes, and opinions is destined to be progressively reduced in the long run. Naively, one expects that stronger and more frequent bursts of collective social phenomena taking place on short timescales may gradually result in more homogeneous cultural traits in the long term. However, the emergence of group boundaries has preserved cultural diversity for millenia across populations (25–27) and survived many technological revolutions that extended the length and range of social interaction (18–21). Moreover, recent studies of online behavior suggest that, even in the virtual world, social influence massively emerges (6) and determines large-scale collective phenomena (7), but at the same time is observed to coexists with the systematic presence of long-lived communities of people sharing similar traits (28–30).

The above paradox is reinforced by the fact that similar mechanisms are believed to be among the key driving forces of both collective social behavior (1, 6) and cultural convergence (31). Various simplified models have been introduced to quantitatively simulate the fate of cultural diversity and the dynamics of opinions in large groups (2, 11). Rather than conveying a detailed picture of reality, models of social dynamics aim at understanding the effects that different mechanisms proposed in social science may have when combined together and when taking place at a large-scale level. In particular, the popular model proposed by Axelrod (31) studies the combined effects of two key phenomena: *social influence* [the observed tendency of social interactions to favor convergence and consensus (32)] and *homophily* [the observed tendency of culturally similar individuals to interact more than dissimilar ones (33–35)]. The importance of the model resides in showing that social influence and homophily do not necessarily reinforce each other and determine a culturally homogeneous society. In fact, the model suggests that the persistence of cultural diversity in the long term is ensured by the inhibition of influence among dissimilar individuals, even if socially tied. These results are enhanced by the introduction of an impor-

Author contributions: D.G. designed research; A.A. and D.G. performed research; L.V., F.P., and D.G. analyzed data; and D.G. wrote the paper.

The authors declare no conflict of interest.

*This Direct Submission article had a prearranged editor.

[1]To whom correspondence may be addressed. E-mail: garlaschelli@lorentz.leidenuniv.nl.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1109514109/-/DCSupplemental.

tant generalization. According to assimilation-contrast theory (36) [recently revived in the modelling literature under the name of *bounded confidence* hypothesis (2, 37–39)], individuals are likely to accept only moderate opinion changes and to reject extreme ones. Quantitatively, when the opinions or cultural traits of an individual are represented as a scalar or vector variable, the bounded confidence hypothesis results in two individuals being potentially influenced by each other only if the distance between their associated variables is smaller than a certain threshold, representing the level of confidence or *tolerance* (2, 37–39). This threshold, which in a simplified picture is assumed to have the same value ω across the entire population, quantifies how susceptible an individual is to possible cultural influences. According to this scenario, two individuals can only influence each other if they are both socially tied and culturally similar: the effective medium of interaction is the overlap between the social network and the *cultural graph* connecting pairs of sufficiently similar individuals (see SI Text for an extended discussion). If introduced into the Axelrod model, the bounded confidence hypothesis strengthens the persistence of cultural diversity (38, 39) and makes it more robust under destabilizing mechanisms such as cultural drift (40, 41).

However, if plugged into other models of social dynamics (2, 9, 11), bounded confidence prevents information diffusion across culturally disconnected groups, and therefore also implies a reduction of collective social behavior in the short term (we will give an explicit example of this effect by simulating both short- and long-term dynamics on random data). This appears to confirm that the coexistence of long-term cultural diversity and short-term collective behavior is a paradox, which can only be solved by invoking different mechanisms at different timescales. However, this expectation follows from the naive assumption that the details of the initial underlying distribution of traits in cultural space are essentially irrelevant to the final or asymptotic outcome of cultural dynamics. This assumption is often implicit in our way of imagining cultural convergence as taking place on otherwise uniformly distributed traits, will little emphasis on what would change if the distribution was actually heterogeneous. In models, the assumption becomes explicit and indeed individuals are always initially assigned uniformly random traits. While many contributions have studied how the process of cultural convergence under social influence changes if different *dynamical* hypotheses are introduced (2, 31, 37–50), little is known about the effects of different initial *cultural* specifications. To fill this gap, here we first perform an empirical analysis of the structure of high-dimensional cultural vectors in a large set of individuals, and then use such vectors as seeds for dynamical models. We show that, even without postulating more complicated scenarios, the paradox can be parsimoniously explained by an insofar ignored aspect of real multidimensional cultural profiles; i.e., their nearly ultrametric distribution in cultural space.

Results

The Hierarchical Distribution of Individuals in Cultural Space. In order to understand the effects of real, rather than completely random, cultural specifications, we need to consider empirical data which are however not already affected by social influence (which produces correlations among cultural traits). As we discuss in detail in the *SI Text*, this leads us to the choice of the large *Eurobarometer* dataset (51–53), the outcome of an official Europe-wide survey based on questionnaires filled during face-to-face interviews, which allowed us to reconstruct the empirical multidimensional vectors of more than 13,000 individuals across 12 European countries. The nature of the dataset (see *SI Text* for details) ensures that we can obtain representative samples of vectors which are maximally random under the constraint of reflecting a country's overall cultural characteristics, and at the same time unbiased by social influence effects. The residual correlations, which are the main focus of our analysis, can therefore be regarded as an intrinsic property of how individuals are empirically distributed in cultural space, as opposed to completely random assumptions. The multiple-choice nature of the questionnaire allowed us to define, for each individual i in a given country, an F-dimesional vector $\vec{v}_i$ whose kth component $v_i^{(k)}$ represents the answer given by i to question k in the survey ($k = 1,\ldots F$, where the number of questions is $F = 161$). For each pair i, j of individuals we defined a normalized metric distance d_{ij} ($0 \le d_{ij} \le 1$) between $\vec{v}_i$ and $\vec{v}_j$, measuring the difference between the answers given by i and j. We considered 13 groups, each of $N = 500$ individuals: one group for each of the 12 countries, plus a thirteenth group of 500 individuals sampled all across Europe.

In addition to real data, we considered two types of randomized data which represent important null models providing informative benchmarks throughout our analysis. A first type of randomization ("random answers") simply consists in defining N random vectors, each obtained drawing F answers uniformly among the possible alternatives. This simulates N individuals giving completely random answers to the questionnaire, and does not depend on the empirically observed answers. This provides a unique random benchmark against which all sampled groups can be compared, and corresponds to the usual initial specification in Axelrod's model (31) and modifications (2, 38–43, 50). A second type of randomization ("shuffled answers") consists in randomly shuffling, for each of the F questions in the questionnaire, the real answers given by the N individuals of the group considered. In this case, different groups have different randomized benchmarks, each characterized by its own probability distribution (determined by real data) of possible answers. This null model is very important, as it preserves the number of times a particular answer was actually given to each question (so it preserves "more fashioned" answers for each group), but destroys the correlations between answers given by the same individual to different questions*.

We analyzed several properties of real and randomized data, as a preliminary step before studying the impact of the empirical structure on short- and long-term dynamics. In the *SI Text* we report a detailed study of the cross-correlations among opinions within individuals, an information which is not available in one-dimensional studies. We found a pattern analogous to the "likes attract" phenomenon: individuals with more beliefs in common are more likely to agree on other opinions (strong positive correlation), while dissimilar individuals tend to ignore, rather than "repel," each other (weak negative correlation). However, in this case we have an evidence of a deeper mechanism, since we know that individuals in our data are socially unrelated. Therefore, rather than an effect of homophily and social influence, the observed result is a signature of intrinsic interopinion correlations within each individual. We also found that the observed "attraction" among opinions does not imply, as one would naively expect, that the empirical vectors $\{\vec{v}_i\}$ are closer to each other in cultural space than shuffled data. Actually, the distribution of intervector distances has the same average value in real and shuffled data. Rather, the main difference is the variance of the distribution: real distances are much more broadly distributed than shuffled ones.

Higher-order differences between real and randomized data can be characterized by performing a hierarchical clustering algorithm of the vectors $\{\vec{v}_i\}$, which represents the latter as leaves of a dendrogram where culturally closer individuals have a lower common branching point. We find (see Fig. 1*A*) that the dendrogram for real data is well structured in sub-branches nested within

*Note that in more traditional one-dimensional analyses where only a single opinion is considered, the shuffling procedure would be impossible: shuffled data would be equivalent to the original data, and the only possible null model would be the first one we described (or one where the empirical abundances are modified arbitrarily).

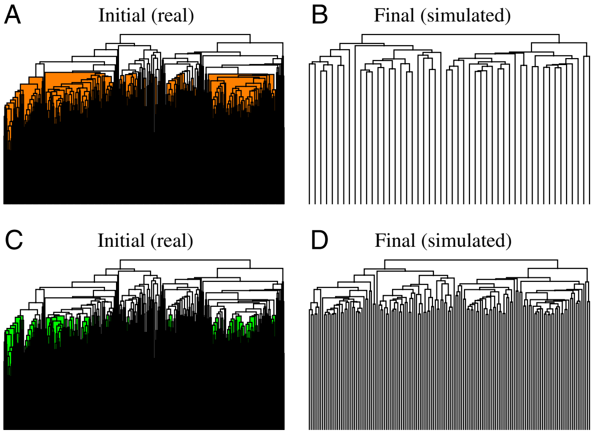

Fig. 1. The hierarchical structure implied by interopinion correlations, and its constraining effects on cultural evolution. (*A*) Dendrograms resulting from the application of an average linkage clustering algorithm to the cultural vectors $\{\vec{v}_i\}$, represented as leaves of the tree along the horizontal axis, for the Germany group; when we simulate the modified Axelrod model (see text), real data are considered as the initial state, and a confidence level (corresponding to the horizontal line below which the shaded region originates) is imposed. (*B*) Due to ultrametricity, in the corresponding final state of the model all the individuals within a common shaded branch in the initial dendrogram collapse to the same cultural vector, with negligible effects on the upper part of the dendrogram. (*C*) The same initial state as (*A*) is considered, but a lower confidence level is imposed. (*D*) Correspondingly, the final state of the model consists of a larger number of distinct cultural vectors, each containing on average less individuals with collapsed vectors. Thus the number of distinct final vectors (the leaves of the final dendrogram) coincides with the number of branches intersecting the horizontal line in the initial dendrogram. If shuffled or random opinions are taken as the initial state of the model (not shown), this is no longer true since the convergence of cultural vectors also affects the dendrogram's structure above the horizontal line, signalling a lack of ultrametricity.

branches, indicating that cultural space is heterogeneously populated by dense communities of similar individuals, separated by sparsely occupied regions. The hierarchical character of this distribution shows that denser regions are iteratively fragmented into denser regions nested within them. This peculiar organization indicates that the original distances are approximately *ultrametric* (54); i.e., the tree-like representation rendered by the dendrogram is not just an artifact of the clustering algorithm, but a natural property of the data. This means that the height of the first branching point connecting two individuals i and j approximately corresponds to the original distance d_{ij} between $\vec{v}_i$ and $\vec{v}_j$. By contrast, the dendrograms for shuffled and real data are trivially structured (see Fig. S2) with no well-defined internal separation between different hierarchical levels. In this case the dendrogram is not representative of the original distribution of vectors, and is merely an uninformative outcome of the algorithm which is forcing nonultrametric data into a tree-like description. In such a situation, the vertical dimension of the dendrogram loses it correspondence with the original intervector distances, and provides a highly distorted image of the latter.

The Effects of Ultrametricity. The ultrametric hierarchy discussed above has important static and dynamic consequences. As we show in the *SI Text*, the branches of the dendrogram "cut" horizontally at a distance ω coincide with the connected components of the ω-dependent cultural graph we defined in the beginning. The (normalized) size of the largest connected component represents a global measure of influence among individuals as a function of ω, while the density of links in the cultural graph represents (for the same value of ω) a purely local meaure of influence. We found that real data are always characterized by higher levels of both local and global influence than shuffled and random data (see *SI Text*). Different groups also show differences among themselves, and their global levels of influence remain significantly different even after standardizing (controlling for) their local influence level. Therefore any process which depends on the cultural distance between individuals might have very different global outcomes even when taking place on locally identical structures.

All the above results show that even randomly sampled individuals are not characterized by uniformly random cultural vectors. While this is not surprising, the peculiar hierarchical distribution implied by empirical interopinion correlations is highly nontrivial, and unpredictable a priori. As we show in what follows, the dynamics of opinions and culture is strongly dependent on initial conditions, and the observed ultrametricity dramatically alters the predictions of models simulating both short- and long-term dynamics. In particular, it determines a condition for the coexistence of cultural heterogeneity and collective social phenomena, thus solving the apparent paradox. Moreover, as we discuss in detail in the *SI Text*, it constrains the dynamics so severely that its outcome is nearly insensitive to the details of the model considered. This result inverts the frequent expectation that initial conditions are less relevant than model specifications, and actually makes the latter less important (with the advantage of reducing the arbitrariness that, in principle, underlies the mathematical definition of any model). Quite unexpectedly, a range of more complicated scenarios [such as models with different interaction probabilities (2, 31, 38, 39), different network topologies (2, 31, 37, 42), coevolution of networks and opinions (43–47), cultural drift (40, 41, 48), higher-order interactions (2, 38, 49), and external sources of information (50)] lead to results that are essentially equivalent (or directly mapped) to what is obtained simply assuming that individuals are subject to dyadic interactions on a complete graph (see *SI Text*). We will therefore illustrate our results in this simple case, which corresponds to the assumption that social interactions are infinite-ranged, and only limited by the bounded confidence hypothesis.

Short-term Collective Social Behavior. We first study the effects of the empirical structure of real opinions on short-term collective social behavior. We consider a simple prototypic model where, on short timescales, cultural vectors do not evolve but nonetheless determine the choices that individuals make under the influence of each other. To this end, we extend the Cont-Bouchaud (CB) model (9), originally proposed to model herding effects in financial markets, to a more general "coordination model" which incorporates a dependence on real cultural vectors $\{\vec{v}_i\}$ (see *SI Text*). Individuals are assumed to make binary choices such as yes/no, buy/sell, approve/reject, etc. We can represent the choice expressed by the ith individual as $\phi_i = \pm 1$. The effects of mutual influence and bounded confidence are modeled by allowing pairs of individuals whose cultural distance d_{ij} is smaller than a threshold ω (which is the only parameter of the model) to exchange information before making their choices. As a result of this information exchange, all the agents belonging to the same connected component of the resulting ω-dependent cultural graph (see *SI Text*) collectively "agree" on the choice to make. If A labels a connected component of the graph, the choice of all agents belonging to A is the same ($\phi_i = \phi_A \ \forall \ i \in A$), while different connected components make statistically independent choices. We can imagine, even if this is not strictly necessary, that the process is repeated several times and with no memory: ϕ_i may first represent whether i liked or disliked an item, then whether i liked or disliked a different item, etc. The key property is that ϕ_i is short-lived; i.e., it represents any choice made over a timescale much shorter than that required to modify the cultural vector $\vec{v}_i$. It is a spurious variable representing action, not culture.

The overall outcome of a single run of the process (e.g., the net demand for a specific item) is the sum of individual preferences, and can be quantified by the average choice

$$\Phi = \frac{1}{N}\sum_{i=1}^{N}\phi_i = \frac{1}{N}\sum_{A} S_A\phi_A \quad [1]$$

where the second sum runs over all connected components, and S_A is the size of component A. The sign of Φ reflects the choice of the majority, and a key property characterizing the outcome of the model is the probability $P_\omega(\Phi)$ that the average choice takes the particular value Φ, for a given value of ω. Following the procedure described in the *SI Text*, we computed $P_\omega(\Phi)$ for various values of ω (from $\omega = 0$ to $\omega = 1$ in increments of 0.01) and for all the 13 groups in our dataset (both real and shuffled), plus the completely random set. In Fig. 2*A* we report the results for real Germany data. As can be seen, there exists a critical value ω_c (in the case shown, $\omega_c = 0.14 \pm 0.01$) such that, for $\omega < \omega_c$, $P_\omega(\Phi)$ is symmetric about zero (as for $\omega = 0$) and, for $\omega > \omega_c$, $P_\omega(\Phi)$ has two symmetric peaks. Right at $\omega = \omega_c$, $P_\omega(\Phi)$ displays a flattened region. This behavior is typical of symmetry-breaking phase transitions. Here the order parameter of the transition is the most probable value(s) $\Phi_\pm$ of Φ. This is shown in Fig. 2*B*, where we also plot the behavior for shuffled and random data. As we show in the *SI Text*, we found that the critical thresholds for shuffled data (which are equal for all groups) are always larger than those

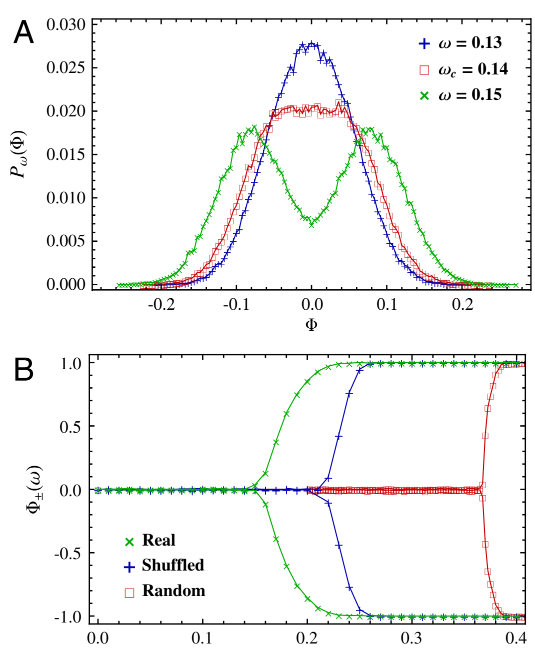

Fig. 2. At a critical confidence level, a spontaneous breaking of choice symmetry occurs. (*A*) When our "coordination model" is simulated on real data (in the example shown, the Germany group), we observe an abrupt change in the probability $P_\omega(\Phi)$ of a collective outcome at a critical confidence value ω_c. For $\omega < \omega_c$, individual choices are uncorrelated and sum up to a vanishing global outcome $\Phi = 0$, at which $P_\omega(\Phi)$ has a single peak. For $\omega > \omega_c$, local interactions result in global correlations that spread across the entire system, and a macroscopically coordinated output, whose probability is peaked about the two nonzero values $\Phi_\pm(\omega)$, emerges. Right at $\omega = \omega_c$, $P_\omega(\Phi)$ displays a flat region typical of critical phenomena. (*B*) The most probable value of Φ is the order parameter of the phase transition. For $\omega < \omega_c$ it is vanishing, while for $\omega > \omega_c$ it branches into the two symmetric values $\Phi_\pm(\omega)$. In addition to the results for real Germany data, here we also show the results for shuffled and random data. Real data always have a lower critical threshold than randomized data, indicating an enhanced possibility to behave collectively. All the other groups show the same behavior.

for real data (which are different across groups), and that the ones for random data are even larger. This shows that empirical interopinion correlations, which are responsible for the ultrametric distribution of individuals in cultural space, strongly facilitate collective social behavior by systematically lowering the resistance to coordination.

This simple model indicates that, depending on the local interaction range, different individual choices or actions can either be "averaged out" and disappear at the macroscopic level or give rise to a collectively coordinated behavior. Understanding this transition in real societies is one of the fundamental open questions of modern social science (1). In economics, this problem is related to whether it is legitimate to use the concept of "representative agent" as an idealized individual that makes the average choice of the society (16). Our simplified model, when simulated on real data, suggests that both regimes are possible, and that a simple local parameter can trigger very different global outcomes. In particular, the variance of the collective outcome, which grows with the separation between the peaks of $P_\omega(\Phi)$, can either decay to zero or be amplified macroscopically. These considerations indicate that a good measure of the level of collective behavior achievable for a given value of ω is the width of $P_\omega(\Phi)$. Thus we can define

$$C(\omega) \equiv \sigma_\omega(\Phi) = \sqrt{\sum_A \left(\frac{S_A}{N}\right)^2_\omega} \quad [2]$$

as a measure of short-term social *coordination*. The latter equality in the above formula (see *SI Text* for a rigorous proof) states that, intriguingly, $C(\omega)$ is uniquely determined by the sizes $\{S_A\}$ of the connected components of the underlying cultural graph obtained for that particular value of ω, and is therefore actually independent of the dynamical model considered. This quantity, which ranges between $C(0) = 0$ (no coordination) and $C(1) = 1$ (complete coordination), will be useful in what follows.

Long-term Cultural Diversity. We now take a dynamic perspective and focus on a longer temporal scale over which the cultural vectors themselves can change. In this case we use a modified version (38, 39) of Axelrod's model (31), which is designed to simulate the evolution of vectors of cultural traits on social networks, again in a way that real data can enter into the model. In an elementary time-step, two individuals i and j are randomly selected. If the normalized overlap $o_{ij} \equiv 1 - d_{ij}$ (where d_{ij} is the above-defined distance between the vectors $\vec{v}_i$ and $\vec{v}_j$) is smaller than or equal to θ, no interaction takes place. Otherwise, with probability equal to o_{ij}, the two individuals interact: a component $v_j^{(k)}$, chosen randomly among the components where $\vec{v}_i$ and $\vec{v}_j$ differ, is changed and set equal to i's corresponding component: $v_j^{(k)} = v_i^{(k)}$. Otherwise nothing happens, and two other individuals are selected. In the allowed final configurations, any two cultural vectors are either completely identical or separated by a distance larger than $\omega \equiv 1 - \theta$, and the average $\langle N_D\rangle_\omega$ (over many realizations) of the number N_D of distinct vectors in the final stage, or equivalently the fraction

$$D(\omega) \equiv \frac{\langle N_D\rangle_\omega}{N} \quad [3]$$

is a convenient way to measure the long-term cultural *diversity* as a function of ω. A more detailed description of the model can be found in the *SI Text*.

We ran several realizations of the model, by taking both real and randomized cultural vectors $\{\vec{v}_i\}$ as the starting configuration. As we show in the *SI Text*, we find that real data are those that achieve the largest level of long-term cultural heterogeneity. Indeed, for real data the realized value of $\langle N_D\rangle$ is the largest pos-

SOCIAL SCIENCES
APPLIED PHYSICAL SCIENCES

sible ($\langle N_D \rangle \approx N_C$) indicating that cultural convergence is confined within the initial connected components, each of which eventually becomes a single cultural domain. By contrast, in randomized data there are less final cultural domains than initial connected components, indicating that the latter often "merge" into larger cultural domains. The reason for the remarkably different behavior of real and randomized data is, once again, the ultrametric character of the former. As illustrated in Fig. 1, ultrametricity implies that the branches obtained cutting the real-data dendrogram at some value of ω will collapse into a single cultural vector. This means that the initial structure of the dendrogram above ω will be "frozen" and unaffected by cultural evolution. This confines cultural convergence locally within the lower branches. By contrast, in randomized data the lack of ultrametricity implies that branches are not well separated, so that the local convergence of vectors within a branch reduce the separation of the branch itself from nearby branches. Thus in this case branches are unstable, and often merge modifying the entire structure globally.

Combining the above findings with our previous results about collective behavior highlights the remarkable effects of ultrametricity. Given a group of individuals, we measured both the (initial) short-term social coordination $C(\omega)$ defined in eq. **2** and the (final) long-term cultural diversity $D(\omega)$ defined in Eq. **3** for various values of ω. Then we plotted $D(\omega)$ versus the value $C(\omega)$ obtained for the same ω, as in Fig. 3. If we look at random data, we retrieve the naive result that the coexistence of cultural heterogeneity and social collective behavior is impossible, since we have either $D \approx 0$ or $C \approx 0$. By contrast, real cultural vectors simultaneously allow high levels of both short-term coordination and long-term diversity, including the approximately balanced regime $C \approx D \approx 1/2$ which in the case shown is achieved for the moderate influence level $\omega \approx 0.17$. Shuffled data follow an intermediate curve, showing that both the heterogeneous frequencies of real opinions and the correlations among the latter play a significant role in enhancing the coexistence of diversity and coordination.

Note that, for real data, as the branches (below the threshold ω) of the dendrogram merge under the effect of cultural convergence, the cultural graph acquires more and more links until its initial connected components eventually become fully connected cliques. However, the number and sizes of such connected components (obtained for that value of ω) will remain unchanged[†]. Therefore, if we repeat the short-term coordination process at any frequency during the long-term cultural process, we will always obtain the same value of $C(\omega)$: the amplitude of collective social phenomena will not be decreased by cultural convergence. Seen from another perspective, this indicates that the persistence, or even an increased frequency, of collective behavior does not directly imply a reduction of cultural diversity. Thus, surprisingly, we find that empirical hierarchical correlations simultaneously enhance collective behavior and sustain cultural heterogeneity. The incompatibility of these two phenomena, which holds for random data, breaks down for real-data.

The above results were obtained assuming infinite-ranged social interactions. Online platforms bypassing offline social ties are modifying the traditional scenario and leading us closer to this extreme setting. Nonetheless, as we discuss in the *SI Text*, our results are surprisingly robust to changes in the dynamical rules of the model considered. This robustness suggests that, even if online interactions are tremendously expanding the substrate for short-term collective social behavior, they are not necessarily going to suppress long-term cultural diversity. It also indicates that, despite the fact that the structure of both offline (55–57) and online (28–30) social networks is intensively studied, our understanding of cultural dynamics is still severely limited by our little knowledge of the large-scale organization of real societies in multidimensional cultural space. These findings highlight the scarce predictive power of models that consider random specifications, and show the unique importance of empirical analyses of high-dimensional cultural vectors offering the possibility to explore cross-correlations among opinions and their consequences.

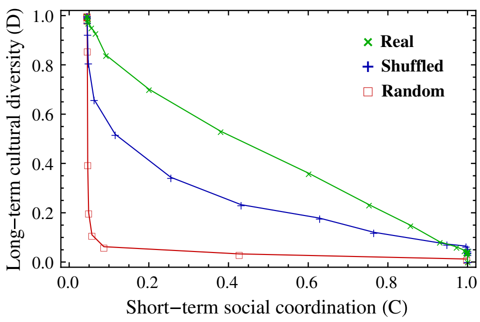

Fig. 3. Phase diagram summarizing our results. The long-term cultural diversity D is shown as a function of short-term social coordination C for real, shuffled and random data. If random cultural vectors are considered, cultural heterogeneity and collective behavior are mutually excluding: one has either $D \approx 0$ or $C \approx 0$. This approximately corresponds to the traditional situation explored when considering a random graph of interaction among individuals. By contrast, real cultural vectors simultaneously allow high levels of both short-term coordination and long-term diversity, including the approximately balanced regime $C \approx D \approx 1/2$. Shuffled data follow an intermediate curve, showing that both the heterogeneous frequencies of real opinions and the correlations among the latter play a significant role in enhancing the coexistence of diversity and coordination in the real world.

Conclusions

By using a large and detailed dataset, we have characterized the empirical properties of the large-scale distribution of individuals in multidimensional cultural space. We found that real interopinion correlations organize individuals hierarchically and ultrametrically in cultural space, a result which is not retrieved when randomized or one-dimensional opinions are considered. By using simple models, we showed that ultrametricity has profound and nontrivial consequences on short- and long-term cultural dynamics. In the short term, we found the existence of a symmetry-breaking phase transition where collective behavior arises out of purely local interactions. The critical threshold of this transition is remarkably lower in real data than in randomized cases, indicating that ultrametricity enhances short-term collective behavior. However, in the long term the same ultrametric property suppresses cultural convergence by restricting it within disjoint domains, implying a strong sensitivity to initial conditions. These opposite effects imply that, whereas in random data the coexistence of short-term coordination and long-term diversity is unfeasible, in real data it is strongly enhanced and can be achieved in a broad region of parameter space. Indeed, real data appear to optimize the coexistence of both phenomena. As another effect of ultrametricity, these results are surprisingly robust to changes in the details of the particular model considered: even more complicated theoretical scenarios would lead to a similar outcome. Thus the apparent paradox of the coexistence of short-term collective social behavior and long-term cultural diversity might have, as a simple and parsimonious explanation, the empirically observed hierarchical distribution of individuals in cultural space.

ACKNOWLEDGMENTS. All authors thank Maria I. Loffredo for her support and guidance. A.A. thanks all her colleagues who have contributed to the series

[†]Since we expect ultrametricity to be a robust property of real cultural traits also in other situations, it would be very interesting to test this prediction on real data reporting the evolution of culture. If we imagine to track the dynamics backwards in time, we also expect that the number and sizes of connected components will be preserved. Since social influence only modifies the lower branches of the dendrogram, the upper structure must be inherited from previous stages. This structure is exactly what we interpret as the residual cultural correlations that are entangled with social groups from the beginning of human history (see *SI Text*).

of Eurobarometer surveys on science and technology over the years. D.G. warmly thanks Felix Reed-Tsochas for stimulating discussions at the CABDyN Complexity Centre (University of Oxford) and acknowledges financial support through the European Commission FP7 Future and Emerging Technologies Open Scheme Project ICTeCollective (Contract 238597). This work was also supported by the Dutch Econophysics Foundation (Stichting Econophysics, Leiden, Netherlands) with funds from Duyfken Trading Knowledge BV, Amsterdam, Netherlands.

1. Buchanan M (2007) *The Social Atom* (Bloomsbury).
2. Castellano C, Fortunato S, Loreto V (2009) Statistical physics of social dynamics. *Rev Mod Phys* 81:591–646.
3. Lazer D, et al. (2009) Life in the network: The coming age of computational social science. *Science* 323:721–723.
4. Fortunato S, Castellano C (2007) Scaling and universality in proportional elections. *Phys Rev Lett* 99:138701.
5. Ratkiewicz J, Fortunato S, Flammini A, Menczer F, Vespignani A (2010) Characterizing and modeling the dynamics of online popularity. *Phys Rev Lett* 105:158701.
6. Onnela J-P, Reed-Tsochas F (2010) Spontaneous emergence of social influence in online systems. *Proc Natl Acad Sci USA* 107:18375–18380.
7. Crane R, Sornette D (2008) Robust dynamic classes revealed by measuring the response function of a social system. *Proc Natl Acad Sci USA* 105:15649–15653.
8. Mantegna R, Stanley HE (2000) *An Introduction to Econophysics: Correlations and Complexity in Finance* (Cambridge University Press, Cambridge, U.K.).
9. Cont R, Bouchaud J-P (2000) Herd behavior and aggregate fluctuations in financial markets. *Macroecon Dyn* 4:170–196.
10. Sinha S, Chatterjee A, Chakraborti A, Chakrabarti B (2010) *Econophysics: An introduction* (Wiley-VCH Verlag GmbH, New York).
11. Chakrabarti B, Chakraborti A, Chatterjee A (2006) *Econophysics and Sociophysics: Trends and Perspectives* (Wiley-VCH Verlag GmbH, New York).
12. Galam S (2007) From 2000 Bush-Gore to 2006 Italian elections: Voting at fifty-fifty and the contrarian effect. *Qual Quant* 41:579–589.
13. Onnela J-P, et al. (2007) Structure and tie strengths in mobile communication networks. *Proc Natl Acad Sci USA* 104:7332–7336.
14. Bohannon J (2006) Tracking people's electronic footprints. *Science* 314:914–916.
15. Fowler J, Christakis N (2010) Cooperative behavior cascades in human social networks. *Proc Natl Acad Sci USA* 107:5334–5338.
16. Kirman AP (1992) Whom or what does the representative individual represent? *J Econ Perspect* 6:117–136.
17. Greig JM (2002) The end of geography? Globalization, communications, and culture in the international system. *J Conflict Resolut* 46:225–243.
18. Brady D, Beckfield J, Zhao W (2007) The consequences of economic globalization for affluent democracies. *Annu Rev Sociol* 33:15. 1–15.22.
19. Rogers EM (1971) *Communication of Innovations: A Cross-Cultural Approach* (Free Press, New York).
20. Rogers EM (1995) *Diffusion of Innovations* (Free Press, New York).
21. Blossfeld HP, Klijzing E, Mills M, Kurz K (2005) *Globalization, Uncertainty and Youth in Society* (Routledge, New York).
22. McNeill W (1988) Diffusion in history. *The Transfer and Transformation of Ideas and Material Culture*, eds Peter J Hugill and D. Bruce Dickson (Texas A&M Press, College Station), pp 75–90.
23. Foster R (1991) Making national cultures in the global ecumene. *Annu Revi Anthropol* 20:235–260.
24. Elkins D (1997) Globalization, telecommunications, and virtual ethnic communities. *Int Polit Sci Rev* 18:139–52.
25. Barth Fredrik (1969) *Ethnic and Group Boundaries* (Waveland, Long Grove, IL), Reprint 1998.
26. Durkheim E (1912) *The Elementary Forms of Religious Life* (Oxford University Press, New York), Reprint 2001.
27. Boyd R, Richerson PJ (2005) *The Origin and Evolution of Cultures* (Oxford University Press, New York).
28. Backstrom L, Huttenlocher D, Lan X, Kleinberg J (2006) Group formation in large social networks: Membership, growth, and evolution. *Proceedings of the 12th ACM SIGKDD International Conference on Knowledge Discovery in Data* (ACM Press, New York), pp 44–54.
29. Szell M, Thurner S (2010) Measuring social dynamics in a massive multiplayer online game. *Soc Networks* 32:313–329.
30. Szell M, Lambiotte R, Thurner S (2010) Multirelational organization of large-scale social networks in an online world. *Proc Natl Acad Sci USA* 107:13636–13641.
31. Axelrod R (1997) The dissemination of culture. *J Conflict Resolut* 41:203–226.
32. Festinger L, Schachter S, Back K (1963) *Social Pressures in Informal Groups: a Study of Human Factors in Housing* (Stanford University Press, Palo Alto, CA).
33. Homans GC (1962) *Sentiments and Activities* (Free Press, New York).
34. McPherson JM, Smith-Lovin L (1987) Homophily in voluntary organizations: Status distance and the composition of face-to-face groups. *Am Sociol Rev* 52:370–79.
35. McPherson JM, Smith-Lovin L, Cook J (2001) Birds of a feather: Homophily in social networks. *Annu Rev Sociol* 27:415–44.
36. Sherif M, Hovland CI (1961) *Assimilation and Contrast Effects in Communication and Attitude Change* (Yale University Press, New Haven, CT).
37. Weisbuch G (2004) Bounded confidence and social networks. *Eur Phys J B* 38:339–343.
38. Flache A, Macy M (2007) Local Convergence and Global Diversity: The Robustness of Cultural Homophily., Arxiv preprint physics/0701333.
39. De Sanctis L, Galla T (2009) Effects of noise and confidence thresholds in nominal and metric Axelrod dynamics of social influence. *Phys Rev E* 79:046108.
40. Klemm K, Eguíluz VM, Toral R, San Miguel M (2003) Global culture: A noise-induced transition in finite systems. *Phys Rev E* 67:045101(R).
41. Klemm K, Eguíluz VM, Toral R, San Miguel M (2005) Globalization, polarization and cultural drift. *J Econ Dyn Control* 29:321–334.
42. Guerra B, Poncela J, Gómez-Gardenes J, Latora V, Moreno Y (2010) Dynamical organization towards consensus in the Axelrod model on complex networks. *Phys Rev E* 81:056105.
43. Centola D, González-Avella JC, Eguíluz VM, San Miguel M (2007) Homophily, cultural drift and the co-evolution of cultural groups. *J Conflict Resolut* 51:905–929.
44. Zanette DH, Gil S (2006) Opinion spreading and agent segregation on evolving networks. *Physica D* 224:156–165.
45. Santos FC, Pacheco JM, Lenaerts T (2006) Cooperation prevails when individuals adjust their social ties. *PLoS Comput Biol* 2:e140.
46. Garlaschelli D, Capocci A, Caldarelli G (2007) Self-organized network evolution coupled to extremal dynamics. *Nat Phys* 3:813–817.
47. Kozma B, Barrat A (2008) Consensus formation on adaptive networks. *Phys Rev E* 77:016102.
48. Mas M, Flache A, Helbing D (2010) Individualization as driving force of clustering phenomena in humans. *PLoS Comput Biol* 6(10):e1000959.
49. Centola D, Macy M (2007) Complex contagions and the weakness of long ties. *Am J Sociol* 113:702–734.
50. González-Avella JC, Cosenza MG, Tucci K (2005) Nonequilibrium transition induced by mass media in a model for social influence. *Phys Rev E* 72:065102(R).
51. INRA, Brussels (Europe) and Report International for the European Commission, Directorate-General XIIScience Research and Development. *Europeans, Science and Technology—Public Understanding and Attitudes*, Available at http://eceuropa.eu/public_opinion/archives/ebs/ebs_076_en.pdf.
52. Gaskell G, et al. (2000) Biotechnology and the European public. *Nat Biotechnol* 18:935–938.
53. Gaskell G, et al. (2011) The 2010 Eurobarometer on the life sciences. *Nat Biotechnol* 29:113–114.
54. Rammal R, Toulouse G, Virasoro M (1986) Ultrametricity for physicists. *Rev Mod Phys* 58:765–788.
55. Granovetter M (1973) The strength of weak ties. *Am J Sociol* 78:1360–80.
56. Wasserman S, Faust K (1994) *Social Network Analysis: Methods and Applications* (Cambridge University Press, Cambridge, U.K.).
57. Caldarelli G (2007) *Scale-Free Networks: Complex Webs in Nature and Technology* (Oxford University Press, New York).