Preferential attachment in multiple trade networks

Rachele Foschi*
LIME, IMT School for Advanced Studies, Lucca, Italy

Massimo Riccaboni†
LIME, IMT Schoool for Advanced Studies, Lucca, Italy and DMSI, K.U. Leuven, Leuven, Belgium

Stefano Schiavo‡
School of International Studies and Department of Economics and Management, University of Trento, Trento, Italy and OFCE-DRIC, France
(Received 10 November 2013; revised manuscript received 7 August 2014; published 27 August 2014)

In this paper we develop a model for the evolution of multiple networks which is able to replicate the concentrated and sparse nature of world trade data. Our model is an extension of the preferential attachment growth model to the case of multiple networks. Countries trade a variety of goods of different complexity. Every country progressively evolves from trading less sophisticated to high-tech goods. The probabilities of capturing more trade opportunities at a given level of complexity and of starting to trade more complex goods are both proportional to the number of existing trade links. We provide a set of theoretical predictions and simulative results. A calibration exercise shows that our model replicates the same concentration level of world trade as well as the sparsity pattern of the trade matrix. We also discuss a set of numerical solutions to deal with large multiple networks.

DOI: 10.1103/PhysRevE.90.022817 PACS number(s): 89.65.Gh, 89.75.Hc

I. INTRODUCTION

The analysis of global systems has progressively moved from considering single networks such as trade, finance, transportation, and communication to analyzing them jointly as multiple interdependent networks. Our work contributes to this line of research by developing a new description of international trade as a set of interdependent multiple networks. The work combines two streams of the recent literature: the former studies the empirical regularities characterizing international trade flows, especially those that are puzzling for standard economic models [1–5]; the latter relates to the increasing use of network concepts to describe economic systems [6,7]. Our approach is able to replicate a set of properties of the world trade web, including the large fraction of missing trade relationships one observes across countries.

The sparse nature of trade data, resulting in a large proportion of zero-trade flows, has received a good deal of attention in recent years [8]. Evidence put forward in Ref. [4] shows that over the period 1970–1997 only half of all possible country-pair links are ever activated (either in one or the other direction). The percentage of missing links grows even larger when we consider multiple trade networks, one for each product traded: 82% of potential product-partner trade flows are actually zero for U.S. trade data, when export product categories are defined according to the 10-digit Harmonized System (HS) [2]. Such a percentage is even larger for U.S. imports (92%). A similar behavior is observed for all countries with a percentage of zero-trade flows that ranges between 69% and 99.5%, with a mean value of 96% (based on UN-Comtrade data at the HS-6 level [9]).

*rachele.foschi@imtlucca.it
†Corresponding author: m.riccaboni@imtlucca.it; Piazza San Ponziano, 6-55100 Lucca, Italy.
‡stefano.schiavo@unitn.it

A second interesting regularity, partly related to the previous one, pertains to the large disparities in trade participation. Trade concentration can be appreciated by looking at the distribution of the number of destinations served by each exporting country, the number of products shipped, or the number of product-destination pairs. Several studies have documented the skewness of such distribution: This feature appears to be stable across the different data sets used and over time [9–13].

In the network literature, very skewed connectivity distributions are found to characterize many real-world applications besides trade (e.g., the Internet, worldwide air transportation, mobile communication, and interbank payments, to name just a few), so a network approach appears well placed to account for the two features of international trade data discussed so far. One of the most successful null models to account for the power-law connectivity distribution of real-world networks is the preferential attachment (PA) growth model [14]. However, to generate skewed connectivity and sparse network structures, the PA regime must be complemented by a constant inflow of new nodes. Such a model does not fit well with the international trade network, where the set of nodes (i.e., countries) is almost constant in time. To solve this puzzle, one approach is to rely on some node-specific attributes, such as, for instance, country gross domestic product and population in a gravity approach. However, this approach does not account for differences across product-specific trade networks, unless one assumes that countries have some intrinsic fitness in the product space [15,16].

In this paper, we go back to a purely stochastic approach to propose a generalization of the PA model which is able to describe the topological structure of bilateral trade flows across countries. Since the number of traded products is large relative to the number of countries, in order to re-create a process matching the large number of zeros observed in the data, the trade network has to be decomposed into subnetworks of different dimensions. Each submatrix represents a trade

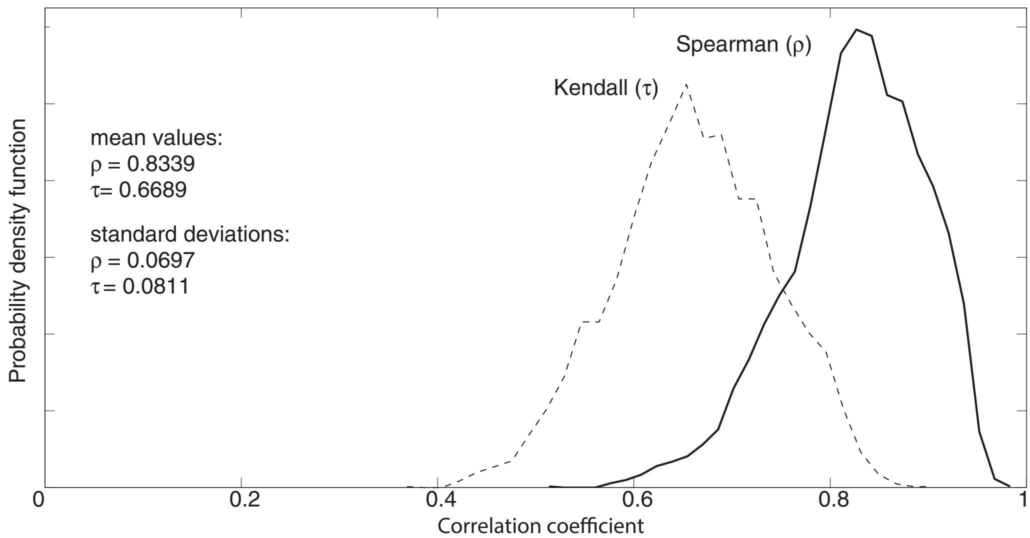

FIG. 1. Rank correlations between out-degree of each country across (96) different two-digit HS chapters. Data refer to 2001: Correlations range between 0.369 and 0.983, with mean values equal to 0.8339 (Spearman) and 0.6689 (Kendall).

network where some specific products are traded and not all countries are simultaneously active in trading all goods. This formal treatment of the problem suggests a learning process whereby many countries trade the most basic products, whereas only a few of them manage to produce and export the most sophisticated manufactured goods [6]. In other words, we keep fixed the set of countries and consider multiple trade networks sorted by the complexity of traded goods. The PA regime is active across all networks and we model the entry probability of a country in high-tech trade as proportional to the total number of trade relationships it has already activated.

The approach we follow entails decomposing the adjacency matrix describing the world trade network in nested submatrices of different dimensions and allocating in each of them a number of links warranting an adequate sparsity structure. The cheaper computational costs of this procedures also reflect an inferior mathematical complexity in writing the distributions of the simulated quantities of interest. This allows us to derive useful analytical properties of the process we want to reproduce.

The paper is structured as follows. Section II illustrates the empirical properties of the real-world trade network that we aim at reproducing. Section III describes our decomposition procedure, and the criteria to establish both the dimensions of the subnetworks and the number of links to allocate in each of them. In Sec. IV we describe some probabilistic properties of our approach whereas in Sec. V the procedure is applied to trade data. Finally, Sec. VI concludes and discusses further research directions.

II. DESCRIPTIVE DATA ANALYSIS

From an empirical point of view, we refer to the BACI data set maintained by CEPII and reporting bilateral trade flows among a large number of countries over the years 1995–2011 [17]. The data are organized according to the HS classification, and each bilateral flow (from source to destination country) is attached a six-digit product code based on the physical characteristics of the goods traded.[1] Hence, each observation in our data set is defined by the source and destination country, product code, and dollar value. Since we are mainly interested in the number of zeros and the connectivity distribution, we disregard the information on the value of trade and reaggregate the data at the country level by counting how many six-digit products are exported from country i to country j. After dropping some small countries and territories, for each year we end up with a 189×189 matrix A, whose (i,j)-th entry K_{ij} represents the number of products exported from i to j. A can be seen as the adjacency matrix of a weighted and directed graph, with integer weights, representing the number of exported and imported products.

From the data, we calculate the percentage of zeros in each trade matrix: The average over the whole 1995–2009 period is about 42%, ranging from 52% in 1995 to 36% percent in 2007.[2] Most countries export a small number of products to a few destinations, while only a few players are extremely connected. Indeed, this is consistent with previous findings pointing to a core-periphery structure of world trade [18]. Data for 2001, for instance, tell that the number of (product-destination) links for each country N_r ranges between 35 and 322 064 (mean 30 075, standard deviation 59 144). We also notice that leading countries tend to dominate trade in every product category: Figure 1 plots the density function of (rank) correlation coefficients among the number of destinations served by

[1]HS is the standard classification for international trade in all major countries and is maintained by the World Customs Organization. Products are initially assigned to 99 broad two-digit categories (e.g., Chapter 87 - Vehicles other than railway), which are then further broken out into more detailed six-digit codes (e.g., code 871110 - Motorcycles, with internal combustion engine not exceeding 50 cc, or code 871120 - Motorcycles, with internal combustion engine between 50 and 250 cc).

[2]Since the structural properties of the resulting network are rather stable over time (see, for instance, Ref. [18]), the specific year analyzed is irrelevant.

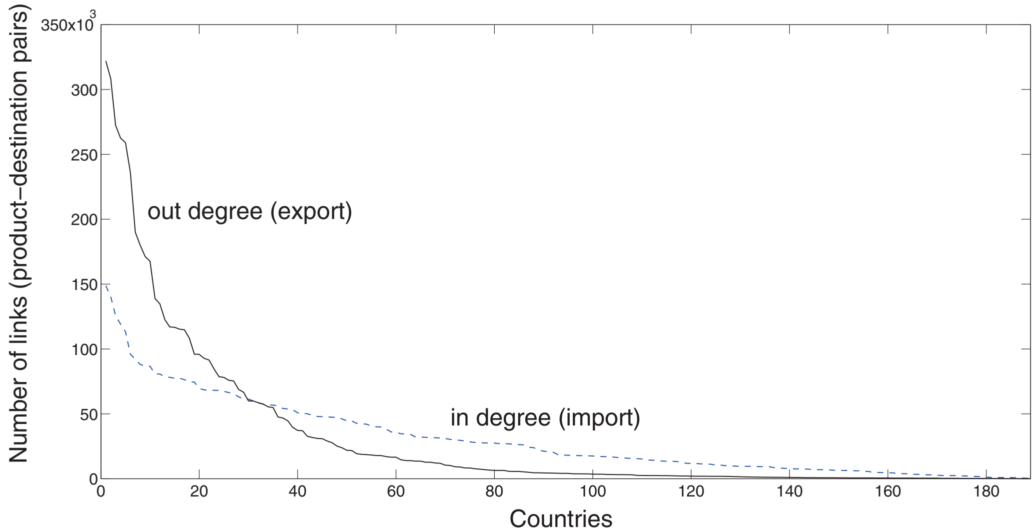

FIG. 2. (Color online) In-degree (import) and out-degree (export) of countries in the world trade web. Each link is a product-specific trade relationship between two countries.

each country across different products (defined as two-digit HS chapters, of which we have 96 in the data) in the year 2001. First, for each of the 189 countries in the sample, and each product, we count the number of destinations served (out-degree), and then we compute the pairwise correlations across the 96 product categories, obtaining 4560 ($96 \times 95/2$) coefficients. The evidence summarized in Fig. 1 supports the view that most central nodes tend to play a relevant role (i.e., serve many destinations) in all product networks.

Finally, it is worth noticing that the connectivity distribution of the trade network is extremely skewed, especially for exports (see Fig. 2).

III. A GENERALIZED PREFERENTIAL ATTACHMENT MODEL FOR MULTIPLE NETWORKS

To replicate the sparsity pattern and the connectivity distribution of world trade we start from a pure PA model [14]: Countries establish new trade links based on the number of connections they already have. Hence, more active exporters are more likely to export new products and/or reach new markets.[3] This mechanism of network formation and growth is consistent with the view that exporting represents a key engine of economic growth thanks to the endogenous forces set in motion by learning effects in the manufacturing sector. A more recent framing of this old idea can be found in Ref. [19], where the search for viable export industries is described as a process of "self-discovery": Attempts to set up new businesses and export new products to new destinations generate valuable public information as they signal profitable opportunities or dead ends. Linked to this is the idea that by producing a given set of goods each country accumulates a number of capabilities: The more capabilities are present, the easier it is to recombine them and put them to novel use [20]. A more micro-based account of PA is offered in Refs. [7] and [21]. The former postulates that firms can establish links with suppliers either at random at via already established connections (meeting friends of friends); the latter assumes that the fixed costs associated with penetrating a foreign market is decreasing in the number of firms already exporting there (from a given source country) due to the presence of (information) spillover effects. Here we do not take a specific position regarding the precise source of PA but rather focus on its effects in terms of the number of zeros and the connectivity distribution.

In our network model we start from a given set of active players (countries) w_0, each trading one product to a single destination. In our application we have $w_0 = 189$, thus there is no entry of new countries. Starting from this 189×189 matrix, $N_{\rm tot}$ trade links are allocated (each representing a product-destination pair), one at each step, according to the following procedure: The outgoing (incoming) link is assigned proportionally to the *export* (*import*) connectivity of countries, that is, the probability of catching a new outgoing (incoming) connection is proportional to the node out-degree (in-degree). This pure PA mechanism with no entry fills up the trade matrix too rapidly (i.e., the share of zeros is too low).[4] Figure 3 shows that the share of zeros decreases as the number of links grows: It goes down very quickly for small values of $N_{\rm tot}$, while it stabilizes for large $N_{\rm tot}$.

In reproducing trade data, we are interested in values of the probability of having a zero entry of A, $P_A[0] \in (0.36,0.52)$. Values in such a range can be obtained by limiting the number of links to be allocated: $N_{\rm tot} \in (55\,000, 130\,000)$. However, this is much smaller than the real number of links observed in the data, which are in the order of $N_{\rm tot} \approx 5\text{–}7 \times 10^6$.

[3]Elsewhere [5,8] we have shown that this setup can effectively replicate the main structural properties of the trade network, such as the skewed connectivity distribution.

[4]Conversely, if we set $w_0 = 0$ and let the country enter with a constant probability α the share of zeros will be too high.

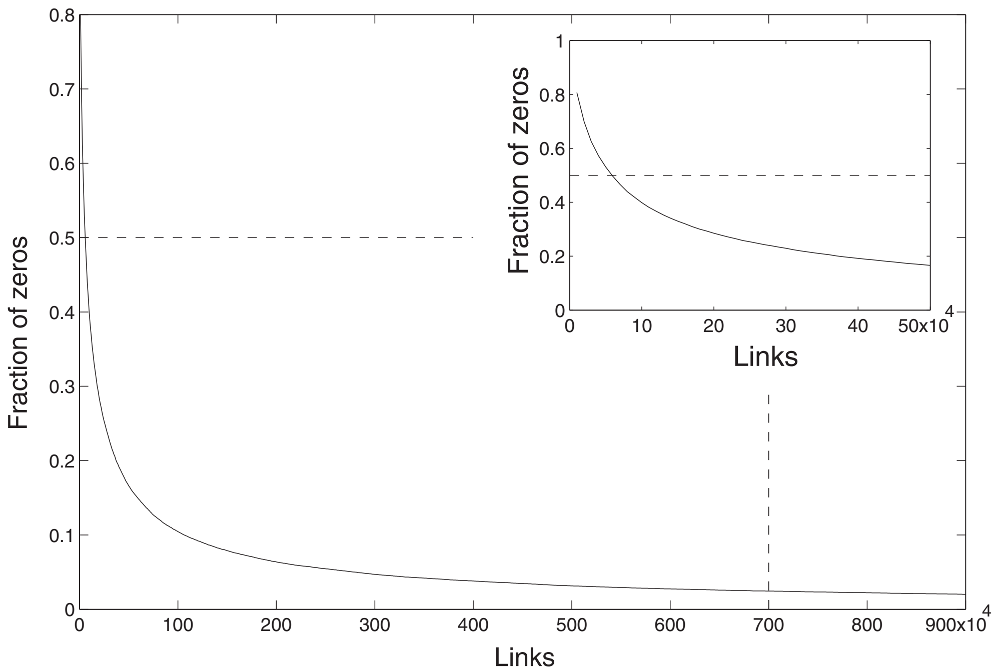

FIG. 3. Probability of missing links in international trade, 189 × 189 country matrix.

To address this issue, we propose a generalized PA mechanism for the growth of multiple networks that consists in allocating products to different subnetworks, i.e., we group products into different categories. More precisely, to implement our method we must decide (a) how to split the total network into product-specific subnetworks, (b) how to determine the number of products to be allocated in each subnetwork, and (c) how to reaggregate subnetworks to obtain the aggregate world trade network. In choosing the different submatrices, we aim at reproducing the allocation dynamics related to different types of products, due, for instance, to their different levels of technological intensity. In this context, products with the lowest complexity are those exported, or generally traded, by all countries, whereas the most sophisticated goods are produced and sold by a small number of countries. This idea is similar to the method of reflections used by Ref. [20] to infer the complexity of a product from the number of countries exporting it and the capabilities of a country from the ubiquity of its export mix: The more products a country exports, the more capabilities it has. The authors describe "each capability as a building block or Lego piece [...], a product is equivalent to a Lego model, and a country is equivalent to a bucket of Legos." Hence, a further connection to our approach comes from the proportionality implied by this view: by exporting more products a country cumulates more capabilities and therefore manages to produce and export even more products. This is exactly the PA mechanism that lies at the heart of our own approach.

Formally, differentiating among goods based on their complexity implies a progressive narrowing of the dimension of the adjacency matrices where we allocate products. Our model generalizes the PA model by differentiating two dimensions: (1) the probability to catch a new trade opportunity for a given product is proportional to the number of links a country already has and (2) the probability to start trading a new product is proportional to the total number of connections a country has across all products. Once we have chosen number and dimensions of the submatrices of the original $N \times N$ matrix describing the trade network, we have to establish the number of products to be allocated in each submatrix. The number of zeros is decreasing with respect to the aggregation operation, as the latter adds items to cells but cannot remove them. This fact offers a first criterion for establishing an upper bound for the number of objects to be allocated in the lowest layer.

Among the possible criteria for choosing number and dimensions of the submatrices and the numbers of products to be allocated in each of them, we use a decomposition method, based on $P(\text{zeros})$ quantiles, keeping (as much as possible) constant the percentage of zeros in the different submatrices. To this aim, we simulate the number of allocations required in order to obtain a given percentage of zeros α in a matrix of a given dimension n (number of countries), $n = 10, \ldots, 189$.

More in detail, we count how many (units of) HS6 products are exported by exactly one country, how many ones by two, and so on (see Fig. 4). In order to obtain a certain percentage of zeros in any matrix, we have to combine the dimension of the matrix and number of objects to be allocated. For example, for filling a 189 × 189 matrix with 50% zeros, we need to allocate about 6×10^4 units. This number is obtained by looking at the number of units of HS6 products exported by exactly x countries, with x ranging between about 160 and 189. By means of such a grouping procedure, we identify overlapping groups of countries with different cardinality. For each group of nations, we count the number of units of HS6 products exported by each country. Such a number coincides with the sum of out-degrees of the countries belonging to the group, where the out-degree of a node is meant as the sum of the weights of the tail end points adjacent to a node.

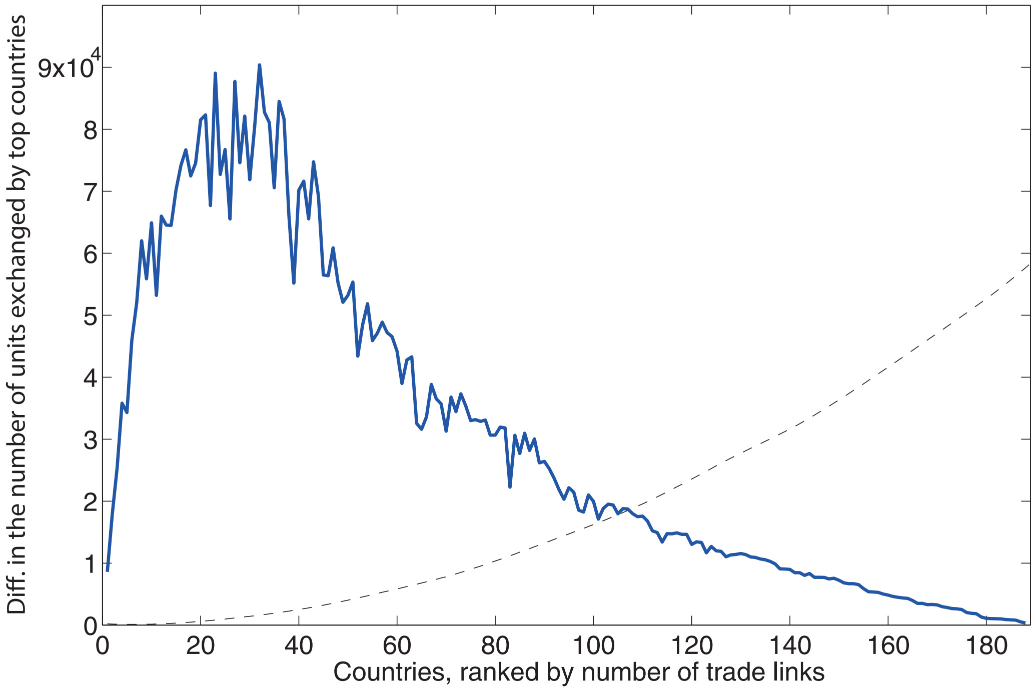

FIG. 4. (Color online) We consider the adjacency matrix A of the world trade web in 2001. There are about 5.7 million entries in the matrix (links) corresponding to different products traded by countries. We sort the matrix by the centrality of countries, as measured by the total number of incoming and outgoing links. Next we count the number of entries in the top-left submatrix of size "dim." The plot reports the number of products added to the network by increasing the size of the submatrix dim, including more peripheral countries. We also report the graphs of $\mathrm{Prod}_\alpha(\mathrm{dim})$, the number of links to be allocated for having a share of zeros $\alpha = 1/2$ in a matrix of size dim.

The idea is the following: Looking at Fig. 4, starting from the right of the horizontal axis, we sum the number of units until such a sum reaches the curve corresponding to the desired share of zeros α. Formally, we have

$$N_1 = N,\ N_{h+1} = \inf\left\{x \middle| \sum_{c=x+1}^{N_h} \mathrm{ex}(c) < \mathrm{Prod}_\alpha(N_h)\right\} \quad (1)$$

and

$$n_h = \sum_{c=N_{h+1}+1}^{N_h} \mathrm{ex}(c),\ h = 1,\ldots,b-1. \quad (2)$$

where $\mathrm{Prod}_\alpha(N_h)$ is the number of links to be allocated for having a share of zeros $\alpha = 1/2$ in a matrix of size N_h and $\mathrm{ex}(c)$ is the number of additional links to be inserted. We obtain a decomposition of the adjacency matrix A into b submatrices $M_1,\ldots,M_b$, whose dimensions are, respectively, $N_1,\ldots,N_b$, and, in each submatrix M_h, we have to allocate n_h links.[5]

The subnetworks must next be aggregated to obtain the world trade matrix. Since we assume that the probability to enter new trade networks is proportional to total connectivity, we aggregate the matrices be means of the operation $\star$, consisting in a sum operation after having ordered columns and rows of the submatrices according to their connectivity. Given that connectivity is strongly correlated across products (see Fig. 1), this aggregation procedure is both theoretically and empirically grounded. Such an ordering concentrates the nonempty entries in the left upper corner and minimizes the probability that a nonempty entry of M_{i+1} is summed to an empty entry of the matrix $M_1 \star \cdots \star M_i$. This method allows us to generate a higher concentration than the classical PA model without entry, which would lead to a too-low share of zeros and to matrices too uniformly filled up. Our aggregation procedure instead makes the most connected countries, i.e., the ones in the left upper corner, benefit from the PA in any technology level where they are active and from a sort of PA across different layers. In other words, such countries "attract" products and, in any market, they "attract" trade opportunities.

Let be $N_1,\ldots,N_b$ the dimensions of the submatrices $M_1,\ldots,M_b$, with b the number of the submatrices and $N_1 > \cdots > N_b$ and $\alpha_1,\ldots,\alpha_b$ the proportions of zeros in any submatrix. We define the aggregated matrix by

$$M = M_1 \star \cdots \star M_b.$$

We denote by $P_M[0](\alpha)$ and $\mathbb{E}_M[0](\alpha)$ the probability of a cell to be empty (missing link) and the expected number of zeros in M, respectively. We provide here some formulas that allow us to compute $P_M[0],\mathbb{E}_M[0]$ in terms of given $b,N_1,\ldots,N_b$, $\alpha_1,\ldots,\alpha_b$. Conversely, these formulas can be used for calibrating the share of zeros $\alpha_1,\ldots,\alpha_b$ (or, for the sake of simplicity, $\alpha = \alpha_1 = \cdots = \alpha_b$), to obtain a given percentage of zeros in the aggregated matrix M. With the following propositions, we provide the analytical expression

[5]Submatrices could be rectangular, whereby few countries export products but all of them import. This extension would complicate the mathematical treatment of the problem without adding much to the current results.

of the function linking the disaggregate share of zeros to the aggregate one.

Proposition 1.

$$\mathbb{E}_M[0](\alpha) = \alpha_1 N_1^2 \prod_{i=2}^{b} \left[1 - (1-\alpha_i)\left(\frac{N_i}{N_1}\right)^2\right]. \quad (3)$$

In this context the probability that a given entry contains a zero can be seen as a binomial probability, therefore it can be simply obtained by its expected value by dividing by the number of trials N_1^2,

$$P_M[0](\alpha) = \alpha_1 \prod_{i=2}^{b} \left[1 - (1-\alpha_i)\left(\frac{N_i}{N_1}\right)^2\right]. \quad (4)$$

The advantage of dealing with expected values in place of probabilities is due to the linearity of the expected value. Equation (4) can be used both to compute $P_M[0](\alpha)$, by assigning $\alpha_1,\ldots,\alpha_b$, and, conversely, to obtain, numerically, the $\alpha_1,\ldots,\alpha_b$ to be assigned in order to get a fixed $P_M[0](\alpha)$. In our application, we consider, for any $i=1,\ldots,b$, $\alpha_i=\alpha$. Notice that $1-(1-\alpha_i)(\frac{N_i}{N_1})^2 < 1$. By replacing $\alpha_i=\alpha$ in Eq. (4) we get $P_M[0](\alpha) < \alpha$. Heuristically, this fact implies that, in order to obtain an aggregated matrix with percentage of zeros $P_M[0](\alpha)$, we have to assign to the disaggregated matrices a higher percentage of zeros α. We provide a formula for the needed percentage α in the following two cases, which is useful to find a suitable α to be inserted in the simulation.

The relation

$$N_i \leqslant (1-\alpha_{i-1})N_{i-1} \text{ for any } i=2,\ldots,b \quad (5)$$

implies that, for any $i=2,\ldots,b$, $(1-\alpha_i)N_i < (1-\alpha_{i-1})N_{i-1}$. In this case, since M_2 is contained in the upper $[(1-\alpha_1)N_1]^2$ block of M_1, any result of the allocations in the sub-matrices $M_2,\ldots,M_b$ does not affect the number of zeros of M_1. Therefore it is sufficient to compute these last ones, amounting to αN_1^2 and trivially giving

$$P_M[0](\alpha) = \alpha. \quad (6)$$

Proposition 2. Let be $\alpha_i=\alpha_1=\alpha$ and $N_i > (1-\alpha)N_{i-1}$ for any $i=2,\ldots,b$. Then

$$P_M[0](\alpha) = 2\alpha - \alpha^2 - \frac{1-\alpha}{N_1^2}\sum_{i=1}^{\bar{b}} \frac{N_i}{2^{i-1}}(N_i - (1-\alpha)N_1). \quad (7)$$

Remark 1. Equation (7) no longer holds if we drop the condition $\alpha_i=\alpha$. In this case, indeed, we are no more able to order the terms $(1-\alpha_i)N_i$ for the different i's.

Actually, objects falling in the upper square occupy an empty entry with a probability greater than 0. In such a case, however, given a greater complexity and variety of situations, we can provide for $P_M[0](\alpha)$ only an upper bound, reflecting on a lower bound for α.

Proposition 3. For any fixed $\alpha \in [0,1]$, let M_1 be the connectivity ordered (square) matrix, with the zeros' percentage α. Let $1-\beta$ be the zeros' percentage in the upper $[(1-\alpha)N_1]^2$ block and $1-\tilde{\beta}$ be the zeros' percentage in the rectangular blocks $(1-\alpha)N_1 \times (N_1-(1-\alpha)N_1)$. Then

$$\begin{aligned} P_M[0](\alpha) \leqslant 1 - \frac{2}{N_1^2}(1-\alpha)\Bigg\{ &\sum_{i=1}^{\bar{b}} (1-\tilde{\beta})^{i-1} N_i^2 \\ &\times \left[\tilde{\beta} + (1-\alpha)\left(\beta - \tilde{\beta}\frac{N_1}{N_i}\right)\right] \\ &+ (1-\alpha)\tilde{\beta}\sum_{i=2}^{\bar{b}} N_i \sum_{h=2}^{i} (1-\tilde{\beta})^{h-2}(1-\beta)^{i-h+1} \\ &\times (N_{i-h+1} - N_{i-h+2})\Bigg\}. \end{aligned} \quad (8)$$

An analog of Eq. (6), when $\beta \neq 1 \neq 2\tilde{\beta}$, can be obtained by dropping in Eq. (8) the part corresponding to the nonzeros added in the rectangular blocks of M_1 and modifying the part corresponding to its upper square block, in the light of the condition $(1-\alpha_i)N_i < N_i < (1-\alpha_{i-1})N_{i-1}$.

IV. PROBABILISTIC ASPECTS AND ANALYTICAL PROPERTIES OF THE ALLOCATION PROCESS

In this section we want to study some analytical properties of the processes involved in the simulation. To this purpose, we give, first, some definitions that will be useful in the following.

Definition 1 (Sufficient statistic). Let $\mathbf{X}$ a sample on $(\Omega,\mathfrak{B},\mathbb{P})$, taking values in $\mathfrak{X}$, and let $\mathcal{F} = \{f_{\mathbf{X}}(\cdot,\theta) : \theta \in \Theta\}$ be a family of probability densities for $\mathbf{X}$, depending on the parameter $\theta \in \Theta$. A statistic $T = T(\mathbf{X})$ is *sufficient* if two functions g,h exist, such that, for any $\theta \in \Theta$ and for almost any $\mathbf{x} \in \mathfrak{X}$,

$$f_{\mathbf{X}}(\mathbf{x},\theta) = h(\mathbf{x})g(T(\mathbf{x}),\theta).$$

Intuitively a sufficient statistic is a function of data containing all the information they can give. For further details and examples, see, e.g., Ref. [22].

Definition 2 (Counting process). A stochastic process $\{N_t\}_{t\in[0,+\infty)}$ is a *counting process* if, for any t, N_t satisfies the following properties:

(i) $N_t \in \mathbb{N}$;

(ii) $P(N_s \leqslant N_t) = 1$ for any $s < t$;

(iii) for any $s<t$, $N_t - N_s$ is the number of events occurred during the interval $(s,t]$.

A counting process is said to be *simple* if $N_0 = 0$ and

$$\lim_{\delta\to 0} P(N_{t+\delta} - N_t > 1) = 0.$$

Definition 3 (Markov process). A stochastic process $\{X_t\}_{t\in[0,+\infty)}$ with discrete state space E is a *Markov process* if, for any $0 \leqslant s_1,\ldots,s_k < s < t$ and for any $i_1,\ldots,i_k,i,j \in E$,

$$\begin{aligned} P(X_t = j | X_s = i, X_{s_1} = i_1,\ldots,X_{s_k} = i_k) \\ = P(X_t = j | X_s = i). \end{aligned} \quad (9)$$

Definition 4. A simple counting process satisfying Eq. (9) (Markov property) is called a *pure birth* process.

Let $\mathfrak{M}_t = (\mathfrak{M}_t^{(1)},\ldots,\mathfrak{M}_t^{(N)}) = (m_t^{(1)},\ldots,m_t^{(N)})$ be the observed configuration of the countries' masses at time $t \in \mathbb{N}$. $\mathfrak{M}_t^{(c)}$ is the random variable counting the number of products allocated to country c until time t. The process $\{\mathfrak{M}_t^{(c)}\}_{t\in\mathbb{N}}$, by construction, is a simple counting process (i.e., it cannot have multiple jumps at a time). The value $m_t^{(c)}$ is a realization of the random variable $\mathfrak{M}_t^{(c)}$. According to the implemented procedure, we suppose $m_0^{(c)}=1$ for any $c\in\{1,\ldots,N\}$.

Let us consider trade flows directionally, that is, let us consider only export or import flows, that is, $\mathfrak{M}_t^{(i)}=N_c^{(i)}(t)$ or $\mathfrak{M}_t^{(i)}=N_r^{(i)}(t)$. In such a way, assigning a product to the country c means inserting it into some entry on the c-th row of the exchange matrix. Let $P_c(+1,t)$ be the probability that, at the instant t (i.e., at the t-th iteration of the procedure), the country c gets the newly inserted product and let $P_c(+k,[t_1,t_2])$ denote the probability that the country c gets k new products in the time interval $[t_1,t_2]$.

Proposition 4. For any c and t,

$$P_c(+1,t)=P_c(+1,1)=\frac{1}{n}. \tag{10}$$

Corollary 1. $\{\mathfrak{M}_t^{(c)}\}_{t\in\mathbb{N}}$ has stationary increments.

Remark 2. Since at any instant exactly one product is drawn, for any t and for any $k>1$, $P_c(+k,t)=0$. Consequently, $P_c(+k,[s,t])>0$ only if $k\leqslant t-s+1$.

Let us assume now that a general initial configuration $\mathfrak{M}_0$ is allowed. In general, $P_c(+1,t)$ depends now on the country c for which it has to be computed and on $\mathfrak{M}_{t-1}$, only through $m_{t-1}^{(c)}$. We have

$$P_c(+1,1|\mathfrak{M}_0)=\frac{m_0^{(c)}}{\sum_{j=1}^n m_0^{(j)}}.$$

Furthermore, the following can be proven.

Proposition 5. For any $s<t$,

$$P_c(+1,t|\mathfrak{M}_s)=P_c\big(+1,t|\mathfrak{M}_s^{(c)}\big)=\frac{m_s^{(c)}}{\sum_{j=1}^n m_s^{(j)}}$$
$$=\frac{nm_s^{(c)}}{\sum_{j=1}^n m_s^{(j)}}P(+1,t).$$

Remark 3. The previous results hold by replacing in them $P_c(+1,t)$ with $P_c(+1,t|\mathfrak{M}_s)$, that is, by multiplying $P_c(+1,t)$ by $\frac{nm_s^{(c)}}{\sum_{j=1}^n m_s^{(j)}}$, where $\mathfrak{M}_s$ is the last observed mass configuration.

Corollary 2. For any k,t,s,s' such that $k<t-s'$, $s<s'$,

$$P(+k,[s',t]|\mathfrak{M}_s)=\prod_{i=0}^{k-1}\frac{m_s^{(c)}+i}{\sum_{j=1}^N m_s^{(j)}+i}\times\prod_{i=k}^{t-s'}\frac{\sum_{j=1,j\neq c}^N m_s^{(j)}+i-k}{\sum_{j=1}^N m_s^{(j)}+i}. \tag{11}$$

Remark 4. $\{\mathfrak{M}_t^{(c)}\}_{t\in\mathbb{N}}$ has no independent increments. In fact, $P(\mathfrak{M}_{t+s}^{(c)}-\mathfrak{M}_s^{(c)}=k|\mathfrak{M}_s^{(c)}=h)=\frac{h}{n+s}P(+k,[1,t])\neq P(+k,[1,t])=P(\mathfrak{M}_{t+s}^{(c)}-\mathfrak{M}_s^{(c)}=k)$. Therefore

$$P\big(\mathfrak{M}_{t+s}^{(c)}-\mathfrak{M}_s^{(c)}=k,\mathfrak{M}_s^{(c)}=h\big)\neq P\big(\mathfrak{M}_{t+s}^{(c)}-\mathfrak{M}_s^{(c)}=k\big)P(\mathfrak{M}_s^{(c)}=h).$$

Theorem 1. For any $c\in\{1,\ldots,N\}$, $\{\mathfrak{M}_t^{(c)}\}_{t\in\mathbb{N}}$ is a homogeneous Markov process, i.e.,

$$P_c\big(+1,t|\mathfrak{M}_{t-1}^{(c)},\ldots,\mathfrak{M}_0^{(c)}\big)=P_c\big(+1,t|\mathfrak{M}_{t-1}^{(c)}\big).$$

Theorem 2. For any $c\in\{1,\ldots,N\}$, $\{\mathfrak{M}_t^{(c)}\}_{t\in\mathbb{N}}$ is a submartingale, i.e.,

$$\mathbb{E}\big[\mathfrak{M}_{t+1}^{(c)}|\mathfrak{M}_t^{(c)},\ldots,\mathfrak{M}_0^{(c)}\big]\geqslant\mathfrak{M}_t^{(c)}.$$

Some consequences of the formulas provided in the theorems and propositions above concern the probability of assigning to a same country products in a certain number of categories, i.e., allocating products in a same row of a certain number of submatrices, as well as the analytical expression for the zeros' probability.

Let us again start with a uniform configuration (one product per country in any matrix). The probability that one product for each of the r lowest levels of technology is assigned to a same fixed country is

$$\left(\frac{1}{N}\right)\left(\frac{1}{N_2}\frac{N_2}{N}\right)\left(\frac{1}{N_3}\frac{N_2}{N}\frac{N_3}{N_2}\right)\cdots=\frac{1}{N^r}, \tag{12}$$

where, in any factor $\frac{1}{N_l}\frac{N_2}{N}\frac{N_3}{N_2}\cdots\frac{N_l}{N_{l-1}}$, $\frac{1}{N_l}$ is the probability that, in the submatrix M_l, the product is assigned to the given

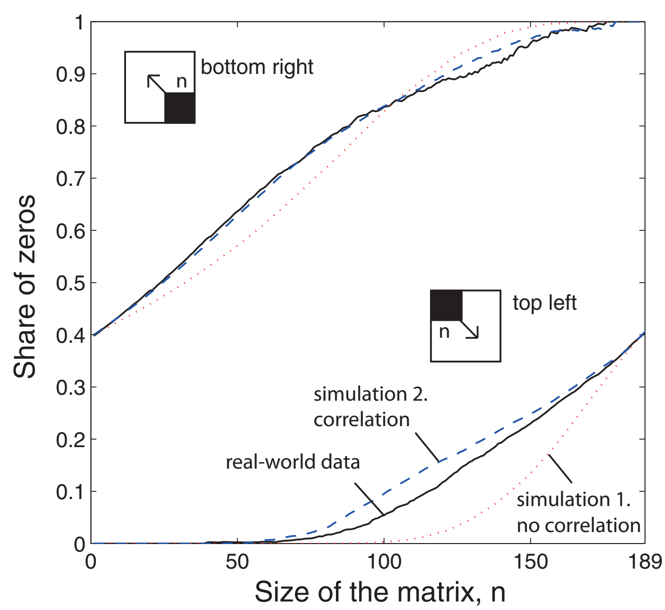

FIG. 5. (Color online) Share of zeros in the real-world data (solid lines) and simulation results (dashed and dotted lines). Real and simulated adjacency matrices have been reordered according to the centrality of countries (most central nodes are at the bottom-left of the matrix). We report the share of zeros in the squared matrices of size n at the top-left and bottom-right of the adjacency matrix. The dotted line shows the results of the simulation when there is no correlation among products. In such a case the total share of zeros is well reproduced, but large departures are detected in the sparsity pattern of the matrix both in the core and the periphery of the network. A far better result is obtain in the second simulation regime (dashed lines) when we impose the same cross-product correlation as in the real data. In such a case, both the share of zeros and the sparsity pattern are well replicated. $\alpha=1/2$.

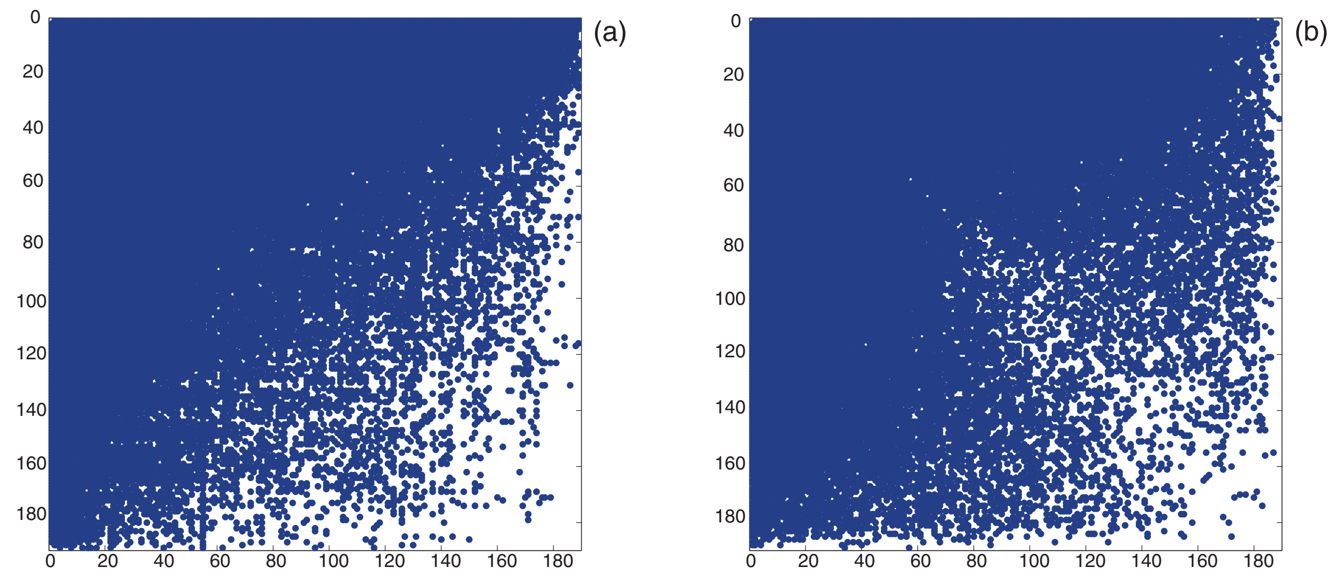

FIG. 6. (Color online) Adjacency matrix of the real-world trade network in the year 2001 (a), and the simulated network (b). The simulated network is based on our model with the same correlation level as in the real data (simulation 2). The two networks have a similar sparsity structure but the simulated one is slightly less disassortative than the real one [−0.30 (real network) vs −0.24 (simulated network)].

country, while any term $\frac{N_h}{N_{h-1}}$ is the probability that that country is active in the technology level h, given that it is active in the technology level $h-1$. Since $\frac{1}{N_l}\frac{N_2}{N}\frac{N_3}{N_2}\cdots\frac{N_l}{N_{l-1}}=\frac{1}{N}$, Eq. (12) is also the probability that one product for each of r given levels of technology is assigned to the same country.

The analytical expression for the zeros' probability can be computed in two different (consistent) ways.

At the first step of the allocation procedure,

$$\begin{aligned}P(K_{ij}=0)&\\ &=[\text{Prob(selecting a row } \neq i)+\text{Prob(selecting row } i)\\ &\quad\times \text{Prob(selecting a column } \neq j|\text{row } i)]P(K_{ij}(0)=0)\\ &=\left(\frac{N-1}{N}+\frac{1}{N}\frac{N-2}{N-1}\right)\left(1-\frac{1}{N}\right).\end{aligned}$$

Such a probability can also be computed as the one of the complementary events of observing a product to be allocated in the entry (i,j), i.e.,

$$P(K_{ij}=0)=\left(1-\frac{1}{N}\right)\left(1-\frac{1}{N}\frac{1}{N-1}\right),$$

$$P(K_{ij}=0|[0,t])=\left(1-\frac{1}{N}\right)\left(1-\frac{1}{N}\frac{1}{N-1}\right)^t.$$

V. SIMULATION RESULTS

Since it is not possible to invert Eq. (7), in order to compute a suitable α for our simulation, we have to (a) start from an arbitrary value of α, representing the percentage of zeros in each submatrix; (b) find a suitable number and dimensions of the submatrices given α; (c) check which condition, between $N_i=(1-\alpha_{i-1})N_{i-1}$ and $N_i>(1-\alpha)N_{i-1}$ for any $i=2,\ldots,b$, our data satisfy; (d) compute the share of zeros, as it results from Eq. (7); and (d) accept or not such a value and possibly repeat the procedure.

Applying it to trade data for 2001, we obtain a decomposition in 156 submatrices, each with a percentage of zeros $\alpha=0.575$. The actual share of zeros in the aggregate matrix amounts to 0.437, instead of the expected $P_M[0](0.575)=0.4$. This is a consequence of the fact that, contrary to what happens in the data, in the simulations the connectivity reordering concentrates nonzero entries in the upper square of each matrix but is not able to fully move zeros in other blocks. However, as the simulated data show, we can assume in a good approximation $\beta=1$, $\tilde{\beta}=1/2$ and use, for a lower bound on α, Eq. (7). Actually, Eq. (7) provides an imprecise lower bound for α but a locally optimal solution: The share of zeros further decreases to 0.34 for $\alpha=0.625$, while, for $\alpha=0.75$, it rises again to 0.35.[6]

In our simulation exercises we consider two scenarios. The first assumes the absence of any within country cross-correlation in trading different products (simulation 1). As shown in Fig. 5 this approach is able to reproduce the share of missing links but fails in reproducing the sparsity pattern of the trade network. Therefore we run another simulation exercise in which we apply the same correlation among products as observed in the data. Figure 5 shows that this is indeed a crucial aspect to generate the same sparsity structure of the real network. This evidence is further confirmed by inspecting

---

[6]We also compared the two solutions in terms of the deviation of the resulting connectivity distribution from the empirical one. In terms of relative frequencies, y_r^i, $y_r^i(0.575)$, $y_r^i(0.75)$, computed on a histogram with 100 bins, we have

$$\sum_{i=1}^{189}|y_r^i-y_r^i(0.575)|=\sum_{i=1}^{189}|y_r^i-y_r^i(0.75)|=66.$$

The reason for such a similarity is that the smaller number of objects allocated in each submatrix for $\alpha=0.75$ is counterbalanced by a more refined partition of the matrix (173 instead of 156 submatrices).

the adjacency matrices in Fig. 6. All in all, our model does a good job of replicating some of the main structural properties of the world trade data, especially if we consider that we do not account for some crucial features of the real-world trade web, such as distance among countries. However, we also noticed that our model generates a network which is less disassortative than the real one. Further work is needed to fully resolve this aspect. In particular, one major limitation of our approach is that countries have the same likelihood to start importing and exporting a new product. As is evident in Fig. 2, this is not the case in real data: To export a large number of products is clearly more difficult than to import them. Better results could be obtained by differentiating import from export in future work.

VI. CONCLUDING DISCUSSION

In this paper we presented a generalized version of the PA model [14] which is able to account for the sparsity structure of the world trade network. Our model is based on the idea that almost every country takes part in the trade of low-tech products, while only a few of them have the capabilities to export sophisticated goods. We define a lower bound for the share of zeros in trade networks by considering no correlation between countries' capabilities in trading different products. However, since we know that trade in different products is correlated [20], we also consider more realistic assumptions about the product space. Our approach is able to replicate the sparsity structure of the trade network by combining a generalized version of the preferential attachment model to the case of multiple networks with more realistic assumptions about the probability to jointly trade related products. From a computational point of view we also contribute to the existing literature by providing a new method to generate large sparse networks. Indeed, our methodology allows us to generate in parallel multiple (product-specific) networks, thus reducing the computational time to simulate the evolution of large trade networks. The decomposition greatly reduces the complexity of the procedure and allows for a reaggregation of different layers to obtain the desired aggregate properties. Different assumptions about cross-layer correlations can be implemented by modifying the aggregation function.

ACKNOWLEDGMENTS

The authors acknowledge funding from the Italian Ministry of Education, University and Research (MIUR) via the projects Tradenetworks (PRIN 2009), CRISIS Lab, and ViWaN (Virtual Water Networks, FIRB 2012). M.R. acknowledges funding from the Multiplex FP7 project (Foundational Research on MULTIlevel comPLEX networks and systems).

APPENDIX A: AN OVERVIEW ON DECOMPOSITION AND RE-AGGREGATION METHODS

The choice of both the decomposition and reaggregation methods is not unique. Our choice derives from having experimented further possible methods.

The most natural reaggregation method consists in simply summing up the entries of the different submatrices. However, we should apply the matrices' sum operation to any possible permutation of the entries of the submatrices. In our case, we would obtain $N^2! \times N_2^2! \times \cdots \times N_b^2!$ different ways of summing up. This process would have huge computational costs. However, all these situations are bounded by two of them: the reaggregation method representing the greatest positive dependence among the different levels of technology and the one representing independence.

The first is just the one based on the connectivity ordering, which we treat in detail in the paper. The second method is the most randomized, consisting in randomly permuting rows and columns of each M_i and them summing them up. We expect that such a method underestimates both the number of zeros and the cells' values and that therefore it can provide a lower bound for such quantities. Actually, it gives a trivial lower bound in that the zeros completely vanish and the values are uniformly distributed in the cells. This fact is also confirmed by the shape of the resulting distribution of export connectivity, which is much less concentrated than the empirical one (see Fig. 7).

Remark 5. We did not consider instead, as a lower bound, a summing rule representing negative dependence among different levels of technology. In fact it is not possible that more than two variables are all pairwise negative dependent. This fact is also confirmed by the trivial lower bound obtained in the case of independence.

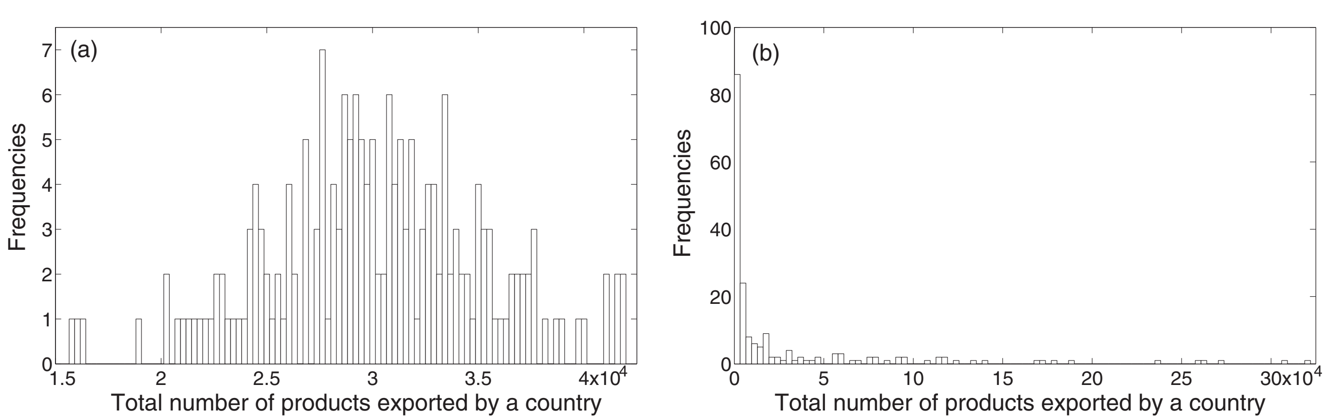

FIG. 7. Export connectivity distribution obtained by reaggregation via randomly permuting rows and columns of each M_i (a) and the empirical export connectivity distribution (b).

The bad performance of the random permutation method is also confirmed by the fact that, in calibrating α, summing the submatrices obtained by the simulation without any transformation, that is, by considering $\star = +$, we obtain $P_M[0](\alpha) < \frac{\alpha}{2}$. This means that, even by taking a high value for α, we obtain a too-low $P_M[0]$.

One of the criteria adopted in the choice of the decomposition method is a kind of "connectivity-based" distance. Indeed, we could consider the decomposition in $b = N$ sub-matrices in such a way that $N_h = b + 1 - h$, $h = 1,\ldots,b$ and compute n_h by (2). We cannot use Eq. (7) for establishing whether such a uniformly spaced decomposition performs better or worse than the one we have adopted, because it holds under the hypothesis, no more warranted, of $\alpha_h = \alpha$ for any $h = 1,\ldots,b$. Furthermore, computing such α_h's would require simulation of the allocation in all the N matrices.

The cheapest procedure we can adopt is to simulate the allocation in M_1, reorder it, and compute $(1-\alpha_1)N_1$ and then continuing the simulation to the matrix $M_{\bar{b}}$ such that $N_{\bar{b}} = b + 1 - \bar{b} \leqslant (1-\alpha_1)N_1$. We can now stop the procedure, since, from this point onward, the probability of filling a zero entry is nil.

Another decomposition, made attractive by the use of the reaggregation operator $\star$, according to condition (5), consists in choosing $N_2 \leqslant (1-\alpha_1)N_1$. Since the behavior of the allocation in the upper block $[(1-\alpha_1)N_1]^2$ does not affect the number of zeros of the aggregate matrix, we could even choose $b = 2$. In this case, the zeros' probability is 0.3351, even if we start with $\alpha_1 = 0.55$. When we simulate the allocation in M_1, in fact, the number of products is too low to guarantee $\tilde{\beta} \approx \beta$, so the allocation in M_2 affects the number of zeros by filling up all the zeros in the upper square block of M_1. Furthermore, the resulting connectivity distribution would not be as smooth as the empirical one. More precisely, in terms of relative frequencies, the deviation from the empirical connectivity distribution is 94; in terms of absolute frequencies, it amounts to 5 676 742; hence it is worse than the ones resulting from our simulation method of decomposition.

We could make it smoother by choosing $b \in \{3,\ldots,N_2+1\}$ and/or controlling the (decreasing) sequence of α_h's. However, the decomposition associated with each b would be not unique. Considering, for each b, all the possible strictly decreasing sequences $\{N_3,\ldots,N_b\}$, $N_3 < N_2$, $N_b \geqslant 1$, is too expensive. We could choose the decomposition obtained by imposing some regularity conditions on the sequence of the α_h's.

We should solve the following nonlinear system of equations:

$$\begin{cases} N_2 = \lfloor (1-\alpha_1)189 \rfloor, & \\ \alpha_i = \alpha_1 \frac{b-i}{b-1}, & i = 2,\ldots,b-1, \\ N_i = N_2 - \left\lfloor N_2 - \frac{N_b}{b-2} \right\rfloor (i-2), & i = 3,\ldots,b-1, \\ n_i = \mathrm{Prod}_{\alpha_i}(N_i), & i = 1,\ldots,b-1, \\ n_b = N_{\mathrm{tot}} - \sum_{i=1}^{b-1} n_i, & \end{cases} \tag{A1}$$

where $\mathrm{Prod}_\alpha(\mathrm{N_i})$ is the number of units of products to be allocated in a $N_i \times N_i$ matrix, needed to have a probability of zeros equal to α, while N_{tot}, and α_1 are exogenously given. We also need to fix N_b.

In spite of the formalization we provide, the choice of such a partition is not unique. In fact, in order to bring the number of unknowns to be equal to the number of equations, some of the conditions we impose in Eq. (A1) are arbitrary.

Therefore, we test only the two cases associated with a unique decomposition, $b = 2$ and $b = N$, and the further relevant one $b = N_2 + 1$, where we set $N_2 = \lfloor (1-\alpha_1)N_1 \rfloor$. As a consequence of this position, the decomposition is unique.

APPENDIX B: PROOFS

Proposition 1. Let us separately consider the matrices $M_1,\ldots,M_b$. The expected number of zeros any M_i contains is $\alpha_i N_i^2$ and therefore any M_i contains $(1-\alpha_i)N_i^2$ nonzero entries. Aggregating by using the coordinatewise matrices' sum plus (i.e., without reordering in advance the elements of the submatrices) the first two matrices, M_1, M_2, any nonzero entry of M_2 has a probability α_1 to occupy a zero entry of M_1. The expected number of zeros of $M_1 + M_2$ is

$$\begin{aligned} \mathbb{E}_{M_1+M_2}[0](\alpha) &= \alpha_1 N_1^2 - (1-\alpha_2)N_2^2\alpha_1 \\ &= \alpha_1 N_1^2\left[1 - (1-\alpha_2)\left(\frac{N_2}{N_1}\right)^2\right]. \end{aligned} \tag{B1}$$

A nonzero entry of M_3 has a probability $\frac{\mathbb{E}_{M_1+M_2}[0](\alpha)}{N_1^2}$ to occupy a zero entry of $M_1 + M_2$, so

$$\begin{aligned} \mathbb{E}_{M_1+M_2+M_3}[0](\alpha) &= \alpha_1 N_1^2\left[1 - (1-\alpha_2)\left(\frac{N_2}{N_1}\right)^2\right] \\ &\quad\times\left[1 - (1-\alpha_3)\left(\frac{N_3}{N_1}\right)^2\right]. \end{aligned} \tag{B2}$$

By iteration, we obtain the thesis. ■

Proposition 2. Let us consider the increasing sequence $\mathfrak{N} = \{(1-\alpha)N_b,(1-\alpha)N_{b-1},\ldots,N_b,\ldots,(1-\alpha)N_1,N_{\bar{b}}\ldots,N_2,N_1\}$, where $\bar{b} = \max\{i|N_i > (1-\alpha)N_1\}$ and the partition of the $N_1 \times N_1$ matrix determined by the Cartesian product $\mathfrak{N}^2$.

All the squares contained in the upper square $(1-\alpha)N_1 \times (1-\alpha)N_1$ are nonzero with probability 1. All the squares contained in the lower square $\alpha N_1 \times \alpha N_1$ are zero with probability 1.

We focus on the squares generated by the Cartesian product $\{(1-\alpha)N_b,(1-\alpha)N_{b-1},\ldots,N_b,\ldots,(1-\alpha)N_1\} \times \{(1-\alpha)N_1,N_{\bar{b}}\ldots,N_2,N_1\}$. The rectangle $\alpha N_1 \times (1-\alpha)N_1$, belonging only to the largest matrix, M_1, has nonzero entries with probability 1/2. The rectangle $N_2 - (1-\alpha)N_1 \times (1-\alpha)N_2$ is also contained in the matrix M_2. We already observe therein a probability 1/2 of nonzero entries due to the allocation of objects in the matrix M_1. The allocation process in the matrix M_2 generates in such a rectangle a nonzero with probability $(1/2)^2$. By iteration, we obtain that the number of nonzeros of the aggregate matrix M is given by

$$(1-\alpha)^2 N_1^2 + (1-\alpha)\sum_{i=1}^{\bar{b}} \frac{N_i}{2^{i-1}}(N_i - (1-\alpha)N_1).$$

■

Proposition 3. We divide M_1 into two regions: the square where the probability of nonzeros is β and the rectangles where the probability of nonzeros is $\tilde{\beta}$.

In each rectangle, the number of nonzeros added at any step i, by adding the (ordered) matrix M_i, is proportional to the number of entries of the rectangle, amounting to $(1-\alpha)N_i[N_i-(1-\alpha)N_1]$; each of these entries is "filled" with probability $\tilde{\beta}$ conditionally on this entry remaining empty to the i-th step; this conditioning event has probability $(1-\tilde{\beta})^{i-1}$. The already-filled entries have been counted recursively at the previous steps.

The behavior in the square block is more complicated. The submatrices M_i, $i=2,\dots,\bar{b}$, contain the block and each of them adds, in the (nested) square $[(1-\alpha)N_i]^2$, nonzero entries with probability β conditionally on this entry having remained to the i-th step; this conditioning event has probability $(1-\beta)^{i-1}$.

From this block, we divide the rectangular block of each matrix M_i, where the nonzeros' percentage is $\tilde{\beta}$, in strips of width $(1-\alpha)(N_{i-h+1}-N_{i-h+2})$, $i=2,\dots,\bar{b}$, $h=2,\dots,i$. In each strip, the nonzeros probability is $\tilde{\beta}(1-\tilde{\beta})^{h-2}(1-\beta)^{i-h+1}$. The superposition of the submatrices $M_i, i=2,\dots,\bar{b}$ increases the number of nonzeros in the rectangular blocks of

$$2\tilde{\beta}(1-\alpha)^2\sum_{i=2}^{\bar{b}}N_i\sum_{h=2}^{i}(1-\tilde{\beta})^{h-2}(1-\beta)^{i-h+1}\times(N_{i-h+1}-N_{i-h+2}).$$

By summarizing the amounts of zeros in the different regions and dividing by the total number of entries of the aggregate matrix, we obtain Eq. (8).

For $i,j=\bar{b}+1,\dots,b$, we no longer know the relation between N_j and $(1-\alpha)N_i$ and therefore are no longer able to compute the expected number of added nonzeros. Thus Eq. (8) comes from an underestimation of the nonzeros' probability, that is, an overestimation of the zeros' probability, and therefore it gives a lower bound for α. ■

Proposition 4. Since allocations are proportional to the initial masses, $P_c(+1,1)=\frac{1}{n}$ obviously follows.

We prove the second equality, $P_c(+1,t)=P_c(+1,1)$, by induction.

For $t=2$, Eq. (10) holds. In fact,

$$\begin{aligned}P_c(+1,2)&=P_c(+1,2|+0,1)P_c(+0,1)+P_c(+1,2|+1,1)\\&\quad\times P_c(+1,1)=\frac{1}{n+1}\frac{n-1}{n}+\frac{1}{n}\frac{2}{n+1}=\frac{1}{n}.\end{aligned}$$

Let us now suppose Eq. (10) to hold for $t-1$ and let us prove it for t.

We want to use the total probabilities formula by conditioning on events of the kind

$$\bigcap_{s=1}^{t-2}\{(+\omega_s,s)\},$$

where $\omega_s\in\{0,1\}$, $\omega_0=1$ for all the countries. We have a different event for each different disposition $\omega=(\omega_1,\dots,\omega_{t-2})$, amounting to a number of 2^{t-2}. Actually we are interested in the sufficient statistic of ω, $\mathcal{S}(\omega)=\sum_{s=1}^{t-2}\omega_s$, consisting in the number of 1's contained in the vector ω. In other words, it is not important to know the times when a country gets a product, i.e., it is not relevant the product's age but only the mass of the country. We will condition then on events of the kind $\{\mathcal{S}(\omega)=v\}$. By the inductive hypothesis,

$$P(\mathcal{S}(\omega)=v)=\binom{t-2}{v}\frac{1}{n^v}\left(\frac{n-1}{n}\right)^{t-2-v}$$

and

$$\begin{aligned}P(+1,t-1)&=\sum_{v=0}^{t-2}P(+1,t-1|\mathcal{S}(\omega)=v)P(\mathcal{S}(\omega)=v)\\&=\sum_{v=0}^{t-2}\binom{t-2}{v}\frac{1}{n^v}\left(\frac{n-1}{n}\right)^{t-2-v}\\&\quad\times\frac{v}{n+t-2}=\frac{1}{n},\end{aligned}$$

$$\begin{aligned}P(+1,t)=&\sum_{v=0}^{t-2}[P(+1,t|+0,t-1,\mathcal{S}(\omega)=v)\\&\times P(+0,t-1|\mathcal{S}(\omega)=v)+P_c(+1,t|+1,t-1,\\&\times\mathcal{S}(\omega)=v)P(+1,t-1|\mathcal{S}(\omega)=v)]P(\mathcal{S}(\omega)=v)\end{aligned}$$

$$\begin{aligned}&=\sum_{v=0}^{t-2}\binom{t-2}{v}\frac{(n-1)^{t-2-v}}{n^{t-2}}\left[\frac{v+1}{n+t-1}\right.\\&\quad\left.\times\frac{v}{n+t-2}+\frac{v}{n+t-1}\left(1-\frac{v}{n+t-2}\right)\right]\\&=P(+1,t-1)=\frac{1}{n}.\end{aligned}$$

Hence $P_c(+1,t)=P_c(+1,1)$. ■

Corollary 1. It follows from Proposition 4. Notice that, for any c, $\mathfrak{M}_t^{(c)}=\sum_{s=1}^{t}\omega_s$ is a sufficient statistic of the history until time t of the attributions of products to the country c,

$$P\left(\mathfrak{M}_t^{(c)}=k+1\right)=P_c(+k,[1,t])=\binom{t}{k}\frac{1}{n^k}\left(1-\frac{1}{n}\right)^{t-k},$$

$$\begin{aligned}P_c(+k,[s+1,t+s])&=P\left(\mathfrak{M}_{t+s}^{(c)}-\mathfrak{M}_s^{(c)}=k\right)\\&=P\left(\sum_{q=1}^{t+s}\omega_q-\sum_{q=1}^{s}\omega_q=k\right)\\&=P\left(\sum_{q=s+1}^{t+s}\omega_q=k\right)=\binom{t+s-(s+1)+1}{k}\frac{1}{n^k}\\&\quad\times\left(1-\frac{1}{n}\right)^{t-k}.\end{aligned}$$

■

Theorem 1.

$$\begin{aligned}P_c\left(+1,t|\mathfrak{M}_{t-1}^{(c)},\dots,\mathfrak{M}_0^{(c)}\right)&=P_c\left(+1,t|m_{t-1}^{(c)},\dots,m_0^{(c)}\right)\\&=\frac{P_c\left(m_{t-1}^{(c)}-m_{t-2}^{(c)}+1,[t-1,t]|m_{t-2}^{(c)},\dots,m_0^{(c)}\right)}{P_c\left(m_{t-1}^{(c)}-m_{t-2}^{(c)},t-1|m_{t-2}^{(c)},\dots,m_0^{(c)}\right)}.\end{aligned}$$

By iteration, we obtain

$$\frac{P_c\big(m_{t-1}^{(c)} - m_0^{(c)} + 1,[1,t]|m_0^{(c)}\big)}{P_c\big(m_{t-1}^{(c)} - m_0^{(c)},[1,t-1]|m_0^{(c)}\big)}.$$

By Corollary 2, we get

$$\frac{\prod_{i=0}^{m_{t-1}^{(c)}-m_0^{(c)}} \frac{m_0^{(c)}+i}{\sum_{j=1}^N m_0^{(j)}+i} \prod_{i=m_{t-1}^{(c)}-m_0^{(c)}+1}^{t-1} \frac{\sum_{j=1,j\neq c}^N m_0^{(j)}+i-m_{t-1}^{(c)}-1+m_0^{(c)}}{\sum_{j=1}^N m_0^{(j)}+i}}{\prod_{i=0}^{m_{t-1}^{(c)}-m_0^{(c)}-1} \frac{m_0^{(c)}+i}{\sum_{j=1}^N m_0^{(j)}+i} \prod_{i=m_{t-1}^{(c)}-m_0^{(c)}}^{t-2} \frac{\sum_{j=1,j\neq c}^N m_0^{(j)}+i-m_{t-1}^{(c)}+m_0^{(c)}}{\sum_{j=1}^N m_0^{(j)}+i}}.$$

By changing the index in the product in the denominator, we finally obtain

$$\frac{m_{t-1}^{(c)}}{\sum_{j=1,j\neq c}^N m_0^{(j)} + m_{t-1}^{(c)}} \times \frac{\sum_{j=1,j\neq c}^N m_0^{(j)} + m_{t-1}^{(c)}}{\sum_{j=1}^N m_0^{(j)} + t - 1} = \frac{m_{t-1}^{(c)}}{\sum_{j=1}^N m_{t-1}^{(j)}}.$$

■

Theorem 2. First, in view of the Markov property,

$$\mathbb{E}\big[\mathfrak{M}_{t+1}^{(c)}|\mathfrak{M}_t^{(c)},\ldots,\mathfrak{M}_0^{(c)}\big] = \mathbb{E}\big[\mathfrak{M}_{t+1}^{(c)}|\mathfrak{M}_t^{(c)}\big].$$

Therefore, we can more easily compute

$$\mathbb{E}\big[\mathfrak{M}_{t+1}^{(c)}|\mathfrak{M}_t^{(c)}\big] = \mathfrak{M}_t^{(c)} P\big(+0,t+1|\mathfrak{M}_t^{(c)}\big) + \big(\mathfrak{M}_t^{(c)}+1\big) P\big(+1,t+1|\mathfrak{M}_t^{(c)}\big) = \mathfrak{M}_t^{(c)} + P\big(+1,t+1|\mathfrak{M}_t^{(c)}\big) > \mathfrak{M}_t^{(c)}.$$

■

[1] R. Armenter and M. Koren, Am. Econ. Rev. **104**, 2127 (2014).
[2] R. Baldwin and J. Harrigan. National Bureau of Economic Research, Working Paper 13214, 2007 (unpublished), http://www.nber.org/papers/w13214.
[3] W. Easterly, A. Reshef, and J. Schwenkenberg. World Bank, Policy Research Working Paper Series 5081, 2009 (unpublished), http://elibrary.worldbank.org/doi/pdf/10.1596/1813-9450-5081.
[4] E. Helpman, M. Melitz, and Y. Rubinstein, Q. J. Econ. **123**, 441 (2008).
[5] M. Riccaboni and S. Schiavo, New J. Phys. **12**, 1 (2010).
[6] C. A. Hidalgo, B. Klinger, A.-L. Barabasi, and R. Hausmann, Science **317**, 482 (2007).
[7] T. Chaney, Am. Econ. Rev. (to be published).
[8] M. Riccaboni and S. Schiavo (2014), J. Complex Networks 2014, doi:10.1093/comnet/cnu025.
[9] W. Easterly and A. Reshef (unpublished).
[10] G. Fagiolo, S. Schiavo, and J. Reyes, Physica A **387**, 3868 (2008).
[11] D. Garlaschelli and M. Loffredo, Phys. Rev. Lett. **93**, 188701 (2004).
[12] D. Garlaschelli and M. Loffredo, Physica A **355**, 138 (2005).
[13] A. Serrano and M. Boguñá, Phys. Rev. E **68**, 015101(R) (2003).
[14] A.-L. Barabási and R. Albert, Science **286**, 509 (1999).
[15] G. Caldarelli, A. Capocci, P. De Los Rios, and M. A. Muñoz, Phys. Rev. Lett. **89**, 258702 (2002).
[16] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero. Sci. Rep. **2**, 723 (2012).
[17] www.cepii.fr/anglaisgraph/bdd/baci.htm.
[18] G. Fagiolo, J. Reyes, and S. Schiavo, Phys. Rev. E **79**, 036115 (2009).
[19] R. Hausmann and D. Rodrik, J. Dev. Econ. **72**, 603 (2003).
[20] C. A. Hidalgo and R. Hausmann, Proc. Natl. Acad. Sci. USA **106**, 10570 (2009).
[21] S. Krautheim, J. Int. Econ. **87**, 27 (2012).
[22] F. Spizzichino. *Subjective Probability Models for Life-Times* (Chapman & Hall/CRC, Boca Raton, FL, 2001).