Statistical mechanics of the international trade network

Agata Fronczak[*] and Piotr Fronczak

Faculty of Physics, Warsaw University of Technology, Koszykowa 75, PL-00-662 Warsaw, Poland

(Received 12 April 2011; revised manuscript received 31 January 2012; published 17 May 2012)

Analyzing real data on international trade covering the time interval 1950–2000, we show that in each year over the analyzed period the network is a typical representative of the ensemble of maximally random weighted networks, whose directed connections (bilateral trade volumes) are only characterized by the product of the trading countries' GDPs. It means that time evolution of this network may be considered as a continuous sequence of equilibrium states, i.e., a quasistatic process. This, in turn, allows one to apply the linear response theory to make (and also verify) simple predictions about the network. In particular, we show that bilateral trade fulfills a fluctuation-response theorem, which states that the average relative change in imports (exports) between two countries is a sum of the relative changes in their GDPs. Yearly changes in trade volumes prove that the theorem is valid.

I. INTRODUCTION

In recent years, an extensive research effort has been devoted to analyzing the structure, function, and dynamics of the international trade network (ITN) [1–8] from a complex network perspective [9]. The knowledge of the topological properties of this network and its evolution over time is not only important per se (e.g., because it enhances our descriptive knowledge of the stylized facts pertaining to the ITN), but it may also be relevant to a better explanation of macroeconomic dynamics [10–17].

Here, we use quantitative and numerical (data-driven) methods originating from statistical mechanics to describe and predict the behavior of the ITN. We analyze a set of year-by-year trade relationships between all countries of the world, covering the time interval 1950–2000. Although the total number of countries and the overall economic conditions influencing the network change over the course of the period, in each year the ITN is shown to be a typical representative of the ensemble in which every network G is assigned the probability [18,19] $P(G) \propto e^{-H(G)}$, where $H(G) = \sum_{i,j} \theta_{ij} w_{ij}$ plays the role of network Hamiltonian; w_{ij} represents the volume of trade between two countries, i and j; $\theta_{ij} \propto (x_i x_j)^{-1}$ is the field parameter conjugated to this trade connection; and $x_i x_j$ corresponds to the product of the GDPs of the trade partners.

Beyond the descriptive power of our approach (which has been confirmed in a number of tests reported in this article consisting in a comparison of GDP-driven Monte Carlo simulations of the trade network with real data on the ITN), it is also important to stress the predictive abilities of the model. In particular, we show here that bilateral trade fulfills a simple fluctuation-response theorem [20]. Supported by the well-known qualitative findings about economic crises, we argue that the theorem provides valuable quantitative insights into the dynamics of the ITN in the time of crises, and, in the future, may be used to uncover mechanisms underlying the emergence of worldwide crises.

II. DATA DESCRIPTION

The results reported in this paper are based on the trade data collected by Gleditsch [21] that contains, for each world country in the period 1950–2000, the detailed list of bilateral import and export volumes. The data are employed to build a sequence of matrices $\mathbf{W}(t)$ corresponding to snapshots of the weighted directed ITN in the consecutive years $t = 1950, \ldots$. In the network, each country is represented by a node and the direction of links follows that of wealth flow. The entry $w_{ij}(t)$ of the trade matrix $\mathbf{W}(t)$ represents the weight of the directed connection. From the point of view of the country denoted by i, $w_{ij}(t)$ refers to the volume of export to j, while, from the point of view of the country labeled by j, it is seen as the volume of import from i. Precisely due to differences in reporting procedures between countries, when analyzing the data one often encounters small deviations between exports from i to j and imports from i to j. To overcome the problem, in our analysis we define $w_{ij}(t)$ as the arithmetic average of the two values.

In this article, apart from trade matrices, which contain complete but often excessively detailed information about the ITN, we also use several other quantities that make a theoretical description of the network possible. In particular, to characterize the economic performance of a country we use its total GDP value $x_i(t)$. To get the whole set of total GDPs, $\{x_i(t)\}$, we simply multiply the GDP *per capita* by the population size of each country [22]. Furthermore, to describe the intensity of the trade relationships of a country we define the so-called outstrength $s_i^{\text{out}}(t)$ and instrength $s_i^{\text{in}}(t)$ of the corresponding node. The quantities are calculated as the total weight of connections (outgoing and incoming, respectively) that are attached to the node, and they represent total volumes of exports and imports of the considered country in a given year t.

All the data used in this paper are given in millions of contemporary US dollars. The disturbing effects of inflation are ruled out in a natural way by the fact that whenever the variables, $x_i(t)$ or $w_{ij}(t)$, are used in the calculations, they are intrinsically divided by the normalization constant that equals the sum of all variables of a given type, i.e., $X(t) = \sum_i x_i(t)$ or $T(t) = \sum_{i,j} w_{ij}(t)$, respectively. (In what follows,

_____
[*]agatka@if.pw.edu.pl

whenever there is no confusion we will often omit the explicit time dependence of the quantities). As a byproduct, the above observation allows one to present the main results of this article in a very concise way, with the help of relative quantities defined as follows:

$$\xi_i = \frac{x_i}{X}, \quad \sigma_i^{\text{out}} = \frac{s_i^{\text{out}}}{T}, \quad \sigma_i^{\text{in}} = \frac{s_i^{\text{in}}}{T}, \quad v_{ij} = \frac{w_{ij}}{T}. \quad (1)$$

III. EXPONENTIAL RANDOM GRAPHS: STRUCTURAL HAMILTONIAN OF THE INTERNATIONAL TRADE NETWORK

The first contributions dealing with international trade from a complex network perspective used a binary-network approach, in which one has assumed that a (possibly directed) link between any two countries is either present or not, depending on whether the trade volume that it carries is larger than a given threshold [1–3]. With reference to this line of research we would like to highlight the paper by Garlaschelli and Loffredo [3]. In the paper, the authors used the same real-world data to analyze an unweighted and undirected version of the ITN, i.e., a network of partnership in trade. They have shown that the total GDP of a country, x_i, can be identified with the fitness variable [23] that, once a form of the probability of the trade connection between two countries is introduced, completely determines the expected structural properties of this network. This in turn implies that the ITN viewed as a binary network is a typical representative of exponential random graphs [18]. Furthermore, one can expect that the same holds true for the weighted version of this network [19,24].

To verify this conjecture, we start by considering the most general ensemble of directed weighted networks, which is described by the following Hamiltonian,

$$H(G) = \sum_i \sum_{j \neq i} \theta_{ij} w_{ij}, \quad (2)$$

with a separate parameter θ_{ij} coupling to each weighted connection. Our aim is to check whether the Hamiltonian is correct and, if so, how the parameters $\{\theta_{ij}\}$ depend on different indicators characterizing the global economy. To do this, we first examine the ensemble as it stands.

Thus, given that w_{ij} is a real number greater than 0 (as is true for trade volumes), the partition function of this ensemble can be written as

$$Z(\{\theta_{ij}\}) = \prod_i \prod_{j \neq i} \int_0^\infty e^{-\theta_{ij} w_{ij}} dw_{ij} = \prod_i \prod_{j \neq i} \frac{1}{\theta_{ij}}. \quad (3)$$

This allows us to rewrite the probability of a network as $P(G) = e^{-H(G)}/Z = \prod_i \prod_{j \neq i} p_{ij}$, where

$$p_{ij}(w_{ij}) = e^{-\theta_{ij} w_{ij}} \theta_{ij} \quad (4)$$

is the probability that there is a directed link of weight w_{ij} from i to j. The expression for p_{ij} that we arrive at is the exponential distribution. Its mean value,

$$\langle w_{ij} \rangle = 1/\theta_{ij}, \quad (5)$$

can be used to calculate the average values of a node's outstrength and instrength:

$$\langle s_i^{\text{out}} \rangle = \sum_{j \neq i} \langle w_{ij} \rangle = \sum_{j \neq i} \frac{1}{\theta_{ij}} \quad \text{and} \quad \langle s_i^{\text{in}} \rangle = \sum_{j \neq i} \frac{1}{\theta_{ji}}. \quad (6)$$

At this stage one can start to make comparisons of theoretical predictions with the empirical data on international trade. With good reason, it is convenient to begin by putting together Eq. (6) and the corresponding empirical relations [see Figs. 1(a)–1(d)]:

$$\langle s_i^{\text{out}} \rangle = A x_i \quad \text{and} \quad \langle s_i^{\text{in}} \rangle = A x_i, \quad (7)$$

where A is the time-dependent parameter having the same value for both the outstrength and instrength of the nodes. Analyzing the expressions, one finds that the simplest way to merge the theoretical approach with real data is to assume a multiplicative form of the parameter θ_{ij}, i.e.,

$$\theta_{ij} = \theta_i \theta_j, \quad (8)$$

where θ_i and θ_j represent some single-node characteristics controlling for the potential ability of the two nodes to be connected. One should note that the symmetric expression for θ_{ij}, Eq. (8), is consistent with observations made by other authors, showing the symmetric character of bilateral trade relations (see, e.g., [3,8]).

To calculate explicit values of all the parameters $\{\theta_i\}$, one just has to insert Eq. (8) into the theoretical formula for $\langle s_i^{\text{out}} \rangle$, Eq. (6), and then equate the obtained relation to the empirical one, Eq. (7). (The analogous calculations can be done for $\langle s_i^{\text{in}} \rangle$.) As a result, one gets the following expression: $\langle s_i^{\text{out}} \rangle = \sum_j \theta_j^{-1}/\theta_i = A x_i$, which, when summed over i, yields an important relation between theoretical and empirical quantities describing the ITN: $T = (\sum_i \theta_i^{-1})^2 = AX$, from which it follows that

$$\theta_i = \frac{1}{\sqrt{T}} \frac{X}{x_i} = \frac{1}{\sqrt{T}} \frac{1}{\xi_i} \quad \text{and} \quad \theta_{ij} = \frac{1}{T} \frac{1}{\xi_i \xi_j}, \quad (9)$$

where ξ_i has been introduced earlier, in Eq. (1). Now, one can insert (9) into (5). As a result one gets a stochastic version of the gravity model of trade [25–27],

$$\langle w_{ij} \rangle = \frac{T}{X^2} x_i x_j. \quad (10)$$

In comparison with the standard gravity equation for bilateral trade volume between i and j,

$$w_{ij} = \text{const} \frac{x_i x_j}{f(D_{ij})}, \quad (11)$$

where $f(D_{ij})$ represents transport costs in international trade, which depend on the geographic distance between the two trading countries.

The expressions, Eqs. (9), together with other relative parameters defined in Eq. (1), can be used to rewrite the most important results of this approach in a very concise way. In particular, as described in terms of trade, the average outstrength and instrength of a node, Eq. (6), when divided by the world's trade volume T turns out to be equal to the country's share in the world's GDP X, i.e.,

$$\langle \sigma_i^{\text{out}} \rangle = \langle \sigma_i^{\text{in}} \rangle = \xi_i. \quad (12)$$
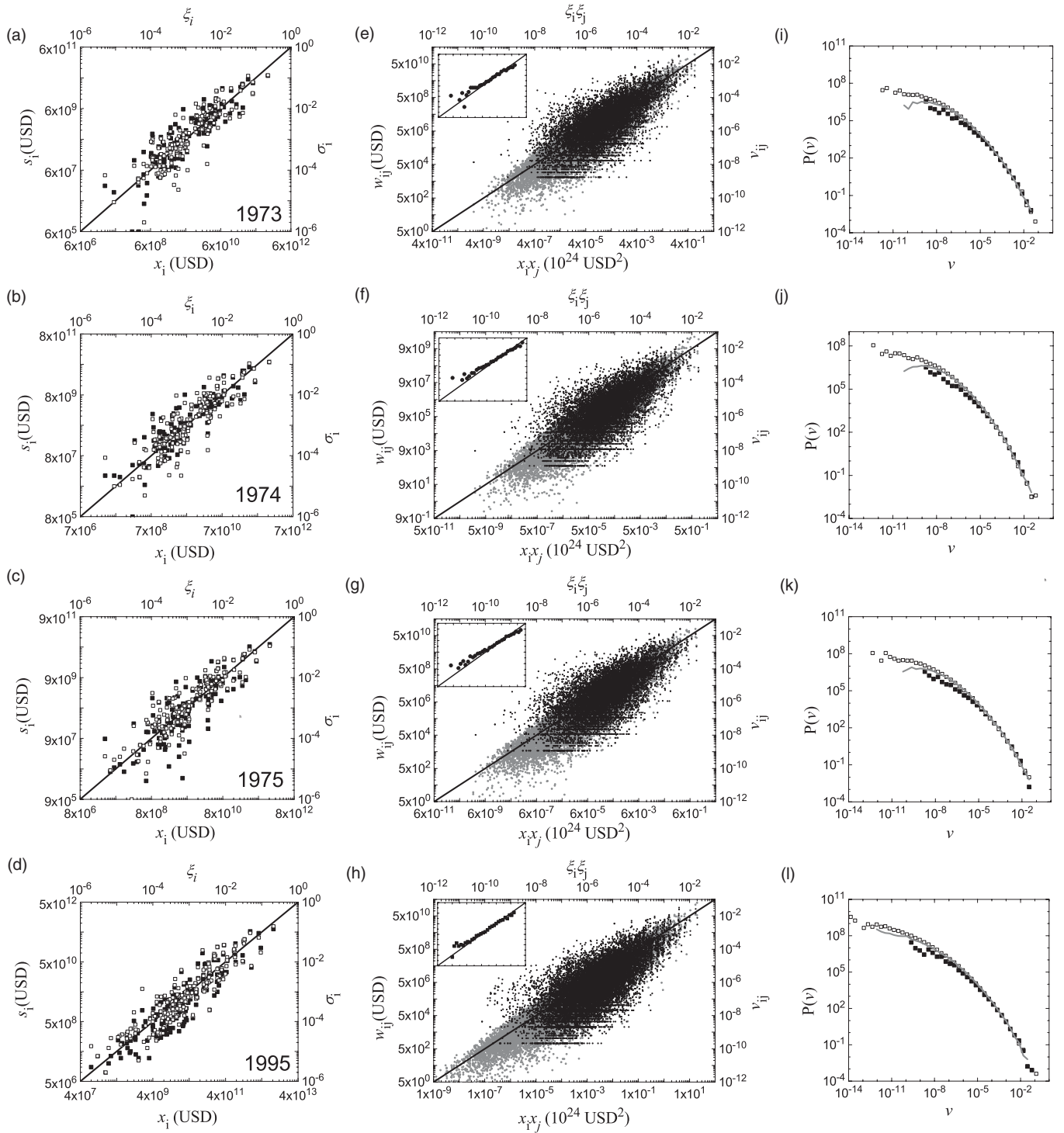
FIG. 1. Structural properties of single snapshots of the ITN. Diagrams placed in the same row refer to the same year. The data are shown in two ways, by using regular and relative quantities; cf. Eq. (1). (a)–(d) Total import and export volumes of all world countries in 1973, 1974, 1975, and 1995 versus GDP (filled and open points correspond to imports and exports, respectively) and their comparison with the expected values described by Eqs. (7) and (12) (solid straight lines). (e)–(h) Bilateral trade flows in the following years versus the product of the trading countries' GDPs (points) as compared with their theoretical predictions based on Eqs. (5) and (13) (lines). Black points correspond to real data, while gray points represent trade volumes obtained from GDP-driven Monte Carlo simulations. In the insets of all the panels [(e)–(h)], comparisons of the expected theoretical values (straight lines) to mean values of the real data are shown. Since trade flows smaller than a given threshold are rarely specified in economic reports (in particular, the considered data set [21] does not contain trade volumes smaller than 1000 USD), the clouds of black points cover smaller areas than the ones corresponding to numerical simulations. (i)–(l) Distributions of trade volumes in the considered years. Filled and open squares correspond to real and simulated data, respectively. The solid lines represent distributions of expected trade flows which, for each pair of countries in a given year, can be calculated using Eq. (5) or (13).

In a similar fashion, the average weight of a directed connection when divided by T, is given by

$$\langle v_{ij} \rangle = \xi_i \xi_j = 1/\eta_{ij}. \tag{13}$$

Finally, the structural network Hamiltonian, Eq. (2), when written in relative variables has the following form:

$$H(G) = \sum_i \sum_{j \neq i} \eta_{ij} v_{ij}. \tag{14}$$

To verify the correctness of the assumed network Hamiltonian, a series of data-driven Monte Carlo simulations employing the Metropolis algorithm has been performed. The

obtained results for four years, 1973, 1974, 1975, and 1995, are shown in Fig. 1 (diagrams placed in the same row refer to one year marked in the left chart). The first three years correspond to the 1973 oil crisis which caused high inflation and a global recession which affected all aspects of living within the 1970s. The figure shows that global economic crises, such as the one triggered by the 1973 oil crisis, do not affect the quality of our approach. In particular, in Figs. 1(e)–1(h), sets of all bilateral trade volumes recorded in a given year versus the product of the trading countries' GDPs are compared with the corresponding set of weights of directed connections in a typical network of the considered ensemble. Although the two sets (clouds)
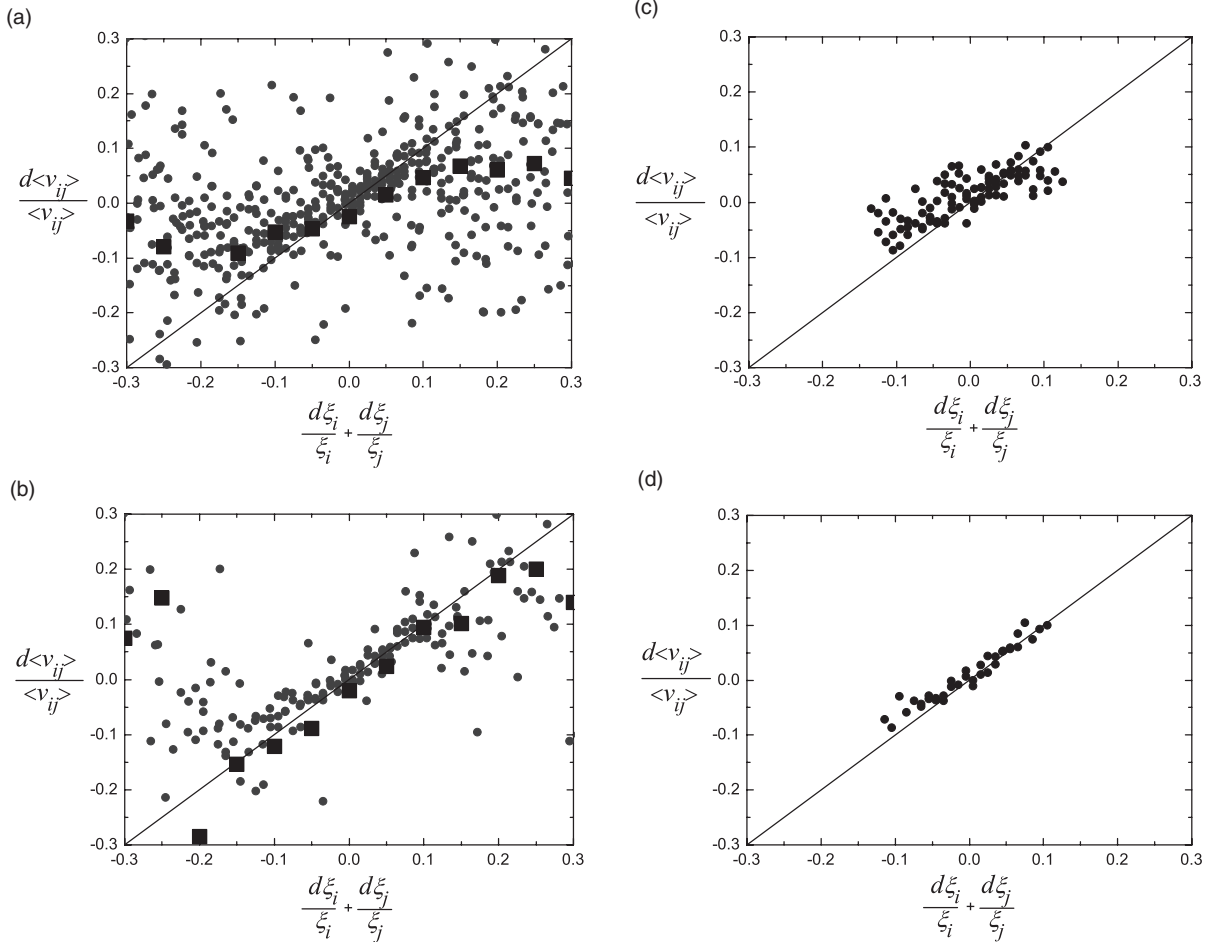

FIG. 2. Fluctuation-response theorem for the ITN. To confirm validity of the theorem, for each connection in the trade network at the turns of 1950 to 2000, one has calculated quantities corresponding to both sides of Eq. (16): the relative change in normalized trade volume, $dv_{ij}/v_{ij} \simeq [v_{ij}(t+1) - v_{ij}(t)]/v_{ij}(t)$, and the sum of relative changes in the GDP of trading countries, $d\xi_i/\xi_i + d\xi_j/\xi_j$, where $d\xi_i/\xi_i \simeq [\xi_i(t+1) + \xi_i(t)]/\xi(t)$. Then, the results were grouped according to similarities in both: local fields conjugated to the corresponding trade flows, η_{ij}, Eq. (13), and their year-by-year changes, $d\xi_i/\xi_i + d\xi_j/\xi_j$, characterizing the initial economic conditions influencing trade and the magnitude of the applied perturbation, respectively. Each group of connections, $V(m,n)$, was characterized by two integers, m and n. The trade flow v_{ij} was classified as belonging to $V(m,n)$ if $m - 1 \leqslant \ln \eta_{ij} < m$ and $n - 1 \leqslant 100 \, d\eta_{ij} < n$. [Note that the grouping with respect to m is in fact logarithmic binning with respect to the expected trade volume, Eq. (13), while the parameter m describes linear binning with respect to the sum of percentage changes in relative GDPs of trade partners.] The gray circles shown in the figure correspond to geometric averaging of the results pertaining to single connections of the ITN. The averaging was employed over the predefined groups $V(m,n)$. The black squares correspond to gray points averaged over the horizontal axis. More precisely, the upper left panel (a) shows all the data such as they are described above. The upper right panel (b) presents the same data with the additional condition: it illustrates the fluctuation-response theorem as applied to trade volumes whose expected share $\langle v_{ij} \rangle$, Eq. (13), in the global trade was greater than 10^{-4}. The lower left panel (c) presents those points from the panel (a), for which the number of single trade volumes contributing to the geometric average was greater than 1000. Finally, in the lower right panel (d), the gray circles which meet the additional conditions specified in panels (b) and (c) are shown.

of points are quite dispersed, they overlap significantly, and their shape is well described by Eq. (13). Moreover, as shown in Figs. 1(i)–1(l), the distributions of trade volumes within these clouds fit very well with each other and agree with the distribution of expected trade flows, $P(\langle v_{ij}\rangle)$, testifying in favor of our simple, yet realistic, approach.

IV. SUSCEPTIBILITY OF TRADE VOLUME TO CHANGES IN GDP: FLUCTUATION-RESPONSE RELATION FOR THE ITN

Extensive comparisons between real data on international trade and its GDP-driven Monte Carlo simulations show that, although the total number of world's countries, $N(t)$, and their GDPs $\{x_i(t)\}$ change over the analyzed period of 50 years, the ITN is continuously well characterized by the same Hamiltonian. This means that the time evolution of this network may be considered as a continuous sequence of equilibrium states (i.e., a quasistatic process) that is yearly sampled by the national reporting procedures. Furthermore, since differences between snapshots of the ITN in the consecutive years are rather small, one can expect that they could be described with the help of linear response theory, of which the simplest (but not yet trivial) result is the fluctuation-response theorem [20].

In the case of exponential random graphs with the Hamiltonian given by Eq. (14), the fluctuation-response theorem has the following form:

$$\langle v_{ij}^2\rangle - \langle v_{ij}\rangle^2 = -d\langle v_{ij}\rangle/d\eta_{ij}. \tag{15}$$

The left-hand side of this expression describes fluctuations in relative weight, $v_{ij} = w_{ij}/T$, of the directed connection from i to j, whereas its right-hand side characterizes susceptibility of this link to its conjugated local field η_{ij}. The susceptibility is defined as the derivative of $\langle v_{ij}\rangle$ with respect to η_{ij} and describes what happens with the expected trade volume $\langle v_{ij}\rangle$ when one changes the parameter η_{ij}, which determines the external conditions for the bilateral exchange.

Taking into account the exponential distribution of weights, $p(v_{ij}) = e^{-\eta_{ij}v_{ij}}\eta_{ij}$, which follows from Eq. (4) and whose variance equals the square of the mean, the fluctuation-response theorem, Eq. (15), can be transformed into the formula

$$\frac{d\langle v_{ij}\rangle}{\langle v_{ij}\rangle} = -\frac{d\eta_{ij}}{\eta_{ij}} = \frac{d\xi_i}{\xi_i} + \frac{d\xi_j}{\xi_j}. \tag{16}$$

Written in such a way, the theorem states that relative changes in normalized (i.e., divided by T) bilateral trade volumes can be estimated on the basis of changes in the GDP of trade partners. Yearly changes in import/export volumes between different countries prove that the fluctuation-response theorem for the ITN is correct (see Fig. 2). Although, we have obtained Eq. (16) as the fluctuation-response relation for exponential

random graphs, it is obvious that it can be also obtained as the logarithmic derivative of Eq. (13).

Relying on Eq. (16) one can, for example, expect that a decline of, say, 2% in the relative GDP of a country, given that its trade partners do not change their share of the world's GDP, will translate into a similar decline in all its bilateral trade volumes. The example shows that the theorem can be used to make simple predictions about the world-wide diffusion of trade-based economic perturbations. Furthermore, assuming that the structure of the ITN is a proxy for meaningful financial linkages between countries, the expression may also help one to understand how financial ripples originating in one country propagate to other countries, giving rise (or not) to global financial contagion. In particular, in the light of this theorem, it becomes apparent that a crisis is amplified if the epicenter country is better integrated into the trade network [15]. This happens because the decline of its GDP, through the proportional downward effect on trade and financial linkages, affects more countries. It also becomes clear that the impact of a crisis on any target country is cushioned if the country in question is better integrated into the ITN, due to the fact that, in the case of such countries, the decline of only one bilateral trade volume does not significantly influence GDP.

V. CONCLUDING REMARKS

The current economic crisis illustrates a critical need for new and fundamental understandings of the structure and dynamics of economic networks. This sentence opens the perspective article from the special issue of *Science* [9] on complex systems and networks entitled "Economic networks: The new challenges" by Schweitzer *et al*. In their article [11], the authors summarize what we know and what we need to know about different economic networks (including the ITN) to reduce the risk of global depression and to design effective strategies to promote economic recovery [28]. Our approach to the ITN is in line with this challenging research area. Having the mathematically tractable yet realistic model of the ITN introduced here, and given its quasistatic time evolution, we believe that, apart from the fluctuation-response theorem, Eq. (16), other well-known results of nonequilibrium statistical physics [29] may be applied to estimate recession (or economic growth) impact on international trade.

The exponential random graphs are in common use within the statistics and social network analysis communities as a practical tool for modeling networks for a few decades. Although such modeling gives a good qualitative description of the data, to our best knowledge, there is no evidence so far that with such an approach one could achieve excellent quantitative accuracy and reproducibility of real world phenomena.

[1] M. A. Serrano and M. Boguñá, Phys. Rev. E **68**, 015101(R) (2003).

[2] X. Li, Y. Y. Jin, and G. Chen, Physica A **328**, 287 (2003).

[3] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **93**, 188701 (2004).

[4] M. A. Serrano, M. Boguñá, and A. Vespignani, J. Econ. Interac. Coord. **2**, 111 (2007).

[5] K. Bhattacharya *et al*., J. Stat. Mech. (2008) P02002.

[6] G. Fagiolo, J. Reyes, and S. Schiavo, Phys. Rev. E **79**, 036115 (2009).

[7] A. Garas *et al*., New J. Phys. **12**, 113043 (2010).

[8] G. Fagiolo, J. Reyes, and S. Schiavo, J. Evolutionary Econ. **20**, 479 (2010).

[9] *Complex Systems and Networks*, special issue of Science **325**, 357 (2009).

[10] R. Kali and J. Reyes, J. Int. Bus. Stud. **38**, 595 (2007).

[11] F. Schweitzer *et al*., Science **325**, 422 (2009).

[12] F. Schweitzer *et al*., Adv. Complex Syst. **12**, 407 (2009).

[13] J. Reichardt and D. R. White, Eur. Phys. J. B **60**, 217 (2007).

[14] J. Reyes, S. Schiavo, and G. Fagiolo, Adv. Complex Syst. **11**, 685 (2008).

[15] R. Kali and J. Reyes, Econ. Inquiry **48**, 1072 (2010).

[16] D. Garlaschelli *et al*., Eur. Phys. J. B **57**, 159 (2007).

[17] D. Wen-Qi and S. Bo-Liang, Chin. Phys. Lett. **26**, 098902 (2009).

[18] J. Park and M. E. J. Newman, Phys. Rev. E **70**, 066117 (2004).

[19] D. Garlaschelli and M. I. Loffredo, Phys. Rev. Lett. **102**, 038701 (2009).

[20] A. Fronczak, P. Fronczak, and J. A. Hołyst, Phys. Rev. E **73**, 016108 (2006).

[21] K. S. Gleditsch [http://privatewww.essex.ac.uk/~ksg/exptradegdp. html].

[22] A. Heston, R. Summers, and B. Aten, *World Table Version 6.1* (Center for International Comparisons of Production, Income and Prices at the University of Pennsylvania, 2002).

[23] G. Caldarelli, A. Capocci, P. DeLosRios, and M. A. Munoz, Phys. Rev. Lett. **89**, 258702 (2002).

[24] In Ref. [16], Garlaschelli et al. presented a class of generalized statistics that describes the structural correlations of weighted networks. Although the equations we derive can be recognized as a specific example of this class, in our paper these equations emerge from empirical observations of real data, not from *ad hoc* theoretical assumptions [cf. Eq. (3) in Ref. [17]].

[25] J. E. Anderson, Amer. Econ. Rev. **69**, 106 (1979).

[26] J. H. Bergstrand, Rev. Econ. Stat. **67**, 474 (1985).

[27] R. Feenstra, *Advanced International Trade: Theory and Evidence* (Princeton University Press, Princeton, 2004).

[28] T. Lux and F. Westerhoff, Nat. Phys. **5**, 2 (2009).

[29] D. Ruelle, Phys. Today **57**(5), 48 (2004).