INSTITUTE
OF ECONOMICS

Scuola Superiore
Sant'Anna

LEM | Laboratory of Economics and Management

Institute of Economics
Scuola Superiore Sant'Anna

Piazza Martiri della Libertà, 33 - 56127 Pisa, Italy
ph. +39 050 88.33.43
institute.economics@sssup.it

LEM
WORKING PAPER SERIES

Intensive and extensive biases in economic networks: reconstructing world trade

Rossana Mastrandrea *
Tiziano Squartini °
Giorgio Fagiolo *
Diego Garlaschelli °

* Institute of Economics, Scuola Superiore Sant'Anna, Pisa, Italy
° Instituut-Lorentz for Theoretical Physics, University of Leiden ,The Netherlands

2014/06 February 2014

ISSN (online) 2284-0400

Intensive and extensive biases in economic networks: reconstructing world trade

Rossana Mastrandrea
Institute of Economics and LEM, Scuola Superiore Sant'Anna, 56127 Pisa (Italy)

Tiziano Squartini
Instituut-Lorentz for Theoretical Physics, University of Leiden, 2333 CA Leiden (The Netherlands)

Giorgio Fagiolo
Institute of Economics and LEM, Scuola Superiore Sant'Anna, 56127 Pisa (Italy)

Diego Garlaschelli
Instituut-Lorentz for Theoretical Physics, University of Leiden, 2333 CA Leiden (The Netherlands)
(Dated: February 17, 2014)

In economic and financial networks, the strength (total value of the connections) of a given node has always an important economic meaning, such as the size of supply and demand, import and export, or financial exposure. Constructing null models of networks matching the observed strengths of all nodes is crucial in order to either detect interesting deviations of an empirical network from economically meaningful benchmarks or reconstruct the most likely structure of an economic network when the latter is unknown. However, several studies have proved that real economic networks are topologically very different from networks inferred only from node strengths. Here we provide a detailed analysis for the World Trade Web (WTW) by comparing it to an enhanced null model that simultaneously reproduces the strength and the number of connections of each node. We study several temporal snapshots and different aggregation levels (commodity classes) of the WTW and systematically find that the observed properties are extremely well reproduced by our model. This allows us to introduce the concept of extensive and intensive bias, defined as a measurable tendency of the network to prefer either the formation of new links or the reinforcement of existing ones. We discuss the possible economic interpretation in terms of trade margins.

I. INTRODUCTION

Over the last fifteen years, there has been a dramatic rise of interest in the understanding of the mechanisms for network formation [1–3]. One of the reasons for this interest is the fact that the dynamics of a wide range of important phenomena, including the spread of disease and information diffusion, is strongly affected by the topology of the underlying network that mediates the interactions. In particular, economic networks are responsible for the phenomenology of many processes of societal relevance, such as globalization, economic integration, financial contagion and the build-up of systemic risk [4].

A. Null models of economic networks

In order to identify the statistically significant structural properties in a real network, one needs the appropriate definition and implementation of a *null model*. Null models of networks are often built through a randomization or reshuffling process that generates an ensemble of graphs which preserve part of the observed topology - the *constraint(s)* - while all the other properties are random. Comparing a real network with its null model allows one to detect the emergence of some empirical stylized facts that depart from the random ensemble. A lot of effort has been devoted to the introduction of null models for graphs [5–13].

In economics, the use of purely random models is not new and spans from industrial agglomeration (see for example [14, 15]) to International Trade (e.g. the "bins-and-balls" model of trade [16]). The identification of the observed properties of real economic networks that carry useful and non-trivial information allows one to select the target quantities that meaningful economic models should aim at explaining or reproducing. If one enforces only local (e.g. node-specific) network properties and obtains a very close agreement between observed and expected higher-order features, it means that the chosen constraints alone are able to explain such features. On the contrary, a bad prediction of the null model indicates the need for the introduction of additional information or mechanisms explaining the observed structure. The quantities that are interesting from an economic point of view are precisely those that call for additional explanations. Null models therefore represent an intermediate approach between statistical null hypotheses and mechanistic models.

An important economic case study that has been investigated in great detail is the World Trade Web (WTW), or International Trade Network (ITN), where nodes are world countries and links represent international trade relationships. Several authors have focused especially on the binary version of the network [17–19], showing the presence of a disassortative pattern and a stable negative correlation between node degrees and clustering coefficients. The relevant role played by the topology on the whole network structure is undeniable. In fact, the observed topological properties turn out to be important in explaining macroeconomics dynamics: Kali and Reyes [20, 21] observed that country positions in the trade network (e.g. in terms of their node degrees) has substantial implications for economic growth and a good potential for explaining episodes of financial contagion. Furthermore, network position appears to be a substitute for physical capital but a complement for human capital.

The introduction of null models in the analysis of WTW seems quite natural. It allows assessing whether the network formalism is really conveying additional, non-trivial information with respect to traditional international-economics analyses, which instead explain the empirical properties of trade in terms of country-specific macroeconomic variables alone. Indeed, the standard economic approach to the empirics of international trade [22] has traditionally focused on the statistical properties of country-specific indicators like total trade, number of trade partners, etc., that can be easily mapped to what, in the jargon of network analysis, one denotes as local properties or first-order node characteristics.

A recent series of studies [23–25] showed that much of the binary WTW architecture, where countries are connected by binary (directed) links if there is a (directed) trade relationship (irrespective of the magnitude of the latter), can be reproduced by a null model controlling for the (in- and out-)degrees of all countries. These results hold both at the aggregate and product-specific level. From a theoretical point of view, this means that it is in principle possible to reproduce the topology of the WTW starting from purely local information: the number(s) of trade partners of each country. This has important consequences for economic modelling, since most macroeconomic theories of trade do not consider the number(s) of partners as a relevant target quantity to explain. By contrast, the aforementioned results imply that, if the degrees of all countries are not reproduced, it is quite difficult to reproduce the large-scale structure of the network as well.

Despite its fundamental role, the binary version of the WTW suffers from an important limitation: it does not account for link heterogeneity and so gives only partial information about the network. Hence, it seems quite natural to analyse the WTW as a weighted network, where (directed) links are now weighted by the magnitude of the observed (directed) trade relationships, and the (in- and out-) strength of a country corresponds to its total (import and export) trade. Indeed Fagiolo *et al.* [26, 27], Fagiolo [28] showed that the binary and weighted version of the WTW are very different from a statistical point of view. For example, the strength distribution is highly left-skewed, indicating that a few intense trade connections co-exist with a majority of low-intensity ones [29, 30]; moreover WTW countries holding many trade partners (and/or very intense trade relationships) are also the richest and most (globally) central. They typically trade with many partners, but very intensively with only few of them (which turn out to be themselves very connected), and form few but intensive-trade clusters (triangular trade patterns) Fagiolo *et al.* [27].

Given the importance of weights in the WTW, Fagiolo *et al.* [24], Squartini *et al.* [31] extended the analysis performed on [23] to the weighted version of the network, now randomizing the latter while preserving the (in- and out-)strength of each country. In contrast with the binary analysis, they found not only a very bad agreement between observed and expected values of higher-order weighted properties, but also a bad prediction of the bare topology, since the rewired networks turn out to be systematically much denser than the observed one. Other works using different methodologies obtained similar results [12, 30, 32–34]. The standard interpretation of these findings is the existence of higher-order mechanisms shaping the structure of the WTW as a weighted network, or equivalently the impossibility to reconstruct the WTW starting from purely country-specific information. Again, this interpretation has important consequences for economic modelling, quite opposite to those reached in the binary analysis. In this case, the strength (total trade) of a country is indeed one of the main targets of macroeconomic theories, but the above findings imply that reproducing the observed strengths is by no means enough in order to explain the structure of the network as a whole. Even if for the opposite reason, this conclusion calls again for a change of perspective in the way economic models approach the international trade system.

B. Network reconstruction

There is another attractive reason for using null models in empirical network studies: the possibility to reconstruct a network from its local properties. In the economic literature, a typical example of increasing interest is that of interbank networks. Generally, it is relatively easy to know the total exposures of each bank, however privacy issues make it much more difficult to know *who* is lending to *whom*, and *how much* [13, 35, 36]. When the available information is just local, one only knows $O(N)$ quantities (e.g. the degrees of all nodes) instead of the total $O(N^2)$ ones (e.g. all entries of the adjacency matrix) fully describing the network. This makes the network reconstruction problem very challenging, since the number of missing variables is still $O(N^2)$. In this respect, the aforementioned results about the WTW [23, 24, 31] suggest that purely weighted local properties (strengths) are much less informative than binary ones (degrees). This would imply that, while the reconstruction of some binary graphs can be achieved successfully, that of weighted networks may be inherently more problematic and practically unfeasible.

However, we recently introduced an enhanced method to build ensembles of networks that simultaneously reproduce the strength and the degree of each node [37]. The application of this method allowed us to show that, for many real networks where the specification of the strengths alone give very poor results, the joint specification of strengths and degrees can reconstruct the original network to a great degree of accuracy. This result completely reverses the standard interpretation about the reconstruction of weighted networks: the knowledge of purely local information can indeed be enough in order to infer the structure of the whole network, *provided that both weighted and purely binary local properties are specified*.

While we already analysed one (aggregated and static) snapshot of the WTW as part of the analysis described above [37], this is not enough in order to conclude that, for this particular economic system, those results can be straightforwardly extended to different temporal snapshots and different levels of resolution. Given the importance of the problem for the more general understanding of economic networks, in this paper we carry out an in-depth investigation of the WTW spanning several years and multiple layers of trade (i.e. commodity classes). Following the mentioned recent works in this direction [23, 24, 31], we use our network reconstruction method to understand if the network approach is really conveying additional and non-trivial information with respect to standard international trade theories.

Our results confirm that, taken separately, the total trade (strength) and the number of trade partners (degree) of all world countries are both uninformative about the weighted structure of the whole WTW. However, when considered together, these local quantities are enough in order to reproduce many higher-order properties of the network, for all levels of disaggregation and all temporal snapshots in our analysis. In order to fully explain the structure of the WTW, binary constraints must therefore be added to the weighted ones. This suggests that, remarkably, additional economic mechanisms besides those accounting for the joint evolution of degrees and strengths are not really necessary in order to explain the structure and dynamics of the WTW.

C. Extensive and intensive margins

One of the added values of our analysis is the fact that, from an economic point of view, we can relate these findings to the so-called *extensive* and *intensive margins* of trade. These two concepts, firstly introduced by Ricardo *et al.* [38], are widely used in economics. In the context of international trade between countries, the extensive and intensive margins refer to the birth (or death) of trade connections and to the growth (or decrease) of their weight, respectively. They can therefore be defined employing a network approach at the worldwide level De Benedictis and Tajoli [39], even if most contributions typically focus only on selected cases and use different (sometimes conflicting) perspectives.

Indeed, even if both margins are known to be relevant, in the economic literature there is neither a systematic treatment of their role in the prediction of the international trade relationships, nor a unified agreement on their relative importance. In fact, some works agree on the relevance of extensive margins: Hummels and Klenow [40] show a cross-country analysis revealing that extensive margin accounts for the 60% of exports for the larger economies; Evenett and Venables [41] find that increasing in the extensive margin has a fundamental role for augmented exports for developing countries.

On the contrary, a large body of work stresses the relevance of intensive margin. Felbermayr and Kohler [42] and Helpman *et al.* [43] show that intensive margin represents a fundamental factor in the period 1970-1990, while Amiti and Freund [44] focus on its impact on China's growth in exports in the period 1992-2005. Eaton *et al.* [45] claim that one half of Colombian countries exporting are "new" in any considered year. Besedeš and Prusa [46] find that the majority of trade growth is due to the intensive margins rather than the extensive one, stressing the importance to concentrate on a dynamical comparison against a standard static approach. Indeed, they introduce the concepts of "survival" and "deepening" to characterize export relationships. Moreover they claim that the controversial results existing in the literature are due to the different levels the two margins are examined: some works define extensive margins at the country-product level, others at the product level, and finally others at the country level.

This problem, together with the composite effect of international changes on trade, could determine mixed and contradictory results. For example, trade liberalization affects trade flows in two ways. On the one hand, since trade is less costly, the trade quantities increase (intensive margin at product level), on the other hand, more firms trade more and more goods are traded (extensive margin both at product and country-product level). Chaney [47] observes that also the elasticity of substitution should be taken into account because it has opposite effects on each margin: high elasticity makes intensive margin more sensitive to changes in trade barriers (trade costs), whereas the extensive margin results less sensitive in this case.

Our work complements the existing literature by introducing the general concepts of *extensive and intensive bias*. While extensive and intensive *margins* are defined at an intrinsically dynamic level, we define extensive and intensive *biases* as purely static notions. Indeed, we focus on cross-sectional data and evaluate whether, at a given point in time, pairs of countries are 'shifted' along the intensive or extensive direction as compared to an appropriate null model defined for the specific snapshot in consideration. We do *not* explicitly establish whether the WTW evolves along the extensive or intensive margin in the traditional way, i.e. by accounting for the variation over time of trade connections and their weights. Rather, our methodology allows us to identify a bias toward either the extensive or the intensive limit in a fundamentally novel way, by exploiting a mathematical property of the null model that specifies both strengths and degrees.

We show that, for any economic network, the specification of the strengths without the separate specification of the degrees corresponds to the assumption that, at a static level, the system is neutral with respect to the two tendencies, i.e. the extensive and intensive biases are perfectly balanced. By contrast, if the degrees are also specified, then for every pair of vertices there is a specific tendency to favour one of the two directions. The fact that the latter model reproduces the real WTW very well, while the former performs very bad, is then a clear indication that the network is not neutral with respect to the two biases. Moreover, for each pair of countries we can measure the entity of the bias towards the dominating direction. Rather than assuming aggregate or average effects, our approach allows different pairs of countries to be characterized by opposite tendencies, exploiting the entire observed complexity of interactions at a static level.

Despite its static character, our analysis allows us to draw some interesting implications on the predictive power of trade margins also from a dynamic perspective. In particular, our study of the temporal evolution of the WTW allows us to suggest relationships between the extensive/intensive biases and the extensive/intensive margins of trade.

The rest of this paper is organized as follows. In Sec. II we introduce the data and briefly summarize the methodology we use to specify both strengths and degrees in weighted networks [37, 48]. In Sec. III we apply the methodology to several temporal snapshots and different aggregation levels (commodity classes) of the WTW. In Sec. IV we discuss our results and their general implications, as well as their economic interpretation in terms of intensive and extensive biases of trade.

II. DATA AND METHODOLOGY

A. Data

We employed international trade data provided by the United Nations Commodity Trade Database (UN COMTRADE[49]) in order to build a time sequence of binary and weighted networks. The sample refers to 11 years, 1992-2002, represented in current U.S. dollars. The choice of this time span, for which we can construct a consistent and rather comprehensive panel of 162 countries and 97 commodity classes, allows us to disaggregate trade flows between most world countries and at various levels of resolution. We can therefore investigate whether local properties are sufficient to explain higher-order quantities for different levels of matrix sparseness.

We chose the classification of trade values into $C = 97$ possible commodities listed according to the Harmonized System 1996 (HS1996 [50]). For each year t and each commodity c, the starting data are represented as a matrix whose elements are the trade flows directed from each country to all other countries [51]. The matrix elements are therefore $e^c_{ij}(t) > 0$ whenever there is an export of good c from country i to country j, and $e^c_{ij}(t) = 0$ otherwise. Rows and columns stand for exporting and importing countries respectively. The value of $e^c_{ij}(t)$ is in current U.S. dollars (USD) for all commodities.

Given the commodity-specific data $e^c_{ij}(t)$, we can compute the total (aggregate) value of exports $e^{AGG}_{ij}(t)$ from country i to country j summing up over the exports of all $C = 97$ commodity classes:

$$e^{AGG}_{ij}(t) \equiv \sum_{c=1}^{C} e^c_{ij}(t) \qquad (1)$$

The particular aggregation procedure described above, introduced in [52], allows us to compare the analysis of the C commodity-specific networks with a $(C+1)-$th aggregate network, avoiding possible inconsistencies between aggregated and disaggregated trade data.

We put special emphasis on the 14 particularly relevant commodities identified in Barigozzi *et al.* [52] and reported in table I. They include the 10 most traded commodities ($c = 84, 85, 27, 87, 90, 39, 29, 30, 72, 71$ according to the HS1996) in terms of total trade value (following the ranking in year 2003, Barigozzi *et al.* 52), plus 4 classes ($c = 10, 52, 9, 93$ according to the HS1996) which are less traded but still important for their economic relevance. Taken together, the 10 most traded commodities (see table I) account for 56% of total world trade in 2003; moreover, they also feature the largest values of trade value per link (also shown in the table). The 14 top commodities account together for 57% of world trade in 2003. As an intermediate level of aggregation between individual commodities and fully aggregate trade, we also consider the network formed by the sum of these 14 commodities. In this way we can also draw conclusions about the robustness of our methodology with respect to the sparseness of the network.

In our analyses, we will focus on the undirected (symmetrized) representation of the network for obvious reasons of simplicity, even if the extension to the directed case is straightforward once the method in ref. [37] is appropriately generalized. In any case, several works have shown that the percentage of reciprocated interactions in the WTW is steadily high [23, 26, 31], giving us reasonable confidence that we can focus on the temporal series of undirected networks. We therefore define the symmetric matrices

$$\begin{aligned} \tilde{w}^c_{ij}(t) \equiv & \quad \lfloor \tfrac{e^c_{ij}(t)+e^c_{ji}(t)}{2} \rceil \\ \tilde{w}^{AGG}_{ij}(t) \equiv & \quad \lfloor \tfrac{e^{AGG}_{ij}(t)+e^{AGG}_{ji}(t)}{2} \rceil \end{aligned} \qquad (2)$$

where $\lfloor x \rceil$ represents the nearest integer to the nonnegative real number x [53]. The above matrices define an undirected weighted network where the weight of a link

HS Code	Commodity	Value (USD)	Value per link (USD)	% of aggregate trade
84	Nuclear reactors, boilers, machinery and mechanical appliances; parts thereof	5.67×10^{11}	6.17×10^{7}	11.37%
85	Electric machinery, equipment and parts; sound equipment; television equipment	5.58×10^{11}	6.37×10^{7}	11.18%
27	Mineral fuels, mineral oils & products of their distillation; bitumin substances; mineral wax	4.45×10^{11}	9.91×10^{7}	8.92%
87	Vehicles, (not railway, tramway, rolling stock); parts and accessories	3.09×10^{11}	4.76×10^{7}	6.19%
90	Optical, photographic, cinematographic, measuring, checking, precision, medical or surgical instruments/apparatus; parts & accessories	1.78×10^{11}	2.48×10^{7}	3.58%
39	Plastics and articles thereof.	1.71×10^{11}	2.33×10^{7}	3.44%
29	Organic chemicals	1.67×10^{11}	3.29×10^{7}	3.35%
30	Pharmaceutical products	1.4×10^{11}	2.59×10^{7}	2.81%
72	Iron and steel	1.35×10^{11}	2.77×10^{7}	2.70%
71	Pearls, precious stones, metals, coins, etc	1.01×10^{11}	2.41×10^{7}	2.02%
10	Cereals	3.63×10^{10}	1.28×10^{7}	0.73%
52	Cotton, including yarn and woven fabric thereof	3.29×10^{10}	6.96×10^{6}	0.66%
9	Coffee, tea, mate & spices	1.28×10^{10}	2.56×10^{6}	0.26%
93	Arms and ammunition, parts and accessories thereof	4.31×10^{9}	2.46×10^{6}	0.09%
ALL	Aggregate (all 97 commodities)	4.99×10^{12}	3.54×10^{8}	100.00%

TABLE I. The 14 most relevant commodity classes (plus aggregate trade) in year 2003 and the corresponding total trade value (USD), trade value per link (USD), and share of world aggregate trade. Source: Barigozzi *et al.* [52].

is the average of the trade flowing in either direction between two countries.

In order to wash away trend effects and make data comparable over time, we normalized our weights according to the total trade volume for each year:

$$w_{ij}^{c}(t) \equiv \frac{\tilde{w}_{ij}^{c}(t)}{\tilde{w}_{TOT}^{c}} \quad \text{and} \quad w_{ij}^{AGG}(t) \equiv \frac{\tilde{w}_{ij}^{AGG}(t)}{\tilde{w}_{TOT}^{AGG}} \qquad (3)$$

where $\tilde{w}_{TOT}^{c} = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\tilde{w}_{ij}^{c}$ and $\tilde{w}_{TOT}^{AGG} = \sum_{i=1}^{N}\sum_{j=i+1}^{N}\tilde{w}_{ij}^{AGG}$. We end up with deflated and adimensional weights that allow proper comparisons over time and consistent analyses of the evolution of network properties.

B. Methodology

Given a network with N vertices, there are several ways to generate a family of randomized variants of it. Most of them suffer from severe limitations and give biased results. Here we use a recent unbiased method based on the maximum-likelihood estimation of maximum-entropy models of graphs [48]. We briefly recall the main steps of this procedure and of the enhanced network reconstruction method [37] that can be derived from it.

Firstly, we specify a set of constraints $\{C_i(G)\}$, where G denotes one particular graph in the ensmeble of possible networks. These constraints are the network properties that we want to preserve during the randomization procedure, according to the specific network and research question. In general they are *local* constraints, such as the degree sequence (the resulting null model is known as the Configuration Model, CM) or the strength sequence (Weighted Configuration Model, WCM), but this methodology can also account for non-local constraints [54, 55]. In order to construct an ensemble of weighted networks where both the degree sequence and the strength sequence are specified [37], we choose

$$\{C_i(G^*)\} \equiv \{k_i(G^*), s_i(G^*)\} \qquad (4)$$

where k_i stands for the i-th node degree, s_i for the i-th node strength and G^* for the observed network. We refer to this model as the "Mixed Configuration Model" (MCM) [37].

Secondly, we need to find the analytical expression for the probability $P(G)$ that maximizes the Shannon-Gibbs entropy

$$S(G) \equiv -\sum_{G} P(G)\ln(P(G) \qquad (5)$$

with the constraints $\langle C_i\rangle \equiv \sum_G P(G)C_i(G) = C_i(G^*)$ for all i, and $\sum_G P(G) = 1$. Note that $P(G)$ stands for the occurrence probability of the graph G in the ensemble of all possible graphs, and the sums are over all such graphs. The formal solution [48] of this constrained maximization problem can be written as a function of the *graph mamiltonian* $H(G,\vec{\theta})$ and the *partition function* $Z(\vec{\theta})$:

$$P(G|\vec{\theta}) \equiv \frac{e^{-H(G,\vec{\theta})}}{Z(\vec{\theta})} \tag{6}$$

where $H(G,\vec{\theta}) \equiv \sum_i \theta_i C_i(G)$ and $Z(\vec{\theta}) \equiv \sum_G e^{-H(G,\vec{\theta})}$. The Hamiltonian is a linear combination of the constraints, with the coefficients θ_i being the Lagrangian multipliers introduced in the constrained-maximization problem. For the MCM, it is possible to show [37] that

$$P(W|\vec{x},\vec{y}) = \prod_{i<j} q_{ij}(w_{ij}|\vec{x},\vec{y}) \tag{7}$$

where $\vec{x}$ and $\vec{y}$ are two N-dimensional Lagrange multipliers (N stands for the number of nodes) controlling for the expected degrees and strengths respectively (with $x_i \geq 0$ and $0 \leq y_i < 1\ \forall i$) and $q_{ij}(w|\vec{x},\vec{y})$ is the conditional probability to observe a link of weight w between nodes i and j. The latter is

$$q_{ij}(w|\vec{x},\vec{y}) = \frac{(x_i x_j)^{\Theta(w)}(y_i y_j)^w (1 - y_i y_j)}{1 - y_i y_j + x_i x_j y_i y_j} \tag{8}$$

The third step of our procedure prescribes to find the values of the Lagrange multipliers $\vec{x}^*, \vec{y}^*$ that maximize the log-likelihood

$$\mathcal{L}(\vec{x},\vec{y}) \equiv \ln P(G^*|\vec{x},\vec{y}) = \textstyle\sum_{i<j} \ln q_{ij}(w^*_{ij}|\vec{x},\vec{y}) \tag{9}$$

representing the logarithm of the probability to observe the empirical graph G^*. The maximization of the likelihood is equivalent to the requirement that the desired constraints are satisfied on average by the ensemble of networks [56], i.e. in this case $\langle k_i\rangle = k_i(G^*)$ and $\langle s_i\rangle = s_i(G^*)$ for all i.

As a final step, one can use the Lagrange multipliers $\vec{x}^*, \vec{y}^*$ to compute the expected value $\langle X\rangle$ of any (higher-order) network property $X(G)$:

$$\langle X\rangle \equiv \sum_G X(G) P(G|\vec{x^*},\vec{y^*}) \tag{10}$$

Comparing $\langle X\rangle$ with the observed value $X(G^*)$ allows us to verify whether the 'reconstructed' value of the property is indeed close to the empirical one.

We will also compare the predictions of the MCM with those of the WCM, that can be obtained by setting $\vec{x}^* = \vec{1}$ and maximizing the likelihood with respect to $\vec{y}$ alone [37].

III. RESULTS

We first report detailed results for the 2002 for the aggregated network, and then consider its temporal evolution. Finally, the commodity-specific analysis is presented. In all cases we show the outcomes for both the binary and weighted part.

First, we are interested in assessing to what extent the MCM is able to replicate the higher-order properties characterizing the WTW over time. We focus on the Average Nearest Neighbor Degree and Strength (respectively indicated by k_i^{nn} and s_i^{nn}) and the Binary and Weighted Clustering Coefficient (respectively indicated by c_i and c_i^W). We document the assortative (disassortative) nature of the WTW over time and its tendency to form triangles both in binary and weighted case. We measure the (Pearson) correlation coefficient between degree and higher-order binary properties (k_i^{nn}, c_i) and between strength and higher-order weighted properties (s_i^{nn}, c_i^W). We show this correlation coefficient for each temporal snapshot, together with 95% confidence intervals.

Secondly, we are interesting in exploring the performance of the MCM at different levels of disaggregation. We analyse both binary and weighted higher-order properties for selected commodities in the 2002 snapshot. This also allows us to investigate the effectiveness of local constraints in replicating higher-order properties according to different levels of network density.

A. Binary and weighted aggregated network

We recall the analytical expressions and briefly describe the aforementioned higher-order quantities that we want to study. Moreover, for the sake of clarity we write down the explicit formula for computing the same quantities overall the ensemble according to the general expression in (10).

If we indicate with A the adjacency matrix and with W the weighted matrix representing our network, we can compute the Average Nearest Neighbor Degree as:

$$k_i^{nn}(W) \equiv \frac{\sum_{j\neq i} a_{ij} k_j}{k_i} = \frac{\sum_{j\neq i}\sum_{k\neq j} a_{ij} a_{jk}}{\sum_{j\neq i} a_{ij}} \tag{11}$$

where $k_i = \sum_{j\neq i} a_{ij}$ stands for the i-th node degree.

This quantity averages the degrees of the partners of a node with degree k, in other terms it gives a measure of the "activity" of node partners just looking at the number of their edges.

The Binary Clustering Coefficient has the following expression:

$$c_i(W) \equiv \frac{\sum_{j\neq i}\sum_{k\neq i,j} a_{ij} a_{jk} a_{ki}}{\sum_{j\neq i}\sum_{k\neq i,j} a_{ij} a_{ki}} \tag{12}$$

It measures the tendency to form triangles of each node: it counts how many closed triangles are attached to each node with respect to all the possible triangles.

The corresponding weighted quantities are the Average Nearest Neighbor Strength, defined as

$$s_i^{nn}(W) \equiv \frac{\sum_{j\neq i} a_{ij} s_j}{k_i} = \frac{\sum_{j\neq i}\sum_{k\neq j} a_{ij} w_{jk}}{\sum_{j\neq i} a_{ij}} \tag{13}$$

where $s_i = \sum_{j \neq i} w_{ij}$ stands for the i-th node strength, and the Weighted Clustering Coefficient [57] defined as

$$c_i^W(W) = \frac{\sum_{j \neq i} \sum_{k \neq i,j} (w_{ij} w_{jk} w_{ki})^{1/3}}{\sum_{j \neq i} \sum_{k \neq i,j} a_{ij} a_{ki}} \quad (14)$$

The s_i^{nn} measures the average strength of the neighbors of a given node with strength s_i. Similarly to the binary counterpart, it reveals the "intensity" of activity of a node looking at the strengths of its partners. The weighted clustering coefficient measures the propensity of node i to be involved in triangular relations taking into account also the edge-values.

The related expected values can be obtained by simply replacing a_{ij} with the occurrence probabilities p_{ij} for the binary case:

$$\langle k_i^{nn}(W) \rangle \equiv \frac{\sum_{j \neq i} p_{ij} \langle k_j \rangle}{\langle k_i \rangle} = \frac{\sum_{j \neq i} \sum_{k \neq j} p_{ij} p_{jk}}{\sum_{j \neq i} p_{ij}} \quad (15)$$

$$\langle c_i(W) \rangle \equiv \frac{\sum_{j \neq i} \sum_{k \neq i,j} p_{ij} p_{jk} p_{ki}}{\sum_{j \neq i} \sum_{k \neq i,j} p_{ij} p_{ki}} \quad (16)$$

where

$$p_{ij} \equiv \frac{x_i x_j y_i y_j}{1 - y_i y_j + x_i x_j y_i y_j} \quad (17)$$

and, by construction, $\langle k_i \rangle \equiv k_i$, $\forall i$.

For the weighted case, we have:

$$\langle s_i^{nn}(W) \rangle \equiv \frac{\sum_{j \neq i} p_{ij} \langle s_j \rangle}{\langle k_i \rangle} = \frac{\sum_{j \neq i} \sum_{k \neq j} p_{ij} \langle w_{jk} \rangle}{\sum_{j \neq i} p_{ij}} \quad (18)$$

where, $\langle w_{ij} \rangle \equiv \dfrac{x_i x_j y_i y_j}{(1 - y_i y_j)(1 - y_i y_j + x_i x_j y_i y_j)}$ and, for assumption, $\langle s_i \rangle \equiv s_i$, $\forall i$.

For the expected value of the c^W one should be more careful, indeed it is necessary to compute the expected product of (powers of) distinct matrix entries:

$$c_i^w(W) = \frac{\sum_{j \neq i} \sum_{k \neq i,j} (w_{ij} w_{jk} w_{ki})^{1/3}}{\sum_{j \neq i} \sum_{k \neq i,j} a_{ij} a_{ki}} \quad (19)$$

In general we observe that:

$$\left\langle \sum_{i \neq j \neq k, \ldots} w_{ij}^{\alpha} \cdot w_{jk}^{\beta} \cdot \ \ldots \right\rangle = \sum_{i \neq j \neq k, \ldots} \langle w_{ij}^{\alpha} \rangle \cdot \langle w_{jk}^{\beta} \rangle \cdot \ \langle \ldots \rangle \quad (20)$$

with the generic term:

$$\langle w_{ij}^{\gamma} \rangle = \sum_{w=0}^{+\infty} w^{\gamma} q_{ij}(w | \vec{x}^*, \vec{y}^*) = \frac{x_i^* x_j^* (1 - y_i^* y_j^*) \mathrm{Li}_{-\gamma}(y_i^* y_j^*)}{1 - y_i^* y_j^* + x_i^* x_j^* y_i^* y_j^*} \quad (21)$$

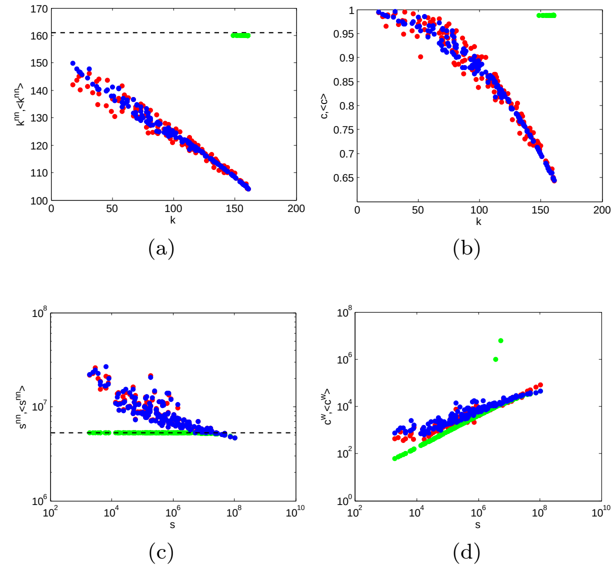

FIG. 1. Comparison between the observed undirected binary and weighted properties (red points) and the corresponding ensemble averages of the WCM (green points) and the MCM (blue points) for the aggregated WTW in the 2002 snapshot. (a) Average Nearest Neighbor Degree k_i^{nn} versus degree k_i.; (b) Binary Clustering Coefficient c_i versus degree k_i; (c) Average Nearest Neighbor Strength s_i^{nn} versus strength s_i; (d) Weighted Clustering Coefficient c_i^W versus strength s_i.

where $\mathrm{Li}_n(z) \equiv \sum_{l=1}^{+\infty} z^l / l^n$ is the nth polylogarithm of z. The simplest cases $\gamma = 1$ and $\gamma = 0$ yield the expected weight $\langle w_{ij} \rangle$ and the connection probability p_{ij}, respectively.

In figure 1 we show the higher-order binary quantities versus the node degree and the weighted ones versus the node strength, for the 2002 snapshot. We plot together the observed values (red points), the corresponding quantities predicted by the WCM (green points) and by the MCM (blue points).

From a theoretical point of view, it emerges a very close agreement between the observed values and the expected ones computed on the maximum-entropy ensemble generated by the MCM.

In recent papers Squartini *et al.* [23, 31] have shown that, at a binary level, the degree correlations and clustering structure of the ITN are excellently reproduced by the Configuration Model, i.e., using only the knowledge of the degree sequence. By contrast, when the WCM is implemented as a *natural* extension of the CM for valued graphs, the binary quantities and also the corresponding weighted quantities are very different from the predicted counterparts. The authors have also proved that these results are very robust and hold true over time and for various resolutions (i.e., for different levels of aggregation of traded commodities).

This outcome, here emerging from the comparison between red and green points, perfectly illustrates that the naïve expectation that weighted quantities are *per se* more informative than the corresponding binary ones is fundamentally incorrect. On the contrary, the MCM performs excellently both for the binary and weighted versions of the WTW. Firstly, it reveals a slightly improved agreement for the binary trends. In fact, the monotonic trend predicted by the degree sequence only now follows, in a closer way, the observed cloud of points, while the prediction by WCM are concentrated far from the observed (red) points. Secondly, we also find a much better agreement between the observed and the randomized weighted trends.

Despite the apparent good agreement between the observed weighted clustering coefficient and its expected value given by the WCM (fig. 1 (d)), we will show later that total level of clustering is in general higher than the one predicted by WCM over time and when we disaggregate for commodities.

Note the difference with the WCM predictions: the expected values for the binary and the weighted quantities are similar to those for a fully connected topology:

$$\langle k_i \rangle_{WCM} \simeq N-1 \tag{22}$$

$$\langle k_i^{nn} \rangle_{WCM} \simeq N-1 \tag{23}$$

$$\langle c \rangle_{WCM} \simeq 1 \tag{24}$$

$$\langle s_i^{nn} \rangle_{WCM} \simeq \frac{\sum_{j\neq i} p_{ij} s_j}{\langle k_i \rangle} \simeq \frac{\sum_i s_i}{N-1} \simeq \frac{2W_{TOT}}{N-1} \tag{25}$$

Where N stands for the number of nodes in the network, while W_{TOT} is total trade volume for the considered year.

Predictions (23) and (25) are represented by the black dashed line respectively in fig. 1 (a), (c), while the (23) corresponds to the black top line in fig.1 (b). So the unconstrained quantities, badly reproduced by the strength sequence alone, become now consistent with the prediction of a null model using the binary information also to predict the weighted structure itself. This implies that the weighted structure alone does not allow a deep understanding of the topology, representing an irreducible piece of information to be accounted for from the beginning.

From an economic point of view k_i^{nn} and c_i give information about indirect interactions respectively of lengths 2 and 3 (the terms $a_{ij}a_{jk}$ and $a_{ij}a_{jk}a_{ki}$ are involved in their computation). Plotting them versus the degree distribution helps in understanding the structural organization of the web. Therefore, the Pearson correlation coefficients allows inspecting whether these 2/3-paths are a simple outcome of the concatenation of two independent edges.

In accordance with the existing literature we find a dissortative pattern for the WTW and a decreasing trend of c_i versus k_i. This confirms that it is very likely to find nodes with many trade partners connected with nodes with small degree (and vice-versa), while trade partners of poorly connected nodes are highly interconnected. Similar considerations hold true when we introduce weights, indeed s_i^{nn} and c_i^W are related with

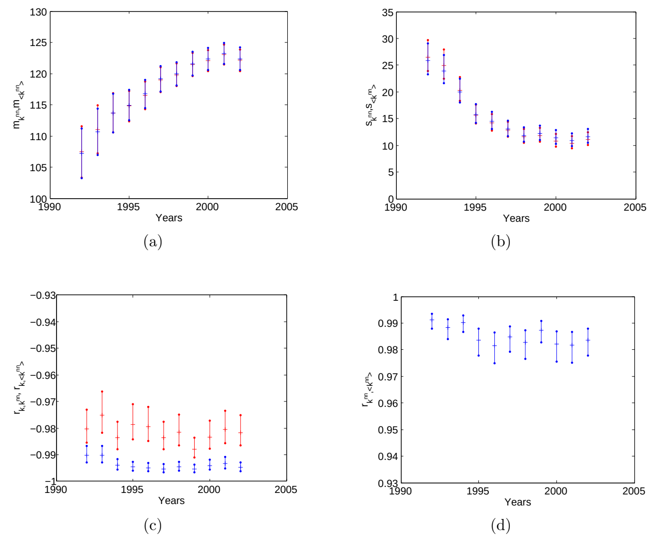

FIG. 2. Temporal evolution of the properties of the ANND k_i^{nn} in the 1992-2002 snapshots of the observed undirected WTW and of the corresponding maximum-entropy ensembles with specified degrees and strengths: (a) average of k_i^{nn} across all vertices (red: obs., blue: randomized); (b) standard deviation of k_i^{nn} across all vertices; (c) correlation coefficient between k_i^{nn} and k_i; (d) correlation coefficient between k_i^{nn} and $\langle k_i^{nn} \rangle$. Red points stands for observed values, blue for the randomized ones; the 95% confidence intervals of all quantities are represented as vertical bars.

indirect paths of length 2 and 3, respectively, but now they summarize mixed information about topology and weights (the term $a_{ij}w_{jk}$ are determinant in this sense). Again by plotting these quantities versus the strength we gather signals that countries highly involved in the ITN are connected with poorly trading countries, confirming a dissasortative pattern (even if less prominent) for the weighted network. Interestingly, these patterns are perfectly reproduced by the quantities predicted using the MCM. This implies that the knowledge of both the number of trade partners of each node and the total amount of trade flowing through each country is maximally informative about the higher-order and non local dynamics of the whole network.

To further investigate this issue, in next section we explore the evolution of the same properties over time.

B. Evolution of the aggregated-network properties

In this section we want to test the robustness of our results performing an over time study of the same network properties. In order to show in a more compact way the whole analysis for the desired period, we consider separately the four network quantities, eq. (11)-(14). We compare the observed values and the expected ones computed on the ensemble generated by the MCM.

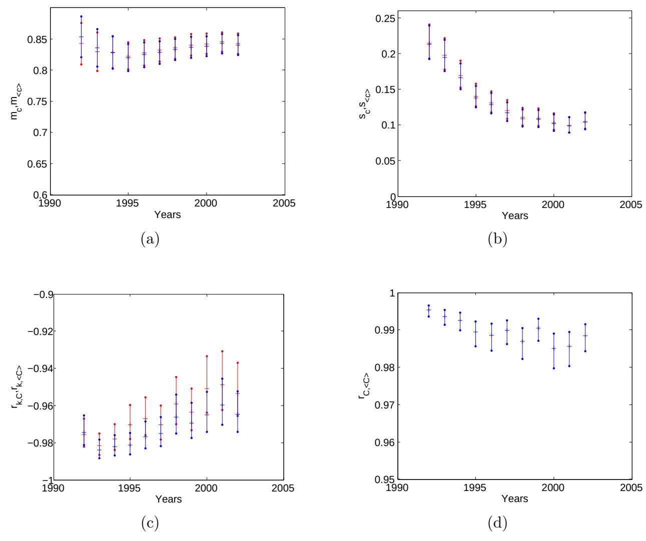

FIG. 3. Temporal evolution of the properties of the BCC c_i in the 1992-2002 snapshots of the observed undirected WTW and of the corresponding maximum-entropy ensembles with specified degrees ans strengths: (a) average of c_i across all vertices (red: obs., blue: randomized); (b) standard deviation of c_i across all vertices; (c) correlation coefficient between c_i and k_i; (d) correlation coefficient between c_i and $\langle c_i \rangle$. Red points stands for observed values, blue for the randomized ones; the 95% confidence intervals of all quantities are represented as vertical bars.

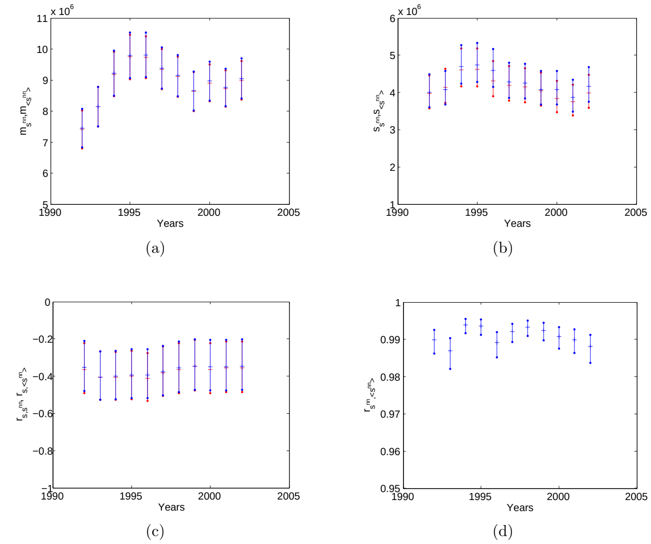

FIG. 4. Temporal evolution of the properties of the ANNS s_i^{nn} in the 1992-2002 snapshots of the observed undirected WTW and of the corresponding maximum-entropy ensembles with specified degrees ans strengths: (a) average of s_i^{nn} across all vertices (red: obs., blue: randomized); (b) standard deviation of s_i^{nn} across all vertices; (c) correlation coefficient between s_i^{nn} and s_i; (d) correlation coefficient between s_i^{nn} and $\langle s_i^{nn} \rangle$. Red points stands for observed values, blue for the randomized ones; the 95% confidence intervals of all quantities are represented as vertical bars.

Basically, for each network property we take the series of observed values, e.g. $\{k_i^{nn}\}$, and the series of its expected values, e.g. $\{\langle k_i^{nn} \rangle\}$. Then, we compute four quantities to perform an over time comparison: the mean and the standard deviation of both lists, the correlation coefficient between the two lists and the correlation coefficient between the analyzed property and the related constraint, e.g. k_i (for assumption $k_i \equiv \langle k_i \rangle$). All these quantities are plotted together with the associated 95% confidence interval.

Figures 2 and 3 show results perfectly in line with the outcome of the CM for the same data-set [23]. This implies that by simultaneously preserving degrees and strength, the MCM does not affect the ability of the CM to predict the topology of the WTW. We obtain a very close agreement between observed and expected values over time as confirmed by the correlation coefficient around 1 in figures 2 (d), 3 (d). Moreover we gather information about the decreasing trend of the Average Nearest Neighbor Degree and the almost stable trend of the clustering, perfectly replicated by the MCM prediction over time (see figs. 2 (a), (b), 3 (a), (b)). Finally, also the relation of those properties with the node degree, besides being very stable over time, is in excellent agreement with the randomized counterpart.

Also for the weighted network properties we observe an excellent agreement between observed quantities and the corresponding averages over the MCM-ensemble for the whole period. Indeed the correlation coefficients between observed and randomized properties is almost 1 all the time (figs.4 (d) and 5 (d)). Moreover the MCM is able to capture the slightly increasing, then quite stable, trend of the s_i^{nn} (fig. 4 (a), (b)) and the stable trend of the clustering coefficient (fig.5 (a), (b)). For concluding, also the correlation of these properties with the node strength is well explained by the MCM in the whole period (figs. 4 (c), 5 (c)).

Once again, we can conclude that the addition of a purely binary information as the number of node partners, makes the MCM very powerful in predicted the WTW higher-order properties, independently on the considered temporal snapshot.

C. Commodity-specific binary and weighted networks

We complete our analysis of the ITN as an undirected network by studying whether the picture changes when one considers the individual networks formed by imports and exports of single commodities.

This application allows us also to gather information about the MCM' s ability to predict different networks according to their level of sparseness. Indeed, we know that the undirected WTW is a highly dense network (density ~ 0.5) and we have already observed that some randomization techniques work only under specific con-

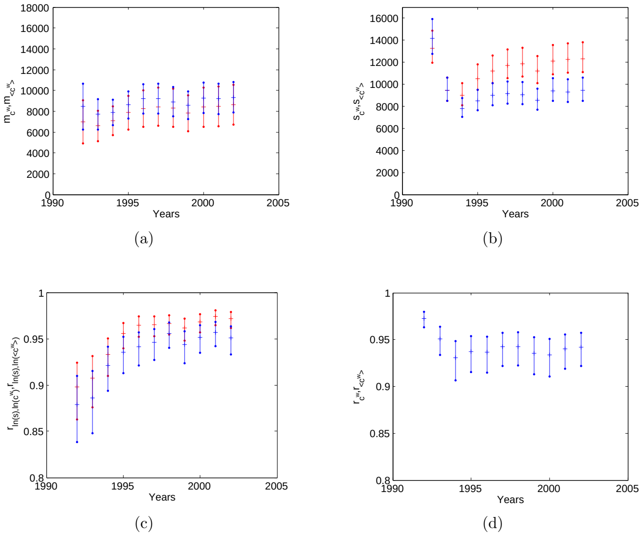

FIG. 5. Temporal evolution of the properties of the WCC c_i^W in the 1992-2002 snapshots of the observed undirected WTW and of the corresponding maximum-entropy ensembles with specified degrees ans strengths: (a) average of c_i^W across all vertices (red: obs., blue: randomized); (b) standard deviation of c_i^W across all vertices; (c) correlation coefficient between c_i^W and s_i; (d) correlation coefficient between c_i^W and $\langle c_i^W \rangle$. Red points stands for observed values, blue for the randomized ones; the 95% confidence intervals of all quantities are represented as vertical bars.

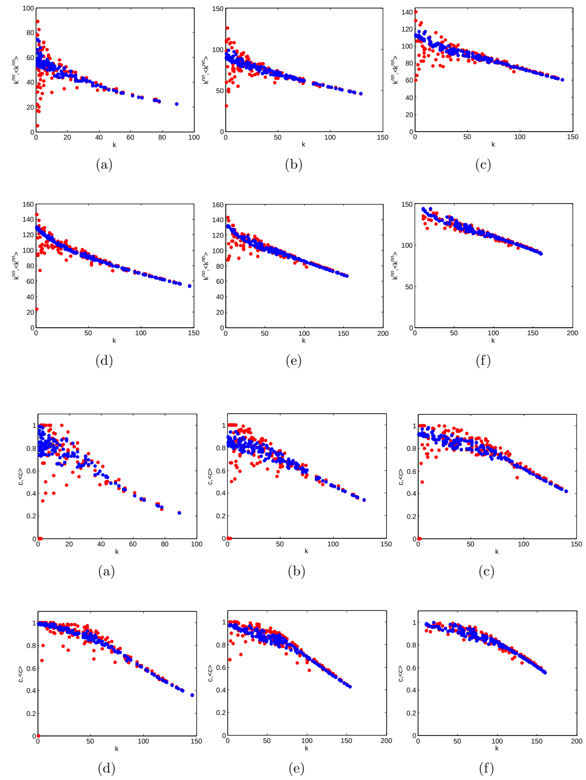

FIG. 6. ANND k_i^{nn} and BCC c_i versus degree k_i in the 2002 snapshots of the commodity-specific (disaggregated) versions of the observed binary undirected WTW (red points), and corresponding average over the maximum entropy ensemble with specified degrees and strengths (blue points): a) commodity 93; b) commodity 09; c) commodity 39; d) commodity 90; e) commodity 84; f) aggregation of the top 14 commodities (see table I for details). From a) to f), the intensity of trade and level of aggregation increases.

ditions. Indeed, the commodities have been chosen and ordered according to the intensity of trade and level of aggregation. We selected the two least traded commodities in the set ($c = 93, 9$), two intermediate ones ($c = 39, 90$), the most traded one ($c = 84$), plus the network formed by combining all the top 14 commodities. The last subnetwork represents an intermediate level of aggregation between single commodities and the completely aggregated data ($c = 0$), which corresponds to the case explored in section III A. For brevity we just show the scatter plot between binary and weighted higher-order properties and the related constraints, respectively k_i and s_i.

We find that the results obtained in our aggregated study also hold for individual commodities, independently on the level of aggregation. We recognize a small improvement in the prediction according to the increase of network density (this is especially true for the weighted case), nevertheless the agreement is always very good.

From an economic point of view, we can just point out a slight growth of dissortativity when less traded commodities are considered. Moreover, we observe more sparse scatter plot associated with less traded good and this is even more pronounced for the weighted quantities.

While the binary results confirm again the outcome of the work by Squartini *et al.* [23], the excellent agreement between observed and randomized weighted properties also for the commodity-specific case is surprising. The case of weighted clustering coefficient is really interesting in this sense. Indeed, in the aggregated case also the WCM seemed to show a good prediction of this quantity, but this outcome is not robust to disaggregation. On the contrary, figure 7 shows that the MCM is not affected by this limit neither for c_i^W nor for any other network quantities.

D. Theoretical comparison between WCM and MCM

The last step of our analysis consists in the comparison between the new enhanced model and the ordinary WCM in term of trade-off between accuracy of the results and parsimony in the use of constraints. Indeed, even if it is evident that the MCM performs better than the WCM in replicating WTW properties, we want to check if the former over-fits the network, i.e., if the introduction of degrees is redundant.

As we mentioned, the WCM can be obtained as a

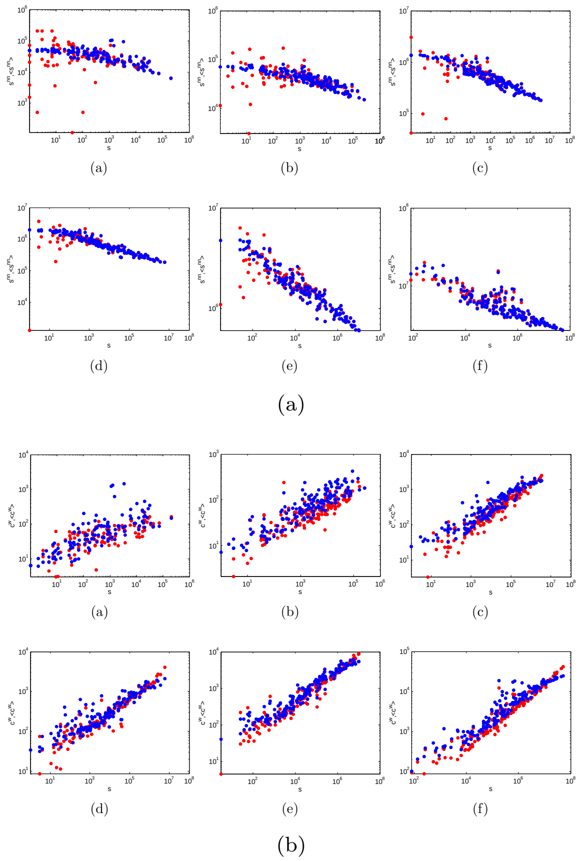

FIG. 7. ANNS s_i^{nn} and WCC c_i^W versus degree k_i in the 2002 snapshots of the commodity-specific (disaggregated) versions of the observed binary undirected WTW (red points), and corresponding average over the maximum entropy ensemble with specified degrees and strengths (blue points): a) commodity 93; b) commodity 09; c) commodity 39; d) commodity 90; e) commodity 84; f) aggregation of the top 14 commodities (see table I for details). From a) to f), the intensity of trade and level of aggregation increases.

particular case of the MCM by setting $x_i^* = 1$, $\forall i$, by "switching" off the Lagrange parameters controlling for the degrees. The log-likelihood of the WCM is therefore the reduced function $\mathcal{L}(\vec{1}, \vec{y})$ of N variables, and is maximized by a new vector $\vec{y}^{**} \neq \vec{y}^*$, where $(\vec{x}^*, \vec{y}^*)$ stands for the solution of the MCM and $\vec{y}^{**}$ the solution of the WCM for the same observed network.

Information-theoretic criteria exist [58] to assess whether the increased accuracy of a model with more parameters implies an excessive loss of parsimony. The most popular choice is the Akaike's Information Criterion (AIC), showing that the optimal trade off between accuracy and parsimony is achieved by discounting the number of free parameters from the maximized likelihood (more details about this criterion can be found in Mastrandrea *et al.* 37). For our two competing null models, following the suggestion in Burnham and Anderson [58] we implement the corrected version of AIC, i.e.:

$$\text{AICc}_{MCM} \equiv -2\mathcal{L}(\vec{x}^*, \vec{y}^*) + 4N + \frac{8N(2N+1)}{N^2 - 5N - 2} \quad (26)$$

$$\text{AICc}_{WCM} \equiv -2\mathcal{L}(\vec{1}, \vec{y}^{**}) + 2N + \frac{4N(N+1)}{N^2 - 3N - 2} \quad (27)$$

The additional term provides the correction to the test when the number of parameters is not negligible with respect to the sample cardinality (more quantitatively, when $n/k < 40$, n being the sample cardinality and k being the number of parameters Burnham and Anderson [58]), thus further reducing the probability of overfitting. Notice that as long as $n >> k^2$, the additional term converges to 0, recovering the standard form of AIC. Precisely for this reason, AICc should be always employed regardless. However, if the AIC difference is small, the two models will still be comparable. To correctly interpret the quantitative AIC differences, it is important to introduce the so-called Akaike Weights, which in our case read

$$w_{MCM}^{AICc} \equiv \frac{e^{-AICc_{MCM}/2}}{e^{-AICc_{MCM}/2} + e^{-AICc_{WCM}/2}} \quad (28)$$

$$w_{WCM}^{AICc} \equiv 1 - w_{MCM}^{AICc} \quad (29)$$

and to quantify the weight of evidence in favour of a model, i.e., the probability that the model is the best one among the (two) models considered.

Given a real network, a low value of w_{MCM}^{AICc} will indicate that the addition of the degree sequence is redundant (the relevant local constraints effectively reduce to the strength sequence, so the "standard" WCM is preferable), while a high value of w_{MCM}^{AICc} will indicate that the local constraints are irreducible to the strength sequence (so the degrees must be separately specified).

We stress that the result of this procedure is not predictable a priori (it depends on the numerical values of $\{s_i\}$ and $\{k_i\}$) and can only be achieved after a comparison with the MCM. Thus, even in cases when the WCM turns out to be the best model, our introduction of the MCM is still a necessary step making the whole approach self-consistent.

TABLE II. AICc and BIC values, AICc and BIC weights for the considered null models applied to the WTW in 2002.

	AICc	**BIC**	$\mathbf{w^{AICc}}$	$\mathbf{w^{BIC}}$
WCM	$209,972$	$211,179$	0	0
MCM	$165,731$	$168,137$	1	1

In table II we show the results for the two competing model. We also used the Bayesian Information Criterion (BIC), Burnham and Anderson [58], that puts a higher penalty on the number of parameters, but is very similar to the AIC formulation.

Both criteria confirm that addition of the degree sequence to the WCM is non-redundant and extremely informative for the prediction of the WTW properties.

IV. DISCUSSION

In economic and financial networks, the total strength of the connections reaching a node has generally an important meaning, such as the size of supply and demand, import and export, or financial exposure. Hence, generating random ensembles of networks matching the observed strengths of all nodes is crucial in order to detect interesting deviations of a known empirical network from economically meaningful benchmarks, to reconstruct the most likely structure of an unknown network, or finally to define a model of economic networks specified by purely local (node-specific) information.

Our results show that, in order to correctly reproduce the whole structure of the WTW as a weighted network, the degree sequence must be constrained in addition to the strength sequence. From the general point of view of network reconstruction, these findings consolidate and widely extend the results in [37]. We confirmed the effectiveness of the MCM in reproducing the higher-order properties of the WTW starting from local constraints, and succesfully tested the robustness of the model with respect to several temporal snapshots and levels of aggregation. So, while the strength sequence (a weighted constraint) turns out to be uninformative about the binary topology of the WTW, the degree sequence (a binary constraint) plays a fundamental role in reproducing its weighted structure. This asymmetric role of binary and weighted constraints is a non-trivial result.

From an economic perspective, the fact that purely local information is enough in order to reproduce the large-scale structure of the WTW implies that parsimonious models of international trade can largely discard additional mechanisms besides those accounting for the number of partners and total trade of world countries. The importance of reproducing and/or explaining the degrees of all world countries, first pointed out in [23], is confirmed by our study, and shown to hold even when one considers the weighted representation of the WTW. This strengthens the view that theories and models of trade, if aiming at explaining the network structure of international trade, should seriously focus on the number of trade partners of countries as an important target quantity to replicate.

We now show that our results have additional and important interpretations in terms of the intensive and extensive margins of trade. As we mentioned in the Introduction, the economic literature has mostly tried to quantify the extent to which international trade has evolved along each of the two margins, with the purpose of identifying, for selected case studies, the most important direction of trade growth. In addition to this dynamic approach to the characterization of trade margins, our results naturally suggest a novel, intrinsically static perspective.

To see this, we note that the MCM specified by eq.(8) has an important property. It is mathematically equivalent to a network formation process where the connection between any two (initially disconnected) nodes i and j is first established, with probability p_{ij} given by eq.(17), via a link of unit weight, and then (if the previous attempt is successful) strengthened with probability $y_i y_j$ by the addition of another unit of weight. For each pair of vertices, such weight-increasing attempts are iterated with the same probability $y_i y_j$ if the previous attempt was successful, and stop as soon as the previous attempt fails. This means that the probability of establishing a unit link for the first time is p_{ij}, while that of reinforcing an existing link by a unit amount is $y_i y_j$. It is easy to show that $p_{ij} > y_i y_j$ if and only if $x_i x_j > 1$. Therefore, if $x_i x_j > 1$ ($x_i x_j < 1$) the creation a link of unit weight between nodes i and j has a larger (smaller) probability than the reinforcement of the same link by a unit of weight. This feature makes the model particularly appropriate to study the extensive/intensive dichotomy in a novel sense, as the value of $x_i x_j$ can bias the network, at a purely static level, towards either the extensive ($x_i x_j > 1$) or the intensive ($x_i x_j < 1$) direction.

More in general, in the network formation process the probability of establishing a link of weight w between two previously disconnected vertices (irrespective of possibile further reinforcements) is $p_{ij}(y_i y_j)^{w-1}$, while that of adding a weight w (again, irrespective of possible further increases) to an already existing connection is $(y_i y_j)^w$. In this case as well, the former probability is larger than the latter if and only if $x_i x_j > 1$. So, independently of the value of w, $x_i x_j > 1$ implies a tendency towards the extensive direction, while $x_i x_j < 1$ signals a preference for the intensive one. For this reason, we denote $x_i x_j$ as the 'extensive bias' for the pair i, j.

If $x_i x_j = 1$ for all i, j, then the network 'is indifferent' with respect to link creation and link reinforcement. Now, it should be noted that this is precisely what is obtained in the WCM (where only the strengths are specified), as the latter can be regarded as a particular case of the MCM where $x_i = 1$ for all i [37]. So the WCM assumes that the network is neutral with respect to the two biases, as there is no preference between the extensive and intensive direction. By contrast, the MCM assumes that, for each pair of nodes, there can be a different bias towards one of the two limits.

Since we found that the WCM and the MCM perform very bad and very good respectively, we have a strong empirical indication that the WTW is not neutral with respect to the two biases. The extensive bias $x_i^* x_j^*$ measured on a particular snapshot/layer of the WTW indicates the preference of a specific pair of countries for the dominant direction. The notion of extensive or intensive bias as indicated by the value of $x_i^* x_j^*$ should therefore not be interpreted in the same sense as the extensive or intensive margin, i.e. as a preferred direction for the dynamical evolution of the network, but in terms of the 'static' deviation of the real network (well reproduced by the MCM) from the neutral topology expected under the WCM. In this sense, the WCM is serving as a null model indicating how an economic network would look like if the extensive and intensive biases were balanced. Note that, since the extensive bias is a product of two country-specific values, it is not possible to determine, on the basis of the value of x_i^* for a single country, whether the dominant bias for that country is the extensive or the intensive one. Thus the preference for one bias turns out to be an inherently dyadic property.

These considerations lead us to interpret that, in order to reproduce the observed structure of the WTW, we need to enforce realistic extensive and intensive biases as detected by the MCM through the additional knowledge of the degrees. From the strengths alone, it is indeed impossible to infer the bias towards a specific direction.

As a final consideration we note that, to the best of our knowledge, in the economic literature there has been no systematic analysis of the predictive power of extensive and intensive trade margins so far. Starting from one snapshot of the international trade network, is the knowledge of the growth of trade along the intensive and/or extensive margin enough to predict the structure of the network at a later time?

Even if our results cannot fully answer such question, they suggest a plausible scenario. We first note that a change in the degree (number of partners) of a country implies that the network is evolving along the extensive margin of trade. On the other hand, a change in the strength (total volume) can be either be due to changes in the number of partners or to changes of the amount of trade for existing links. This means that, while changes in the degree only reflect the extensive margin, changes in the strength reflect both the extensive and intensive margins: this is a second asymmetry between the different pieces of information encoded into the degrees and the strengths. It is also another indication that the WCM, by enforcing the strengths alone, cannot distinguish between the two margins, while the MCM can isolate the extensive information (degrees) from the combined one (strengths).

Our findings imply that, if the structure of the international trade network is known at time t, and if the growth (or decrease) of both strengths and degrees from time t to time $t + \Delta t$ is also known, then it is possible to predict the structure of the network at time $t + \Delta t$ with great accuracy. By contrast, if only the growth of the strengths is known, the future structure of the network cannot be satisfactorily predicted. These results can then be interpreted in terms of the fact that a combined knowledge of intensive and extensive margins (in this case, the change of the strengths) does not allow us to correctly model the network, while if the extensive margin (change of the degrees) is also separately specified (thus indirectly controlling for the residual intensive margin as well), then the model can successfully explain the data. Although these considerations require further verification, it is encouraging that these arguments derived from a dynamical interpretation of trade margins are perfectly consistent with the role of extensive and intensive biases that we have characterized through a completely static analysis.

V. CONCLUDING REMARKS

In this paper we employed a maximum-entropy approach to economic networks. We illustrated the accuracy of this method in reproducing the higher-order properties of the WTW starting from the knowledge of local constraints, provided that these constraints include both strenghts and degrees. Our results are robust for both the binary and weighted representations of the network, for different levels of disaggregation, and for several temporal snapshots.

From a theoretical point of view, our findings completely reverse the standard results concerning the reconstruction of weighted networks. We proved that it is indeed possible, using only local information, to reproduce at highly satisfactory level several higher-order binary and weighted properties for the WTW, for various temporal snapshots and at different levels of aggregation.

Economically speaking, these and previous results [23, 24, 31] allow making some considerations in relation to the extensive and intensive margins of trade. In particular, they suggest that different pairs of countries have different intrinsic biases towards either the extensive of the intensive direction. If such biases are not taken into account, it appears impossible to explain the observed structure of the WTW. This is presumably the reason why the strengths alone, by assuming balanced biases, fail in reproducing the real network. Since the effectiveness of the MCM has been shown for various other networks including non-economic ones [37], the importance of separately specifying the extensive and intensive biases might actually be a very general result. Important future steps in this direction include the verification of the relationship between trade margins and extensive/intensive biases through the exploration of a complementary, explicitly dynamic framework.

[1] R. Albert and A.-L. Barabási, Reviews of Modern Physics **74**, 47 (2002).

[2] M. O. Jackson, G. Demange, S. Goyal, and A. V. D. Nouwel, in *In Group Formation in Economics: Networks, Clubs and Coalitions* (Cambridge University Press, 2003).

[3] A. Watts, in *Networks and Groups* (Springer, 2003) pp. 337–345.

[4] F. Schweitzer, G. Fagiolo, D. Sornette, F. Vega-Redondo, A. Vespignani, and D. R. White, Science **325**, 422

(2009).
[5] L. Katz and J. H. Powell, Ann. Math. Stat. **28**, 442 (1957).
[6] P. Erdős and A. Rényi, in *Publication of the Mathematical Institute of the Hungarian Academy of Sciences* (1960) pp. 17–61.
[7] P. Holland and S. Leinhardt, Sociological Methodology **7**, 1 (1976).
[8] A. Rao, R. Jana, and S. Bandyopadhyay, Sankhya: Indian Journal of Statistics **58**, 225 (1996).
[9] J. Roberts, Social Networks **22**, 273 (2000).
[10] M. E. J. Newman, S. H. Strogatz, and D. J. Watts, Physical Review E **64**, 026118 (2001).
[11] S. Maslov, K. Sneppen, and A. Zaliznyak, Physica A: Statistical and Theoretical Physics **333**, 529 (2004).
[12] G. Ansmann and K. Lehnertz, Physical Review E **84**, 026103 (2011).
[13] L. Bargigli and M. Gallegati, Journal of Economic Behavior and Organization **78**, 396 (2011).
[14] G. Ellison and E. L. Glaeser, Journal of Political Economy **105**, 889 (1997).
[15] M. Rysman and S. Greenstein, Economics Letters **86**, 405 (2005).
[16] R. Armenter and M. Koren, *A Balls-and-Bins Model of Trade*, Working Paper DP7783 (CEPR, 2010).
[17] A. Serrano and M. Boguñá, Physical Review E **68**, 015101(R) (2003).
[18] D. Garlaschelli and M. Loffredo, Physical Review Letters **93**, 188701 (2004).
[19] D. Garlaschelli and M. Loffredo, Physica A **355**, 138 (2005).
[20] R. Kali and J. Reyes, Journal of International Business Studies **38**, 595 (2007).
[21] R. Kali and J. Reyes, Economic Inquiry **48**, 1072 (2010).
[22] R. Feenstra, *Advanced international trade : theory and evidence* (Princeton University Press, 2004).
[23] T. Squartini, G. Fagiolo, and D. Garlaschelli, Physical Review E **84**, 046117 (2011).
[24] G. Fagiolo, T. Squartini, and D. Garlaschelli, Journal of Economic Interaction and Coordination , 1 (2013).
[25] T. Squartini and D. Garlaschelli, in *Self-Organizing Systems* (Springer, 2012) pp. 24–35.
[26] G. Fagiolo, S. Schiavo, and J. Reyes, Physica A **387**, 3868 (2008).
[27] G. Fagiolo, S. Schiavo, and J. Reyes, Physical Review E **79**, 036115 (2009).
[28] G. Fagiolo, Journal of Economic Interaction and Coordination **5**, 1 (2010).
[29] K. Bhattacharya, G. Mukherjee, J. Sarämaki, K. Kaski, and S. Manna, Journal of Statistical Mechanics: Theory Exp. A **2**, P02002 (2008).
[30] K. Bhattacharya, G. Mukherjee, J. Saramäki, K. Kaski, and S. S. Manna, Journal of Statistical Mechanics: Theory and Experiment **2008**, P02002 (2008).
[31] T. Squartini, G. Fagiolo, and D. Garlaschelli, Physical Review E **84**, 046118 (2011).
[32] M. Á. Serrano and M. Boguñá, arXiv preprint cond-mat/0501750 (2005).
[33] M. Á. Serrano, M. Boguñá, and R. Pastor-Satorras, Physical Review E **74**, 055101 (2006).
[34] A. Fronczak and P. Fronczak, Physical Review E **85**, 056113 (2012).
[35] S. Wells, (2004).
[36] G. Caldarelli, A. Chessa, F. Pammolli, A. Gabrielli, and M. Puliga, Nature Physics **9**, 125 (2013).
[37] R. Mastrandrea, T. Squartini, G. Fagiolo, and D. Garlaschelli, arXiv preprint arXiv:1307.2104 (2013).
[38] D. Ricardo, E. C. K. Gonner, and Q. Li, *The principles of political economy and taxation* (World Scientific, 1819).
[39] L. De Benedictis and L. Tajoli, The World Economy **34**, 1417 (2011).
[40] D. Hummels and P. J. Klenow, *The variety and quality of a nation's trade*, Tech. Rep. (National Bureau of Economic Research, 2002).
[41] S. J. Evenett and A. J. Venables, (2009).
[42] G. J. Felbermayr and W. Kohler, Review of World Economics **142**, 642 (2006).
[43] E. Helpman, M. Melitz, and Y. Rubinstein, The Quarterly Journal of Economics **123**, 441 (2008).
[44] M. Amiti and C. Freund, in *China's Growing Role in World Trade* (University of Chicago Press, 2010) pp. 35–56.
[45] J. Eaton, M. Eslava, M. Kugler, and J. Tybout, "The margins of entry into export markets: Evidence from colombia," (2008).
[46] T. Besedeš and T. J. Prusa, Journal of Development Economics **96**, 371 (2011).
[47] T. Chaney, The American Economic Review **98**, 1707 (2008).
[48] T. Squartini and D. Garlaschelli, New Journal of Physics **13**, 083001 (2011).
[49] Http://comtrade.un.org/.
[50] Http://unstats.un.org/.
[51] We employ the flow as reported by the importer [26] since the importer and exporter records do not always match.
[52] M. Barigozzi, G. Fagiolo, and D. Garlaschelli, Physical Review E **81**, 046104 (2010).
[53] Rounding to integers is required by the randomization procedure (for more details see [48] and [37]).
[54] T. Squartini, F. Picciolo, F. Ruzzenenti, and D. Garlaschelli, Nat. Sci. Rep. **3** (2013), 10.1038/srep02729.
[55] F. Picciolo, T. Squartini, F. Ruzzenenti, R. Basosi, and D. Garlaschelli, in *Signal Image Technology and Internet Based Systems (SITIS), 2012 Eighth International Conference on* (IEEE, 2012) pp. 784–792.
[56] D. Garlaschelli and M. I. Loffredo, Physical Review E **78**, 015101 (2008).
[57] G. Fagiolo, Physical Review E **76**, 026107 (2007).
[58] K. P. Burnham and D. R. Anderson, Sociological methods & research **33**, 261 (2004).