Demography-based adaptive network model reproduces the spatial organization of human linguistic groups

José A. Capitán[1,*] and Susanna Manrubia[2]

[1]*Departamento de Matemática Aplicada, Universidad Politécnica de Madrid, Av. Juan de Herrera 4, 28040 Madrid, Spain*
[2]*Centro Nacional de Biotecnología (CSIC), C/ Darwin 3, 28049 Madrid, Spain*

The distribution of human linguistic groups presents a number of interesting and nontrivial patterns. The distributions of the number of speakers per language and the area each group covers follow log-normal distributions, while population and area fulfill an allometric relationship. The topology of networks of spatial contacts between different linguistic groups has been recently characterized, showing atypical properties of the degree distribution and clustering, among others. Human demography, spatial conflicts, and the construction of networks of contacts between linguistic groups are mutually dependent processes. Here we introduce an adaptive network model that takes all of them into account and successfully reproduces, using only four model parameters, not only those features of linguistic groups already described in the literature, but also correlations between demographic and topological properties uncovered in this work. Besides their relevance when modeling and understanding processes related to human biogeography, our adaptive network model admits a number of generalizations that broaden its scope and make it suitable to represent interactions between agents based on population dynamics and competition for space.

PACS number(s): 89.75.Hc, 87.23.Kg, 87.10.−e

I. INTRODUCTION

Adaptive networks, where the dynamics of nodes is coupled to the dynamics of network links, have received considerable attention in the last decade [1]. This kind of network represents not only a natural extension of models where either dynamics on complex networks or the origin of nontrivial topology of networks itself had been the focus of attention, but is in its own right a field of interest. Indeed, in many natural systems node dynamics and network dynamics are intimately coupled, and their interplay captures important aspects that would be missed if both processes are not taken simultaneously into account. Examples are, among many others [2], neural networks, where neuron activity affects synaptic strength [3], ecological networks, where population dynamics is coupled to food web structure [4], catalytic networks, where the appearance of autocatalytic sets formed by sufficiently abundant chemical species is essential for the maintenance of the system [5] and where the explicit introduction of space leads to segregation of parasitic species that may otherwise disrupt network structure [6], or generic models, where distinct populations of nodes separate when connection strength is allowed to vary [7].

The coupling between node and link dynamics is especially relevant in social networks, where nodes are individuals, companies, human groups, or countries, e.g., and links represent social contacts of various kinds. In many of these networks, agents can actively change their interactions, thus causing a systematic modification of network topology. A well-studied case is that of epidemics, where susceptible individuals may suppress their links with infected neighbors, leading to networks assortative in degree and to first-order transitions between healthy and endemic states [8] and even to infection suppression [9]. Similar situations hold in socioeconomic contexts, where adaptive networks display an interesting phenomenology that includes phase transitions and hysteresis between dissimilar states of agents [10] and self-organization leading to broad wealth distributions [11]. In a broader scenario, it has been shown that changes in the state of nodes coupled to rewiring of links systematically causes network fragmentation [12]. This fact seems to be enhanced by the spatial embedding of many social networks, which constraints interactions [13] and may induce the spatial separation of different socioeconomic classes [14].

In this work we address the relationship between the demographic dynamics of human linguistic groups in the last millennium and the topology of their networks of spatial contacts. We present a model for an adaptive spatial network where neighboring relationships are determined by the growth of groups and their concomitant attempt to modify the total area they occupy. The model is based on two previous and independent observations regarding the organization of human groups. In Ref. [15], a mean-field model was introduced to reproduce the population-area relationship observed in human languages; the spatial structure of groups did not play any essential role in explaining that observation, and thus was disregarded. As a consequence, that model cannot describe the complex topology of the network of contacts that was later uncovered [16]. Networks of contacts between linguistic groups reflect their spatial embedding and display a set of properties previously unseen in spatial networks. Among others, those networks have high intervality, a property shared with food webs [17,18]. Inspired by this latter system, and in niche models, which had successfully captured that property, a nichelike algorithm was proposed and shown to recover most topological features of language networks [16].

Human demographic dynamics and the spread of populations on space, which determines their intergroup contacts, are two coupled processes. As we report in this paper, their mutual dependence is behind observed correlations between

---

*[*]ja.capitan@upm.es

the population of a group and the number of spatial neighbors it has, and is needed to explain the appearance of assortative properties in empirical language networks. Results from an adaptive network model that we here introduce are compared with several world regions and to the topology of the network of linguistic contacts in each of them. The paper is structured as follows. In Sec. II we introduce the adaptive model for language networks. Previous relevant results are summarized for the sake of completeness and clarity. Section III reviews data on human linguistic groups as used in this study, particularly emphasizing the meaning of model parameters. In Sec. IV, we fit the model to empirical data and show how the adaptive model qualitatively and quantitatively, in most cases, reproduces the population-area relationship, the degree and shortest-path distributions of empirical language networks (in addition to other topological features), and nontrivial correlations between demography and topology. The paper finishes with an overall discussion and some proposals for model extensions and future research.

II. ADAPTIVE MODEL FOR LANGUAGE NETWORKS

The adaptive network model yields a dynamic network of interactions among groups arising from explicit demographic dynamics and competition for space. As the size of groups varies, their neighboring relationships are modified and possible conflicts with different groups sharing boundaries may ensue. The precise example used is the development of human linguistic groups in the last thousand years. First, we define demographic dynamics following current knowledge on the world population growth and suitable rules for intergroup contacts and conflicts. Second, the network of contacts between groups is updated in the light of changes in the areas they occupy.

A. Demography and conflict rules

The modeling of demographic dynamics is based on Ref. [15]. Dynamics relies on a stochastic multiplicative process of the form $P_i(t + i) = \alpha_i(t)P_i(t)$ for the size of each population P_i, where the distribution of α_i values is estimated from empirical data. This process describes the growth of linguistic groups [19] and reproduces the observation of a log-normal distribution of the number of speakers per language [20]. Subsequently, demographic changes are coupled to variations in the area over which groups are spread. That model was devised with the goal of explaining the population-area (P-A) relationship observed in human linguistic groups, which follows $A \propto P^z$ [15].

Relevant model parameters have been derived from world population estimations, as follows. There were about $P_0 = 3.1 \times 10^8$ humans in year 1000 [21], while in year 2000 the world population reached $P_T = 5.7 \times 10^9$ [22]. Assuming an exponential growth in the last ten centuries, an average annual growth rate $\alpha \simeq 1.0029$ is obtained, and a dispersion $\sigma_\alpha = 0.096$ can be associated to the process [19,23]. The simplest distribution for the stochastic growth rate α_i is a uniform distribution of average α and mean-square dispersion σ_α [15,19]. A constant number of languages in this time interval, equal to the current estimated linguistic diversity (6900 languages) [22], is

considered. Though some languages may have appeared in the last millennium, and many others have disappeared, in this model we disregard language birth or extinction for the sake of simplicity. In a previous model that constitutes the basis for the demographic dynamics here implemented, it has been numerically shown that those two processes did not affect the statistical results [19]. As initial condition, we take uniform populations $[P_i(0) = 3.1 \times 10^8/6900]$, and areas $A_i = 1$, in arbitrary units. Numerical simulations show that changes in the initial condition do not affect in a significant way the final distribution of group sizes (see also Ref. [23]). Dynamics are run for 1000 time steps to compare with current available data. In the scenario described, population dynamics are defined so as to agree with empirical observations. Therefore, parameters α and σ_α implicitly contain information on all processes that may have potentially affected demographic changes in the last millennium (that is births and deaths, but also casualties due to wars or pandemic diseases, for example). This is also the reason to couple in a directed fashion population dynamics to areas. Notice that the units of area remain undetermined to a multiplicative factor.

The log-normal distributions of language sizes P_i [19,20] and areas A_i [15] imply that the log-transformed variables $p_i = \ln P_i$ and $a_i = \ln A_i$ for each linguistic group i follow Gaussian distributions. As a result, the stochastic multiplicative process in the original variables can be cast in the form of a stochastic additive process in p_i and a_i [15]. The logarithmic number of individuals in a group therefore follows

$$p_i(t + 1) = p_i(t) + \beta_i(t), \tag{1}$$

where t is measured in years, and β_i is randomly drawn at each time step from a uniform distribution $\Pi(\beta; \epsilon, \eta)$ in the interval $(\epsilon - \eta, \epsilon + \eta)$,

$$\Pi(\beta; \epsilon, \eta) = \frac{1}{2\eta}\{\Theta[\beta - (\epsilon - \eta)] - \Theta[(\epsilon + \eta) - \beta]\}, \tag{2}$$

where $\Theta(x)$ is the Heaviside step function, and mean value $\epsilon = -0.00186$ and half width $\eta = 0.169$ are obtained when the original multiplicative process is mapped to an additive one [15,24]. Similarly, the logarithmic area a_i is assumed to obey

$$a_i(t + 1) = a_i(t) + \xi_i(t). \tag{3}$$

The evolution of p_i and a_i is coupled following two rules:

(i) The area covered by a group shrinks when its population decreases: if $\beta_i(t) < 0$, then $\xi_i(t)$ is randomly drawn from a uniform interval $[-r|\beta_i(t)|, 0]$.

(ii) Increases in the population size lead to conflict between group i and one of its neighbors on the network of contacts between groups (see below). A neighbor j of node i is chosen at random; if its growth rate is smaller than that of i, the area of i grows, and vice versa. Specifically:

(a) If $\beta_j(t) < \beta_i(t)$, $\xi_i(t)$ is drawn from $[0, w\beta_i(t)]$;

(b) If $\beta_j(t) \geqslant \beta_i(t)$, $\xi_i(t)$ is drawn from $[-w\beta_i(t), 0]$.

The spontaneous retreat parameter r measures to what extent log-areas spontaneously shrink when populations decrease. The outcome of conflicts is weighted through w, which determines the associated benefit for the population with the faster growth and is, in general, different from r. Actually, mean-field fits to actual values of p_i and a_i at present have

revealed a sublinear relationship between population decrease and area reduction and a larger increase in areas as a result of conflicts, yielding $w > 1 > r$ for the six world regions analyzed in Ref. [15]. It remains to be seen whether this constraint remains in other world regions here analyzed, and whether the values of w and r in the current adaptive network model significantly deviate from mean-field results.

B. Network dynamics

Space is effectively introduced in the form of a network of neighbors that coevolves with the demographic dynamics just described. Although networks of contacts between linguistic groups are embedded in a two-dimensional space, their clustering coefficient is close to that of one-dimensional regular networks [16]. In addition, they exhibit a significant degree of intervality [16], which implies that a few fundamental variables can account for their structure. Thus, most of the structural properties of language networks are captured by using a single variable in a one-dimensional space. Therefore, the construction of the network is inspired in a static algorithm that used a given distribution of areas and contained the rules to construct a network of contacts between groups [16]. Now, instead, the network is continuously updated taking into consideration the area associated to each node, as obtained in the previous step. Neighboring relationships between nodes arise from an assumption on perimeter contact. Based on geometric constraints, it can be assumed that the perimeter of node i is comparable to the sum of perimeters of its potential neighbors up to a multiplicative factor,

$$A_i^{1/2} \simeq f_i \sum_{j \in \mathrm{nn}(i)} A_j^{1/2}, \tag{4}$$

where the perimeter overlap $f_i > 0$ measures the average fraction of perimeter of each neighbor that is shared with node i. In general, f_i—as defined in Eq. (4)—is a node-dependent quantity, but for simplicity we assume an effective value all across the network, such that f_i will be substituted by its network average $f = N^{-1} \sum_i f_i$, where N is the number of nodes (languages) in the network.

The network of contacts is generated in two steps:

(i) *Directed network generation.* Given the (arbitrarily ordered) set of areas $\{A_1(t), A_2(t), \ldots, A_N(t)\}$ at time step t, where $A_i(t) = e^{a_i(t)}$, we draw directed links between each node i and nodes at positions $i \pm 1, i \pm 2, \ldots$, until the upper bound of the right-hand side of Eq. (4) is first exceeded. The first neighbor j_0 is either $i + 1$ or $i - 1$ with equal probability, and subsequent nodes are chosen following the rules

$$j_{2n+1} = \begin{cases} i - n - 1, & j_0 = i + 1, \\ i + n + 1, & j_0 = i - 1, \end{cases} \tag{5}$$

and

$$j_{2n} = \begin{cases} i + n + 1, & j_0 = i + 1, \\ i - n - 1, & j_0 = i - 1, \end{cases} \tag{6}$$

for $n = 0, 1, 2, \ldots$. Periodic boundary conditions have been assumed when $j_n < 0$ or $j_n > N$.

(ii) *Transformation to an undirected network.* Since spatial neighboring relationships are undirected, the previous network should be transformed to an undirected one. Links may be added or removed so as to guarantee the symmetry of the adjacency matrix. For this purpose we introduce a symmetrization parameter $0 \leqslant q \leqslant 1$. If a directed link $i \to j$ does not have a reverse counterpart, $j \nrightarrow i$, we draw a uniformly distributed random number x in $(0,1)$ and add the missing link $j \to i$ to the network if $x < q$. If $x \geqslant q$, the original link $i \to j$ is removed. In any case, the relationship between i and j has been symmetrized after the process. Note that this process affects neighboring relationships as defined in Eq. (4), so it will be important to assess its eventual effect in the demographic and topological properties we aim at reproducing.

The use of a one-dimensional array of areas to construct the network is analogous to the procedure used in ecological niche models, where a single variable suffices to reproduce most topological properties of food webs and where the explicit consideration of population dynamics is not essential. In the case of networks of contacts between linguistic groups, their local structure was shown to be equivalent to that of almost regular, one-dimensional networks, with the area playing the role of the niche variable [16].

Figure 1 illustrates some important properties of the model just described. Figure 1(a) exemplifies the dynamics of logarithmic areas a_i and populations p_i, as well as the number of neighbors of group i—its degree k_i—for 500 years. At the end of the simulation [Fig. 1(b)], for $t = 1000$, the degree distribution $p(k)$ is calculated. It presents a well-defined average value and a significant tail to large k values. Finally, in Fig. 1(c) we illustrate one of the quantities that could not
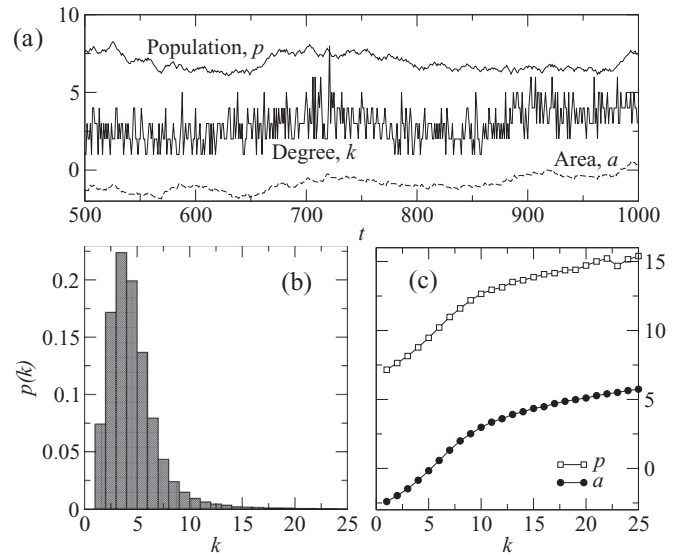

FIG. 1. Model dynamics. (a) Time series corresponding to a realization of the model showing the dynamics of the natural logarithm of area and population size for a single linguistic group, and the number of neighboring languages it has. (b) Degree distribution obtained at the end of the realization for a system with 1000 interacting groups. (c) Relationship between the number of neighbors k and the total area or population for the same ensemble. Parameters are $r = 1$, $w = 1.5$, $f = 0.2$, and $q = 0.3$, and averages over 100 and 500 independent realizations have been performed in (b) and (c), respectively.

be reproduced by models based only on demography [15] or on a nichelike algorithm to construct the network [16], namely, the relationship between population or area and degree of a linguistic group. Among others, these quantities will be compared to those measured in current linguistic groups with the purpose of establishing whether the dynamical niche model is able to reproduce the observations and of determining the value of the model parameters that best fits the latter. Summarizing, the dynamical niche model is characterized by four parameters: the spontaneous retreat r, the outcome of conflicts w, the average perimeter overlap f, and the symmetrization parameter q.

III. DATA ON HUMAN LINGUISTIC GROUPS

Data on linguistic groups stems from a collection by SIL International (http://www.ethnologue.com/) and a map developed by Global Mapping International (World Language Mapping System, http://www.gmi.org/wlms/index.htm). A detailed description of the database appears in the Ethnologue [22], from which information on 6900 extant languages, including their spatial distribution and number of speakers, can be found. Each language is characterized by a centroid, which is a point in latitude-longitude coordinates that represents its average location. Centroids are the nodes of linguistic groups. Two nodes are linked if the groups they represent share spatial borders in any of the domains where a language is spoken (note that the speakers of a language may occupy disconnected domains, a situation that is relatively frequent). The interested reader can find details on network construction in Ref. [16].

In the present study, we are not taking into account languages that are widespread as a result of colonization. Languages such as English, Spanish, or Portuguese in the Americas, or Mandarin Chinese in Asia, are in several senses outsiders: they percolate across continental regions and act as hubs in networks of contacts between linguistic groups, enhancing the formation of large connected components in language networks [16]. The number of neighbors of widespread languages (that is, their degree k_i) is severalfold higher than that of other languages in the same network, thus significantly deviating from the bulk degree distribution. In this sense, widespread languages can be considered the "dragon kings" of languages [25], and the dynamical processes that underlie their spread are different from the basic demographic dynamics implemented in our model. Widespread languages constitute a small fraction of world languages. The 50 largest languages (with 24×10^6 or more speakers) represent only about 0.7% of the data points here considered. The elimination of widespread languages causes the fragmentation of otherwise connected networks in some world regions, remarkably in continental North America. This effect is not seen if, for instance, the largest languages are eliminated in the network corresponding to continental Africa, whose largest connected component remains essentially unchanged.

The main properties of 12 networks of linguistic groups selected for the current study are reported in Table I. They correspond to five continental regions (Africa, Asia, Europe, and North and South America), though in the case of North America no large connected component can be identified:

TABLE I. Largest connected components of networks of contacts between linguistic groups obtained for each continent. The number of nodes N and the number of links L are shown. Relevant quantities that the model intends to reproduce are the exponent z, the correlation ρ, and the average perimeter overlap f. The deviation of the distribution of f_i values in each network is σ_f. CC: Connected component; C Africa: continental Africa; C Asia: continental Asia; New Guinea, Sulawesi, and Luzon are islands. C Europe: continental Europe; Mex1: Mexico (1); Yucatan: Yucatan peninsula; Mex2: Mexico (2); CS America: continental South America; ABP: ABP borders.

CC	N	L	z	ρ	f	σ_f
C Africa	2126	6154	0.87	0.63	0.13	0.11
C Asia	1370	3967	0.65	0.72	0.10	0.10
New Guinea	663	1543	0.62	0.42	0.22	0.16
Australia	99	176	0.72	0.52	0.28	0.15
Sulawesi	64	121	0.66	0.77	0.25	0.19
Luzon	56	140	0.44	0.79	0.13	0.08
C Europe	231	547	0.60	0.65	0.13	0.31
Mex1	68	120	0.82	0.59	0.32	0.24
Yucatan	50	111	1.22	0.61	0.27	0.57
Mex2	39	71	0.67	0.73	0.32	0.24
CS America	234	399	0.40	0.56	0.28	0.17
ABP	33	59	0.66	0.76	0.37	0.42

the three largest networks are found in or around Mexico, and named Mex1, Mex2, and Yucatan. In addition, we study the networks of Australia, New Guinea, Sulawesi and Luzon islands, as well as an additional small network found in the borders shared by Argentina, Bolivia, and Paraguay (ABP).

An example of some model quantities and empirical properties of the continental Africa network are represented in Fig. 2. In Fig. 2(a) a part of the whole network is shown, emphasizing the area A_i of a given linguistic domain i and its neighboring relationships. As can be seen, the perimeter overlap depends on each pair of groups in contact. In this example, language i shares boundaries with eight different languages, so it has a degree $k_i = 8$. In practice, f_i is calculated from its definition, $f_i \simeq A_i^{1/2} / \sum_{j \in \mathrm{nn}(i)} A_j^{1/2}$ for each node, and then averaged over the whole network to obtain the value f reported in Table I. For completeness, the last column of Table I summarizes, for each network in the data set, the standard deviation σ_f of the distribution of f_i values.

A. Population-area relationship

The relationship between the logarithm of the size of a linguistic group and the logarithm of the area over which its speakers are spread follows an allometric relationship that has a counterpart in ecology, where the abundance of a species and its home range are similarly related [26,27]. It has been shown that area a and population p fulfill $a = zp + c$ for the whole world, where c is a constant, for six large continental regions, and also for groups of hunter gatherers [15]. Since language sizes and areas follow log-normal distributions, the
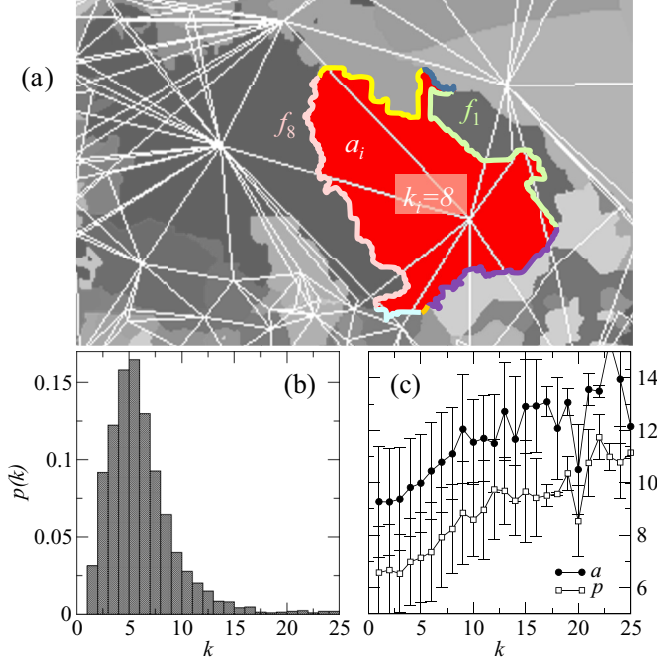
FIG. 2. (Color online) Properties of African largest connected component, with 2126 nodes. (a) Detail of the local structure of contacts between groups. Gray regions represent different domains where a language is spoken; one language might be spoken in disconnected domains. The centroid corresponds to the whole of a language, and therefore might fall even outside a particular domain; (b) Degree distribution; (c) Dependence of the logarithmic area a and population p of linguistic groups on the number of neighbors k in the network of spatial contacts. Error bars stand for the standard deviation of a and p values at each fixed k.

transformed logarithmic variables a and p are well fitted by Gaussian distributions, and their joint distribution can be approximated by a bivariate normal distribution. This joint distribution is characterized by z, which is the slope of the major ellipse axis of the scatter plot containing all (p_i, a_i) pairs, and a coefficient ρ that quantifies how correlated a and p are.

Let us define, for each network, the average logarithmic area $\langle a \rangle$ and the average logarithmic population $\langle p \rangle$,

$$\langle a \rangle = N^{-1} \sum_i a_i, \quad \langle p \rangle = N^{-1} \sum_i p_i, \qquad (7)$$

and the corresponding standard deviations

$$\sigma_a^2 = N^{-2} \sum_i (a_i - \langle a \rangle)^2, \quad \sigma_p^2 = N^{-2} \sum_i (p_i - \langle p \rangle)^2. \quad (8)$$

The covariance matrix C of a and p is

$$\mathsf{C} = \begin{pmatrix} \sigma_a^2 & \rho \sigma_a \sigma_p \\ \rho \sigma_a \sigma_p & \sigma_p^2 \end{pmatrix} \qquad (9)$$

with $\rho \sigma_a \sigma_p = N^{-2} \sum_i (a_i - \langle a \rangle)(p_i - \langle p \rangle)$. The eigenvectors of matrix C can be written as $(1, z)$, $(-z, 1)$, where z corresponds to the exponent relating both quantities. The value of ρ determines the degree of correlation between a and p: The larger ρ, the more correlated the two variables are. Values

of z and ρ obtained through this procedure for the networks here analyzed are reported in Table I. The interested reader can find example plots of this relationship for empirical data in Ref. [15].

B. Topological properties

Networks of contacts between linguistic groups hold a number of nontrivial topological properties [16]. They are an example of quasi-interval graphs, a property they share with food webs [17,18,28]. The dependence between the clustering coefficient and the linkage density $2L/N$ reveals that language networks are akin to one-dimensional regular networks at the local level [16]. Together with intervality, this property supports the existence of a configuration space of low dimensionality, and partly explains the success of a nichelike algorithm to account for several of the topological properties of language networks. Two additional properties that we will analyze and compare with model results are the shortest-path length and the degree distributions.

A representative example of the degree distribution $p(k)$ is shown in Fig. 2(b). Most language networks analyzed so far present degree distributions compatible with log-normal functions [16]. In this work, one of our goals is to find out how likely it is that the adaptive network model generates degree distributions compatible with observations. The same applies to the distribution of shortest-path lengths $p(d)$. The latter have a complex shape that depends on the particular network, as will be shown. We have chosen these two distributions because of their presumable relevance regarding intergroup dynamics. For example, the degree distribution is related to the likelihood of entering into conflict with different linguistic (or cultural) groups as a result of shared boundaries, but may also affect linguistic evolution due to frequent contacts with dissimilar languages. The shortest-path length distribution may play a role in the dissemination of cultural innovations, under the reasonable assumption that intragroup spread of novelties is significantly faster than intergroup spread: the fewer intermediates, the faster the propagation.

C. Dependence between demographic and topological variables

Demographic and topological features of language networks are not independent. For instance, Fig. 2(c) illustrates the empirical dependence of the logarithmic area a and population p on the degree k. This relation is qualitatively similar to the dependence yielded by the adaptive network model, see Fig. 1. In forthcoming sections we will make this relation quantitative by optimizing the values of model parameters to fit empirical observations. There is a final observation as yet unexplained, which is the appearance of population-population, area-area, and degree-degree correlations between neighboring nodes in empirical networks (see below).

IV. MODEL PARAMETERS: FITS TO EMPIRICAL DATA

With the aim of quantitatively reproducing demographic and topological features of language groups and networks, we analyze which values of the model parameters r, w, f, q best fit each of the 12 empirical networks considered. Specific

TABLE II. Optimal mean-field parameters (obtained with the model in Ref. [15]) and adaptive network model parameters r, w for $f^\star = 0.1$ and $q^\star = 0.1$.

CC	r_{MF}	w_{MF}	r	w
C Africa	0.651	2.23	0.72 ± 0.02	2.31 ± 0.02
C Asia	0.951	1.54	1.00 ± 0.01	1.56 ± 0.01
Australia	0.125	2.20	0.16 ± 0.05	2.27 ± 0.06
Sulawesi	1.27	1.25	1.32 ± 0.07	1.23 ± 0.08
Luzon	1.04	0.735	1.07 ± 0.14	0.70 ± 0.16
C Europe	0.603	1.69	0.64 ± 0.02	1.71 ± 0.02
Mex1	0.422	2.24	0.48 ± 0.07	2.30 ± 0.07
Yucatan	0.695	2.79	0.84 ± 0.12	2.92 ± 0.13
Mex2	1.16	1.16	1.11 ± 0.08	1.49 ± 0.09
CS America	0.195	1.47	0.20 ± 0.01	1.48 ± 0.02
ABP	1.20	1.33	1.27 ± 0.10	1.28 ± 0.11

goals are to reproduce the empirical parameters z and ρ characterizing the (logarithmic) population-area relationship, and two topological features: the degree distribution and the distribution of shortest-path lengths. We will finally evaluate how the dynamical model with the so-obtained parameters reproduces the relationship between demographic and topological variables, as well as the appearance of correlations in node properties (area, population, and degree).

Quantitative values of z and ρ obtained with the adaptive network model fixing values of $f^\star$ and $q^\star$ do not differ substantially from those obtained in the mean-field approximation used in Ref. [15], see Table II. The mean-field coupled dynamics of growth and conflict cannot be fitted to New Guinea island, contrary to what was observed in Ref. [15]. Such discrepancy is due to the fact that we are only considering here connected networks, disregarding isolated languages that were taken into account in Ref. [15] when calculating the empirical values of z and ρ. Additionally, in that reference the mean-field dynamics could not be fitted to the pooled set of North American languages, for which the correlation value was sensibly smaller than the rest ($\rho = 0.40$). Here the same phenomenon occurs for New Guinean connected languages ($\rho = 0.42$, see Table I). Consequently, the adaptive network model cannot be fitted to New Guinea network; hence, from now on, we will reduce our analysis to the remaining 11 networks.

The relative difference between the surfaces $z(r,w,f^\star,q^\star)$ and $\rho(r,w,f^\star,q^\star)$ and their mean-field counterparts $z_{\mathrm{MF}}(r,w)$ and $\rho_{\mathrm{MF}}(r,w)$ has been measured as

$$n_z = \max_{i,j} \left| \frac{z(r_i,w_j,f^\star,q^\star) - z_{\mathrm{MF}}(r_i,w_j)}{z_{\mathrm{MF}}(r_i,w_j)} \right| \quad (10)$$

and

$$n_\rho = \max_{i,j} \left| \frac{\rho(r_i,w_j,f^\star,q^\star) - \rho_{\mathrm{MF}}(r_i,w_j)}{\rho_{\mathrm{MF}}(r_i,w_j)} \right| \quad (11)$$

for discretizations (r_i,w_j) of the (r,w) parameter subspace, with $r_i \in (0,1.5)$, $w_j \in (0.5,3)$, $f^\star = 0.1$ and $q^\star = 0.1$ [29]. Comparison with mean-field values yields $n_z = 0.12$ and $n_\rho = 0.019$, which implies that maximum relative differences are around 12% and 2% for z and ρ, respectively, when

network dynamics is explicitly considered. Since the mean-field model does not depend on f and q, this result suggests a weak dependence of the population-area relationship (through variables z and ρ) on the latter parameters. Deeper numerical explorations show that ρ is almost independent of f and q, whereas z varies moderately in the region $q \approx 0$, becoming almost constant for $q > 0.3$.

Assuming that the parameter subspace (r,w) is mostly uncoupled to the subspace (f,q), our fits will be performed in two steps. First, we fit (r,w) to empirical values of z and ρ, keeping $f^\star = 0.1$ and $q^\star = 0.1$ fixed (see Table I).

Second, we obtain estimates for f and q by imposing that simulated degree and shortest-path length distributions keep close (in a precise sense to be defined) to empirical distributions. Finally, we check that the estimates of z and ρ still reproduce empirical values for the f and q obtained.

A. Fitting r and w to data

As initial guesses for (r,w), we use the mean-field values $(r_{\mathrm{MF}},w_{\mathrm{MF}})$ reported in Table II. Subsequently, the adaptive network model is simulated in square neighborhoods of $(r_{\mathrm{MF}},w_{\mathrm{MF}})$, keeping $f^\star = 0.1$ and $q^\star = 0.1$ fixed. We use model networks with the same sizes of empirical ones and average over 1000 model realizations. For each point (r_i,w_j) of the grid, each model network can be used to estimate the parameters $z(r_i,w_j)$ and $\rho(r_i,w_j)$ that characterize the population-area relationship. Then we estimate the averages over realizations $z_{ij} = \langle z \rangle (r_i,w_j)$ and $\rho_{ij} = \langle \rho \rangle (r_i,w_j)$, as well as the corresponding standard deviations σ_{ij}^z and σ_{ij}^ρ.

Let us define $\mathbf{y} = (z,\rho)^{\mathrm{T}}$ and $\mathbf{x} = (r,w)^{\mathrm{T}}$. In a local neighborhood of each point of the grid we expect an approximate (two-dimensional) linear dependence between $\mathbf{y}$ and $\mathbf{x}$,

$$\mathbf{y} \approx \mathsf{M}\mathbf{x} + \mathbf{b} \quad (12)$$

for a constant 2×2 (Jacobian) matrix $\mathsf{M} = (m_{ij})$ and a vector $\mathbf{b} = (b_1,b_2)^{\mathrm{T}}$ to be determined. We estimate the required coefficients by means of a two-dimensional, weighted least-squares fit to simulated data, i.e.,

$$\begin{aligned} z_{ij} &= m_{11}r_i + m_{12}w_j + b_1, \\ \rho_{ij} &= m_{21}r_i + m_{22}w_j + b_2. \end{aligned} \quad (13)$$

Fit's weights are chosen in the usual way, as $1/\sigma_{ij}^2$, provided that standard deviations for z and ρ are known. Note that the least-squares method provides estimates for standard errors of m_{ij} and b_i.

Finally, r and w estimates come from

$$\mathbf{x} = \mathsf{M}^{-1}(\mathbf{y} - \mathbf{b}), \quad (14)$$

where $\mathbf{y} = (z_e,\rho_e)^{\mathrm{T}}$ is the pair of empirical values that characterize the population-area relationship (see Table I). The errors of r and w have been calculated using standard error propagation according to Eq. (14). Results are listed in Table II, where they can be compared to mean-field estimates. There are some quantitative differences regarding previous results in different world regions [15]. The value of the spontaneous retreat r is not always below 1, implying that the reduction in area caused by a decrease in the population size is not sublinear in all cases. The four exceptions coincide with the smallest networks in our data set (Sulawesi, Luzon, Mex2,

and ABP), so it cannot be discarded that this effect reflects a limited statistical power. The relationship $r < w$ holds in most cases, with the exception of Sulawesi and Luzon islands.

B. Fitting f and q to data

Now we proceed with the fit of f and q to topological quantities. For each network's size, we simulate 2000 model realizations keeping r and w fixed and equal to the estimates previously obtained. The estimation of f and q proceeds through two different approaches: (i) minimizing the separation between the empirical and the simulated degree distributions; (ii) jointly adjusting the degree and the shortest-path length distributions.

1. Optimization based on the degree distribution

We determine f and q as the values that minimize the Hellinger distance between the empirical degree distribution and the simulated degree distribution. For two arbitrary discrete distributions $\mathbf{g} = (g_i)$ and $\mathbf{h} = (h_i)$, the Hellinger distance [30] is defined as

$$d_{\mathrm{H}}(\mathbf{g},\mathbf{h}) = \frac{1}{\sqrt{2}} \sum_i (\sqrt{g_i} - \sqrt{h_i})^2 = \frac{1}{\sqrt{2}} \| \sqrt{\mathbf{g}} - \sqrt{\mathbf{h}} \|_2 , \quad (15)$$

i.e., d_{H} is proportional to the Euclidean norm of the difference of square-root vectors. We have used the Hellinger distance because it measures the similarity between two probability distributions. For discrete distributions it can be shown that

$$d_{\mathrm{H}}^2(\mathbf{g},\mathbf{h}) \leqslant \tfrac{1}{2} \| \mathbf{g} - \mathbf{h} \|_1 \leqslant \sqrt{2} d_{\mathrm{H}}(\mathbf{g},\mathbf{h}), \quad (16)$$

which implies that minimizing the Hellinger distance is tantamount to minimizing the 1-norm of the difference of distributions. We choose the pair (f,q) that minimizes d_{H} for all networks here considered, where $g_k = p_{\mathrm{e}}(k)$ is the empirical degree distribution, and $h_k = p_{\mathrm{s}}(k)$ is the simulated degree distribution.

Minimization has been carried out in two steps: first we perform a parameter screening in $f \in [0.05,1]$ and $q \in [0,1]$. This yields an estimate of the pair (f,q) that minimizes d_{H}. Second, we use the estimation as initial guess for a standard algorithm of numerical minimization.

For large values of f and small values of q, the adaptive network model yields disconnected graphs. This is due to the fact that large f values imply a small number of neighbors (therefore a low connectivity), while small q values tend to eliminate all unpaired links. Therefore, the number of nodes of the giant component can be well below the size of the empirical network. Since we fix the network size to simulate model networks, we have to restrict the range of feasible f and q in the minimization procedure in a way that avoids small connected components and the sizes of empirical and model networks significantly match. In order for the sizes of empirical and model networks to be comparable, the range of f and q values is restricted by the requirement

$$\langle N \rangle \geqslant 0.9 N_{\mathrm{e}}, \quad (17)$$

N_{e} being the empirical network size. Parameter values yielding network sizes outside of this region are not taken into account during minimization. This process yields the estimates listed

TABLE III. Fitted model parameters f and q obtained through minimization of the distance between simulated and empirical degree distributions.

CC	f	q	d_{H} (degree)	d_{H} (path)
C Africa	0.14 ± 0.01	0.30 ± 0.01	0.09	0.16
C Asia	0.15 ± 0.01	0.30 ± 0.01	0.12	0.56
Australia	0.12 ± 0.01	0.05 ± 0.01	0.08	0.17
Sulawesi	0.20 ± 0.01	0.20 ± 0.01	0.14	0.17
Luzon	0.11 ± 0.01	0.00 ± 0.01	0.13	0.44
C Europe	0.15 ± 0.01	0.18 ± 0.02	0.15	0.25
Mex1	0.14 ± 0.01	0.07 ± 0.01	0.15	0.37
Yucatan	0.46 ± 0.01	0.71 ± 0.01	0.21	0.23
Mex2	0.49 ± 0.02	0.84 ± 0.03	0.24	0.19
CS America	0.17 ± 0.01	0.13 ± 0.01	0.12	0.16
ABP	0.15 ± 0.01	0.12 ± 0.01	0.14	0.10

in Table III for f and q, and Hellinger's distance for the degree and the shortest-path length distributions.

2. Optimization based on the degree and shortest-path length distributions

We now apply a joint minimization of Hellinger's distance to the degree and shortest-path length distributions. The sum of both distances is used as the objective function to minimize,

$$s(f,q) = d_{\mathrm{H}}(\mathbf{d},\mathbf{e}) + d_{\mathrm{H}}(\mathbf{g},\mathbf{h}), \quad (18)$$

where $\mathbf{d} = (p_{\mathrm{e}}(d))$ is the empirical distribution of shortest-path length and $\mathbf{e} = (p_{\mathrm{s}}(d))$ is its simulated counterpart, and similarly for the empirical and simulated degree distributions $\mathbf{g} = (p_{\mathrm{e}}(k))$ and $\mathbf{h} = (p_{\mathrm{s}}(k))$. Results are listed in Table IV. The restriction given by Eq. (17) also applies here.

Note that too large values of f or too small values of q might cause a transition from a large connected component to a mostly disconnected ensemble of small networks when the network dynamics step is applied. The effect of decreasing f and/or increasing q from sufficiently low values (where nodes are disconnected) eventually causes a percolation transition comparable to that described in the standard Erdős-Renyi model [31] as the number of links increases. The fitted values

TABLE IV. Fitted model parameters f and q using the joint minimization scheme.

CC	f	q	d_{H} (degree)	d_{H} (path)
C Africa	0.11 ± 0.01	0.14 ± 0.01	0.13	0.11
C Asia	0.09 ± 0.01	0.25 ± 0.01	0.28	0.06
Australia	0.11 ± 0.01	0.01 ± 0.01	0.10	0.07
Sulawesi	0.21 ± 0.01	0.20 ± 0.02	0.15	0.14
Luzon	0.13 ± 0.01	0.29 ± 0.06	0.31	0.15
C Europe	0.11 ± 0.01	0.12 ± 0.01	0.19	0.04
Mex1	0.67 ± 0.03	0.85 ± 0.03	0.18	0.19
Yucatan	0.11 ± 0.03	0.06 ± 0.04	0.26	0.08
Mex2	0.61 ± 0.01	0.97 ± 0.01	0.26	0.15
CS America	0.16 ± 0.01	0.12 ± 0.02	0.13	0.11
ABP	0.16 ± 0.01	0.11 ± 0.01	0.15	0.06

we have obtained seem to balance such that the resulting networks are connected. For example, in Table III we observe that low f values correspond to mostly low q values (as in Luzon, Australia, continental Africa, and Mex1), while high f values are associated to high q values (as in Yucatan and Mex2). This association is also observed in Table IV, with interesting, consistent inversions of the correspondence seen in Mex1 and Yucatan.

C. Performance of the model

1. Degree distributions

Figure 3 shows the comparison between empirical and simulated degree distributions. Results in Fig. 3(a) have been obtained through minimization of Hellinger's distance between degree distributions, whereas in Fig. 3(b) the result of the joint minimization procedure is shown. In the former case,
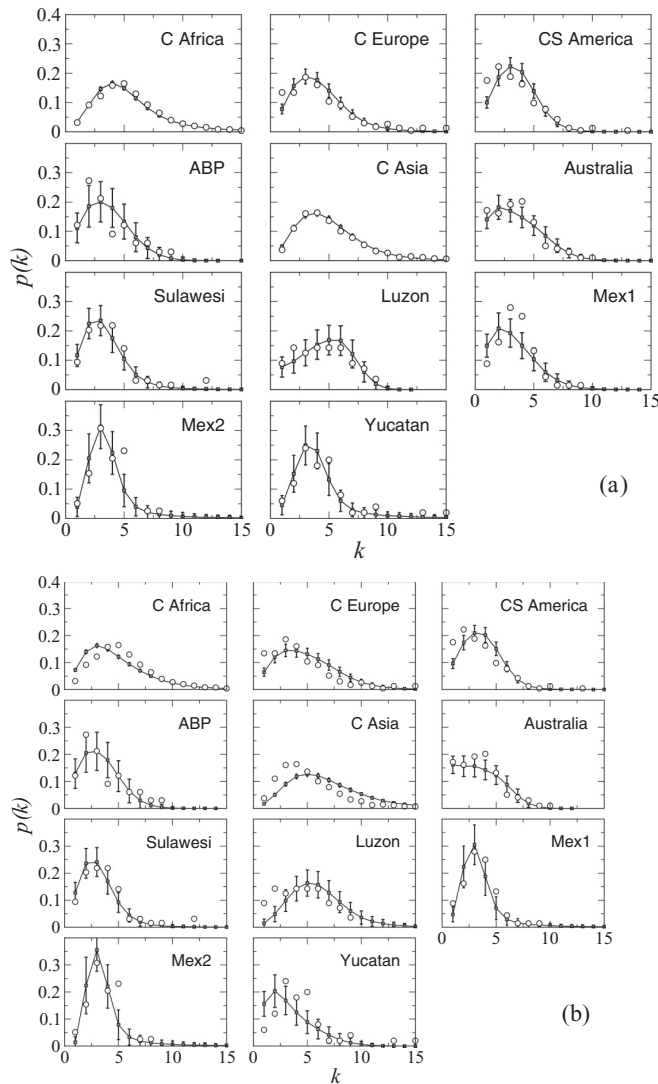
FIG. 3. Empirical degree distributions (open circles) vs simulated distributions (linked squares) averaged over 2000 model realizations. Error bars correspond to standard deviations of model-simulated degree distributions. (a) Minimization based on the degree distribution. (b) Joint minimization. Ranges of axes are the same in all plots.

the agreement with empirical data is very good, even when the statistics of the original data (i.e., the network size, see Table I) is poor. The Hellinger distances for the degree distribution obtained with the joint minimization procedure (cf. Table IV) are, as expected, larger than the minimum values reported in Table III.

2. Shortest-path distributions

Figure 4 shows the results of the minimization of Hellinger's distance for the degree distribution [Fig. 4(a)] and for the degree and shortest-path distribution jointly [Fig. 4(b)]. In the former case, the agreement between empirical and simulated distributions is poor in several cases (and especially in continental Asia), but there are some exceptions where the empirical distribution is reasonably reproduced, for example in ABP borders, continental South America (the third largest network), or Sulawesi island. The likelihood of the null hypothesis that the model can generate networks whose average shortest-path length $\langle d \rangle$ is compatible with the empirical value $\langle d_e \rangle$ has been statistically tested using minimization of Hellinger's distance based only on degree distributions. We have calculated the p values of the null hypothesis, $\Pr(\langle d \rangle \geqslant \langle d_e \rangle)$, which are listed in Table V. At a 99% confidence level, the null hypothesis is rejected only for continental Asia, Luzon, and Mex1. Therefore, even if the distribution of shortest-path lengths is not explicitly considered to estimate model parameters, the adaptive network model is not statistically rejected to reproduce average path lengths in most empirical networks.

The joint minimization significantly improves the fit to empirical shortest-path length distributions, yielding low values of Hellinger's distance in most cases, see Table IV. Though the joint fit to the degree and the shortest-path length distributions worsens the performance of the fit regarding the degree distribution, the overall fit to both distributions is significantly improved, as can be seen by comparing the sum d_H (degree) + d_H (path) in Tables III and IV.

3. Consistency check

Parameters r and w were obtained at constant values of the perimeter overlap f and the symmetrization value q, namely $(f^\star, q^\star) = (0.1, 0.1)$. To test the consistency of our estimation procedure, we check now whether the use of final estimated values in Tables III and IV substantially modify the performance of the model regarding z and ρ. Additional simulations for each (r, w, f, q) set of parameter values have been carried out, and averages for exponent z and correlation ρ have been calculated. The results are summarized in Table VI, and should be compared with empirical data in Table I. As can be seen, all empirical values lie within the error bars, thus validating *a posteriori* the methodology used.

4. Demographic and topological variables

Figures 5 and 6 depict the correlation between averaged logarithmic areas and averaged logarithmic populations, respectively, and degree k. Simulation data have been produced with the set of parameters obtained under both minimization schemes. For visualization purposes, model results have been displaced in the vertical axis through the addition of an
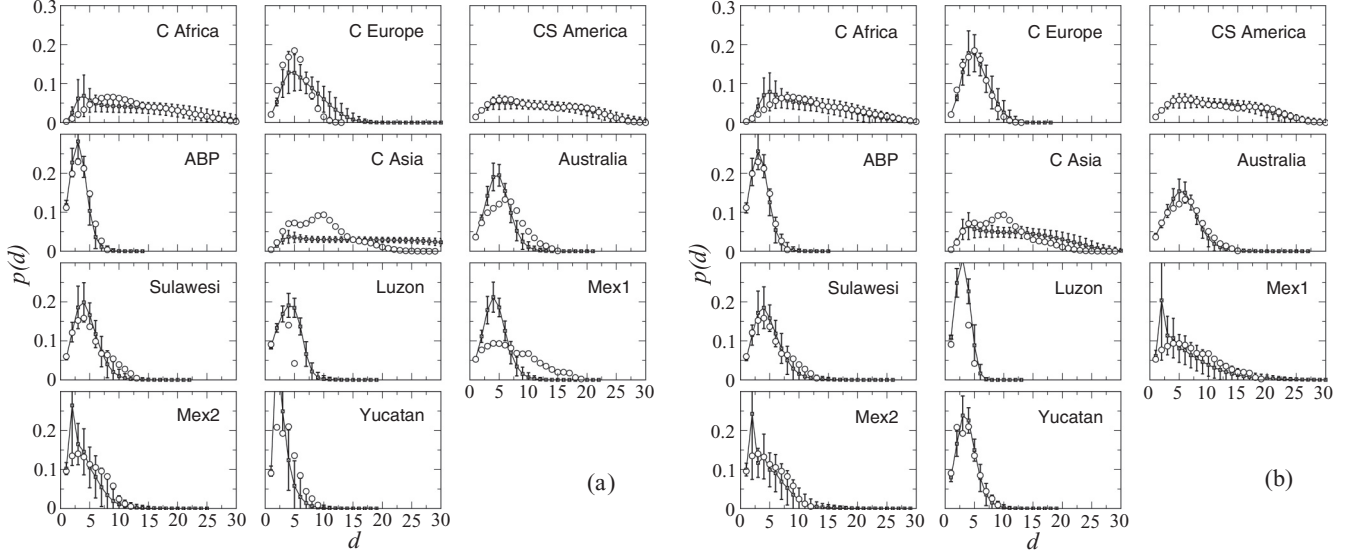
FIG. 4. Empirical shortest-path length distributions (open circles) vs simulated distributions (linked squares) averaged over 2000 model realizations. Error bars correspond to standard deviations of model-simulated shortest-path length distributions. (a) Minimization based on the degree distribution. (b) Joint minimization. Ranges of axes are the same in all plots.

arbitrary constant (recall that area and population units are defined up to a constant). Except for the ABP network (which is the smallest one, with $N = 33$ nodes and therefore a poor statistical power), empirical logarithmic areas and populations monotonically increase with k. Though these functions are not explicitly considered to obtain model parameters, simulations reproduce with remarkable accuracy the empirical observations.

The coupling between the stochastic process that determines population and areas and the update of the network of contacts under the perimeter overlap rule leads to the emergence of correlations between the area, population, and degree of neighboring nodes. These autocorrelations over the network are measured as

$$\tau_i = \frac{1}{2} \sum_{j=1}^{N} \frac{v_j}{n_{j,i}} \sum_{|k-j|=i} v_k, \quad (19)$$

TABLE V. Empirical average path lengths for language networks and p values of the null hypothesis corresponding to the adaptive network model with minimization of Hellinger's distance on degree distributions.

CC	$\langle d_e \rangle$	p value
C Africa	12.9	0.46
C Asia	9.5	$<10^{-2}$
Australia	6.1	0.02
Sulawesi	5.2	0.13
Luzon	2.6	$<10^{-2}$
C Europe	5.0	0.13
Mex1	7.2	$<10^{-2}$
Yucatan	3.7	0.14
Mex2	4.8	0.18
CS America	11.9	0.33
ABP	3.4	0.17

where v_j stands for any of the node properties a_j, p_j, or k_j, and the second sum runs over the nodes that are at distance i from node j (v_j are normalized to satisfy $\sum_j v_j = 1$ so that autocorrelations for different variables are independent of their natural scales and can be mutually compared). Distances are measured as shortest-path lengths between nodes. We normalize the product $v_j v_k$ by $n_{j,i}$, that is by the number of nodes that are at distance i from node j. The $1/2$ factor takes into account that all links are double counted, since the sum runs over the whole network.

Autocorrelations over the network have been plotted in Fig. 7. Correlations decay as the separation between nodes increases, demonstrating that the network is assortative regarding area, population and node degree. These correlations cannot be observed if demography and topology are

TABLE VI. For each parameter set obtained through degree distribution minimization, 2000 model realizations yield the estimates z_s and ρ_s, and similarly for the joint minimization. Subindex s stands for simulation results; upper-index j indicates joint minimization. Both estimates compare well with the empirical values reported in Table I.

| CC	z_s	ρ_s	z_s^j	ρ_s^j
C Africa	0.91 ± 0.03	0.63 ± 0.01	0.88 ± 0.02	0.63 ± 0.01
C Asia	0.66 ± 0.02	0.72 ± 0.01	0.67 ± 0.02	0.72 ± 0.01
Australia	0.67 ± 0.12	0.52 ± 0.08	0.66 ± 0.13	0.52 ± 0.08
Sulawesi	0.63 ± 0.07	0.75 ± 0.06	0.63 ± 0.07	0.75 ± 0.06
Luzon	0.43 ± 0.05	0.78 ± 0.05	0.44 ± 0.05	0.79 ± 0.05
C Europe	0.59 ± 0.05	0.66 ± 0.04	0.60 ± 0.05	0.65 ± 0.04
Mex1	0.75 ± 0.14	0.58 ± 0.08	0.87 ± 0.15	0.60 ± 0.08
Yucatan	1.34 ± 0.27	0.63 ± 0.09	1.16 ± 0.26	0.60 ± 0.09
Mex2	0.69 ± 0.11	0.74 ± 0.08	0.69 ± 0.11	0.73 ± 0.08
CS America	0.38 ± 0.04	0.56 ± 0.04	0.38 ± 0.04	0.56 ± 0.05
ABP	0.64 ± 0.10	0.76 ± 0.08	0.64 ± 0.10	0.75 ± 0.08
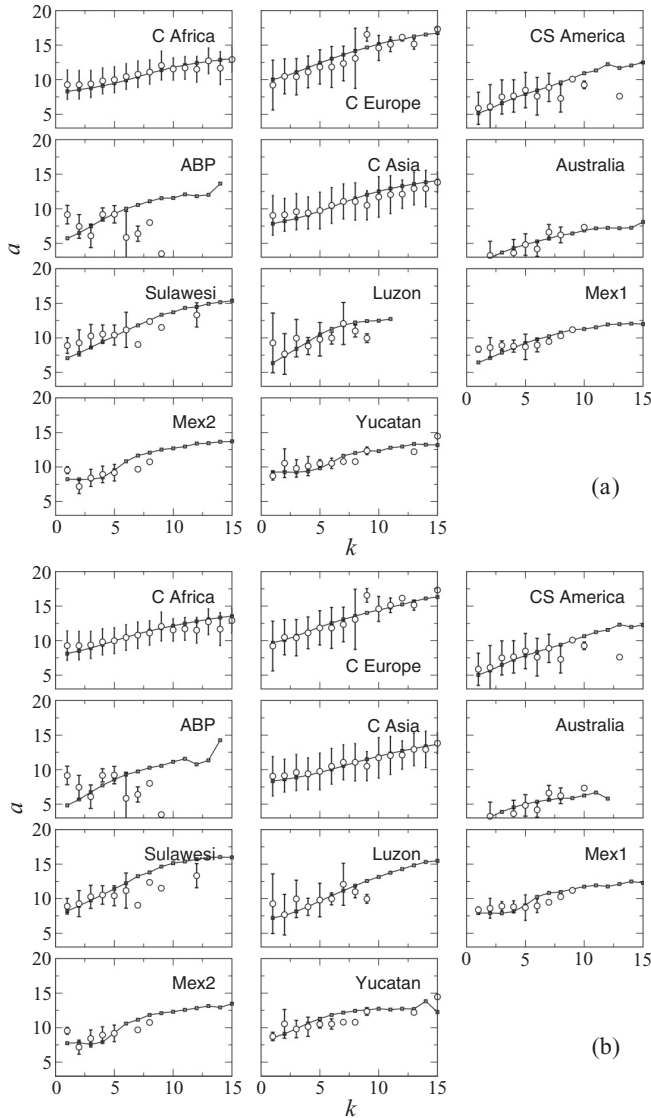
FIG. 5. Logarithmic area a vs node's degree k. Black circles correspond to empirical data; error bars are the standard deviation of a for each degree. Linked squares are averages over 2000 realizations of the adaptive network model. (a) Minimization based on the degree distribution. (b) Joint minimization. Ranges of axes are the same in all plots.

uncoupled: Demographic dynamics without an underlying network of contacts correspond to a mean-field model without spatial structure; a static algorithm that reproduces network topology is unable to account for correlations in population sizes, a variable disregarded in the algorithm. Model results compare only qualitatively with empirical networks (see Fig. 7 for an example). Following an initial rapid decay, real networks exhibit an intermediate range of node separations where autocorrelations are roughly constant. For large distances, however, correlations decay as predicted by the mode.

V. DISCUSSION AND CONCLUSIONS

The coevolution of population demography and spatial contacts represents an example of an adaptive network in the social sciences. By means of a model coupling both processes, we have shown that previous known properties of this system are robustly reproduced: population-area relationships and language network topology. Besides, a number of features relating demography and topology, as well as certain assortative properties of those networks, are consistently obtained in the adaptive network approach here introduced. Assortativity in node population, area, and degree are byproducts induced by subsequent cycles of population change and modification of topological neighborhoods, which cannot be obtained in scenarios where these two processes are decoupled. Remarkably, the agreement between several empirical and simulated quantities is obtained through fits of just four model parameters. We believe this is due to the deep meaning of model rules, which are by themselves sufficient to explain the qualitative properties of linguistic groups and their associated spatial networks. There is an important exception that cannot be recovered by the model, in particular the population-area relationship, which fails to be reproduced already by the mean-field approach [15]: New Guinea Island. Since this is an often studied example of a region with an extremely high linguistic diversity, it is worth mentioning that the demographic and conflict rules we implement do not suffice to yield the ρ value empirically measured. In general, the mean-field model is not able to reproduce the set of values $\rho \lesssim 0.5$ for the area-population correlation. Here, the set of languages in the giant connected component for New Guinea (note that this is a subset of the New Guinean languages used in Ref. [15]) yields $\rho = 0.42$, which cannot be accounted for with the proposed dynamical rules. Similarly, the correlation for all North American languages annotated in the Ethnologue was not reproduced by the mean-field model proposed in Ref. [15].

The adaptive network model here presented admits a number of extensions. First, the introduction of additional factors may make it more realistic. In this study, we have not considered the appearance of new languages or the death of existing ones. The origination of new languages can be easily implemented by splitting an existing language. Following our rules, a new set of neighbors and an independent evolution of either population appear in a straight way. Death of languages can also be considered, for instance, by eliminating those groups whose population falls below a prescribed level (one individual, for instance). Quantities such as the average lifetime of languages could be studied in this scenario. Second, model rules could be modified to consider factors such as language attractiveness or frequency of conflicts dependent on degree or population size [32].

Language attractiveness is an important driver in the disappearance of minority languages [33], and could be implemented through a migration of population from one language to any of its neighbors. In this way, population sizes would be modified through processes different from stochastic growth. Extreme versions of migration mechanisms might account for the growth of widespread languages [34], and perhaps explain the emergence of dragon kings in linguistic groups [25]. In the model here used, every group enters into conflict with a neighbor once per time step. This rule could be modified to a likely more realistic version where conflict frequency is proportional to the number of links, and not to the number of nodes, and the outcome of conflicts could
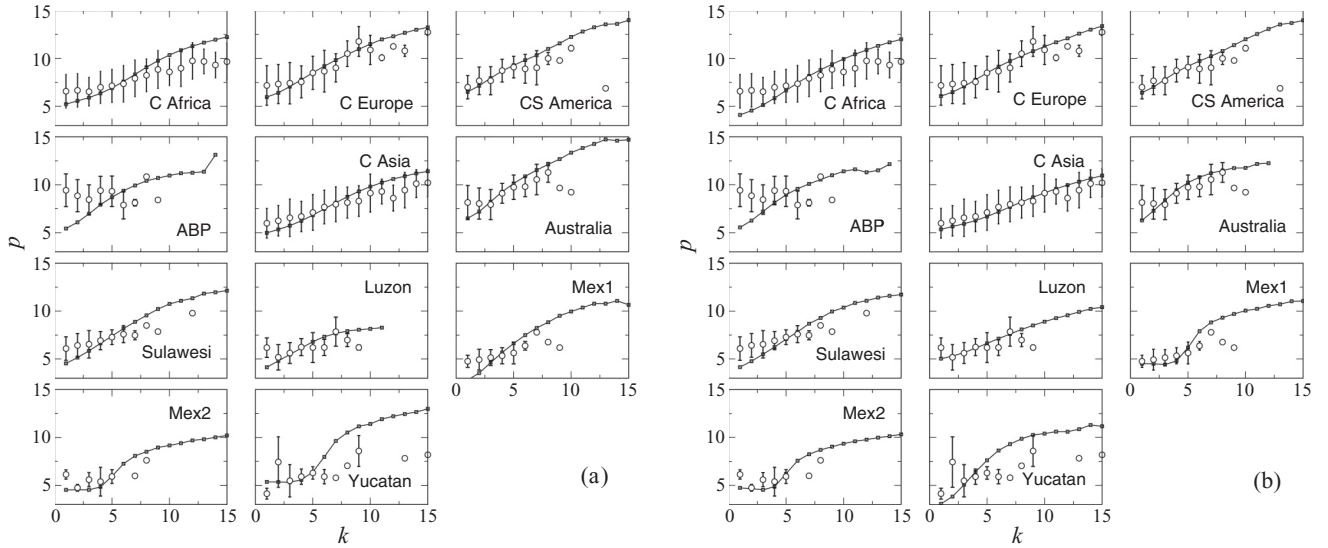
FIG. 6. Logarithmic population p vs node's degree k. Black circles correspond to empirical data; error bars are the standard deviation of p for each degree. Linked squares are averages over 2000 realizations of the adaptive network model. (a) Minimization based on the degree distribution. (b) Joint minimization. Ranges of axes are the same in all plots.

also consider the relative population of the involved parties. Also additional cultural markers, such as political, linguistic or religious similarities, might modify the frequency and strength of conflicts. The results of these modifications are difficult to foresee. Third, the formation of links is now homogeneous and does not consider the structure of human settlements in relation to the landscape. The introduction of preferential attachment depending on stylized landscape features might help explaining the appearance of a low dimensional niche space in language networks [16,35].
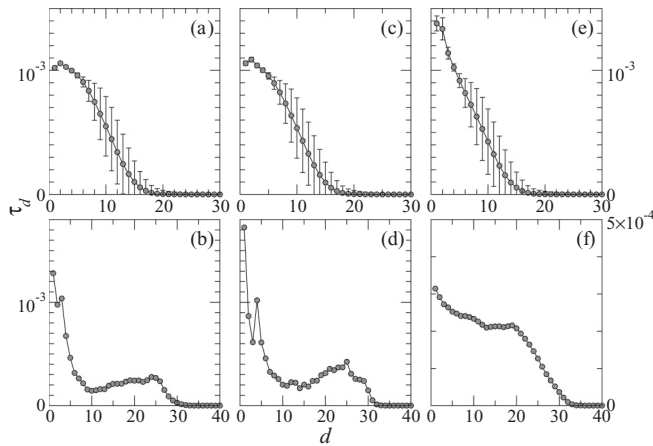

FIG. 7. Autocorrelation τ_d over the network, as defined by Eq. (19), as a function of the separation d between nodes, measured in the model for (a) log-area, (c) log-population, and (e) degree. Model parameters are $r = 0.5$, $w = 2$, $f = 0.1$, and $q = 0.1$ for networks with $N = 500$ nodes. Quantities have been averaged over 500 independent realizations. For the sake of comparison, we depict the autocorrelations for (b) log-area, (d) log-population, and (f) degree, obtained for the empirical network of continental Africa. Note the different range of the vertical axis in (f).

The competition for areas between neighboring populations is a form of demographic conflict. In the scenario here devised, these conflicts do not affect population sizes and by definition occur at a characteristic time scale of the order of one year. There is a body of literature that has addressed the frequency and distribution of conflicts with the number of casualties in terrorist attacks [36], wars [37], or fatal quarrels in general [38] as the main variable. Those events might have frequencies measured in days and have been often modeled as processes of fragmentation and coalescence of groups [39]. It would be interesting to integrate the dynamical network perspective of our study with the fast evolution of groups dynamics and its effect on population sizes of these other conflict analyses with the goal of devising more complete models for cultural and political clashes.

Finally, we believe that the model could be applied to other model systems with analogous node and network dynamics. One such example is ecology, where an explicit competition for space of species occupying the same niche is known to occur. Further, the applicability of the model to that system is supported by a relationship between population sizes and ranges of occupation functionally equivalent to the population-area law followed by human linguistic groups. Demographic dynamics similar to those used here, perhaps with the addition of temporal biases to grow or decrease, might represent the dynamics of agents such as companies or religious groups, for example. Suitable modifications of how links are established might shed light on the distribution of group sizes and on the relevance of competition and intergroup conflicts.

[1] *Adaptive Networks*, edited by Th. Gross and H. Sayama (Springer, Berlin, 2009).

[2] Th. Gross and B. Blasius, J. R. Soc. Interface **5**, 259 (2008).

[3] J. Gjorgjieva, C. Clopath, J. Audet, and J.-P. Pfister, Proc. Natl. Acad. Sci. USA **108**, 19383 (2011).

[4] B. Drossel, P. G. Higgs, and A. J. McKane, J. Theor. Biol. **208**, 91 (2001).

[5] S. Jain and S. Krishna, Proc. Natl. Acad. Sci. USA **98**, 543 (2001)

[6] S. C. Manrubia and J. F. Poyatos, Europhys. Lett. **64**, 557 (2003).

[7] J. Ito and K. Kaneko, Phys. Rev. Lett. **88**, 028701 (2001).

[8] T. Gross, Carlos J. Dommar D'Lima, and B. Blasius, Phys. Rev. Lett. **96**, 208701 (2006).

[9] D. H. Zanette and S. Risau-Gusmán, J. Biol. Phys. **34**, 135 (2008).

[10] G. C. M. A. Ehrhardt, M. Marsili, and F. Vega-Redondo, Phys. Rev. E **74**, 036106 (2006).

[11] Z. Burda, A. Krzywicki, and O. C. Martin, Phys. Rev. E **78**, 046106 (2008).

[12] J. L. Herrera, M. G. Cosenza, K. Tucci, and J. C. González-Avella, Europhys. Lett. **95**, 58006 (2011).

[13] E. zu Erbach-Schoenberg, S. Bullock, and S. Brailsford, Soc. Sci. Comp. Rev. **32**, 373 (2014).

[14] S. Metcalf and M. Paich, Proc. 23rd Intl. Conf. of the Syst. Dyn. Soc. (2005).

[15] S. C. Manrubia, J. B. Axelsen, and D. H. Zanette, PLoS ONE **7**, e40137 (2012).

[16] J. A. Capitán, J. B. Axelsen, and S. Manrubia, Proc. R. Soc. Lond. B **282**, 20142947 (2015).

[17] D. B. Stouffer, J. Camacho, and L. A. N. Amaral, Proc. Natl. Acad. Sci. USA **103**, 19015 (2006).

[18] J. A. Capitán, A. Arenas, and R. Guimerà, J. Theor. Biol **334**, 35 (2013).

[19] D. H. Zanette, Int. J. Mod. Phys. C **19**, 237 (2008).

[20] W. J. Sutherland, Nature **423**, 276 (2003).

[21] United Nations, Department of Economic and Social Affairs. Available at www.un.org/esa/population/publications/sixbillion.htm. Accessed August 19, 2014.

[22] R. G. Gordon, *Ethnologue: Languages of the World*, 15th ed. (SIL International, Dallas, 2005).

[23] It can be shown that the dispersion in the distribution of language sizes at the beginning of the stochastic process (1000 years ago) has a negligible contribution to the value of σ_α (at most 0.5%, obtained from the limit situation where each language had one single speaker). The figure here reported stems thus from a multiplicative stochastic process in population growth with arbitrary initial conditions, average value 1.0029, and a mean-square dispersion caused by that of the current distribution of extant languages. See Ref. [19] for a full derivation of parameters α and σ_α.

[24] Note that the distribution of β_i is related to the distribution of the (untransformed) growth rate α_i as $P(\beta)d\beta = Q(\alpha)d\alpha$, with $\beta_i = \log \alpha_i$. Then $\epsilon = \log \alpha$ and η is derived from the relation between the two distributions.

[25] D. Sornette, Int. J. Terraspace Sci. Eng. **2**, 1 (2009).

[26] J. Harte, T. Blackburn, and A. Ostling, Am. Nat. **157**, 374 (2001).

[27] K. J. Gaston and F. He, Proc. R. Soc. Lond. B **269**, 1079 (2002).

[28] J. E. Cohen, Proc. Natl. Acad. Sci. USA **74**, 4533 (1977).

[29] We chose $f^\star = 0.1$ because values around 0.1 or 0.2 can be a reasonable election to represent empirical data, especially for large networks (see Table I). Similarly, using the static algorithm for language networks [16], we observed that large empirical networks can be well fitted by $q \approx 0.1$. This motivated the value chosen for $q^\star$.

[30] E. Hellinger, J. Reine Angew. Math. **136**, 210 (1909).

[31] P. Erdős and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).

[32] M. Lim, R. Metzler, and Y. Bar-Yam, Science **317**, 1540 (2007).

[33] D. M. Abrams and S. H. Strogatz, Nature (London) **424**, 900 (2003).

[34] J. M. Diamond, Nature (London) **389**, 544 (1997).

[35] J. B. Axelsen and S. Manrubia, Proc. R. Soc. Lond. B **281**, 20133029 (2014).

[36] A. Clauset, M. Young, and K. S. Gleditsch, J. Conflict Resolut. **51**, 58 (2007).

[37] L. E. Cederman, Am. Polit. Sci. Rev. **97**, 135 (2003).

[38] L. F. Richardson, J. Am. Stat. Assoc. **43**, 523 (1948).

[39] J. C. Bohorquez, S. Gourley, A. R. Dixon, M. Spagat, and N. F. Johnson, Nature (London) **462**, 911 (2009).