

# Statistical mechanics of the international trade network

Agata Fronczak, Piotr Fronczak\*

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

(Dated: April 18, 2011)

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. quasi-static 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 fluctuation-response theorem, which states that the average relative change in import (export) between two countries is a sum of relative changes in their GDPs. Yearly changes in trade volumes prove that the theorem is valid. Supported by the well-known qualitative findings about economic crises, we argue that the theorem provides valuable quantitative insights into the mechanisms underlying the emergence of worldwide crises.

PACS numbers: 89.65.Gh, 89.75.-k, 05.40.-a, 02.10.Ox

In recent years, an extensive research effort has been devoted to analyzing the structure, function, and dynamics of the international trade network (ITN) [1–6] from a complex network perspective [7]. The knowledge of 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 [8–10]. In particular, it has been suggested that the analysis of ITN may help in recognizing the pattern of economic interdependencies responsible for the propagation of crises across countries [11–13] and that it can be used to explain the role of international trade in spurring the efficiency of economic recovery after recession [14, 15]. The approach to international trade described in this article is a significant step towards clarifying the above-mentioned macroeconomic issues.

Here, we use quantitative and numerical (data-driven) methods originating from statistical mechanics to describe and predict the behavior of ITN. We analyze a set of year-by-year trade relationships between all countries of the world, covering the time interval 1950 – 2000 [16]. Although the total number of countries and the overall economic conditions influencing the network change over the course of the period, in each year ITN is shown to be a typical representative of the ensemble in which every network,  $G$ , is assigned the probability [17, 18]  $P(G) \propto e^{-H(G)}$ ; 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.

Behind the descriptive power of our approach (which has been confirmed in a number of tests reported in this article consisting in comparison of GDP-driven Monte

Carlo simulations of the trade network with real data on 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 [22], which states that the average relative change in trade is a sum of the relative changes in the GDPs of trade partners. We argue that the theorem provides valuable insights into mechanisms underlying the emergence and propagation of worldwide crises. Yearly changes in import/export volumes between countries prove that the theorem is valid.

The results reported in this work are based on the empirical analysis of expanded trade data collected by K. S. Gleditsch [16]. The data set [19] contains, for each world country in the period 1950 – 2000, the detailed list of bilateral import and export volumes. The data on the population size of each country and its GDP *per capita* has been taken from the Penn World Tables (PWT) [20].

The trade data are employed to build a sequence of matrices,  $\mathbf{W}(t)$ , corresponding to snapshots of weighted directed ITN in the consecutive years,  $t = 1950, \dots, 2000$ . 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 ITN, we also use several other quantities that make

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. Furthermore, to describe the intensity of the trade relationships of a country we define the so-called out-strength,  $s_i^{out}(t)$ , and in-strength,  $s_i^{in}(t)$ , of the corresponding node. The quantities are calculated as the total weight of connections (outgoing and incoming, respectively) that are attached to the node, e.g.  $s_i^{out}(t) = \sum_j w_{ij}(t)$ , and they represent total volumes of export and import of the considered country in a given year,  $t$ .

Finally, we would like to stress that both GDP data,  $\{x_i(t)\}$ , and trade matrices,  $\mathbf{W}(t)$ , used in this study are given in millions of contemporary U.S. dollars (that is, in terms of the value of one U.S. dollar in the reported year,  $t$ ). To factor out the effects of inflation and to allow for a direct comparison between snapshots of ITN in different years, one usually expresses the data in standard reference money units (e.g. in 1996 U.S. dollars [14]). In our approach, however, the disturbing effects mentioned 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, whenever there is no confusion we will often omit 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:

$$\xi_i = x_i/X, \quad \sigma_i^{(\dots)} = s_i^{(\dots)}/T \quad \text{and} \quad v_{ij} = w_{ij}/T. \quad (1)$$

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 (2004), published in *Physical Review Letters* [3]. In the paper, the authors used the same real-world data to analyze an unweighted and undirected version of 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 [21] that, once a form of the probability of trade connection between two countries is introduced, completely determines the expected structural properties of this network. This in turn implies that ITN viewed as a binary network is a typical representative of exponential random graphs [17]. Furthermore, one can expect that the same holds true for the weighted version of this network [18].

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  $\theta_{ij}$  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. In order 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,  $P(G) = e^{-H(G)}/Z$ , in the following way

$$P(G) = \prod_i \prod_{j \neq i} p_{ij}(w_{ij}), \quad (4)$$

where

$$p_{ij}(w_{ij}) = e^{-\theta_{ij} w_{ij}} \theta_{ij} \quad (5)$$

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

$$\langle w_{ij} \rangle = 1/\theta_{ij}, \quad (6)$$

can be used to calculate the average values of a node's out- and in-strength

$$\langle s_i^{out} \rangle = \sum_{j \neq i} \langle w_{ij} \rangle = \sum_{j \neq i} \frac{1}{\theta_{ij}} \quad \text{and} \quad \langle s_i^{in} \rangle = \sum_{j \neq i} \frac{1}{\theta_{ji}}. \quad (7)$$

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. (7) and the corresponding empirical relations (see Fig. 1 a):

$$\langle s_i^{out} \rangle = Ax_i \quad \text{and} \quad \langle s_i^{in} \rangle = Ax_i, \quad (8)$$

where  $A$  is the time-dependent parameter having the same value for both out- and in-strength 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  $\theta_{ij}$ , i.e.

$$\theta_{ij} = \theta_i \theta_j, \quad (9)$$

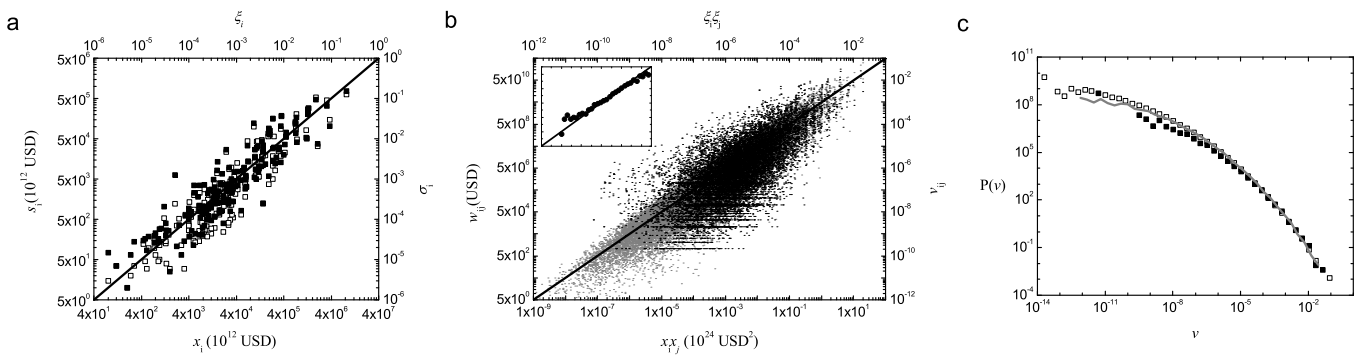


FIG. 1. **Structural properties of a single snapshot of ITN.** The data are shown in two ways, by using regular and relative quantities, cf. Eq. (1). **a**, Total import and export volumes of all world countries in 1995 vs. GDP (filled and open points, respectively) and their comparison with the expected values described by Eqs. (8) and (11) (solid lines). **b**, Bilateral trade flows in 1995 vs. the product of the trading countries' GDPs (points) as compared with their theoretical prediction based on Eqs. (6) and (12) (lines). Black points correspond to real data, while gray points represent trade volumes obtained from GDP-driven Monte Carlo simulations. Since trade flows smaller than a given threshold are rarely specified in economic reports (in particular, the considered data set [19] does not contain trade volumes smaller than 1000 USD), the cloud of black points cover smaller area than the one corresponding to numerical simulations. The effect of unreported exports/imports is also perceptible in the insets of both panels, which show the relationship between the average trade volume and the product of GDPs obtained with logarithmic binning of the latter. In the insets, points represent real data, while the solid lines stand for their theoretical prediction. **c**, Distributions of trade volumes in 1995. 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 Eqs. (6) or (12).

where  $\theta_i$  and  $\theta_j$  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  $\theta_{ij}$ , Eq. (9), is consistent with observations made by other authors, showing the symmetric character of bilateral trade relations ( see e.g. [3, 6]).

To calculate explicit values of all the parameters  $\{\theta_i\}$ , one just has to insert Eq. (9) into the theoretical formula for  $\langle s_i^{out} \rangle$ , Eq. (7), and then equate the obtained relation to the empirical one, Eq. (8). (The analogous calculations can be done for  $\langle s_i^{in} \rangle$ .) As a result, one gets the following expression:  $\langle s_i^{out} \rangle = \sum_j \theta_j^{-1} / \theta_i = A x_i$ , which, when summed over  $i$ , yields an important relation between theoretical and empirical quantities describing ITN:  $T = (\sum_i \theta_i^{-1})^2 = A X$ , from which it follows that

$$\theta_i = \frac{1}{\sqrt{T}} \frac{X}{x_i} = \frac{1}{\sqrt{T}} \frac{1}{\xi_i} \quad \text{and} \quad \theta_{ij} = \frac{1}{T} \frac{1}{\xi_i \xi_j}, \quad (10)$$

where the relative quantity  $\xi_i$  has been introduced earlier, in Eq. (1).

The expressions, Eqs. (10), together with other relative parameters that were 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 out- and in-strength of a node, Eq. (7), 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.

$$\langle \sigma_i^{out} \rangle = \langle \sigma_i^{in} \rangle = \xi_i. \quad (11)$$

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

$$\langle v_{ij} \rangle = \xi_i \xi_j = \frac{1}{\eta_{ij}}. \quad (12)$$

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

$$H(G) = \sum_i \sum_{j \neq i} \eta_{ij} v_{ij}. \quad (13)$$

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 are shown in Figure 1. In particular, in Figure 1 b, the set of all bilateral trade volumes recorded in a given year versus the product of the trading countries' GDPs is compared with the corresponding set of weights of directed connections in a typical network of the considered ensemble. Although the two sets (clouds) of points are quite dispersed, they overlap significantly, and their shape is well-described by Eq. (12). Moreover, as shown in the panel c in the same figure, 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.

In general, extensive comparisons between real data on international trade and its GDP-driven Monte Carlo simulations show that, although the total number of world's

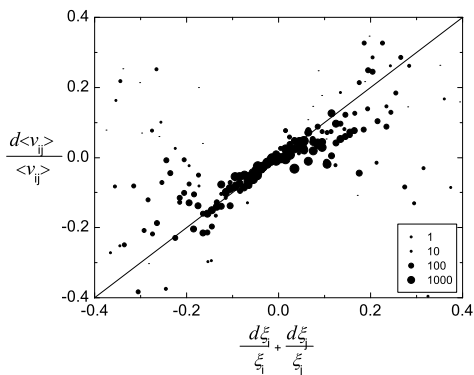


FIG. 2. **Fluctuation-response theorem for ITN.** To confirm validity of the theorem, for each connection in the trade network at the turns of 1950 – 2000, one has calculated quantities corresponding to both sides of Eq. (15): the relative change in normalized trade volume,  $dv_{ij}/v_{ij} \simeq (v_{ij}(t+1) - v_{ij}(t))/v_{ij}(t)$ , and the sum of relative changes in GDP of trading countries,  $d\xi_i/\xi_i + d\xi_j/\xi_j$ , where  $d\xi_i/\xi_i \simeq (\xi_i(t+1) + \xi_i(t))/\xi_i(t)$ . Then, the results were grouped according to similarities in both: local fields conjugated to the corresponding trade flows,  $\eta_{ij}$ , Eq. (12), and their year-by-year changes,  $d\eta_{ij} = d\xi_i/\xi_i + d\xi_j/\xi_j$ , 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  $m - 1 \leq \ln \eta_{ij} < m$  and  $n - 1 \leq 100 d\eta_{ij} < n$ . (Note that the grouping with respect to  $m$  is in fact logarithmic binning with respect to the expected trade volume, Eq. (12), while the parameter  $m$  describes linear binning with respect to the sum of percentage changes in relative GDPs of trade partners.) Points shown in the figure were obtained by averaging results pertaining to single connections of ITN over the predefined groups,  $V(m, n)$ . Sizes of points depend on the number of trade volumes contributing to the average, according to the key in the figure. The figure illustrates the fluctuation-response theorem as applied to trade volumes whose expected share,  $\langle v_{ij} \rangle$ , Eq. (12), in the global trade,  $T$ , is greater than  $10^{-4}$ .

countries,  $N(t)$ , and their GDPs,  $\{x_i(t)\}$ , change over the analyzed period of 50 years, 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. quasi-static process) that is yearly sampled by the national reporting procedures. Furthermore, since differences between snapshots of 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 [22] applying to equilibrium ensembles.

In the case of exponential random graphs with the Hamiltonian of the form  $H(G) = \sum_i \sum_{j \neq i} \eta_{ij} v_{ij}$ ,

Eq. (13), the fluctuation-response theorem is given by

$$\langle v_{ij}^2 \rangle - \langle v_{ij} \rangle^2 = -\frac{d\langle v_{ij} \rangle}{d\eta_{ij}}. \quad (14)$$

The l.h.s. of this expression describes fluctuations in relative weight,  $v_{ij} = w_{ij}/T$ , of the directed connection from  $i$  to  $j$ , whereas its r.h.s. characterizes susceptibility of this link to its conjugated local field,  $\eta_{ij}$ . The susceptibility is defined as the derivative of  $\langle v_{ij} \rangle$  with respect to  $\eta_{ij}$  and describes what happens with the expected trade volume,  $\langle v_{ij} \rangle$ , when one changes the parameter  $\eta_{ij}$ , which determines external conditions for the bilateral exchange. In simple words, the theorem states that in equilibrium the average response of  $v_{ij}$  to changes in  $\eta_{ij}$  is proportional to fluctuations in  $v_{ij}$ .

Taking into account the exponential distribution of weights,  $p(v_{ij}) = e^{-\eta_{ij} v_{ij}} \eta_{ij}$ , which follows from Eq. (5) and whose variance equals the square of the mean, the fluctuation-response theorem, Eq. (14), can be transformed into the formula

$$\frac{d\langle v_{ij} \rangle}{\langle v_{ij} \rangle} = -\frac{d\eta_{ij}}{\eta_{ij}} = \frac{d\xi_i}{\xi_i} + \frac{d\xi_j}{\xi_j}. \quad (15)$$

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 ITN is correct (see Figure 2).

Relying on Eq. (15) one can, for example, expect that a decline of, say, 2 percent 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 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 [13]. 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 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.

*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* [7] on complex systems and networks entitled "Economic networks: The new challenges" by F. Schweitzer et al. In their article [9], the authors summarize what we know and what we need to know about different economic networks (including ITN) to reduce the risk of global depression and to design effective strategies to promote economic recovery [23]. Our approach to ITN, exploiting ideas originating in statistical mechanics, is in line with this challenging research area. Having the mathematically tractable yet realistic model of ITN introduced here, and given the quasi-static time evolution of this network, we believe that, apart from the fluctuation-response theorem, Eq. (15), other well-known results of non-equilibrium statistical physics [24] may be applied to estimate recession (or economic growth) impact on international trade.

This work was supported by the Polish Ministry of Science, Grant No. 496/N-COST/2009/0.

---

\* agatka@if.pw.edu.pl%; <http://www.if.pw.edu.pl/~agatka%>;

- [1] M. Ángeles 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. Loffredo, *Phys. Rev. Lett.* **93**, 188701 (2004).
- [4] K. Bhattacharya, G. Mukherjee, J. Saramäki, K. Kaski, and S.S. Manna, *J. Stat. Mech.* P02002 (2008).
- [5] G. Fagiolo, J. Reyes, and S. Schiavo, *Phys. Rev. E* **79**, 036115 (2009).
- [6] G. Fagiolo, J. Reyes, and S. Schiavo, *J. Evolutionary Econ.* **20**, 479 (2010).
- [7] *Complex systems and networks*, Special issue, *Science*, 325 (2009), pp. 357–504.
- [8] R. Kali and J. Reyes, *J. Int. Bus. Stud.* **38**, 595 (2007).
- [9] F. Schweitzer et al., *Science* **325**, 422 (2009).
- [10] F. Schweitzer et al., *Adv. Complex Syst.* **12**, 407 (2009).
- [11] J. Reichardt and D.R. White, *Eur. Phys. J. B* **60**, 217 (2007).
- [12] J. Reyes, S. Schiavo, and G. Fagiolo, *Adv. Complex Syst.* **11**, 685 (2008).
- [13] R. Kali and J. Reyes, *Econ. Inquiry* **48**, 1072 (2010).
- [14] D. Garlaschelli et al., *Eur. Phys. J. B* **57**, 159 (2007).
- [15] D. Wen-Qi and S. Bo-Liang, *Chin. Phys. Lett.* **26**, 098902 (2009).
- [16] K. Gleditsch, *J. Conflict Resolution* **46**, 712 (2002).
- [17] J. Park and M.E.J. Newman, *Phys. Rev. E* **70**, 066117 (2004).
- [18] D. Garlaschelli and M.I. Loffredo, *Phys. Rev. Lett.* **102**, 038701 (2009).
- [19] <http://privatewww.essex.ac.uk/~ksg/exptradegdp.html>.
- [20] 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.
- [21] G. Caldarelli, A. Capocci, P. De Los Rios, and M.A. Muñoz, *Phys. Rev. Lett.* **89**, 258702 (2002).
- [22] A. Fronczak, P. Fronczak, and J.A. Holyst, *Phys. Rev. E* **73**, 016108 (2006).
- [23] T. Lux and F. Westerhoff, *Nature Physics* **5**, 2 (2009).
- [24] D. Ruelle, *Phys. Today* **57**, 48 (2004).