Complexity in Neural and Financial Systems: From Time-Series to Networks

Lead Guest Editor: Tiziano Squartini
Guest Editors: Andrea Gabrielli, Diego Garlaschelli, Tommaso Gili, Angelo Bifone, and Fabio Caccioli

Complexity in Neural and Financial Systems: From Time-Series to Networks

Complexity in Neural and Financial Systems: From Time-Series to Networks

Lead Guest Editor: Tiziano Squartini
Guest Editors: Andrea Gabrielli, Diego Garlaschelli, Tommaso Gili, Angelo Bifone, and Fabio Caccioli

Editorial Board

Contents

Editorial

Complexity in Neural and Financial Systems: From Time-Series to Networks

Tiziano Squartini [iD],[1] **Andrea Gabrielli,**[2] **Diego Garlaschelli,**[1,3] **Tommaso Gili** [iD],[1] **Angelo Bifone,**[4] **and Fabio Caccioli**[5]

[1]*IMT School for Advanced Studies Lucca, Lucca, Italy*
[2]*Sapienza University of Rome, Rome, Italy*
[3]*Instituut-Lorentz for Theoretical Physics, Leiden Institute of Physics, Leiden, Netherlands*
[4]*Center for Neuroscience and Cognitive Systems, Istituto Italiano di Tecnologia, Rovereto, Italy*
[5]*University College London, London, UK*

Correspondence should be addressed to Tiziano Squartini; tiziano.squartini@imtlucca.it

When can a system be unambiguously defined as "complex"? Although many real-world systems are believed to bear the signature of complexity, the question above remains unanswered. Our special issue aims at contributing to this ongoing discussion by collecting a number of studies tackling two aspects of complexity that have recently gained increasing attention: the *temporal* one and the *structural* one.

The seven papers composing this special issue offer an articulated overview of these topics by proposing novel techniques for the analysis of systems described by multiple time-series (such as functional brain data, stock prices, and market indices) and networked interaction patterns. The choice of focusing on neural and financial systems is dictated by the importance that topics like the identification of precursors of stock market movements, the application of causality-testing techniques to brain data, and the definition of null models for the analysis of correlation matrices (only to mention a few) have gained in recent years.

In what follows, a brief overview of the contributions is provided.

H. Liao et al. contribute to the stream of research on "economic complexity," focusing on the International Trade Network and addressing the problem of forecasting the economic evolution of a country using predictors that go beyond standard economic quantities (such as GDP). In order to identify the best method, the authors compare three different metrics across a dataset ranging from 1962 to 2000. As a result, the "Fitness and Complexity" approach seems to outperform the competing techniques.

B. Podobnik et al. address another timely prediction problem, that is, the rise of EU right-wing populism in response to unbalanced immigration. In particular, the authors analyze the relationship between the percentage of right-wing (RW) populist voters in a given country, the prevalence of immigrants in the population of the same country, and the total immigration inflow into the entire EU over the last three years. They find that the increase in the percentage of RW voters substantially overcomes the percentage of immigration inflow. This result questions the role of EU institutions and calls for a deeper understanding of the EU citizens perception of the "globalization" topic.

A different, yet related, kind of prediction is the "link-prediction" problem, addressed by M.-Y. Zhou et al. The underlying hypothesis of any link prediction algorithm is that the likelihood of any two nodes to establish a connection depends on the number of "common" characteristics. The authors, however, propose a way of resorting link scores which enhances the precision of existing methods by penalizing to a lesser extent nodes characterized by lower similarity scores.

Prediction problems are related to causality problems. Upon defining "causality" in an information-theoretic sense, T. Aste et al. investigate the performance of three methods (G-Lasso, Ridge, and LoGo) to detect causality links when short time-series are considered. A combination of filtering techniques and graphical modelling provides the best performance.

The same kind of filtering techniques is applied to investigate different types of dependency in financial multiplex networks. N. Musmeci et al. consider a four-layer multiplex defined by linear, nonlinear, tail, and partial correlations among a set of financial time-series. The structural evolution of this peculiar kind of network reveals changes associated with periods of financial stress, whose detection is enhanced by the multilayer character of the considered system.

Network techniques to analyze time-evolving, multiagent systems are also employed by A. Lombardi et al. to study the functional connectivity of the human brain. A novel framework is employed to quantify the synchronization of pairs of signals by exploiting the so-called cross-recurrence plots: community detection based on such a metric seems to outperform the usual one, based on the Pearson correlation coefficient.

Networks constitute also the support to explore "evolutionary" game theory. G. Cimini studies two different evolutionary dynamics in order to refine equilibria multiplicity of games of strategic substitutes and complements. When the latter are embedded in complex topologies, different behaviors are found for different classes of games. In particular, when coordination games on infinitely large scale-free networks are considered, equilibria arise for any value of the incentive to cooperate.

Tiziano Squartini
Andrea Gabrielli
Diego Garlaschelli
Tommaso Gili
Angelo Bifone
Fabio Caccioli

Research Article

A Comparative Analysis of the Predictive Abilities of Economic Complexity Metrics Using International Trade Network

Hao Liao [iD][1] **and Alexandre Vidmer** [iD][1,2]

[1]*National Engineering Laboratory for Big Data System Computing Technology, Guangdong Province Key Laboratory of Popular High Performance Computers, College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China*
[2]*Department of Physics, University of Fribourg, 1700 Fribourg, Switzerland*

Correspondence should be addressed to Hao Liao; jamesliao520@gmail.com and Alexandre Vidmer; alexandre@vidmer.com

The complex networks approach has proven to be an effective tool to understand and predict the evolution of a wide range of complex systems. In this work, we consider the network representing the exchange of goods between countries: the international trade network. According to the type of goods they export, the complex networks approach allows inferring which countries will have a bigger growth compared to others. The aim of this work is to study three different methods characterizing the complex networks and study their behaviour on two main topics. Can the method predict the economic evolution of a country? What happens to those methods when we merge the economies?

1. Introduction

The development of countries is a crucial question in economics. In the next ten years, which countries will grow and develop the biggest industries? The GDP per capita is one of the main criteria to assess a country's development. Even though it does not account for the wealth distribution or for the purchasing power of the countries, it is a good indicator of the country's wealth and especially of its industrial sector. Twice a year, in April and in October, the *International Monetary Fund* (IMF) makes projection for the future GDP growth rate of countries. They do not detail the precise procedure; however they indicate that they use many factors to do the predictions. Actually, the prediction is done in a different way for each country and is then aggregated and readjusted in order to harmonize the projections. The predictions performed using complex networks do not outperform the projections made by the IMF, but their goal is to make additional information available that an organization such as the IMF could benefit from by taking into account their factors.

Two methods, the Method of Reflections [1] and Fitness and Complexity [2], were designed to rank the countries

according to their potential growth. The two methods use a complex network approach to perform the ranking. The networks are built according to the revealed comparative advantage (RCA) procedure [3] described in the datasets section. The Method of Reflections proposed the Economic Complexity Index (ECI) to account for the production characteristics of countries, and the authors claimed it to be a better predictor than existing governance, institution, education quality, and economic competitiveness indexes [4]. However, there is also some criticism towards the Method of Reflections; [2, 5, 6] point out that the Method of Reflections always underestimates the important counties with high diversification in exports, such as China and India. Moreover, the ECI definition based on the second eigenvector has been firstly shown in [7]. The Fitness and Complexity method was designed to take this feature into account and was featured in Nature news (http://www.nature.com/news/physicists-make-weather-forecasts-for-economies-1.16963).

In this paper, we study three metrics, Fitness and Complexity, the Method of Reflections, and the simple sum of different exports for each country. We review their potential to predict the economic growth of countries in detail and show

that the Fitness and Complexity is the best method to rank the countries and their exports according to their importance. We push the study further to show one additional aspect of the Fitness and Complexity approach by merging countries' economies. We show that this approach captures additional features of the data that cannot be performed by a simpler approach.

2. Materials and Methods

2.1. International Trade Network. Two different datasets are used in this work. The first one ranges from 1998 to 2014 and has been cleaned by the BACI team [8]. The data were taken from [9]. The second dataset ranged from 1962 to 2000 and is used for the study of countries' merging and separation, as there are several occurrences during this period. This dataset was cleaned and processed by [10]. Both datasets are represented as bipartite in which one type of node represents the countries and the other type represents the products.

1962 to 2000. This dataset is very complete. The documentation lists all the operations that were made on the data, as well as the problems we might encounter. The gathering of the data would seem to be quite easy, as exports are likely to be registered somewhere. However this is not always the case, for three main reasons. First, even if we use the word *product* in order to simplify the discussion, only the *category of products* is specified in the data. Some examples of categories are "domestic, nonelectric, heating, cooking apparatus, and parts," "television receivers and color," and "office and stationary supplies, of base metal." This is a very rough categorization, which can contain products of various qualities. Some cooking apparatus will break after two uses while a fine quality one will last for decades. This is also a problem when the categories change. In 1984, the classification of products was revised, and so new categories appeared in the dataset. In order to avoid troubles linked to new categorization, we only use data from 1984 to 2000.

Secondly, a country (or someone working at the customs) can misreport a product assigning it to the wrong category or just misreport the quantity of export. The choice was made in [10] to give more credibility to the importer reports over the exporter reports, as they believe it would be more accurate. This makes sense if we consider that countries have a tax on imports and so have an interest in an accurate reporting. When the importer report is not available however, the exporter report is then considered. The misreports for categories were treated by creating additional categories called "aggregate categories." In our work, we exclude these aggregates.

The last reason is the absence of many countries' reports. In the paper, 72 countries reported to the UN database. However, we can still find information about the other countries by considering their trade with the 72 base countries. Some information is of course missing, as we know nothing about the trade between two countries that are not part of the 72 that reported. We can choose to include only 72 countries, or take all of them into account. The main trouble that arises if we exclude countries is that some low complexity objects (e.g.,

soy) are considered as high complexity ones because they are only seldom exported by the 72 countries, while being low complexity products (being easy to produce compared to a last-generation smartphone). Nonetheless, previous studies have used data from both countries inside and outside the core and obtained coherent results [11, 12].

The products can be categorized with different number of digits, but we keep only the categories at 4 digits in order to avoid too many unique items. We discard every product that is not in the official classification, as they were aggregately added by the authors to make the numbers more consistent. These products are easy to identify as they end with letters A and X or a 0 in the dataset.

1998 to 2014. This dataset was cleaned using *harmonization* techniques in [8]. Contrary to the previous dataset, the authors try to reconcile the exporter reports with the importer reports. When a trade is reported, it is usually done adding the transport cost on the exporter side, while the importer does not include this cost. The authors then evaluate the transport cost using a gravity equation depending on the distance between the two countries. We use this dataset to study the recent and the future evolution of countries, as the data reported in this dataset is more complete. Since the year 2000, approximately 150 countries have adopted the classification used in these data as legal instrument to report their trade.

RCA. In the international trade network, the information consists of countries' exports to other countries. We take a simplified approach here and only consider the total volume of export in US$ of individual products that a country exports, and by extension we refer to it as its production.

We represent the data as a bipartite network for which one type of nodes is countries while the other one is products. Obviously, if a country exports only a tiny amount of a product, it should not be considered as an exporter of this product. The export of a product should be a reasonable fraction of the economy of the country, and the country should also export a reasonable fraction of the global export of the product. We define a clear boundary to choose whether a country is an exporter of a product or not by considering the "revealed comparative advantage" (RCA) [3]:

$$\text{RCA}_{i\alpha} = \frac{e_{i\alpha} / \sum_j e_{j\alpha}}{\sum_\beta e_{i\beta} / \sum_{j\beta} e_{i\beta}}, \tag{1}$$

where $e_{i\alpha}$ is the export volume of country i for product α in US$. We consider that country i is an exporter of good α only if $\text{RCA}_{i\alpha} \geq 1$. This ratio determines the relative importance of an export for a country and compares it with its relative importance for other countries.

Let us take an intuitive example to illustrate the sense of this metric, based on the values of year 1998 found in the dataset created by [10]. We compare the relative importance of chocolate exports for France and Switzerland in Table 1. Though France exports twice more chocolate than Switzerland, it is still considered to have less advantage in the export of chocolate. This is due to the fact that France exports much more other goods than Switzerland, so it is expected

TABLE 1: Illustration using RCA of chocolate for France and Switzerland. The data are taken from 1998 and the values are given in billions of US$ [10].

	Total of all exports	Total for chocolate	RCA
France	292	0.68	1.71
Switzerland	83	0.29	2.60
World	5209	7.1	-

that its exports are on average higher than those of Switzerland.

After applying the RCA threshold to the international trade, we can build the international trade network. Instead of the numerous pieces of information available before the procedure, we are left with binaries information: *does this country export this product?* Based solely on this simplified data, we show how we are able to extract relevant information in the next sections.

2.2. Methods. One of the first approaches to devise a score to rank countries and products in the international trade network was done in [1]. The algorithm is iterative and is given at step n by

$$
\begin{aligned}
k_i^{(n)} &= \frac{1}{k_i} \sum_{\alpha \in \mathcal{N}(i)} k_\alpha^{(n-1)} \\
k_\alpha^{(n)} &= \frac{1}{k_\alpha} \sum_{i \in \mathcal{N}(\alpha)} k_i^{(n-1)},
\end{aligned}
\tag{2}
$$

where $k_i^{(n)}$ is the score of country i and $k_\alpha^{(n)}$ the score of country α. Both scores are initialized with degree ($k_i^{(0)} = k_i$ and $k_\alpha^{(0)} = k_\alpha$). In the original method, a threshold value is set, and when the total change on the scores is smaller than this value, the computation stops. The choice of the threshold is important, as the scores all converge to a trivial fixed point. The threshold has to be set big enough so that the differences between scores do not exceed machine precision; this point is discussed in [6]. Note that a definition based on eigenvectors is given in [4] and provides similar results but gets rid of the need to define a threshold. However, it has been noted in [13] that stopping the computation after two iterations gives the most coherent results according to our current knowledge of the network, as the correlation between the diversification of a country and its score reduces after additional iterations [6]. We then make the choice of stopping the algorithm after only two iterations. We consider here only the deviations to the average values obtained with the algorithm and label the final scores MR_i^c and MR_α^p for country i and product α, respectively:

$$
\begin{aligned}
\mathrm{MR}_i^c &= \frac{\widetilde{\mathrm{MR}}_i^c - \langle \widetilde{\mathrm{MR}}^c \rangle}{\mathrm{std}\left(\widetilde{\mathrm{MR}}^c\right)} \\
\mathrm{MR}_\alpha^p &= \frac{\widetilde{\mathrm{MR}}_\alpha^p - \langle \widetilde{\mathrm{MR}}^p \rangle}{\mathrm{std}\left(\widetilde{\mathrm{MR}}^p\right)}.
\end{aligned}
\tag{3}
$$

$\widetilde{\mathrm{MR}}_i^c$ is the value of $k_i^{(2)}$ and similarly $\widetilde{\mathrm{MR}}^p$ corresponds to $k_\alpha^{(2)}$. $\langle \widetilde{\mathrm{MR}}^c \rangle$ stands for average over all scores for countries, and $\mathrm{std}(\widetilde{\mathrm{MR}}^c)$ its standard deviation. The score attributed to countries was shown to have a good predictive power with the economic growth compared to more traditional metrics [4].

The Fitness and Complexity metrics have been developed to measure the well-being of countries and their production [2]. The procedure has been shown to perform well in an ecological network, by ranking the species according to their importance in the sustainability of the system [14]. The algorithm has been generalized by exponentiating the product term [15]. The iterative self-consistent set of equations is defined as

$$
\begin{aligned}
F_i^n &= \sum_{\alpha \in \mathcal{C}_i} Q_\alpha^{n-1} \\
Q_\alpha^n &= \frac{1}{\sum_{i \in \mathcal{C}_\alpha} 1/F_i^{n-1}},
\end{aligned}
\tag{4}
$$

where F_i^n is the fitness of country i and Q_α^n the complexity of product α after n iterations. The convergence of the algorithm and its stopping condition were studied in [15].

The idea behind the algorithm is that successful countries export complex products and complex products are only exported by successful countries. Indeed, the fitness of a country in (4) is simply the sum of the complexities of its exported goods. So a successful country should export many products, and these products should be of high complexity in order to achieve a high fitness. The complexity of a product in (4) is defined in a nonlinear way, which makes effectively the algorithm nonlinear. Due to the $1/F_i^{n-1}$ factor in the equation, the complexity of a product is mainly characterized by the fitness of its lowest exporter. For instance, if we take a product α with two exporters i and j with fitness values of 0.1 and 10, respectively, the complexity of the product would be 0.099. If only country i exports the product, the complexity would be 0.1. The two values are close and we clearly see the dependence on the worst exporter. This makes sense for the interpretation of the score as if poor scoring country can export a good; this good should be of low complexity. And if only high scoring countries export a good, it should be hard to produce; otherwise some low fitness country would export it.

3. Results

In addition to the two metrics described previously, we add the degree for comparison (i.e., number of exports having a RCA equal to or above 1). The comparison with degree is lacking in the literature about Fitness and Complexity, and so we are interested in studying the additional information brought by the two iterative approaches. Indeed, previous works have shown a very high correlation between the degree of countries' exports and the fitness of these countries [13].

3.1. GDP Prediction. As mentioned in Introduction, the GDP growth is one of the main pieces of information about the economic performance of a country. In order to study the link between the metrics and the GDP evolution, we show in

Figure 1 the trajectories of the countries in the GDP per capita plane coupled with the three different metrics [2] (degree, Methods of Reflections, and Fitness and Complexity). The data range from 1998 to 2014 [8]. We remark that the countries that lie in the region with high degree but somehow lower GDP per capita (such as China and India) tend to have a trajectory with growing GDP. The remark is also valid for fitness, but for Methods of Reflections it is less visible. The ones in the region with high GDP per capita compared to their degree (such as Norway and Iceland) tend to have a more static trajectory along the GDP per capita axis. This remark was made in [16] for fitness metric. It is hard to distinguish the different features between degree and Fitness on this graph; the same three countries (Czech Republic, China, and India) stand out. The exception is Philippines, which only stands out when looking at the Fitness metric. Its trajectory is quite flat for the moment as it just arrived in the supposed GDP growth region. It is hard to distinguish anything in the middle region of the graph, and so a coarse grained vision is needed in order to be able to see the tendencies.

Similarly to [16], we average the trajectories from 1998 to 2014 in Figure 3 for the three different metrics coupled with GDP per capita. In the degree plot, we see that every country which has a degree superior to about 100 has a significant GDP growth. The longest arrows can be observed in the far right region under the regression line in red, which means that the countries in this region experience the biggest growth. This is even more true in the panel with Fitness in Figure 3(c). The countries in the lower right part of the plot exhibit a growth significantly higher than in the other regions of the plane. In Methods of Reflections however, every arrow seems to point up at the same rate and it is hard to distinguish particular regions. By eye, the Fitness and Complexity metric seems to have more discriminating power than the degree, definitely more than Methods of Reflections. However, it is hard to draw any hard conclusion on the superiority of Fitness compared to degree based on these panels. But we will see in the next section in which domain the Fitness and Complexity metrics outperform the simple degree metric.

3.2. The Selective Predictability Scheme.

In [16], a method based on the *method of analogues* [17, 18] was applied to the prediction of the countries' GDP. This method is named *the selective predictability scheme*. This method is useful in order to have a clear visualization of the tendencies. The average displacement in the plane presented in Figure 2 gives only a general tendency. But we can also ask the question: can we predict where the country will land in the plane in the future? If we can say with high probability where a country will be in the future according to its current position in the plane, it is also an interesting observation, while this may disappear in the averaging procedure used in the arrows picture. The evaluation goes as follows:

(i) The GDP-Fitness space in logarithmic plane is divided into boxes.

(ii) We report the box in which countries are at a given year and the one at which they end ten years later.

(iii) The total number of countries inside box i at year t is N_t^i, and the number of different boxes occupied after ten years' evolution coming from box i is labeled n_t^i.

(iv) For each box, we evaluate the quantity $\mathscr{C}_t^i = (n_t^i/N_t^i - 1/N_t^i)/(1 - 1/N_t^i)$. This quantity is related to the dispersion of the box. Indeed, if the evolution of countries is highly unpredictable, n_t^i is large, and so $\mathscr{C}_t^i$ is large. Conversely, if countries tend to have a predictable evolution, they would end in the same boxes, and so n_t^i would be small, resulting in a smaller value for $\mathscr{C}_t^i$.

The results for the three metrics are displayed in Figure 3. At first glance, it seems that degree and Fitness perform similarly, while Method of Reflections does not show exhibition and predictability. There are more boxes indicating low predictability (red) in the case of degree than for fitness. However there are also more boxes with high predictability (dark green). There are two shortcomings with the method of analogues. The first one is that the results are only graphical; it is then hard to say that a method is more adequate than another. The second one is that the results are dependent on the number of boxes we choose, as well as the range of the axis.

The method of analogues assigns a predictability score for the countries located inside specific boxes. Then, by summing over the predictability scores, the result should account for an overall predictability of the method. However, we need to get rid of the strong dependency on the choice of the number of boxes and their location. With this in mind, we propose to sum over all the boxes, taking into account the number of occupied boxes N_{occ} as a parameter. The score is as follows:

$$\mathscr{P}^S = \mathscr{P} \cdot \frac{2 \log (N_{\text{occ}})}{n}, \tag{5}$$

where $\mathscr{P} = \sum_{i,t} 1 - \mathscr{C}_t^i$ and n is the number of events used during the computation, where an event is defined as the presence of a country in a box. The fact that n is important is because we only take into account the events that are located in a box with at least five events. We choose to take the inverse of the dispersion in the sum, so that a high value of overall predictability $\mathscr{P}$ means that the prediction is accurate. The results are shown in Figure 4 as a function of the number of boxes occupied N_{occ}. The shadowed areas represent the standard deviation over 256 different choices of axis limits. The choice of the axis limits matters as the boxes are shifted depending on which values are comprised in the boxes. For instance, in case (a) the first interval is [0.1,0.2] and in case (b) [0.12,0.22]. A dot located at 0.21 would be inside the first interval in case (b) but not in case (a). It is quite obvious that Methods of Reflections performs bad, especially that the results are highly dependent on the choice of axis' limits. While it is clear that the Method of Reflections has a poor predictability, it is clear that degree is performing better than Fitness and Complexity. We argue that this shows the predictive power of Fitness and Complexity is different from the one of diversification. Indeed, the score was developed to take into account the evolution of outliers. While diversification would grasp a better general trend, Fitness and Complexity method
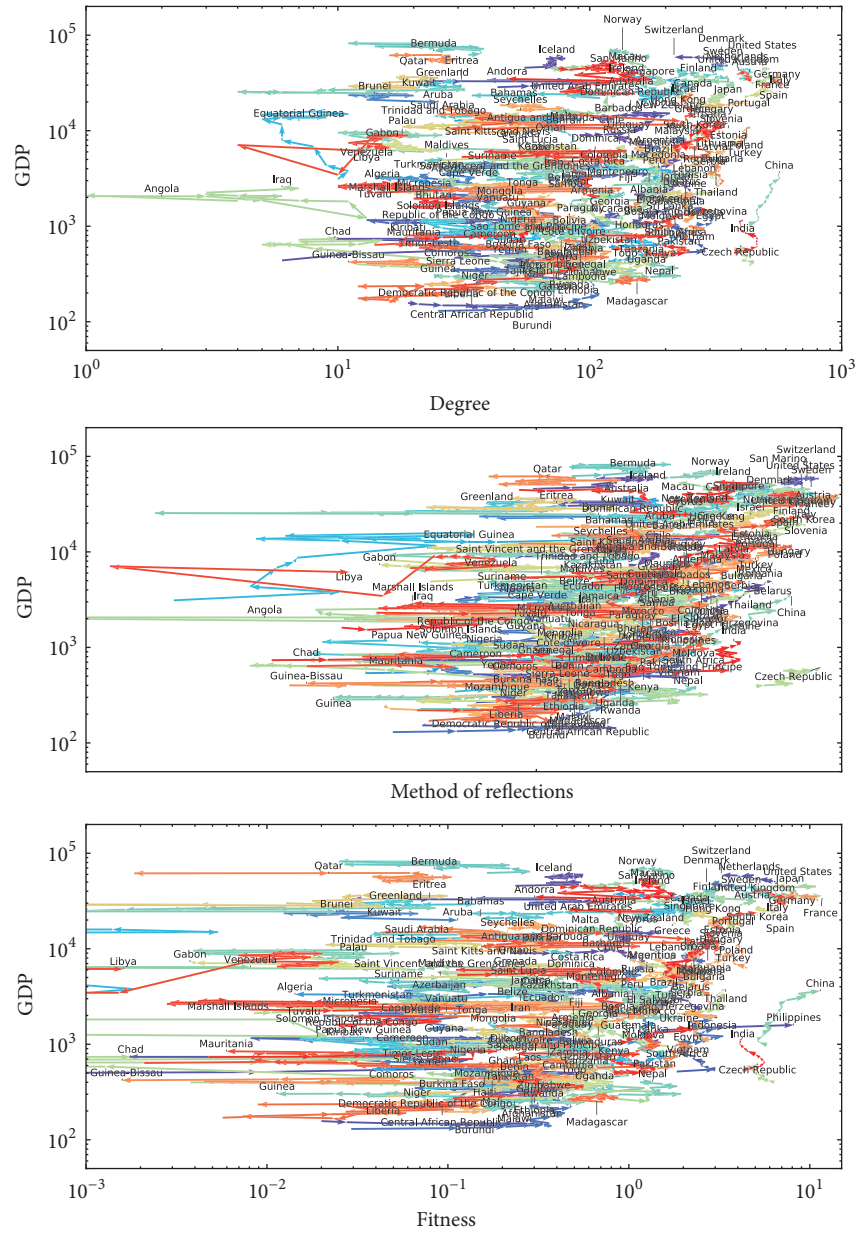
FIGURE 1: Evolution of the countries in the GDP per capita-score plane, where score can be degree, country's score obtained by Methods of Reflections or Fitness. The data we consider range from 1998 to 2014 and countries are assigned random colors for easier spotting of their respective trajectory.

has already proven to be a good predictor for countries in specific location [19].

3.3. Effects of Countries' Unification on the Fitness of Countries. What happens if countries decide to merge their exports? Will their fitness increase or will it decrease? This could go both ways, as the countries would lose some exports due to the RCA threshold: if the total export increases and the export for a specific product remains nearly constant, the RCA value can drop under one. Or in a different case if two countries export a lot to each other, they will lose a lot of exports when unifying. We show that some countries would benefit

from a unification while some would worsen their fitness score. Fortunately, the unification and separation of countries happen several times in the dataset. We take the example of Germany, Czechoslovakia, and Yugoslavia to illustrate how the unification or separation of economies affects the fitness of countries. In a second time, we take artificial example to show that it is hard to predict which country would benefit from a unification and which one would not. This depends mainly on the details of their exports and not solely on their current fitness value.

In the original exports data, which country exports which products to another country is specified. While we chose to

(a)

(b)

(c)
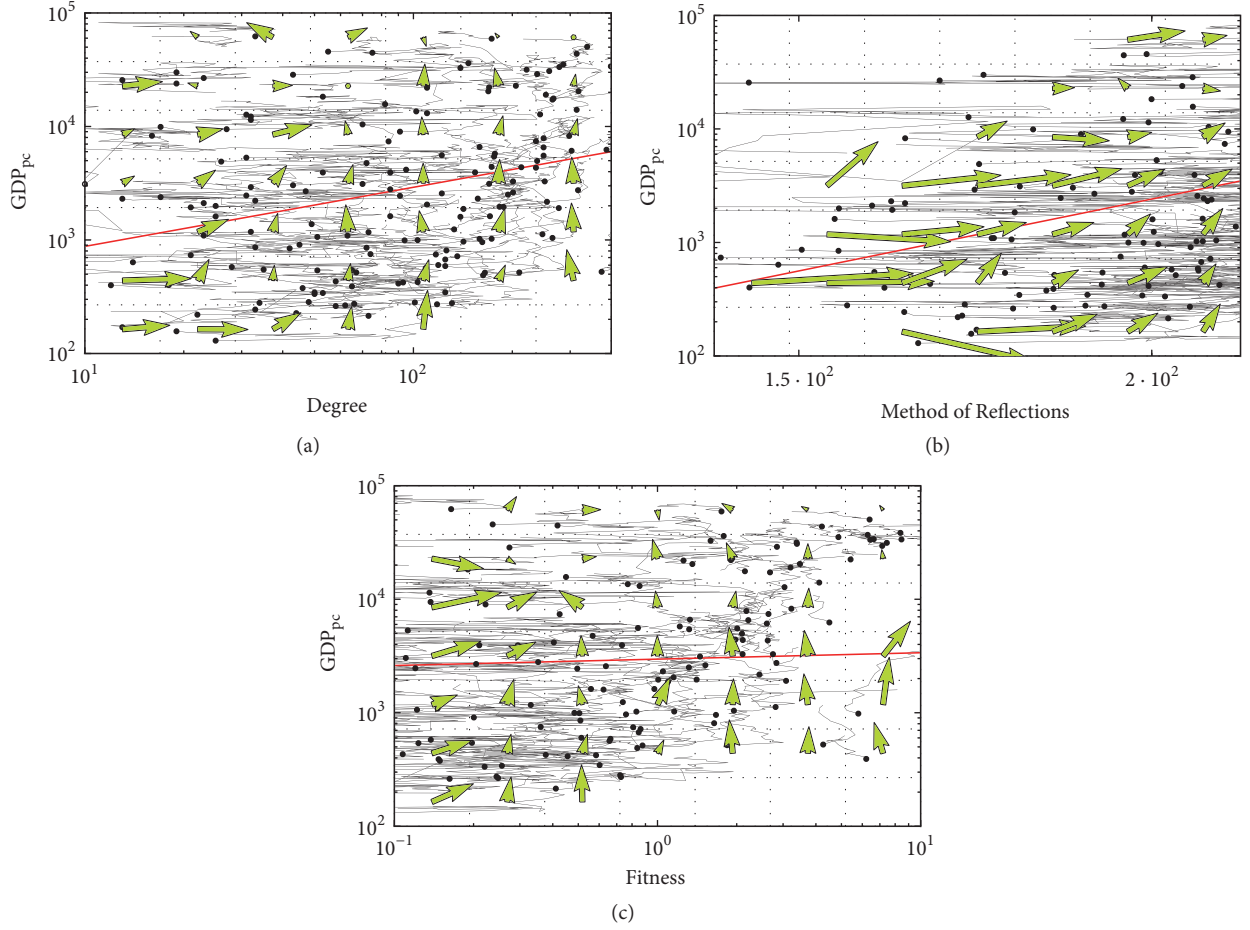
FIGURE 2: Average displacement in the GDP per capita and country score plane. The score of countries is degree in panel (a), Methods of Reflections in panel (b), and Fitness in panel (c). For a better view of the GDP change, the displacement along the GDP-axis is multiplied by 5 for a better visualization. The linear regression is shown by a red line for each metric.

simplify the data and take into account only the total exports for each product in the previous sections, the detailed data now prove useful when unifying the countries, as we are able to subtract the exports between them. We would like to stress before starting the analysis that we can only merge the exports of countries and not their ability to produce goods. Indeed, when two economies unify together, it is possible that their common capabilities on manufacturing can be combined to produce new goods or to increase significantly their export for a product [1, 6].

*3.3.1. Real-World Examples.* It happened several times during the world history that some countries decided to unify or that a country choose to separate into smaller ones. We investigate the case of separation and unification from a purely economic point of view, without considering any other factors. The question of interest here is to know whether the fitness of the countries increases or decreases in the process. If we theoretically unify two countries and the fitness of these two countries increases, the countries would have an economic benefit to unify. On the other hand, if the fitness of the countries decreases when theoretically unified, the

country with the highest fitness would not benefit from the unification.

We study three real-world cases. The first is the reunification of East and West Germany in 1990. Nowadays, Germany is one of the top economies in the world; while we cannot study the imaginary economy that would result if current Germany would split again, we can merge the old East and West Germany. It is well-known that the West was more economically successful than the East in 1990. However, if the two economies are complementary, it is possible that the West could benefit from a merging as well. This is highly probable as the two parts of the countries were totally disconnected and the political systems were completely different. We show in Figure 5(a) the theoretical merge of the two Germanies. Not surprisingly, the fitness of West Germany is higher than the one of East Germany before 1990. We can see that the fitness of Germany starts to rise before 1990; this is due to the fact that data and some exports were referenced before the official classification; however the total export volume is really low. The curve named merge is the theoretical merge of the two Germanies. We observe that it is constantly higher than both the fitness of the West and East Germany. We could
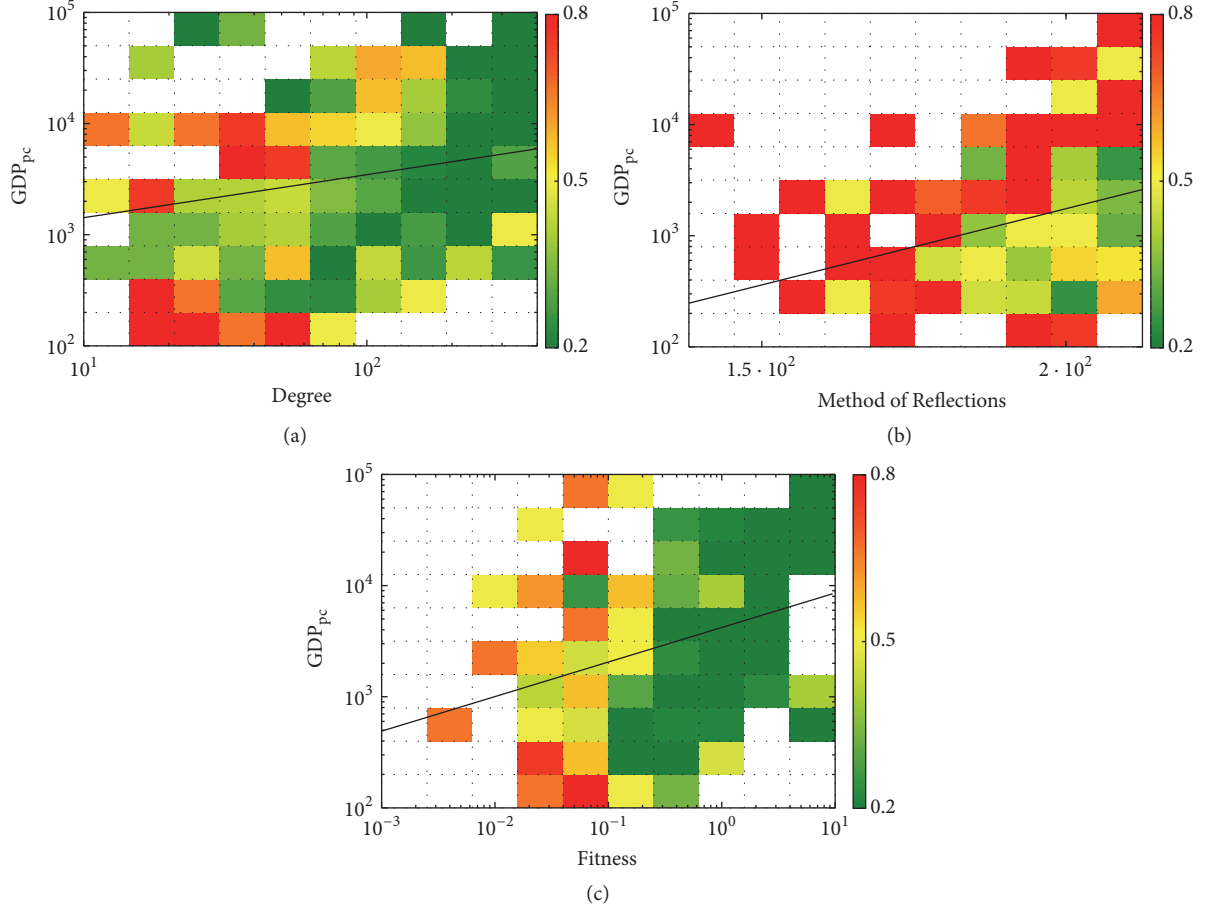
(a)

(b)

(c)

FIGURE 3: The selective predictability scheme with $N = 10$ for the three different metrics. A low value means a low dispersion and so a high predictability, while a high value means unpredictability.

say that if the countries were to merge, they would benefit from it. And eventually they merged and became one of the strongest economies in the world, even the first according to the Fitness and Complexity metric. We do not show any confidence interval; however the robustness to noise is high and, most importantly, the tendency is consistent with years.

Our second example is the dissolution of Czechoslovakia into Czech Republic and Slovakia in 1992. The cause of the dissolution was mainly political, but there were also economic tensions [20]. The Slovaks had the impression that they were *paying for the Czech*, which seems to be true as the leading parties were mostly Czech and promoted Czech interests first. These tensions and the strong nationalism of Slovaks led to the separation of Czechoslovakia. The fitness of Czech Republic and Slovakia, Czechoslovakia before the dissolution, and the theoretical merge after the dissolution of Czechoslovakia are shown in Figure 5(b). In contrary to the Germany case, we see that the Czech Republic scores higher in the fitness metric than the merge of the two states. This is an interesting observation as we could claim that Czech Republic had a strong economic benefit in this dissolution.

The last example is the dissolution of Yugoslavia into Slovenia, Macedonia, Bosnia and Herzegovina, Albania, Federal Republic of Yugoslavia (Serbia and Montenegro), and Croatia. According to [21], The Yugoslavian state, which was created after the first world war, was held together by having common enemies, especially Italy at that time. During the 20th century, there were many tensions in Europe and so the state stayed united in case of an attack. Due to the lack of pressure from the East, the ethnic tensions became stronger than the necessity of staying together against a stronger enemy. From 1991 to 1992, the state gradually separated into a multitude of smaller states. As for the two previous examples, the results are shown in Figure 5(c). Among all the states, only Slovenia benefited from the dissolution in terms of fitness. In a very short term, the Federal Republic of Yugoslavia also benefited from it, but the wars that created the dissolution certainly damaged strongly its economy.

In our three real-world examples, when two countries would benefit from merging, they end up by merging together. And when one country would benefit from a dissolution, they end up separating. We note that these are only anecdotal evidences, and there are factors affecting the dissolution or merging other than the economic factor. However the economy of a country in general reflects well its stability, its dynamism, and its well-being. As observed in the previous section about GDP, fitness does not necessarily add information when compared to degree. It is then natural to
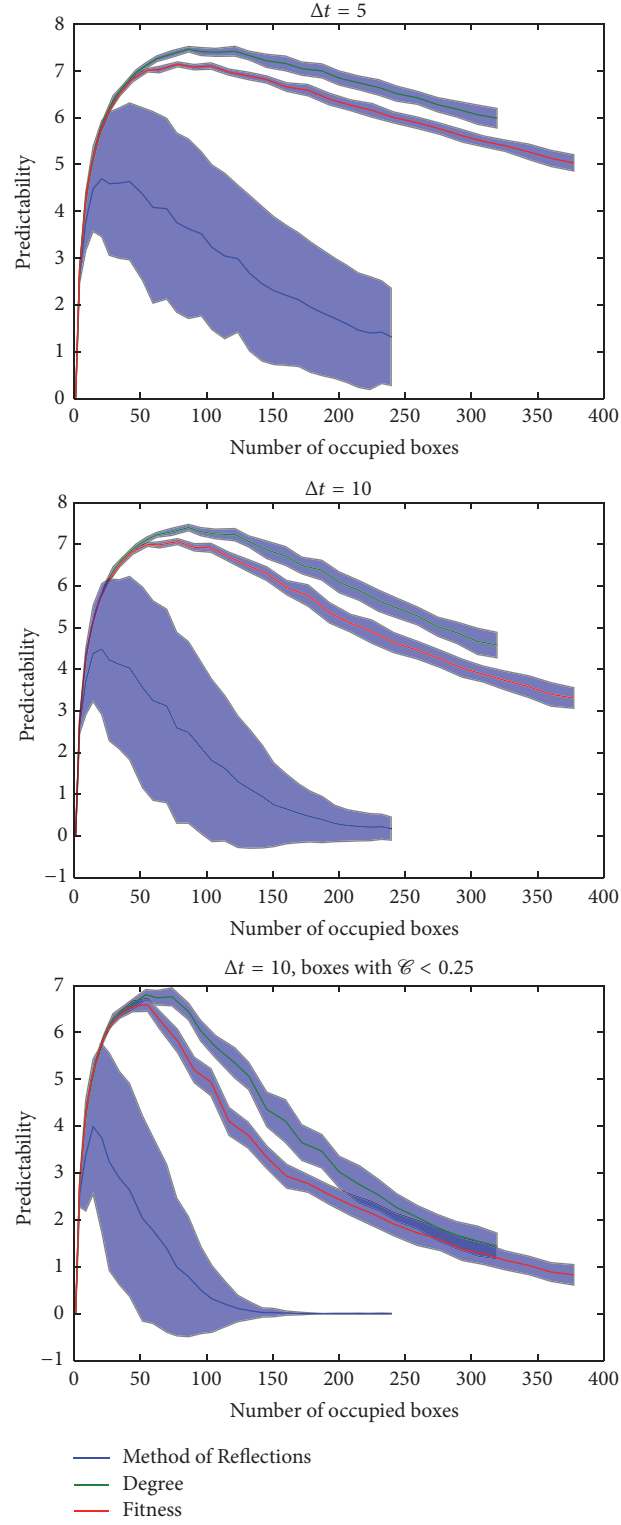
FIGURE 4: Total predictability as a function of the number of boxes occupied N_{occ} along one direction. The standard deviation is shown as shadowed area. The parameter Δt stands for the evolution step; the countries start in a box and we observe in which one they end Δt years later. A high score means a high predictability.

(a)

(b)

(c)
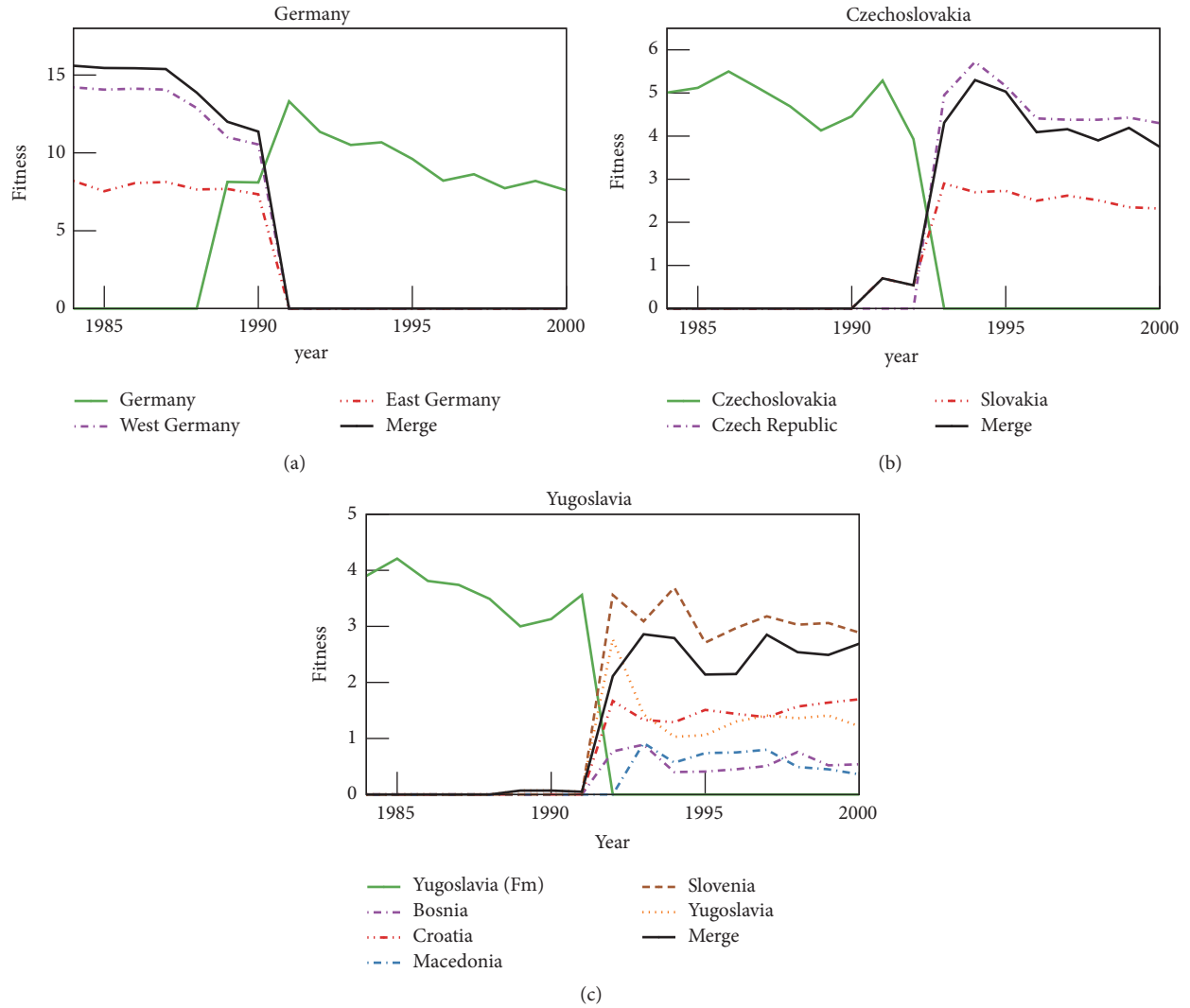
Figure 5: Comparison of countries' fitness values before they merge together or before they dissolve. The cases of Germany, Czechoslovakia, and Yugoslavia are presented in the panels.

ask the question: could we have obtained the same results by using only degree? We show the equivalent with degree in Figure 6. We see that it is in general less pronounced than with fitness and sometimes quite different. For instance, Slovenia performs worse than the merge when considering the degree metric, and West Germany performs very similarly to the merged Germanies. Another important point to note is the fact that the degree is nearly perfectly continuous and does not undergo tremendous variations. Conversely, Fitness shows some quick variations, which shows that it captures additional information that is hidden in a traditional approach.

3.3.2. Theoretical Cases. We saw in the previous section that if we merge economies together, the resulting economy can be either stronger or weaker in terms of fitness. But so far we did not investigate the mechanisms behind this result. In order to find the explanation for this phenomenon, we take fictitious

example by merging one country with every other country in the dataset. By studying the relations between the countries when the fitness goes up, or goes down, we should be able to identify the common property between positive and negative merges.

We present the results of merging for four different countries in Figure 7. Obviously, it is not trivial to see if the fitness score of a country will rise or fall when merging. We see a clear dependency on the fitness of the target country; however it is only a tendency. For instance, Italy would benefit from merging with the two countries with the highest fitness scores (Germany and USA), while Japan would see its fitness decreasing. At the same time, the base countries rarely benefit from merging with a country with lower fitness. This is also interesting as it shows that this is quite rare and our previous example with West and East Germany is a special case. It is also surprising to see in the case of Italy that, even when merging with a country having a higher fitness, the fitness

(a)

(b)

(c)
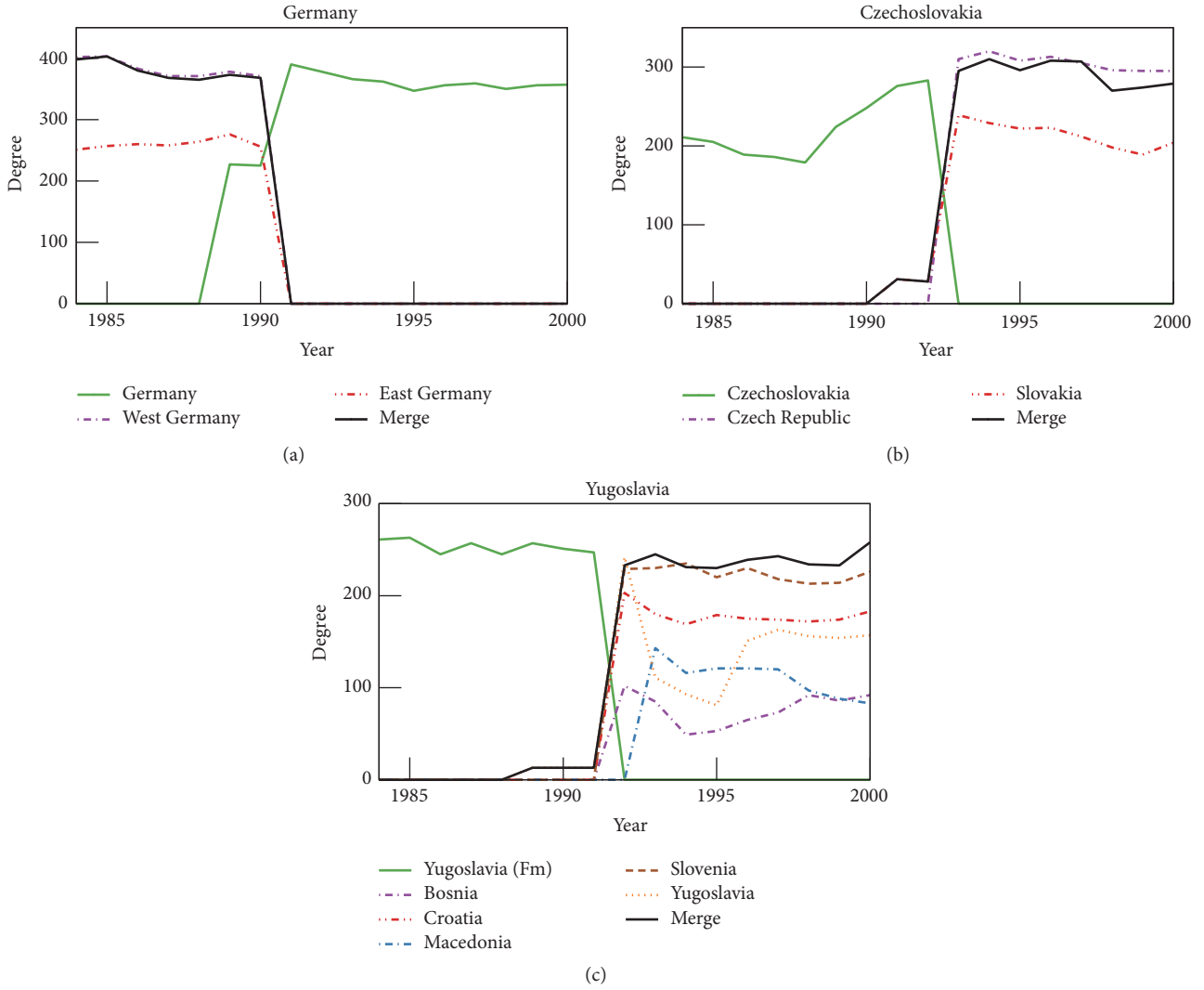
Figure 6: Comparison of countries' degrees before they merge together or before they dissolve. The cases of Germany, Czechoslovakia, and Yugoslavia are presented in the panels.

of Italy decreases significantly. This country that would make the fitness of Italy lower is actually France.

There is of course a dependency on the fitness of the target country, but it is not the only factor. As we see in Figure 7, the share of common exports is not an explanatory factor. However, we see that Japan would in general gain more fitness when merging with another country and that it has less common exports with other countries than the other three base countries. While we found no solid predictor to find if a country would benefit from a merge or not, it would definitely be linked with the rarity of their exports' basket. One possible way to investigate this problem would be to also take into account the complexity scores of the products they have or not in common. If they mostly share low complexity exports and have few high complexity products in common, a merge would add more high complexity products, while if they already share most high complexity products, a merge would not add many new products in the process.

4. Conclusion

Fitness and Complexity was shown to be the top performing algorithm to rank the countries and products according to their importance and allowed the exhibition of additional features compared to other approaches. While it is hard to conclude on its ability to predict economic growth of countries, it is undeniable that it captures additional information. In this work, we showed a way of comparing methods and a potential use of the methods' additional information. As shown in a previous study [22], the Fitness analysis can help the weak economies to understand how to exit the poverty trap, by increasing complexity of their exported products and increasing their diversification. Surprisingly, we saw that it is a good predictor for the economic results in case of the merging and separation of countries. While the analysis of the results remains a difficult task, a deeper understanding of the fitness change when two countries merge together or when a country
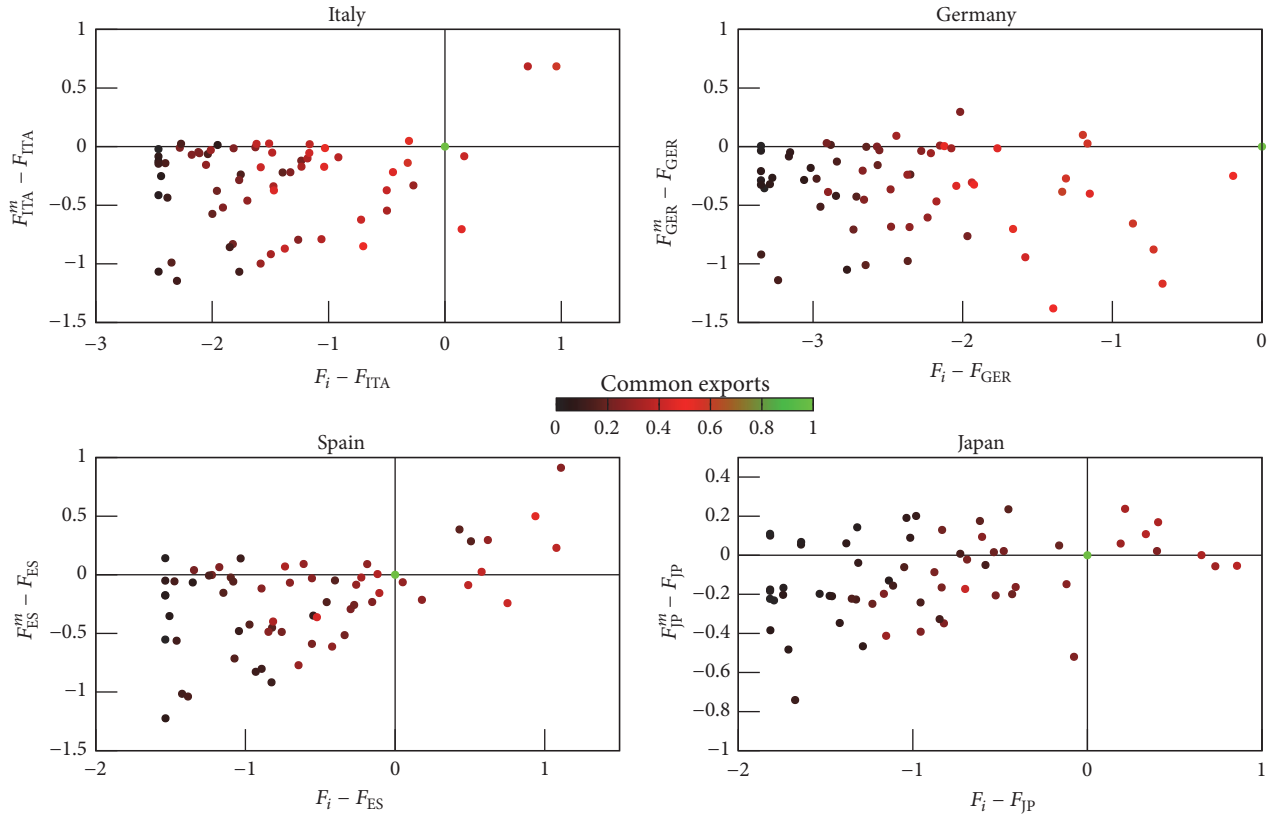
FIGURE 7: Merging of Italy, Germany, Spain, and Japan with all other countries in the dataset. The x-axis is the difference of fitness with the initial country and the target country it is merged with, while the y-axis is the difference of fitness when the initial country is merged with the target country. The color stands for the fraction of common exports between the base and target countries.

separates itself into smaller countries would tell us a lot about the metrics and the economic behaviour of countries.

Conflicts of Interest

There are no conflicts of interest.

Acknowledgments

References

[1] C. A. Hidalgo and R. Hausmann, "The building blocks of economic complexity," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 106, no. 26, pp. 10570–10575, 2009.

[2] A. Tacchella, M. Cristelli, G. Caldarelli, A. Gabrielli, and L. Pietronero, "A new metrics for countries' fitness and products' complexity," *Scientific Reports*, vol. 2, article 723, 2012.

[3] B. Balassa, "Trade Liberalisation and "Revealed" Comparative Advantage," *The Manchester School*, vol. 33, no. 2, pp. 99–123, 1965.

[4] R. Hausmann, C. A. Hidalgo, S. Bustos, M. Coscia, A. Simoes, and M. A. Yildirim, *The Alas of Economic Complexity: Mapping Paths to Prosperity*, MIT Press, 2014.

[5] L. Pietronero, M. Cristelli, A. Gabrielli et al., "Economic Complexity: "Buttarla in caciara" vs a constructive approach," *Economics*.

[6] M. Cristelli, A. Gabrielli, A. Tacchella, G. Caldarelli, and L. Pietronero, "Measuring the intangibles: a metrics for the economic complexity of countries and products," *PLoS ONE*, vol. 8, no. 8, Article ID e70726, 2013.

[7] G. Caldarelli, M. Cristelli, A. Gabrielli, L. Pietronero, A. Scala, and A. Tacchella, "A network analysis of countries' export flows: firm grounds for the building blocks of the economy," *PLoS ONE*, vol. 7, no. 10, Article ID e47278, 2012.

[8] G. Gaulier and S. Zignago, "Baci: International trade database at the product-level. the 1994-2007 version," in *Proceedings of the Working Papers 2010-23*, CEPII, 2010, http://www.cepii.fr/CEPII/fr/publications/wp/abstract.asp?NoDoc=2726.

[9] A. J. G. Simoes and C. A. Hidalgo, "The economic complexity observatory: An analytical tool for understanding the dynamics of economic development," in *Scalable Integration of Analytics and Visualization*, 2011.

[10] R. Feenstra, R. Lipsey, H. Deng, A. Ma, and H. Mo, "World Trade Flows: 1962-2000," National Bureau of Economic Research, 2005.

[11] C. A. Hidalgo, B. Winger, A.-L. Barabási, and R. Hausmann, "The product space conditions the development of nations," *Science*, vol. 317, no. 5837, pp. 482–487, 2007.

[12] R. Hausmann, J. Hwang, and D. Rodrik, "What you export matters," *Journal of Economic Growth*, vol. 12, no. 1, pp. 1–25, 2007.

[13] M. S. Mariani, A. Vidmer, M. Medo, and Y.-C. Zhang, "Measuring economic complexity of countries and products: which metric to use?" *The European Physical Journal B*, vol. 88, no. 11, article 293, pp. 1–9, 2015.

[14] V. Domínguez-García and M. A. Muñoz, "Ranking species in mutualistic networks," *Scientific Reports*, vol. 5, article 8182, 2015.

[15] E. Pugliese, A. Zaccaria, and L. Pietronero, "On the convergence of the Fitness-Complexity algorithm," *The European Physical Journal Special Topics*, vol. 225, no. 10, pp. 1893–1911, 2016.

[16] M. Cristelli, A. Tacchella, and L. Pietronero, "The heterogeneous dynamics of economic complexity," *PLoS ONE*, vol. 10, no. 2, Article ID e0117174, 2015.

[17] E. N. Lorenz, "Atmospheric predictability as revealed by naturally occurring analogs," *Journal of the Atmospheric Sciences*, vol. 26, no. 4, pp. 636–646, 1969.

[18] E. N. Lorenz, "Three approaches to atmospheric predictability," *Bulletin of the American Meteorological Society*, vol. 50, no. 3454, p. 349, 1969.

[19] M. Cristelli, A. Tacchella, L. Pietronero, and C. W. Chen, "The heterogeneous dynamics of economic complexity," *PLoS ONE*, vol. 10, no. 2, Article ID e0117174, pp. 1–15, 2015.

[20] P. S. Hilde, "Slovak nationalism and the break-up of Czechoslovakia," *Europe - Asia Studies*, vol. 51, no. 4, pp. 647–665, 1999.

[21] D. Sekulic, "The creation and dissolution of the multinational state: The case of Yugoslavia," *Nations and Nationalism*, vol. 3, no. 2, pp. 165–179, 1997.

[22] E. Pugliese, G. L. Chiarotti, A. Zaccaria, and L. Pietronero, "Complex economies have a lateral escape from the poverty trap," *PLoS ONE*, vol. 12, no. 1, Article ID e0168540, 2017.

Research Article

Connecting Patterns Inspire Link Prediction in Complex Networks

Ming-Yang Zhou, Hao Liao, Wen-Man Xiong, Xiang-Yang Wu, and Zong-Wen Wei

Guangdong Province Key Laboratory of Popular High Performance Computers, College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China

Correspondence should be addressed to Hao Liao; jamesliao520@gmail.com

Link prediction uses observed data to predict future or potential relations in complex networks. An underlying hypothesis is that two nodes have a high likelihood of connecting together if they share many common characteristics. The key issue is to develop different similarity-evaluating approaches. However, in this paper, by characterizing the differences of the similarity scores of existing and nonexisting links, we find an interesting phenomenon that two nodes with some particular low similarity scores also have a high probability to connect together. Thus, we put forward a new framework that utilizes an optimal one-variable function to adjust the similarity scores of two nodes. Theoretical analysis suggests that more links of low similarity scores (long-range links) could be predicted correctly by our method without losing accuracy. Experiments in real networks reveal that our framework not only enhances the precision significantly but also predicts more long-range links than state-of-the-art methods, which deepens our understanding of the structure of complex networks.

1. Introduction

Modern science and engineering techniques increase our availability to various kinds of data including online social networks, scientific collaboration networks, and power grid networks [1–5]. Many interesting phenomena could be uncovered from these networks. For example, analyzing the data of Facebook and Twitter helps find lost friends by only counting their common friends [6, 7] and recommendation systems in online stores [8, 9]. Restricted by instrument accuracy and other obstacles, we only obtain a small fraction or a snapshot of the complete networks [10, 11], promoting us to filter the information in complex networks [12–14]. Link prediction is a straightforward approach to retrieve networks by predicting missing links and distinguishing spurious links [15–17]. Thus great efforts have been devoted to link prediction in recent years [16, 18]. Link prediction is used in different kinds of networks, including unipartite networks and bipartite networks, where unipartite networks consist of nodes with the same type (e.g., social networks and neural networks) and bipartite networks consist of nodes with two

types (e.g., user-object purchasing networks and user-movie networks) [19, 20].

In classical link prediction approaches, similarity scores are computed first for two disconnected nodes, and then nonexisting links in the top of the score list are predicted as potential ones [16]. Consequently, the key issue is to search effective score-assigning methods that are mainly divided into three categories [16, 21]: similarity based algorithms, Bayesian algorithms, and maximum likelihood algorithms. First, similarity based algorithms [22–24] suppose that similar nodes have a high probability to link together. Similarities are evaluated by common neighbors, random walk resource allocation, and some other local and global indices. Second, Bayesian algorithms [25–27] abstract the joint probability distribution from the observed networks and then utilize conditional probability to estimate the likelihood of a nonexisting link. Third, maximum likelihood algorithms [28, 29] presuppose that some underlying principles rule the structure of a network, with the detailed rules and specific parameters obtained by maximum likelihood estimation. Scores of nonexisting links are acquired through the details of

these principles. Most of these methods favor predicting links with high similarity scores and perform badly in the detection of long-range links with low similarities.

In the aforementioned methods, the basic hypothesis that two nodes with a high similarity score have a high likelihood of connecting together lacks an in-depth illustration. Recent works have demonstrated that long-range links exist extensively in complex networks and play an important role in routing, epidemical diffusion, and other dynamics [30, 31]. However, in practice, the endpoints of a long-range link usually have weak interaction and low similarity [30], which prevents the detection of long-range links by traditional methods [32, 33]. Hence, the structural patterns underlying the networks are of great importance to study.

Our study takes a different but complementary approach to link prediction problem. By analyzing the score distributions of existing and nonexisting links, respectively, we find an interesting phenomenon that the existing and nonexisting links follow different connecting patterns in respective of their similarity scores. Then, inspired by the precision-recall curves [34–36], we propose a metric, named precision-to-noise ratio (PNR), to characterize the ability to distinguish potential links for different scores. PNR describes the local precision of a given set of links with the same score. Based on PNR, a novel framework, which projects one-variable function to adjust the scores of a given method, is put forward. We argue that the framework finds the optimal transforming function that exploits the full capacities of traditional link prediction methods and improves their performance both on precision and on the detection of long-range links. Experiments in six real-world networks demonstrate the effectiveness of our method.

The rest of the paper is organized as follows. In Section 2, we first brief the link prediction problem and then introduce our proposed method. In Section 3, we compare the performances of our method and the classical methods. Finally, the conclusion is given.

2. Materials and Methods

We give the link prediction formulism in Section 2.1 and the baseline method in Section 2.2. Our proposed framework is introduced in Section 2.3.

2.1. Network Formation and Metrics. Given a network $A = (a_{ij})_{N \times N}$, $E = \{(i, j) \mid a_{ij} \neq 0\}$ with $a_{ij} = 1$ if node i connects to j; otherwise, $a_{ij} = 0$. When evaluating the prediction performance, we usually divide the links randomly into $1 - p^H$ training set E^T and p^H probe set E^P ($p^H \in (0, 1)$), with $E^T \bigcap E^P = \emptyset$ and $E^T \bigcup E^P = E$. The goal is to accurately predict the links in probe set only by using the information in training set.

We first assign a score to each nonexisting link and then choose links with the highest top-L scores as potential ones. State-of-the-art similarity evaluation methods could be utilized to carry out link prediction, including common neighbors (CN), Jaccard index (JB), resource allocation index (RA), local path index (LP), and structural perturbation method (SPM) (see the part of *Baseline* and [38]).

There are two popular metrics to characterize the accuracy: area under the receiver operating characteristic curve (AUC) [39] and the precision [40, 41]. AUC can be interpreted as the probability that a randomly chosen missing link (i.e., a link in E^P) has a higher score than a randomly chosen nonexisting link. Then, AUC requires n times of independent comparisons. We randomly choose a real link and a nonexisting link to compare their scores. After n different comparisons, we record n_1 times where real links have higher scores, and n_2 times where the two kinds of links have the same score. The final AUC is calculated as

$$\text{AUC} = \frac{n_1 + 0.5 \times n_2}{n}. \tag{1}$$

If all the scores are given by an independent and identical distribution, then AUC should be around 0.5. A higher AUC is corresponding to a more accurate prediction.

Another metric is precision that characterizes the ratio of correctly predicted links for a given prediction list. That is to say, if the length of prediction list is L, among which L_r links are the right potential links, then the precision is

$$p = \frac{L_r}{L}. \tag{2}$$

Clearly, higher precision means higher prediction accuracy. Intuitively, higher accuracy means higher AUC and higher precision. In the experiments, we will see that precision has little correlation with AUC and that improving the precision may not result in the improvement of AUC.

2.2. Baseline Prediction Methods. There exists a large number of score-assigning approaches in link prediction problem. All these methods could be introduced into our framework. Though we only investigate some state-of-the-art score-assigning approaches, the results and conclusions are also applicable for other score-assigning methods. The five score-assigning approaches [6, 16] are as follows.

(i) Common Neighbor (CN). The metric supposes that if two nodes i and j have more common neighbors, they are more likely to connect together. The neighborhood overlap of the two nodes is as follows:

$$s_{ij}^{\text{CN}} = \left| \Gamma(i) \bigcap \Gamma(j) \right|, \tag{3}$$

where $\Gamma(i)$ is the neighbor set of node i and $|\cdots|$ indicates the size of a set. The drawback of CN is that it favors large-degree nodes. Though the similarity of two large-degree nodes is low, they still have many common neighbors.

(ii) Jaccard Coefficient (JB). Jaccard is a conventional similarity metric that aims to suppress the influence of large-degree nodes, which is

$$s_{ij}^{\text{Jaccard}} = \frac{\left| \Gamma(i) \bigcap \Gamma(j) \right|}{\left| \Gamma(i) \bigcup \Gamma(j) \right|}. \tag{4}$$

Since the similarity is normalized by the size of the union set of the two nodes' neighbors, low similarity still exists between two large-degree nodes even though they may have many common neighbors.

(iii) Resource Allocation (RA). This index is inspired by the resource allocation dynamics in complex networks. Given a pair of unconnected nodes i and j, suppose that the node i needs to allocate some resource to j, using common neighbors as transmitters. Each transmitter (common neighbor) starts with a single unit of resource and then distributes it equally among all its neighbors. The similarity between i and j can be calculated as the amount of resource received from their common neighbors:

$$s_{ij}^{\text{RA}} = \sum_{z \in \Gamma(i) \bigcap \Gamma(j)} \frac{1}{k_z}. \tag{5}$$

Comparing with Jaccard method, RA could also suppress the influence of large-degree nodes, but more specifically. Different neighbors contribute to the similarity differently. If two nodes prefer to connect low-degree nodes, it means that they have a higher probability to share common interests or characteristics. However, many pair-nodes have common high-degree neighborhoods, resulting in that high-degree nodes play a weak role when evaluating similarity. Based on the idea, Adamic-Adar (AA) index is obtained by using $\log(k_z)$ instead of k_z in (5).

(iv) Local Path (LP). CN considers the intersection of neighborhoods, which actually utilizes the one-path neighbors to characterize similarity. LP takes a general consideration of paths by considering two-path neighbors:

$$s^{\text{RA}} = A^2 + \epsilon A^3, \tag{6}$$

where A is the adjacent matrix of a network and ϵ is a small positive number. LP supposes that one-path neighbors contribute more to the similarity than two-path neighbors. LP is the low order parts of Katz method ($s^{\text{Katz}} = A^2 + \epsilon A^3 + \epsilon^2 A^4 + \cdots$), but with much lower computing complexity.

(v) Structural Perturbation Method (SPM). Lü et al. [6] suppose that network structure follows consistency after some random perturbation. In SPM, training set A^T is divided into a small fraction of perturbation set ΔA and the remaining set A^R ($A^T = A^R + \Delta A$). A^T has similar eigenvectors with A^R, but different eigenvalues. For the kth largest eigenvalues of A^T and A^R,

$$\Delta \lambda_k = \lambda_k \left(A^T \right) - \lambda_k \left(A^R \right) \approx \frac{x_k^T \Delta A x_k}{x_k^T x_k}, \tag{7}$$

where x_k is the eigenvector of A^R, corresponding to $\lambda_k(A^R)$. The similarity matrix $s = (s_{ij})_{N \times N}$ is

$$s = \sum_{k=1}^{N} \left(\lambda_k + \Delta \lambda_k \right) x_k x_k^T. \tag{8}$$

SPM first divides a network into training set and probe set and further divides the training set into perturbation set and the remaining set. For a given division of training and probe set, we calculate the average of 10 times independent simulations of (8) as the similarity matrix.

Apart from the five similarity metrics introduced above, for more similarity-evaluating methods, please refer to [42, 43].

2.3. The Proposed Method. We start our framework by reinvestigating the definition of precision. Supposing that s_{ij} is the similarity score of nodes i and j obtained by a prediction method F only based on training set E^T, $p_r(s)$ is the similarity distribution that a randomly chosen existing link in training set has score s, and $p_n(s)$ is the similarity distribution that a randomly chosen nonexisting link in the training set has score s. Due to random division of training set and probe set, links in the probe set should have the same similarity distribution with that of the training set at high confidence according to the law of large numbers [44, 45]. Thus we would not differentiate similarity distribution of existing links in the training and probe sets in the following paper. The assumption is reasonable according to the statistical theory if the size of samples goes to infinity [44, 45]. Since classical methods only predict links with high scores, the estimated precision of the method F is written as

$$p_F^0 = \frac{\left| E^P \right| \int_{c_0}^{s_{\max}} p_r(s) \, ds}{\left| U - E^T \right| \int_{c_0}^{s_{\max}} p_n(s) \, ds}, \tag{9}$$

where $|E^P|$ is the size of E^P, c_0 is a constant, and U is the whole set of all possible links ($|U| = (1/2)N(N-1)$). $s_{\max}$ is the maximum score. In real scenarios, the length of the prediction list is usually the size of the probe set [16], which requires c_0 subjecting to $|U - E^T| \int_{c_0}^{s_{\max}} p_n(s) ds = |E^P|$. If $p_r(s) \ll p_n(s)$ at $s > c_0$, the precision $p_F \to 0$. Otherwise, $p_r(s) \gg p_n(s)$ gives rise to a high precision. Since only links with top-L highest scores are predicted as potential links, precision could be calculated by (2) [6, 16]. Equation (2) is a much easier formula to describe precision than (9).

Most previous link prediction methods only predict links with high similarity scores. We generalize (9) by considering links of different similarities. Supposing that links with scores $s_{ij} \in S = (s_1, s_2) \bigcup (s_3, s_4) \cdots \bigcup (s_{2m-1}, s_{2m})$ are predicted as potential links, the precision is as follows:

$$p_F = \frac{\left| E^P \right| \int_{S=(s_1,s_2) \bigcup (s_3,s_4) \cdots \bigcup (s_{2m-1},s_{2m})} p_r(s) \, ds}{\left| U - E^T \right| \int_{S=(s_1,s_2) \bigcup (s_3,s_4) \cdots \bigcup (s_{2m-1},s_{2m})} p_n(s) \, ds}, \tag{10}$$

where $s_1 < s_2 < s_3 < \cdots s_{2m-1} < s_{2m}$. To confine the length of the prediction list, a precondition requires $|U - E^T| \int_S p_n(s) ds = |E^P|$. Note that, in most previous works, $S = (c_0, s_{\max})$, and equation (10) reduces to (9). Our generalized precision equation (10) considers links with both high and low scores.

The main concern is to select appropriate set S in (10) to maximize the precision. We propose precision-to-noise ratio (PNR) to determine S,

$$\text{PNR}(s) = \frac{p_r(s)}{p_n(s)}, \tag{11}$$
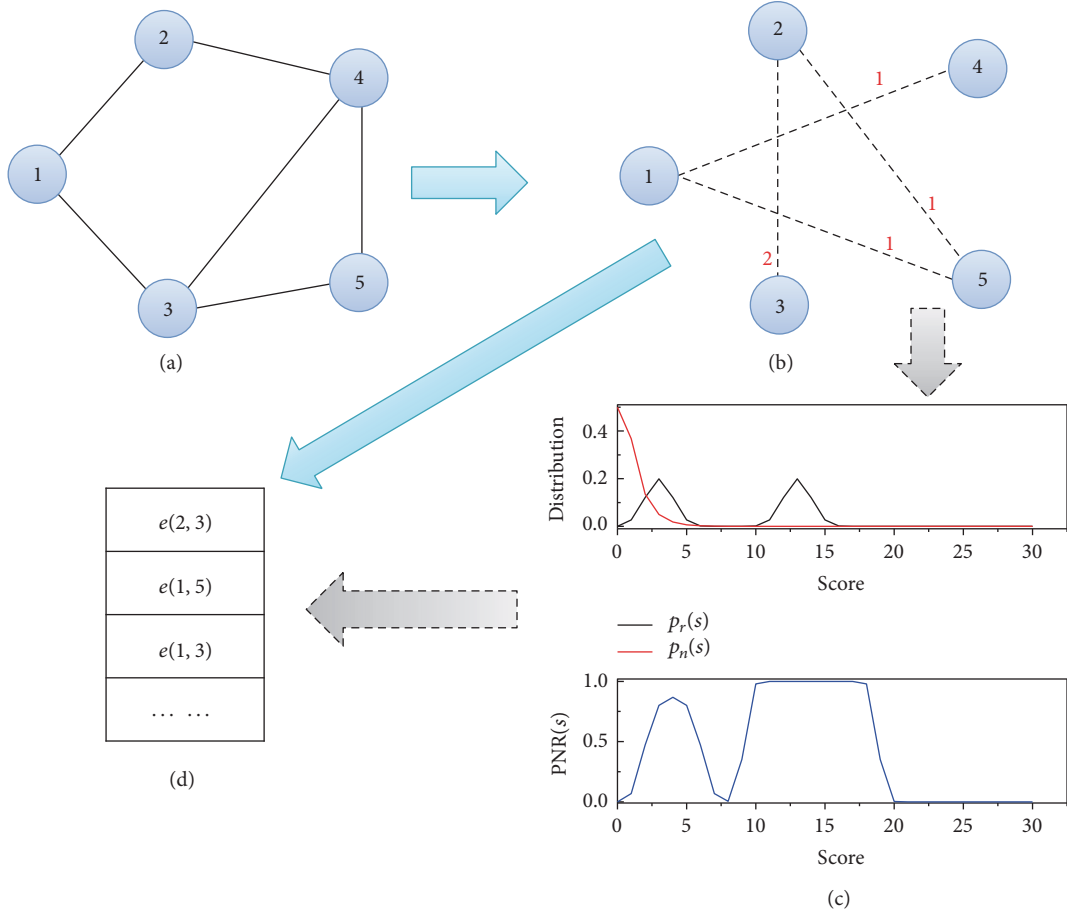
FIGURE 1: Schematic shows the proposed framework based on CN. (a) A snapshot of a large network. (b) Score of nonexisting links calculated by CN method. (c) The top panel is the score distributions of existing and nonexisting links, $p_r(s)$ and $p_n(s)$. The bottom panel is PNR(s) = $p_r(s)/p_n(s)$. (d) Predicted links. State-of-the-art prediction methods follow the path (a)→(b)→(d), while our proposed framework follows the path (a)→(b)→(c)→(d), which has an additional path PNR(s).

where PNR(s) measures the ability to distinguish real links with the same score. Note that a nonexisting link in training set may be an existing link in probe set. Given a nonexisting link in training set with the similarity s_{ij}, the probability that it is an existing link in probe set (i.e., the precision) is $p' = (|E^P| \cdot p_r(s))/(|U - E^T| \cdot p_n(s)) = \alpha$PNR$(s)$, where $\alpha = |E^P|/|U - E^T|$ is a constant.

The central issue of our framework is to use PNR(s) to determine the optimal score set S. We first calculate the similarity scores of all links only based on training set by a traditional method. Second, $p_r(s)$, $p_n(s)$, and PNR(s) are computed. Third, we reassign the scores of each link s'_{ij} = PNR(s_{ij}), where s_{ij} is the original similarity score by the first step. Finally, we sort links in the descending order of s' and links with top-L scores are predicted as potential links [16, 18]. The optimal score set S_{opt} corresponds to the original similarity scores whose reassigned scores rank in the top-L score list.

Different kinds of similarity evaluations could be introduced into the framework. Taking CN similarity method as an example, our framework is as follows:

(1) Divide the links of a network into $1 - p^H$ training set and p^H probe set randomly.

(2) Calculate the similarity scores of all existing and nonexisting links by CN method only according to training set.

(3) Calculate PNR(s). Divide the scores into K uniform bins and count how many existing ($n_{e,i}$) and nonexisting ($n_{n,i}$) links locate in each bin i (i.e., calculate discrete $p_r(s)$, $p_n(s)$). Then we obtain PNR(s_k) = $p_r(s_k)/p_n(s_k)$, $k = 1, 2, \ldots, K$. Note that if $p_n(s_k) = 0$, we define PNR$(s_k) = 0$.

(4) Obtain the readjusting scores of the nonexisting links in training set by $s' = $ PNR(s).

(5) Determine the prediction list by choosing links with $L = |E^P|$ highest scores s', and calculate the precision.

Figure 1 depicts the proposed framework based on CN method. After obtaining the similarity scores of links (Figure 1(a)→1(b)), traditional CN method directly predicts potential links according to the scores (Figure 1(b)→1(d)), while the proposed framework calculates PNR(s) (Figures

TABLE 1: Structural properties of the different real networks. Structural properties include network size (N), link number (E), degree heterogeneity ($H = \langle k^2 \rangle / \langle k \rangle^2$), degree assortativity (r), average clustering coefficient ($\langle C \rangle$), average shortest path length ($\langle d \rangle$), and sparsity.

Network	N	E	H	r	$\langle C \rangle$	$\langle d \rangle$	Sparsity
Email	33696	180811	6.070	−0.060	0.170	4.08	3.2×10^{-4}
PDZBase	161	209	2.263	−0.466	0.001	5.11	1.6×10^{-2}
Euroad	1039	1305	1.228	0.090	0.004	18.39	2.4×10^{-3}
Neural	297	2148	1.81	−0.163	0.292	2.46	4.9×10^{-2}
Roundworm	453	2025	4.485	−0.226	0.647	2.66	2.0×10^{-2}
USair	332	2126	3.464	−0.208	0.625	2.74	3.9×10^{-2}

1(b)→1(c)) and later predicts potential links according to the modified scores (Figures 1(c)→1(d)).

An important property of our framework is that if S is determined according to PNR(s), that is, PNR(x) > PNR(y), $\forall x \in S$, $\forall y \in \mathbb{R} - S$, the precision p_F could exploit full capacity of a given similarity-evaluating method. PNR(s) is the optimal transforming function $f_{opt}(s) = $ PNR(s). It means that no matter how we transform the similarity by other one-variable function, $s'' = f'(s)$, the precision performance of s'' cannot outperform the proposed method by PNR(s). For the proof of the optimal PNR(s), please see part I in the supplementary materials.

3. Experimental Results

We first describe the six real networks in Section 3.1. The precision comparison between our method and the baseline methods is given in Section 3.2. Finally, the characteristics of the predicted links by different methods are investigated in Section 3.3.

3.1. Datasets. To verify the effectiveness of the proposed method, we measure the performance of our framework in six empirical networks from diverse disciplines and backgrounds: (1) email [46]: Enron email communication network covers all the email communication within a dataset of around half million emails; nodes of the network are email addresses and if an address i sent at least one email to address j, the graph contains an undirected link from i to j; (2) PDZBase [47]: an undirected network of protein-protein interactions from PDZBase; (3) Euroad [48]: international E-road network that locates mostly in Europe; the network is undirected, with nodes representing cities and links denoting e-road between two cities; (4) neural [49]: a directed and weighted neural network of *C. elegans*; (5) USair [6]: an directed network of flights between US airports in 2010; each link represents a connection from one airport to another in 2010; (6) roundworm [49]: a metabolic network of *C. elegans*.

Different real networks contain directed or undirected, weighted or unweighted links. To simplify the problem, we treat all links undirected and unweighted. Besides, only the giant connected components of these networks are taken into account. This is because for a pair of nodes located in two disconnected components, their similarity score will be zero according to most prediction methods. Table 1 shows the basic statistics of those networks.
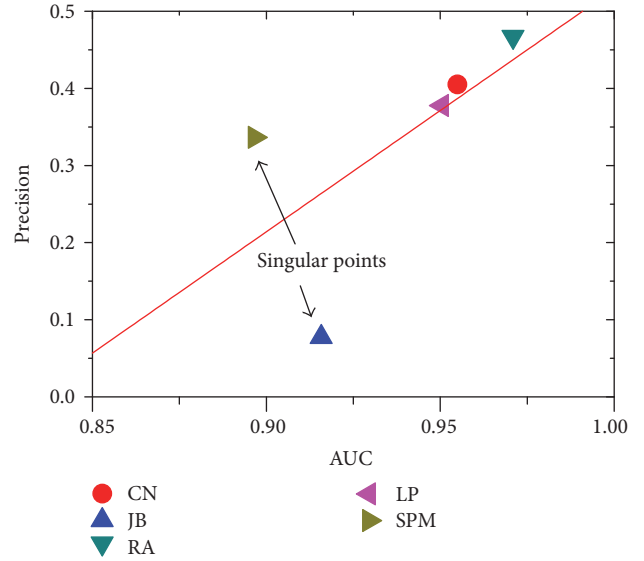
FIGURE 2: AUC and precision of the USair network obtained by five different approaches: common neighbors (CN), Jaccard index (JB), resource allocation index (RA), local path index (LP), and structural perturbation method (SPM). The results are obtained by 50 independent simulations. SPM method achieves high precision, yet low AUC, and JB has low precision, but high AUC (>0.9).

3.2. Precision Evaluation. In the experiments, we set $p^H = 10\%$ that means the networks are randomly divided into 90% training set and 10% probe set. All the experiments are the average of 50 independent simulations.

Figure 2 shows AUC and precision of five different methods in USair network. In Figure 2, CN method achieves low AUC, yet high precision, whereas RA method achieves similar AUC with methods of CN, JB, and SPM, but much lower precision. Apart from USair network, the deviation between AUC and precision also exists in other real-world networks (see FIG. S1 in the supplementary materials). The main reason is that AUC characterizes the score difference between existing and nonexisting links in the whole networks, whereas precision only counts the links with top-L high scores. Specifically, from the perspective of score distributions, AUC $= \int_{-\infty}^{+\infty} p_r(x) \int_{-\infty}^{x} p_n(y) dx\, dy$. Comparing with (10), the definitions of the two metrics are completely different, resulting in little correlation between them.

Figure 3 shows PNR and the score distributions of existing and nonexisting links for USair network by CN method.

TABLE 2: Maximal precision comparison of the proposed methods and traditional high-similarity methods for six real-world networks. Traditional precision is obtained by the maximum of traditional methods, that is, max{CN, Jaccard, RA, LP, SPM}. Proposed precision is obtained by our framework, that is, max{PNR_{CN}, $PNR_{Jaccard}$, PNR_{RA}, PNR_{LP}, PNR_{SPM}}.

	Email	PDZBase	Euroad	Neural	Roundworm	USair
Traditional p_{max}	0.0171	0.0032	0.0052	0.0107	0.2651	0.4670
Proposed p_{max}	**0.0313**	**0.3159**	**0.0674**	**0.0392**	**0.3475**	**0.5309**
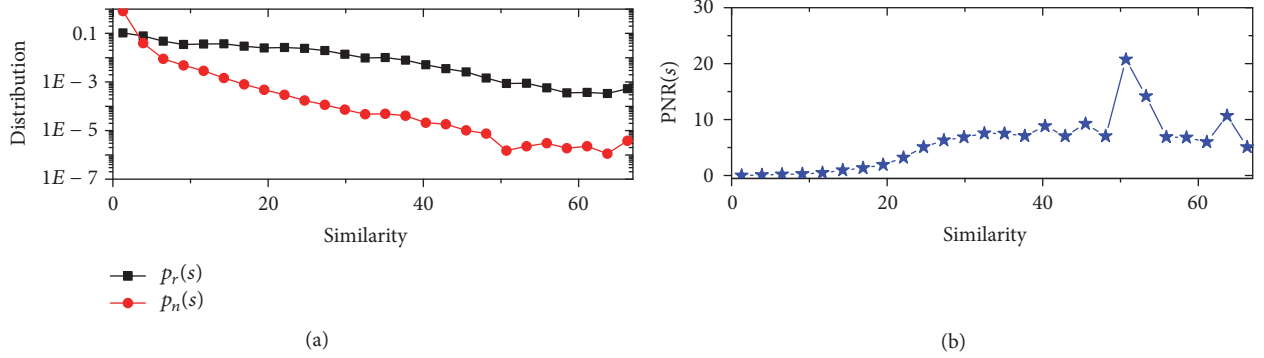
FIGURE 3: Similarity distributions and the corresponding PNR(s) of USair network, where the similarity is obtained by CN method. (a) Similarity distributions of the existing and nonexisting links, $p_r(s)$ and $p_n(s)$, respectively. (b) PNR(s) as a function of similarity in USair network.

In Figure 3(a), the scores of existing and nonexisting links follow power law distribution largely. High scores sometimes correspond to low PNR, especially at Similarity $\approx$ 60 (see Figure 3(b)). Nevertheless, some low scores achieve high PNR, indicating that for a nonexisting link in training set with this particular score, the link is likely to be an existing link in probe set. For a nonexisting link in training set with high score, yet with low PNR, it has a high probability not to be an existing link in probe set. The similar phenomenon also exists in other networks (see FIG. S2 in the supplementary materials). In consequence, the foundation of traditional methods, which suppose that similar nodes have a high likelihood to form links, is confronted with great challenges in precisely predicting links of low similarities.

Figure 6 shows the precision difference between the proposed PNR methods and the baseline methods. Our proposed method enhances precision remarkably compared with the original methods in most cases. Some fluctuation exists in these methods, due to the limited size of networks. Table 2 gives the maximal precision increasement in the six networks. In Table 2, precision is obtained by the maximum of traditional methods and PNR methods, respectively, that is, max{CN, Jaccard, RA, LP, SPM} and max{PNR_{CN}, $PNR_{Jaccard}$, PNR_{RA}, PNR_{LP}, PNR_{SPM}}. Our method outperforms state-of-the-art methods in the six networks. Besides, Figure 4 shows the influence of the probe set size on the precision performance. We find that our method outperforms classical methods when $p^H > 0.85$, except for JB method when $p^H > 0.6$. Other networks have similar results (see FIG. S3 in the supplementary materials). However, according to the theoretical analysis (see the first part in the supplementary materials), our method should perform better than, or at least equally to, the classical
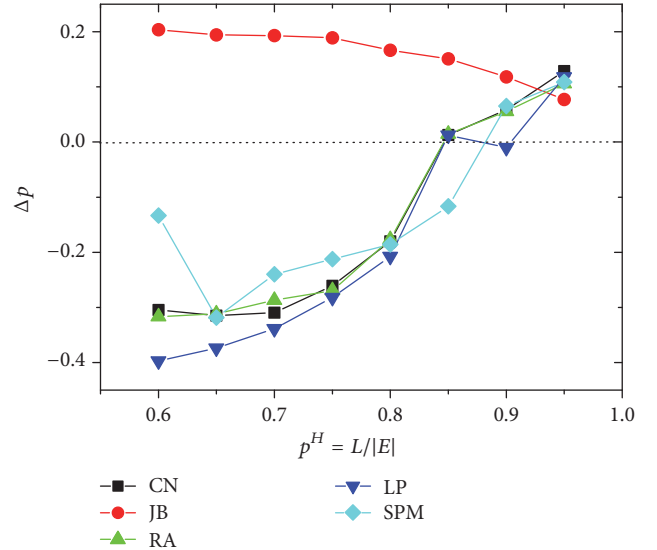
FIGURE 4: The precision difference Δp as a function of probe set size $p^H = L/|E|$ in Usair network. $\Delta p = p_{PNR} - p_{original}$. $\Delta p > 0$ means that our method outperforms the original methods. In the panel, when $p^H > 0.85$, $\Delta p > 0$.

methods. The reason is that we suppose the network structure is not influenced by the random division of training and probe set. Thus, the training subnetwork should have similar structure with the original networks. The assumption is rational when p^H is small. If the size p^H of the probe set is large, the training sets have many differences with the entire networks, which violates the assumption of our method. Therefore, our method performs well when the fraction of the probe set is small.
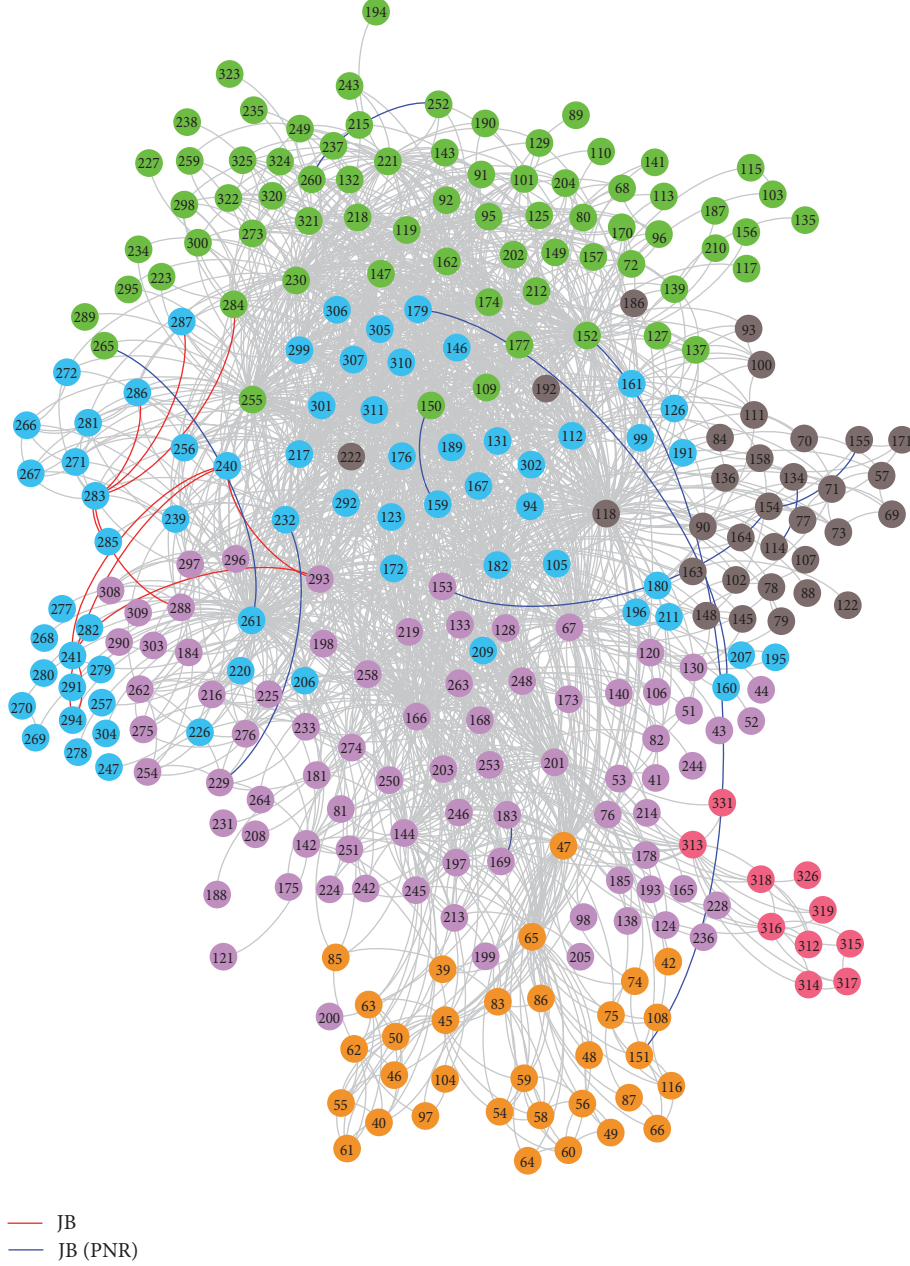
FIGURE 5: The comparison of the predicted edges between JB and the corresponding PNR methods in the Usair network. In the panel, we predict 10 edges for both JB and PNR$_{JB}$ methods. The Usair network is divided into different communities by the method in [37]. Nodes in the same community have the same color and short geographical distances. Our method (blue lines) predicts more edges between faraway nodes in different communities, while the original JB method (red lines) only predicts edges between close nodes.

3.3. Characteristics of the Predicted Links.

Long-range links play an important role in the dynamics of networks and it is of much significance to predict long-range links [32, 50]. Figure 5 gives a comparison of the predicted links between JB and the corresponding PNR methods in the Usair network. In Figure 5, our method predicts more links between faraway nodes in different communities, while the original JB method only predicts links between close nodes. Community detection method in [37] is utilized in Figure 5. However, it is difficult to evaluate long-range links solely based on community divisions. Since long-range links usually have long distances and low similarities, we would investigate the average distance and average similarity of the predicted links by our proposed framework.

The distance d_{ij} of a link e_{ij} is the shortest distance between nodes i and j only based on training set. Since the endpoints of the predicted links do not connect directly, $d_{ij} \geq 2$. The average distance of the predicted links is

$$\overline{d}_{\text{predict}} = \frac{1}{L} \sum_{e_{ij} \in \text{predcited links}} d_{ij}. \tag{12}$$
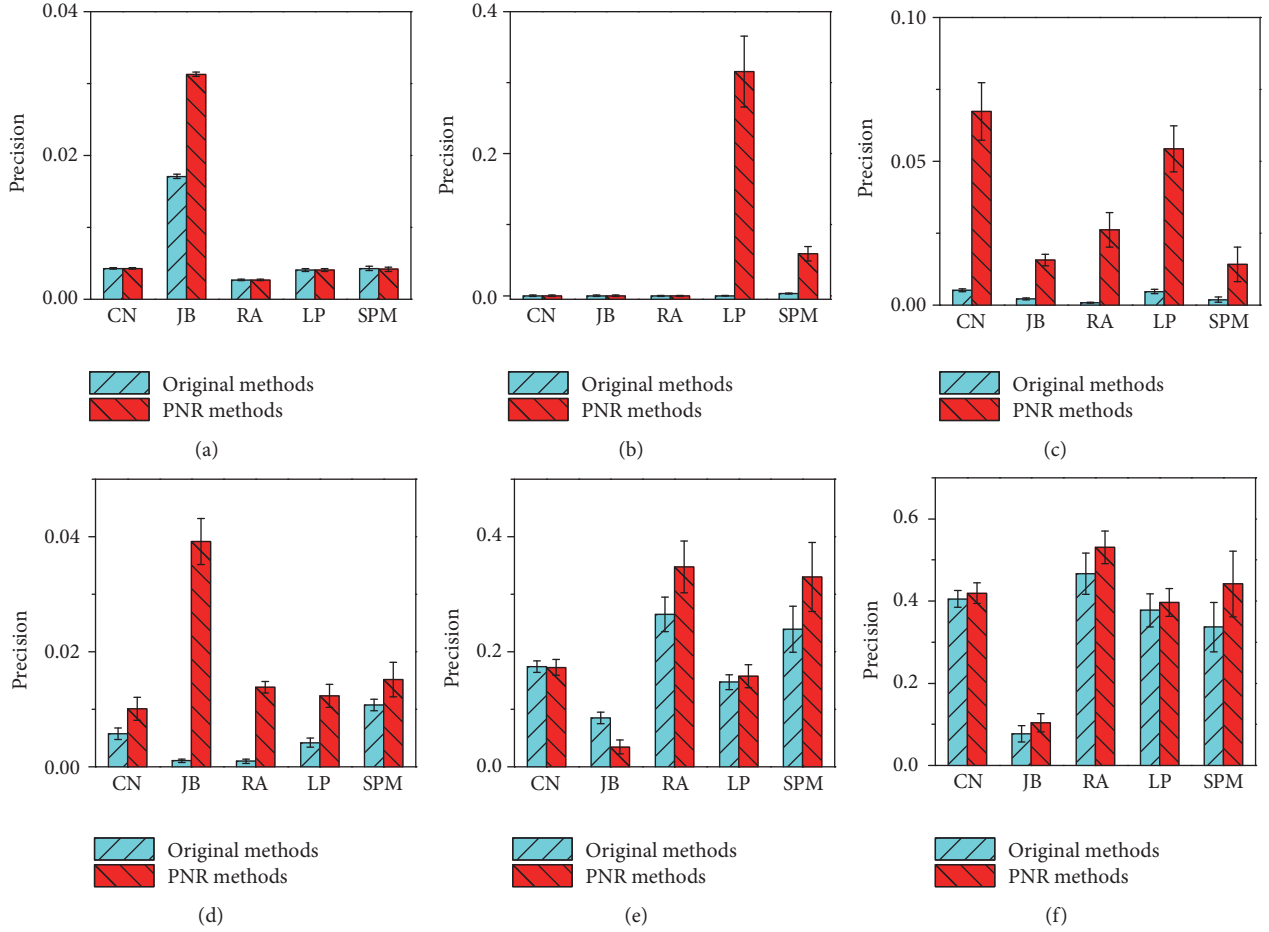
FIGURE 6: Precision comparison of the proposed methods (red) and traditional high similarity based methods (cyan) for six real-world networks. (a) Email network. (b) PDZBase network. (c) Euroad network. (d) Neural network. (e) Roundworm network. (f) USair network. Results are the average of 50 independent simulations. Our proposed framework increases precision in most cases.

Analogously, the average similarity of the predicted links is

$$\bar{s}_{\text{predict}} = \frac{1}{L} \sum_{e_{ij} \in \text{predcited links}} s_{ij}, \qquad (13)$$

where s_{ij} is the similarity of nodes i and j in training set.

Figure 7 shows the difference of the average distances obtained by PNR method and the corresponding original methods. Generally, PNR method achieves a higher average distance than the corresponding original methods in the six networks, especially for SPM in Email network and LP in USair network, whereas for many cases, PNR and the original methods have the same average distance $\bar{d} = 2$. It is because that the distance of most unconnected nodes are 2, revealing that most commonly used methods incline to predict triangle edges. Therefore, our method has little influence on the average distance. However for some sparser networks, such as neural and USair networks, the average distance is improved by our framework, especially for LP in USair network. Previous works show that the two endpoints of a long-range link usually have a high distance or low similarity. Since PNR framework could increase the average distance of the predicted links, it can be conjectured that more long-range

links are predicted. Besides, integrating Figures 6 and 7, we can find that our framework predict more long-range links correctly.

Furthermore, Figure 8 shows the difference of average similarity obtained by PNR method and the corresponding original methods. In Figure 8, PNR method achieves a lower average similarity than the corresponding original methods in the six networks, except RA method in roundworm network. The reason is that PNR has much fluctuations due to the limited size of networks, bringing about the unusual phenomenon of RA in roundworm network. Similar to the analysis of average distance, we show that PNR methods are beneficial to the prediction of long-range links, which agrees with the conclusion from Figure 7.

4. Conclusion

In summary, we systematically study the drawbacks of similarity-based link prediction methods and show that some link prediction methods achieve high AUC, yet low precision. Based on the differences of the similarity distributions of existing and nonexisting links, we propose a metric (PNR) to explain the problem of high AUC and low precision.

Figure 7: Comparison of average distance of the PNR predicted links with that of the corresponding original methods for different networks. (a) Email network. (b) PDZBase network. (c) Euroad network. (d) Neural network. (e) Roundworm network. (f) USair network. Results are the average of 50 independent simulations. Our proposed framework increases the average distance on the whole, which indicates that more long-range links are predicted correctly.

Two nodes with some particular low scores also have a high likelihood of forming links between them. Furthermore, we prove that PNR is the optimal one-variable function to adjust the likelihood scores of links. Experiments in real networks demonstrate the effectiveness of PNR, and the precision is greatly enhanced. Additionally, the proposed framework could also reduce the average similarity and increase the average distance of the predicted links, which indicates that more missing long-range links can be detected correctly.

Though the proposed approach investigates link prediction in unipartite networks, it could also be generalized to bipartite and other kinds of networks. What is more, our method provides a novel way to explore the connecting patterns of real networks that may inspire other better score-assigning methods in the future.

Conflicts of Interest

The authors declare no competing financial interests.

Acknowledgments

Supplementary Materials

In the supplementary materials, we prove that PNR(x) is the optimal transferring function in Section 1. The deviation of AUC and precision in different networks is shown in
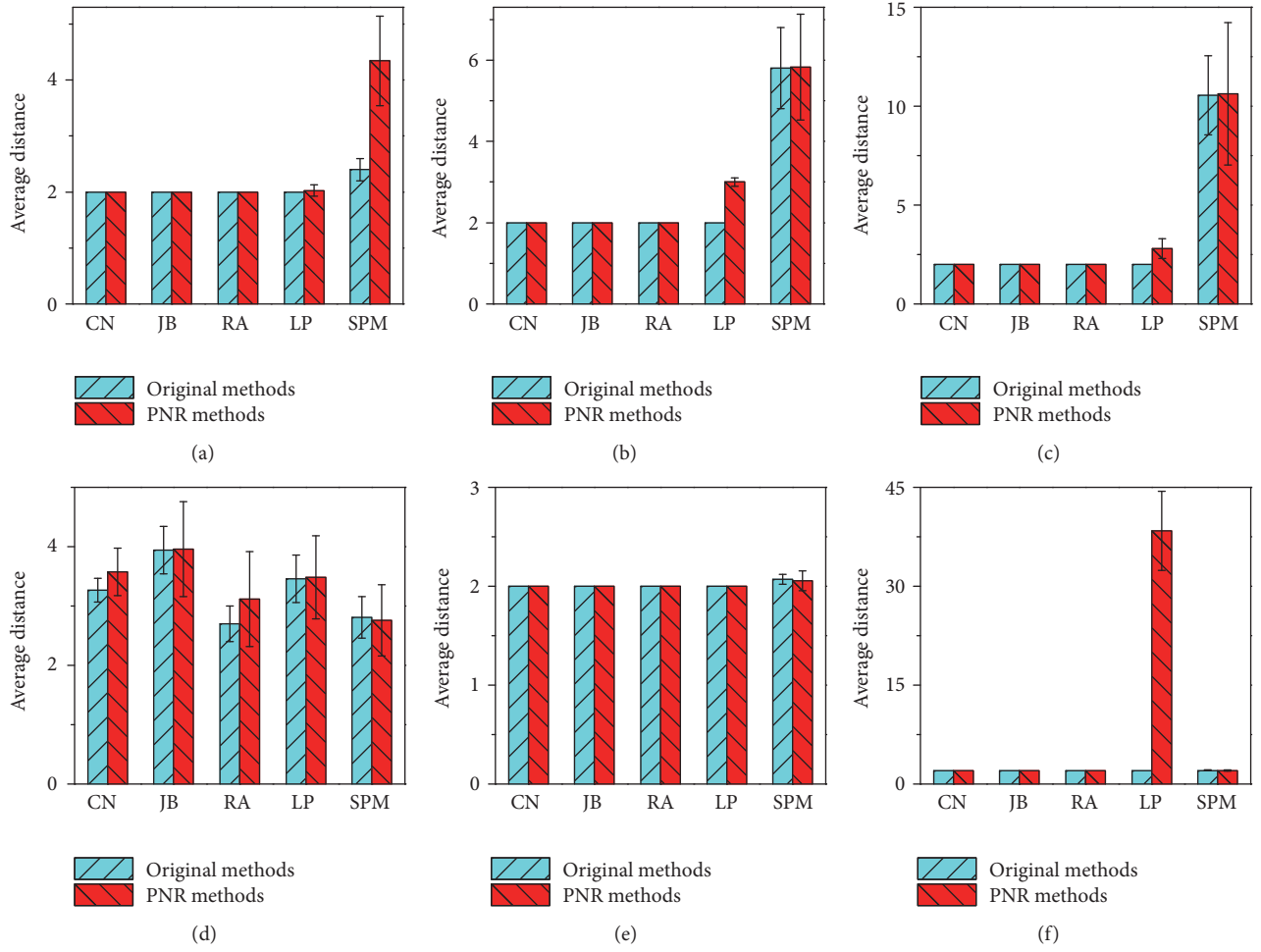
(a)

(b)

(c)

(d)

(e)

(f)
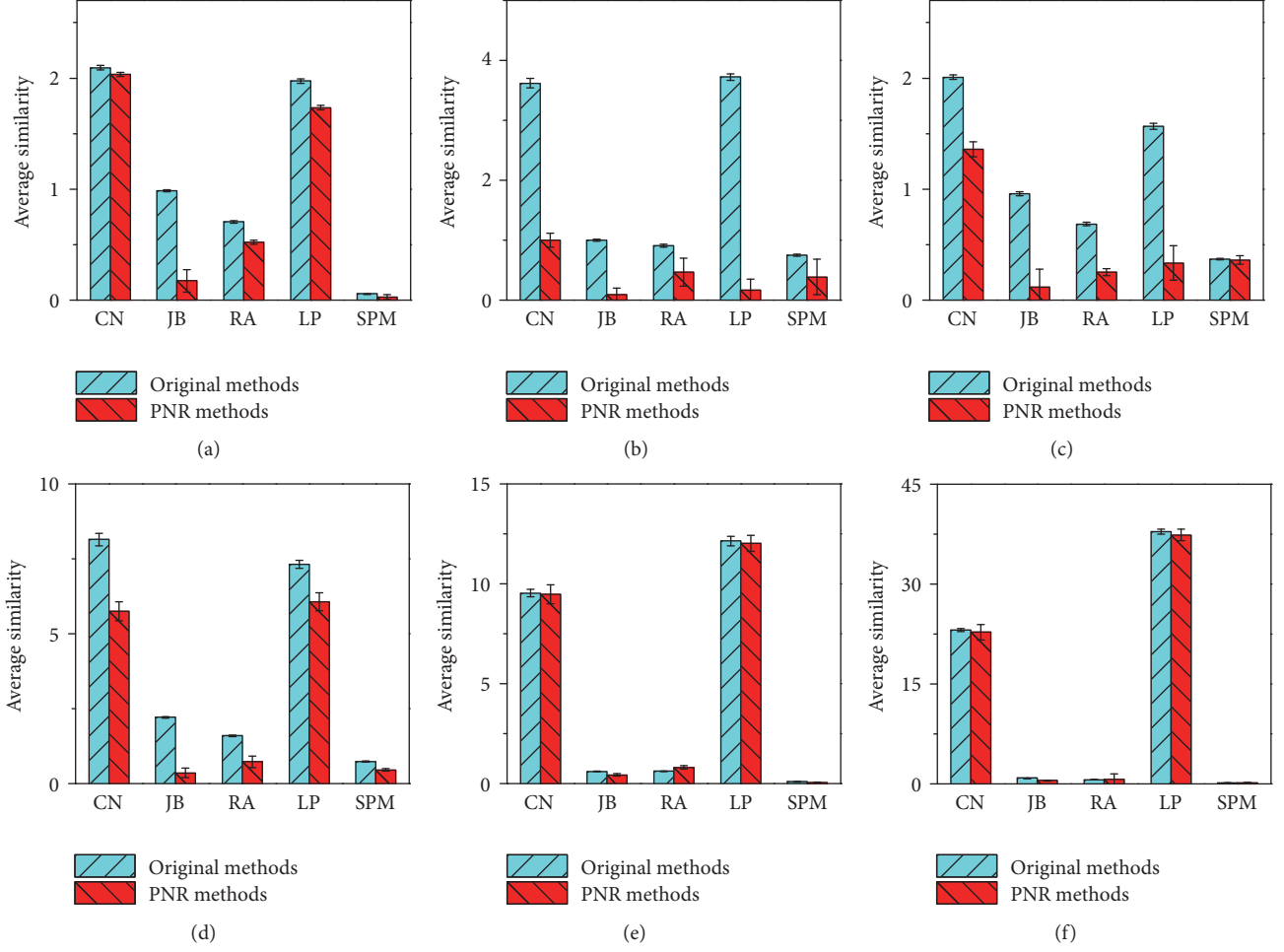
FIGURE 8: Comparison of average similarity of the PNR predicted is linked with that of the corresponding original methods for different networks. (a) Email network. (b) PDZBase network. (c) Euroad network. (d) Neural network. (e) Roundworm network. (f) USair network. Results are the average of 50 independent simulations. Our proposed framework reduces the average similarity on the whole, which indicates that more long-range links are predicted correctly.

Section 2. The PNR performances of different methods in different networks are shown in Section 3. In Section 3, we first plot the PNR(x) by different methods in FIG. S2 and then show the influence of the probe set size on the precision in Fig. S3. FIG. S1 (color online), AUC and precision of six real-world networks (see Table 2) by five different popular approaches. Results are average of 50 independent simulations. In the experiments, $p^H = 0.1$ means that we utilize 90% existing edges as training set to predict the other 10% edges (probe set). FIG. S2 (color online), PNR for six networks by five different methods. (a) Email network. (b) PDZBase network. (c) Euroad network. (d) Neural network. (e) Roundworm network. (f) USair network. Results are the average of 50 independent simulations and are obtained only according to training set. For different methods and different networks, scores are normalized to 0~1 with $s_{new} = (s - s_{min})/(s_{max} - s_{min})$. FIG. S3 (Color online), the precision difference Δp as a function of probe set size $p^H = L/|E|$ in the four networks, where Δp is the difference between the five classical and the corresponding PNR methods, $\Delta p = p_{PNR} - p_{original}$. $\Delta p > 0$ means that our method outperforms

the original methods. In the panels, when $p^H > 0.85$, $\Delta p > 0$. *(Supplementary Materials)*

References

[1] J. Esquivel-Gómez, R. E. Balderas-Navarro, P. D. Arjona-Villicaña, P. Castillo-Castillo, O. Rico-Trejo, and J. Acosta-Elias, "On the emergence of islands in complex networks," *Complexity*, 2017.

[2] W. X. Wang, Y. C. Lai, and C. Grebogi, "Data based identification and prediction of nonlinear and complex dynamical systems," *Physics Reports*, vol. 644, pp. 1–76, 2016.

[3] A.-L. Barabási, "Network science," *Philosophical Transactions of the Royal Society A: Mathematical, Physical & Engineering Sciences*, vol. 371, no. 1987, article no. 0375, 2013.

[4] T. G. Lewis, *Network science: Theory and applications*, vol. 8, John Wiley and Sons, 2011.

[5] G. Chen, R. Mao, and K. Lu, "A parallel computing framework for big data," *Frontiers of Computer Science*, vol. 11, no. 4, pp. 608–621, 2017.

[6] L. Lü, L. Pan, T. Zhou, Y. C. Zhang, and H. E. Stanley, "Toward link predictability of complex networks," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 112, no. 8, pp. 2325–2330, 2015.

[7] T. Wang, M. Y. Zhou, and Z. Q. Fu, "Link prediction in evolving networks based on the popularity of nodes".

[8] L. Sharma and A. Gera, "A survey of recommendation system: Research challenges," *International Journal of Engineering Trends and Technology (IJETT)*, vol. 4, no. 5, pp. 1989–1992, 2013.

[9] J. Lu, D. S. Wu, M. S. Mao, W. Wang, and G. Zhang, "Recommender system application developments: a survey," *Decision Support Systems*, vol. 74, pp. 12–32, 2015.

[10] W.-X. Wang, Y.-C. Lai, C. Grebogi, and J. Ye, "Network Reconstruction Based on Evolutionary-Game Data via Compressive Sensing," *Physical Review X*, vol. 1, no. 2, Article ID 021021, pp. 1–7, 2011.

[11] Z. Shen, W.-X. Wang, Y. Fan, Z. Di, and Y.-C. Lai, "Reconstructing propagation networks with natural diversity and identifying hidden sources," *Nature Communications*, vol. 5, article no. 5323, 2014.

[12] H. Liao, A. Zeng, M. Zhou, R. Mao, and B. Wang, "Information mining in weighted complex networks with nonlinear rating projection," *Communications in Nonlinear Science and Numerical Simulation*, vol. 51, pp. 115–123, 2017.

[13] Y. Wang, J. Wang, H. Liao, and H. Chen, "An efficient semi-supervised representatives feature selection algorithm based on information theory," *Pattern Recognition*, vol. 61, pp. 511–523, 2017.

[14] A. L. Barabási, *Network Science*, vol. 4, Cambridge University Press, 2016.

[15] R. Mao, P. Zhang, X. Li, X. Liu, and M. Lu, "Pivot selection for metric-space indexing," *International Journal of Machine Learning and Cybernetics*, vol. 7, no. 2, pp. 311–323, 2016.

[16] L. Lü and T. Zhou, "Link prediction in complex networks: a survey," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 6, pp. 1150–1170, 2011.

[17] A. Zeng and G. Cimini, "Removing spurious interactions in complex networks," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 85, no. 3, Article ID 036101, 2012.

[18] V. Martínez, F. Berzal, and J.-C. Cubero, "A survey of link prediction in complex networks," *ACM Computing Surveys*, vol. 49, no. 4, article no. 69, 2016.

[19] A. Javari and M. Jalili, "A probabilistic model to resolve diversity–accuracy challenge of recommendation systems," *Knowledge and Information Systems*, vol. 44, no. 3, pp. 609–627, 2015.

[20] A. Javari and M. Jalili, "Accurate and novel recommendations: An algorithm based on popularity forecasting," *ACM Transactions on Intelligent Systems and Technology*, vol. 5, no. 4, 2015.

[21] P. Zhang, X. Wang, F. Wang, A. Zeng, and J. Xiao, "Measuring the robustness of link prediction algorithms under noisy environment," *Scientific Reports*, vol. 6, Article ID 18881, 2016.

[22] Z. Huang, X. Li, and H. Chen, "Link prediction approach to collaborative filtering," in *Proceedings of the the 5th ACM/IEEE-CS joint conference*, p. 141, Denver, CO, USA, June 2005.

[23] D. K. Lin, "An information-theoretic definition of similarity," in *Proceedings of the In 15th International Conference on Machine Learning (ICML '15)*, vol. 98, pp. 296–304, 1998.

[24] E. A. Leicht, P. Holme, and M. E. J. Newman, "Vertex similarity in networks," *Physical Review E: Statistical, Nonlinear, and Soft Matter Physics*, vol. 73, no. 2, Article ID 026120, 2006.

[25] Z. Liu, Q.-M. Zhang, L. Lü, and T. Zhou, "Link prediction in complex networks: a local naïve Bayes model," *EPL (Europhysics Letters)*, vol. 96, no. 4, Article ID 48007, 2011.

[26] D. Heckerman, C. Meek, and D. Koller, "Probabilistic entity-relationship models, prms, and plate models," in *Introduction to Statistical Relational Learning*, pp. 201–238, 2007.

[27] K. Yu, W. Chu, S. Yu, V. Tresp, and Z. Xu, "Stochastic relational models for discriminative link prediction," in *Proceedings of the 20th Annual Conference on Neural Information Processing Systems(NIPS '06)*, pp. 1553–1560, December 2006.

[28] A. Clauset, C. Moore, and M. E. J. Newman, "Hierarchical structure and the prediction of missing links in networks," *Nature*, vol. 453, no. 7191, pp. 98–101, 2008.

[29] C. Zhou, L. Zemanová, G. Zamora, C. C. Hilgetag, and J. Kurths, "Hierarchical organization unveiled by functional connectivity in complex brain networks," *Physical Review Letters*, vol. 97, no. 23, Article ID 238103, 2006.

[30] P. De Meo, E. Ferrara, G. Fiumara, and A. Provetti, "On facebook, most ties are weak," *Communications of the ACM*, vol. 57, no. 11, pp. 78–84, 2014.

[31] S. Aral, "The future of weak ties," *American Journal of Sociology*, vol. 121, no. 6, pp. 1931–1939, 2016.

[32] X. Wang, W. Lu, M. Ester, C. Wang, and C. Chen, "Social recommendation with strong and weak ties," in *Proceedings of the 25th ACM International Conference on Information and Knowledge Management, CIKM 2016*, pp. 5–14, USA, October 2016.

[33] A. L. Kavanaugh, D. D. Reese, J. M. Carroll, and M. B. Rosson, "Weak ties in networked communities," *Information Society*, vol. 21, no. 2, pp. 119–131, 2005.

[34] B. Ozenne, F. Subtil, and D. Maucort-Boulch, "The precision-recall curve overcame the optimism of the receiver operating characteristic curve in rare diseases," *Journal of Clinical Epidemiology*, vol. 68, no. 8, pp. 855–859, 2015.

[35] J. M. Lobo, A. Jiménez-valverde, and R. Real, "AUC: A misleading measure of the performance of predictive distribution models," *Global Ecology and Biogeography*, vol. 17, no. 2, pp. 145–151, 2008.

[36] J. Davis and M. Goadrich, "The relationship between precision-recall and ROC curves," in *Proceedings of the 23rd International Conference on Machine Learning (ICML '06)*, pp. 243–275, ACM, NY, USA, June 2006.

[37] R. Lambiotte, J. C. Delvenne, and M. Barahona, *Laplacian Dynamics and Multiscale Modular Structure in Networks*, Physics, 2008.

[38] H. Liao, M. S. Mariani, M. s. Medo, Y.-C. Zhang, and M.-Y. Zhou, "Ranking in evolving complex networks," *Physics Reports*, vol. 689, pp. 1–54, 2017.

[39] J. A. Hanley and B. J. McNeil, "The meaning and use of the area under a receiver operating characteristic (ROC) curve," *Radiology*, vol. 143, no. 1, pp. 29–36, 1982.

[40] I. R. Dunsmore and S. Geisser, "Predictive Inference: An Introduction.," *Journal of the Royal Statistical Society. Series A (Statistics in Society)*, vol. 158, no. 1, p. 191, 1995.

[41] J. L. Herlocker, J. A. Konstan, L. G. Terveen, and J. T. Riedl, "Evaluating collaborative filtering recommender systems," *ACM Transactions on Information and System Security*, vol. 22, no. 1, pp. 5–53, 2004.

[42] L. Getoor and C. P. Diehl, "Link mining: a survey," *SIGKDD Explorations Newsletter*, vol. 7, no. 2, pp. 3–12, 2005.

[43] M. A. Hasan and M. J. Zaki, "A survey of link prediction in social networks," in *Social Network Data Analytics*, pp. 243–275, Springer, NY, USA, 2011.

[44] C. C. Heyde, "A supplement to the strong law of large numbers," *Journal of Applied Probability*, vol. 12, pp. 173–175, 1975.

[45] K. L. Chung, "The strong law of large numbers," *Selected Works of Kai Lai Chung*, pp. 145–156, 2008.

[46] J. Leskovec, K. J. Lang, A. Dasgupta, and M. W. Mahoney, "Community structure in large networks: Natural cluster sizes and the absence of large well-defined clusters," *Internet Mathematics*, vol. 6, no. 1, pp. 29–123, 2009.

[47] T. Beuming, L. Skrabanek, M. Y. Niv, P. Mukherjee, and H. Weinstein, "PDZBase: A protein-protein interaction database for PDZ-domains," *Bioinformatics*, vol. 21, no. 6, pp. 827-828, 2005.

[48] L. Subelj and M. Bajec, "Robust network community detection using balanced propagation," *The European Physical Journal B*, vol. 81, no. 3, pp. 353–362, 2011.

[49] University of koblenz-Landau, "the koblenz network collection," http://konect.uni-koblenz.de/networks/.

[50] X. F. Wang and G. Chen, "Complex networks: small-world, scale-free and beyond," *IEEE Circuits and Systems Magazine*, vol. 3, no. 1, pp. 6–20, 2003.

Research Article

Sparse Causality Network Retrieval from Short Time Series

Tomaso Aste[1,2,3] **and T. Di Matteo**[1,2,3,4]

[1]*Department of Computer Science, UCL, London, UK*
[2]*UCL Centre for Blockchain Technologies, UCL, London, UK*
[3]*Systemic Risk Centre, London School of Economics and Political Sciences, London, UK*
[4]*Department of Mathematics, King's College London, London, UK*

Correspondence should be addressed to Tomaso Aste; t.aste@ucl.ac.uk

We investigate how efficiently a known underlying sparse causality structure of a simulated multivariate linear process can be retrieved from the analysis of time series of short lengths. Causality is quantified from conditional transfer entropy and the network is constructed by retaining only the statistically validated contributions. We compare results from three methodologies: two commonly used regularization methods, Glasso and ridge, and a newly introduced technique, LoGo, based on the combination of information filtering network and graphical modelling. For these three methodologies we explore the regions of time series lengths and model-parameters where a significant fraction of true causality links is retrieved. We conclude that when time series are short, with their lengths shorter than the number of variables, sparse models are better suited to uncover true causality links with LoGo retrieving the true causality network more accurately than Glasso and ridge.

1. Introduction

Establishing causal relations between variables from observation of their behaviour in time is central to scientific investigation and it is at the core of data-science where these causal relations are the basis for the construction of useful models and tools capable of prediction. The capability to predict (future) outcomes from the analytics of (past) input data is crucial in modelling and it should be the main property to take into consideration in model selection, when the validity and meaningfulness of a model is assessed. From a high-level perspective, we can say that the whole scientific method is constructed around a circular procedure consisting in observation, modelling, prediction, and testing. In such a procedure, the accuracy of prediction is used as a selection tool between models. In addition, the principle of parsimony favours the simplest model when two models have similar predictive power.

The scientific method is the rational process that, for the last 400 years, has mostly contributed to scientific discoveries, technological progress, and the advancement of human knowledge. Machine learning and data-science are nowadays pursuing the ambition to mechanize this discovery process by feeding machines with data and using different methodologies to build systems able to make models and predictions by themselves. However, the automatisation of this process requires to identify, without the help of human intuition, the relevant variables and the relations between these variables out of a large quantity of data. Predictive models are methodologies, systems, or equations which identify and make use of such relations between sets of variables in a way that the knowledge about a set of variables provides information about the values of the other set of variables. This problem is intrinsically high-dimensional with many input and output data. Any model that aims to explain the underlying system will involve a number of elements which must be of the order of magnitude of the number of relevant relations between the system's variables. In complex systems, such as financial markets or the brain, prediction is probabilistic in nature and modelling concerns inferring the probability of the values of a set of variables given the values of another set. This requires the estimation of the joint probability of all

variables in the system and, in complex systems, the number of variables with potential macroscopic effects on the whole system is very large. This poses a great challenge for the model construction/selection and its parameter estimation because the number of relations between variables scales with, at least, the square of the number of variables but, for a given fix observation window, the amount of information gathered from such variables scales, at most, linearly with the number of variables [1, 2].

For instance, a linear model for a system with p variables requires the estimation from observation of $p(p+1)/2$ parameters (the distinct elements of the covariance matrix). In order to estimate $\mathcal{O}(p^2)$ parameters one needs a comparable number of observations requiring time series of length $q \sim p$ or larger to gather a sufficient information content from a number of observations which scales as $p \times q \sim \mathcal{O}(p^2)$. However, the number of parameters in the model can be reduced by considering only $\mathcal{O}(p)$ out of the $\mathcal{O}(p^2)$ relations between the variables reducing in this way the required time series length to $\mathcal{O}(p)$. Such models with reduced numbers of parameters are referred to in the literature as sparse models. In this paper we consider two instances of linear sparse modelling: Glasso [3] which penalizes nonzero parameters by introducing a ℓ_1 norm penalization and LoGo [4] which reduces the inference network to an $\mathcal{O}(p)$ number of links selected by using information filtering networks [5–7]. The results from these two sparse models are compared with the ℓ_2 norm penalization (nonsparse) ridge model [8, 9].

This paper is an exploratory attempt to map the parameter regions of time series length, number of variables, penalization parameters, and kinds of models to define the boundaries where probabilistic models can be reasonably constructed from the analytics of observation data. In particular, we investigate empirically, by means of a linear autoregressive model with sparse inference structure, the true causality link retrieval performances in the region of short time series and large number of variables which is the most critical region—and the most interesting—in many practical cases. Causality is defined in information theoretic sense as a significant reduction on uncertainty over the present values of a given variable provided by the knowledge of the past values of another variable obtained in excess to the knowledge provided by the past of the variable itself and—in the conditional case—the past of all other variables [10]. We measure such information by using transfer entropy and, within the present linear modelling, this coincides with the concept of Granger causality and conditional Granger causality [11]. The use of transfer entropy has the advantage of being a concept directly extensible to nonlinear modelling. However, nonlinearity is not tackled within the present paper. Linear models with multivariate normal distributions have the unique advantage that causality and partial correlations are directly linked, largely simplifying the computation of transfer entropy, and directly mapping the problem into the sparse inverse covariance problem [3, 4].

Results are reported for artificially generated time series from an autoregressive model of $p = 100$ variables and time series lengths q between 10 and 20,000 data points. Robustness of the results has been verified over a wider range

of p from 20 to 200 variables. Our results demonstrate that sparse models are superior in retrieving the true causality structure for short time series. Interestingly, this is despite considerable inaccuracies in the inference network of these sparse models. We indeed observe that statistical validation of causality is crucial in identifying the true causal links, and this identification is highly enhanced in sparse models.

The paper is structured as follows. In Section 2 we briefly review the basic concepts of mutual information and conditional transfer entropy and their estimation from data that will then be used in the rest of the paper. We also introduce the concepts of sparse inverse covariance, inference network and causality networks. Section 3 concerns the retrieval of causality network from the computation and statistical validation of conditional transfer entropy. Results are reported in Section 4 where the retrieval of the true causality network from the analytics of time series from an autoregressive process of $p = 100$ variables is discussed. Conclusions and perspectives are given in Section 5.

2. Estimation of Conditional Transfer Entropy from Data

In this paper causality is quantified by means of statistically validated transfer entropy. Transfer entropy $T(\mathbf{Z_i} \rightarrow \mathbf{Z_j})$ quantifies the amount of uncertainty on a random variable, $\mathbf{Z_j}$, explained by the past of another variable, $\mathbf{Z_i}$, conditioned to the knowledge about the past of $\mathbf{Z_j}$ itself. Conditional transfer entropy, $T(\mathbf{Z_i} \rightarrow \mathbf{Z_j} \mid \mathbf{W})$, includes an extra condition also to a set variables $\mathbf{W}$. These quantities are introduced in detail in Appendix A (see also [11–13]). Let us here just report the main expression for the conditional transfer entropy that we shall use in this paper:

$$
\begin{aligned}
T\left(\mathbf{Z_i} \longrightarrow \mathbf{Z_j} \mid \mathbf{W}\right) &= H\left(\mathbf{Z_{j,t}} \mid \left\{\mathbf{Z_{j,t}^{lag}}, \mathbf{W}_t\right\}\right) \\
&\quad - H\left(\mathbf{Z_{j,t}} \mid \left\{\mathbf{Z_{i,t}^{lag}}, \mathbf{Z_{j,t}^{lag}}, \mathbf{W}_t\right\}\right),
\end{aligned}
\tag{1}
$$

where $H(\cdot \mid \cdot)$ is the conditional entropy and $\mathbf{Z_{j,t}}$ is a random variable at time t, whereas $\mathbf{Z_{i,t}^{lag}} = \{\mathbf{Z_{i,t-1}}, \ldots, \mathbf{Z_{i,t-\tau}}\}$ is the lagged set of random variable "$\mathbf{i}$" considering previous times $t - 1 \cdots t - \tau$ and $\mathbf{W}_t$ are all other variables and their lags (see Appendix A, (A.5)).

In this paper we use Shannon entropy and restrict to linear modelling with multivariate normal setting (see Appendix B). In this context the conditional transfer entropy can be expressed in terms of the determinants of conditional covariances $\det(\mathbf{\Sigma}(\cdot \mid \cdot))$ (see (B.5) in Appendix B):

$$
\begin{aligned}
&T\left(\mathbf{Z_i} \longrightarrow \mathbf{Z_j} \mid \mathbf{W}\right) \\
&= \frac{1}{2} \log \det\left(\mathbf{\Sigma}\left(\mathbf{Z_{j,t}} \mid \left\{\mathbf{Z_{j,t}^{lag}}, \mathbf{W}_t\right\}\right)\right) \\
&\quad - \frac{1}{2} \log \det\left(\mathbf{\Sigma}\left(\mathbf{Z_{j,t}} \mid \left\{\mathbf{Z_{j,t}^{lag}}, \mathbf{Z_{i,t}^{lag}}, \mathbf{W}_t\right\}\right)\right).
\end{aligned}
\tag{2}
$$

Conditional covariances can be conveniently computed in terms of the inverse covariance of the whole set of variables

$\mathbf{Z}_t = \{\mathbf{Z}_{k,t}, \mathbf{Z}_{k,t-1}, \ldots, \mathbf{Z}_{k,t-\tau}\}_{k=1}^{p} \in \mathbb{R}^{p \times (\tau+1)}$ (see Appendix C). Such inverse covariance matrix, $\mathbf{J}$, represents the structure of conditional dependencies among all couples of variables in the system and their lags. Each subpart of $\mathbf{J}$ is associated with the conditional covariances of the variables in that part with respect to all others. In terms of $\mathbf{J}$, the expression for the conditional transfer entropy becomes

$$
\begin{aligned}
T\left(\mathbf{Z_i} \longrightarrow \mathbf{Z_j} \mid \mathbf{W}\right) = -\frac{1}{2} \\
\cdot \log \det\left(\mathbf{J_{1,1}} - \mathbf{J_{1,2}}\left(\mathbf{J_{2,2}}\right)^{-1} \mathbf{J_{2,1}}\right) + \frac{1}{2} \log \det\left(\mathbf{J_{1,1}}\right),
\end{aligned}
\tag{3}
$$

where the indices "**1**" and "**2**" refer to submatrices of $\mathbf{J}$, respectively, associated with the variables $\mathbf{Z}_{j,t}$ and $\mathbf{Z}_{i,t}^{\text{lag}}$.

2.1. Causality and Inference Networks.

The inverse covariance $\mathbf{J}$, also known as precision matrix, represents the structure of conditional dependencies. If we interpret the structure of $\mathbf{J}$ as a network, where nodes are the variables and nonzero entries correspond to edges of the network, then we shall see that any two subsets of nodes that are not directly connected by one or more edges are conditionally independent. Condition is with respect to all other variables.

Links between variables at different lags are associated with causality with direction going from larger to smaller lags. The network becomes therefore a directed graph. In such a network entropies can be associated with nodes, conditional mutual information can be associated with edges between variables with the same lag, and conditional transfer entropy can be associated with edges between variables with different lags. A nice property of this mapping of information measures with directed networks is that there is a simple way to aggregate information which is directly associated with topological properties of the network. Entropy, mutual information, and transfer entropies can be defined for any aggregated subset of nodes with their values directly associated with the presence, direction, and weight of network edges between these subparts.

Nonzero transfer entropies indicating, for instance, variable $\mathbf{i}$ causing variable $\mathbf{j}$ are associated with some nonzero entries in the inverse covariance matrix $\mathbf{J}$ between lagged variables $\mathbf{i}$ (i.e., $\mathbf{Z}_{i,t-\tau}$, with $\tau > 0$) and variable $\mathbf{j}$ (i.e., $\mathbf{Z}_{j,t}$). In linear models these nonzero entries define the estimated *inference network*. However, not all edges in this inference network correspond to transfer entropies that are significantly different from zero. To extract the structure of the *causality network* we shall retain only the edges in the inference network which correspond to statistically validated transfer entropies.

Conditioning eliminates the effect of the other variables retaining only the exclusive contribution from the two variables in consideration. This should provide estimations of transfer entropy that are less affected by spurious effects from other variables. On the other hand, conditioning in itself can introduce spurious effects; indeed two independent variables can become dependent due to conditioning [13]. In this paper we explore two extreme conditioning cases: (i) conditioned to all other variables and their lags; (ii) unconditioned.

In principle, one would like to identify the maximal value of $T(\mathbf{Z_i} \rightarrow \mathbf{Z_j} \mid \mathbf{W})$ over all lags and all possible conditionings $\mathbf{W}$. However, the use of multiple lags and conditionings increases the dimensionality of the problem making estimation of transfer entropy very hard especially when only a limited amount of measurements is available (i.e., short time series). This is because the calculation of the conditional covariance requires the estimation of the inverse covariance of the whole set of variables and such an estimation is strongly affected by noise and uncertainty. Therefore, a standard approach is to reduce the number of variables and lags to keep dimensionality low and estimate conditional covariances with appropriate penalizers [3, 8, 9, 14]. An alternative approach is to invert the covariance matrix only locally on low dimensional subsets of variables selected by using information filtering networks [5–7] and then reconstruct the global inversion by means of the LoGo approach [4]. Let us here briefly account for these two approaches.

2.2. Penalized Inversions.

The estimate of the inverse covariance is a challenging task to which a large body of literature has been dedicated [14]. From an intuitive perspective, one can say that the problem lies in the fact that uncertainty is associated with nearly zero eigenvalues of the covariance matrix. Variations in these small eigenvalues have relatively small effects on the entries of the covariance matrix itself but have major effects on the estimation of its inverse. Indeed small fluctuations of small values can yield to unbounded contributions to the inverse. A way to cure such near-singular matrices is by adding finite positive terms to the diagonal which move the eigenvalues away from zero: $\widehat{\mathbf{J}} = ((1 - \gamma)\mathbf{S} + \gamma\mathbf{I}_N)^{-1}$, where $\mathbf{S} = \text{cov}(\mathbf{Z})$ is the covariance matrix of the set of variables $\mathbf{Z} \in \mathbb{R}^N$ estimated from data and $\mathbf{I}_N \in \mathbb{R}^{N \times N}$ is the identity matrix (where $N = p \times (\tau + 1)$; see later). This is what is performed in the so-called ridge regression [9], also known as shrinkage mean-square-error estimator [15] or Tikhonov regularization [8]. The effect of the additional positive diagonal elements is equivalent to compute the inverse covariance which maximizes the log-likelihood: $\log \det(\widehat{\mathbf{J}}) - \text{tr}(\mathbf{S}\widehat{\mathbf{J}}) - \gamma \| \widehat{\mathbf{J}} \|_2$, where the last term penalizes large off-diagonal coefficients in the inverse covariance with a ℓ_2 norm penalization [16]. The regularizer parameter γ tunes the strength of this penalization. This regularization is very simple and effective. However, with this method insignificant elements in the precision matrix are penalized toward small values but they are never set to zero. By using instead ℓ_1 norm penalization $\log \det(\widehat{\mathbf{J}}) - \text{tr}(\mathbf{S}\widehat{\mathbf{J}}) - \gamma \| \widehat{\mathbf{J}} \|_1$, insignificant elements are forced to zero leading to a sparse inverse covariance. This is the so-called lasso regularization [3, 14, 17]. The advantage of a sparse inverse covariance consists in the provision of a network representing a conditional dependency structure. Indeed, let us recall that in linear models zero entries in the inverse covariance are associated with couples of nonconditionally dependent variables.

2.3. Information Filtering Network Approach: LoGo.

An alternative approach to obtain sparse inverse covariance is by using information filtering networks generated by keeping

the elements that contribute most to the covariance by means of a greedy process. This approach, named LoGo, proceeds by first constructing a chordal information filtering graph such as a Maximum Spanning Tree (MST) [18, 19] or a Triangulated Maximally Filtered Graph (TMFG) [7]. These graphs are built by retaining edges that maximally contribute to a given gain function which, in this case, is the log-likelihood or—more simply—the sum of the squared correlation coefficients [5–7]. Then, this chordal structure is interpreted as the inference structure of the joint probability distribution function with nonzero conditional dependency only between variables that are directly connected by an edge. On this structure the sparse inverse covariance is computed in such a way to preserve the values of the correlation coefficients between couples of variables that are directly connected with an information filtering graph edge. The main advantage of this approach is that inversion is performed at local level on small subsets of variables and then the global inverse is reconstructed by joining the local parts through the information filtering network. Because of this Local-Global construction this method is named LoGo. It has been shown that LoGo method yields to statistically significant sparse precision matrices that outperform the ones with the same sparsity computed with lasso method [4].

3. Causality Network Retrieval

3.1. Simulated Multivariate Autoregressive Linear Process. In order to be able to test if causality measures can retrieve the true causality network in the underlying process, we generated artificial multivariate normal time series with known sparse causality structure by using the following autoregressive multivariate linear process [20]:

$$\mathbf{Z}_t = \sum_{\lambda=1}^{\tau} \mathbf{A}_\lambda \mathbf{Z}_{t-\lambda} + \mathbf{U}_t, \tag{4}$$

where $\mathbf{A}_\lambda \in \mathbb{R}^{p \times p}$ are matrices with random entries drawn from a normal distribution. The matrices are made upper diagonal (diagonal included) by putting to zero all lower diagonal coefficients and made sparse by keeping only a $\mathcal{O}(p)$ total number of entries different from zero in the upper and diagonal part. $\mathbf{U}_t \in \mathbb{R}^p$ are random normally distributed uncorrelated variables. This process produces autocorrelated, cross-correlated, and causally dependent time series. We chose it because it is among the simplest processes that can generate this kind of structured datasets. The dependency and causality structure is determined by the nonzero entries of the matrices $\mathbf{A}_\lambda$. The upper-triangular structure of these matrices simplifies the causality structure eliminating causality cycles. Their sparsity reduces dependency and causality interactions among variables. The process is made autoregressive and stationary by keeping the eigenvalues of $\mathbf{A}_\lambda$ all smaller than one in absolute value. For the tests we used $\tau = 5$, $p = 100$ and sparsity is enforced to have a number of links approximately equal to p. We reconstructed the network from time series of different lengths q between 5 and 20,000 points. To test statistical reliability the process was repeated 100 times with every time a different set of randomly generated matrices

$\mathbf{A}_\lambda$. We verify that the results are robust and consistent by varying sample sizes from $p = 20$ to 200, by changing sparsity with number of links from $0.5p$ to $5p$ and for τ from 1 to 10. We verified that the presence of isolated nodes or highly connected hub nodes does not affect results significantly.

3.2. Causality and Inference Network Retrieval. We tested the agreement between the causality structure of the underlying process and the one inferred from the analysis of p time series of different lengths q, $\mathbf{Z}_t \in \mathbb{R}^p$ with $t = 1 \cdots q$, generated by using (4) We have p different variables and τ lags. The dimensionality of the problem is therefore $N = p \times (\tau + 1)$ variables at all lags including zero.

To estimate the inference and causality networks we started by computing the inverse covariance, $\mathbf{J} \in \mathbb{R}^{N \times N}$, for all variables at all lags $\mathbf{Z} \in \mathbb{R}^{N \times q}$ by using the following three different estimation methods:

(1) ℓ_1 norm penalization (Glasso [14]);

(2) ℓ_2 norm penalization (ridge [8]);

(3) information filtering network (LoGo [4]).

We retrieved the inference network by looking at all couples of variables, with indices $\mathbf{i} \in [1, \ldots, p]$ and $\mathbf{j} \in [1, \ldots, p]$, which have nonzero entries in the inverse covariance matrix $\mathbf{J}$ between the lagged set of $\mathbf{j}$ and the nonlagged $\mathbf{i}$. Clearly, for the ridge method the result is a complete graph but for the Glasso and LoGo the results are sparse networks with edges corresponding to nonzero conditional transfer entropies between variables $\mathbf{i}$ and $\mathbf{j}$. For the LoGo calculation we make use of the regularizer parameter as a local shrinkage factor to improve the local inversion of the covariance of the 4-cliques and triangular separators (see [4]).

We then estimated transfer entropy between couples of variables, $\mathbf{i} \rightarrow \mathbf{j}$ conditioned to all other variables in the system. This is obtained by estimating of the inverse covariance matrix (indicated with an "hat" symbol) by using (C.7) (see Appendix C.2) with

$$
\begin{aligned}
\mathbf{Z_1} &= \mathbf{Z}_{\mathbf{j},t} \\
\mathbf{Z_2} &= \left\{ \mathbf{Z}_{\mathbf{i},t-1} \cdots \mathbf{Z}_{\mathbf{i},t-\tau} \right\} \\
\mathbf{Z_3} &= \left\{ \mathbf{Z}_{\mathbf{j},t-1} \cdots \mathbf{Z}_{\mathbf{j},t-\tau}, \mathbf{W} \right\},
\end{aligned} \tag{5}
$$

with $\mathbf{W}$ a conditioning to all variables $\mathbf{Z}$ except $\mathbf{Z_1}, \mathbf{Z_2}$, and $\{\mathbf{Z}_{\mathbf{j},t-1} \cdots \mathbf{Z}_{\mathbf{j},t-\tau}\}$. The result is a $p \times p$ matrix of conditional transfer entropies $T(\mathbf{Z}_{\mathbf{i},t} \rightarrow \mathbf{Z}_{\mathbf{j},t})$. Finally, to retrieve the causality network we retained the network of statistically validated conditional transfer entropies only. Statistical validation was performed as follows.

3.3. Statistical Validation of Causality. Statistical validation has been performed from likelihood ratio statistical test. Indeed, entropy and likelihood are intimately related: entropy measures uncertainty and likelihood measures the reduction in uncertainty provided by the model. Specifically, the Shannon entropy associated with a set of random variables, $\mathbf{Z_i}$, with probability distribution $p(\mathbf{Z_i})$ is $H(\mathbf{Z_i}) = -\mathbb{E}[\log p(\mathbf{Z_i})]$

(see (B.1)) whereas the log-likelihood for the model $\widehat{p}(\mathbf{Z_i})$ associated with a set of independent observations $\widehat{\mathbf{Z}}_{i,t}$ with $t = 1 \cdots q$ is $\log \mathscr{L}(\widehat{\mathbf{Z}}_{\mathbf{i}}) = \sum_{t=1}^{q} \log \widehat{p}(\widehat{\mathbf{Z}}_{i,t})$ which can be written as $\log \mathscr{L}(\widehat{\mathbf{Z}}_{\mathbf{i}}) = q\mathbb{E}_{\widehat{p}}[\log \widehat{p}(\mathbf{Z_i})]$. Note that q is the total available number of observations which, in practice, is the length of the time series minus the maximum number of lags. It is evident from these expressions that entropy and the log-likelihood are strictly related though this link might be nontrivial. In the case of linear modelling this connection is quite evident because the entropy estimate is $H = (1/2)(-\log|\widehat{\mathbf{J}}| + p\log(2\pi) + p)$ and the log-likelihood is $\log \mathscr{L} = (q/2)(\log|\widehat{\mathbf{J}}| - \mathrm{Tr}(\widehat{\mathbf{\Sigma}}\widehat{\mathbf{J}}) - p\log(2\pi))$. For the three models we study in this paper we have $\mathrm{Tr}(\widehat{\mathbf{\Sigma}}\widehat{\mathbf{J}}) = p$ and therefore the log-likelihood is equal to q times the opposite of the entropy estimate. Transfer entropy and conditional transfer entropy are differences between two entropies: the one of a set of variables conditioned to their own past minus the one conditioned also to the past of another variable. This, in turns, is the difference of the unitary log-likelihood of two models and therefore it is the logarithm of a likelihood ratio. As Wilks pointed out [21, 22] the null distribution of such model is asymptotically quite universal. Following the likelihood ratio formalism, we have $\lambda = qT$ and the probability of observing a transfer entropy larger than T, estimated under null hypothesis, is given by $p_v \sim 1 - \chi_c^2(rqT, d)$ with $r \simeq 2$ and χ_c^2 the chi-square the cumulative distribution function with d degrees of freedom which are the difference between the number of parameters in the two models. In our case the two models have, respectively, $\tau(p_j^2 + 1)$ and $\tau(p_j^2 + 1) + \tau(p_j p_i)$ parameters.

3.4. Statistical Validation of the Network.
The procedures described in the previous two subsections produce the inference network and causality network. Such networks are then compared with the known network of true causalities in the underlying process which is defined by the nonzero elements in the matrices A_λ (see (4)). The overlapping between the retrieved links in the inference or causality networks with the ones in the true network underlying the process is an indication of a discovery of a true causality relation. However some discoveries can be obtained just by chance or some methodologies might discover more links only because they produce denser networks. We therefore tested the hypothesis that the matching links in the retrieved networks are not obtained just by chances by computing the null-hypothesis probability to obtain the same or a larger number of matches randomly. Such probability is given by the conjugate cumulative hypergeometric distribution for a number equal or larger than TP of "true positive" matching causality links between an inferred network of n links and a process network of K true causality links, from a population of $p^2 - p$ possible links:

$$P(X \geq \mathrm{TP} \mid n, K, p) = 1 - \sum_{k=0}^{\mathrm{TP}-1} \frac{\binom{K}{k}\binom{p^2-p-K}{n-k}}{\binom{p^2-p}{n}}. \quad (6)$$

Small values of P indicate that the retrieved TP links out of K are unlikely to be found by randomly picking n edges from $p^2 - p$ possibilities. Note that in the confusion matrix notation [23] we have $n = \mathrm{TP} + \mathrm{FP}$ and $K = \mathrm{TP} + \mathrm{FN}$ with TP number of true positives, FP number of false positives, FN number of false negatives, and TN number of true negatives. The total number of "negatives" (unlinked couples of vertices) in the true model is instead $m = \mathrm{FP} + \mathrm{TN}$.

4. Results

4.1. Computation and Validation of Conditional Transfer Entropies.
By using (4) we generated 100 multivariate autoregressive processes with known causality structures. We here report results for $p = 100$ but analogous outcomes were observed for dimensionalities between $p = 20$ and 200 variables. Conditional transfer entropies between all couples of variables, conditioned to all other variables in the system, were computed by estimating the inverse covariances by using tree methodologies, ridge, Glasso, and LoGo and applying (3). Conditional transfer entropies were statistically validated with respect to null hypothesis (no causality) at $p_v = 1\%$ p value. Results for Bonferroni adjusted p value at 1% (i.e., $p_v = 0.01/(p^2 - p) \sim 10^{-6}$ for $p = 100$) are reported in Appendix E. We also tested other values of p_v from 10^{-8} to 0.1 obtaining consistent results. We observe that small p_v reduce the number of validated causality links but increase the chance that these links match with the true network in the process. Conversely large values of p_v increase the numbers of mismatched links but also of the true links discoveries. Let us note that here we use p_v as a thresholding criteria and we are not claiming any evidence of statistical significance of the causality. We assess the goodness of this choice a posteriori by comparing the resulting causality network with the known causality network of the process.

4.2. Statistical Significance of the Recovered Causality Network.
Results for the contour frontiers of significant causality links for the three models are reported in Figure 1 for a range of time series with lengths q between 10 and 20,000 and regularizer parameters γ between 10^{-8} and 0.5. Statistical significance is computed by using (6) and results are reported for both $P < 0.05$ and $P < 10^{-8}$ (continuous and dotted lines respectively). As one can see, the overall behaviours for the three methodologies are little affected by the threshold on P. We observe that LoGo significance region extends well beyond the Glasso and ridge regions.

The value of the regularizer parameter γ affects the results for the three models in a different way. Glasso has a region in the plane γ-p/q where it has best performances (in this case it appears to be around $\gamma \simeq 0.1$ and $p/q \simeq 2.5$). Ridge appears instead to be little affected with mostly constant performances across the range of γ. LoGo has best performances for small, even infinitesimal, values of γ. Indeed, different from Glasso in this case γ does not control sparsity but instead acts as local shrinkage parameter. Very small values can be useful in some particular cases to reduce the effect of noise but large values have only the effect to reduce information.

4.3. Causality Links Retrieval.
Once identified the parameter regions where the retrieved causality links are statistically
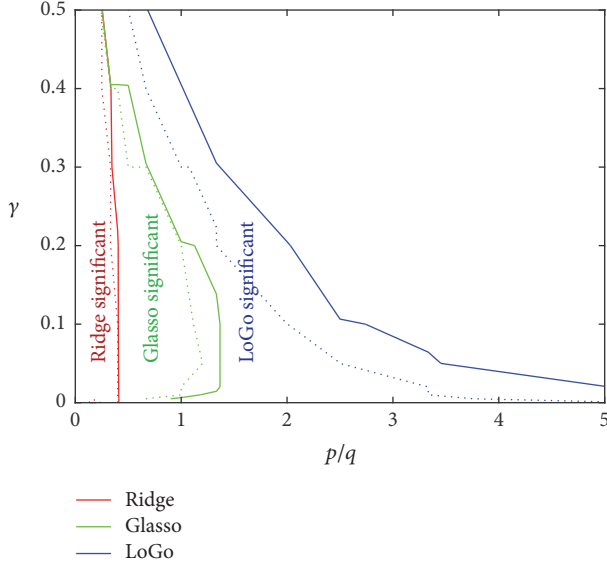
FIGURE 1: *Regions in the p/q-γ space where causality networks for the three models are statistically significant*. The significance regions are all at the left of the corresponding lines. Tick line reports the boundary $P < 0.05$ (see (6)) and dotted lines indicate $P < 10^{-8}$ significance levels (P is averaged over 100 processes). The plots refer to $p = 100$ and report the region where the causality networks are all significant for 100 processes.

significant, we also measured the fraction of true links retrieved. Indeed, given that the true underlying causality network is sparse, one could do significantly better than random by discovering only a few true positives. Instead, from any practical perspective we aim to discover a significant fraction of the edges. Figure 2 shows that the fraction of causality links correctly discovered (true positive, TP) with respect to the total number of causality links in the process (n) is indeed large reaching values above 50%. This is the so-called true positive rate or sensitivity, which takes values between 0 (no links discovered) and 1 (all links discovered). Reported values are averages over 100 processes. We observe that the region with discovering of 10% or more true causality links greatly overlaps with the statistical validity region of Figure 1.

We note that when the observation time becomes long, $p/q \lesssim 0.25$, ridge discovery rate becomes larger than LoGo. However, statistical significance is still inferior to LoGo, indeed the ridge network becomes dense when q increases and the larger discovery rate of true causality links is also accompanied by a larger rate of false links incorrectly identified (false positive FP).

The fraction of false positives with respect to the total number of causality links in the process (n) are reported in Table 1 together with the true positive rate for comparison. This number can reach values larger than one because the process is sparse and there are much more possibilities to randomly chose false links than true links. Note that this is not the false positive rate, which instead is FP/m, and cannot be larger than one. Consistent with Figure 1 we observe that, for short time series, up to $p/q \sim 0.5$, the sparse
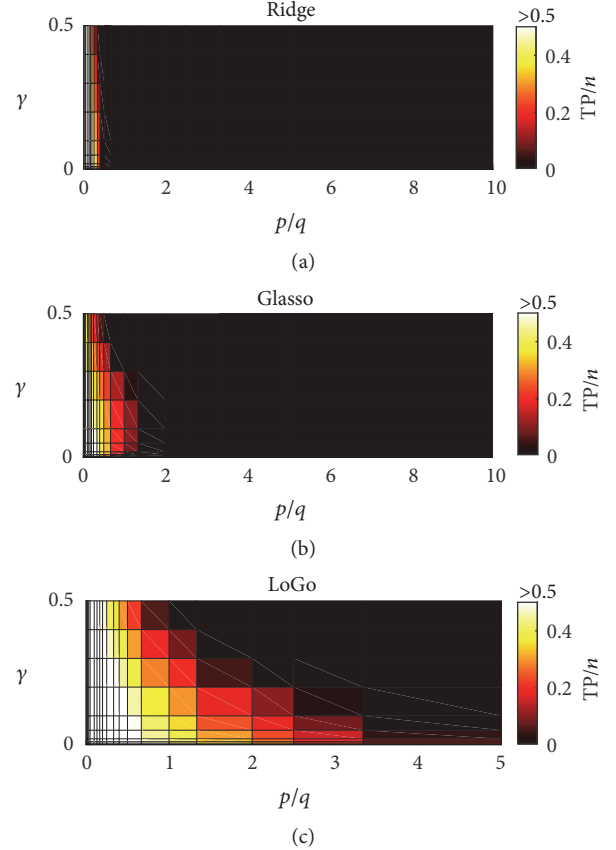
FIGURE 2: *True positive rate: fraction of retrieved true causality links (TP) with respect to the total number of links in the process (n)*. The three panels refer to ridges, Glasso, and LoGo ((a), (b), and (c)). Data are average fractions over 100 processes.

models have better capability to identify true causality links and to discard the false ones with LoGo being superior to Glasso. Remarkably, LoGo can identify a significant fraction of causality links already from time series with lengths of 30 data points only. p value significance, reported in the table with one or two stars, indicates when all values of $P(X \geq TP \mid n, K, p)$ from (6) for all 100 processes have, respectively, $P < 0.05$ or $P < 10^{-8}$. Again we observe that LoGo discovery rate region extends well beyond the Glasso and ridge regions.

4.4. Inference Network. We have so far empirically demonstrated that a significant part of the true causality network can be retrieved from the statistically validated network of conditional transfer entropies. Results depend on the choice of the threshold value of p_v at which null hypothesis is rejected. We observed that lower p_v are associated with network with fewer true positives but also fewer false positives and conversely larger p_v yield to causality networks with larger true positives but also larger false positives. Let us here report on the extreme case of the inference network which contains all causality channels with no validation. For the ridge model this network is the complete graph with all variables connected to each other. Instead, for Glasso and LoGo the inference network is sparse.

TABLE 1: *Causality network validation*. Comparison between fraction of true positive (TP/n) and fraction of false positive (FP/n), statistically validated, causality links for the three models, and different time-series lengths. The table reports only the case for the parameter $\gamma = 0.1$. Statistical validation of conditional transfer entropy is at $p_v = 1\%$ p value. Note that LoGo can perform better than reported in this table for smaller values of γ (see Figures 1 and 2).

q	10	20	30	50	200	300	1000	20000
Ridge TP/n	0.00	0.00	0.00	0.00	0.23**	0.49**	0.76**	0.93**
Ridge FP/n	0.00	0.00	0.00	0.00	0.00	0.10	0.65	1.06
Glasso TP/n	0.00	0.00	0.00	0.13**	0.48**	0.53**	0.62**	0.74**
Glasso FP/n	0.00	0.00	0.00	0.00	0.06	0.10	0.23	0.54
LoGo TP/n	0.00	0.08*	0.21**	0.37**	0.61**	0.65**	0.75**	0.90**
LoGo FP/n	0.00	0.00	0.00	0.01	0.06	0.08	0.15	0.34

*$P < 0.05$; **$P < 10^{-8}$.

TABLE 2: *Inference network validation: comparison between fraction of true positive (TP/n) and fraction of false positive (FP/n)*. Data for ridge are only for comparison because it is a complete graph with all links present. The table reports only the case for the parameter $\gamma = 0.1$.

q	10	20	30	50	200	300	1000	20000
Ridge TP/n	1.00	1.00	1.00	1.00	1.00	1.00	1.00	1.00
Ridge FP/n	97.84	97.84	97.84	97.84	97.84	97.84	97.84	97.84
Glasso TP/n	0.61*	0.74*	0.79*	0.85*	0.87**	0.84**	0.80**	0.80**
Glasso FP/n	28.39	38.11	45.79	53.58	40.61	26.60	1.54	0.92
LoGo TP/n	0.31*	0.50**	0.58**	0.63**	0.75**	0.78**	0.85**	0.93**
LoGo FP/n	4.53	4.27	4.18	4.03	3.72	3.63	3.44	3.21

*$P < 0.05$; **$P < 10^{-8}$.

Results are summarized in Table 2. In terms of true positive rate we first notice that they are all larger than the ones in Table 1. Indeed, the network of statistically validated conditional transfer entropies is a subnetwork of the inference network. On the other hand we notice that the false positive fraction is much larger than the ones in Table 2. Ridge network has a fraction of 1 because, in this case, the inference network is the complete graph.

Glasso also contains a very large number of false positives reaching even 55 times the number of links in the true network and getting to lower fractions only from long time series with $q > 1000$. These numbers also indicate that Glasso networks are not sparse. LoGo has a sparser and more significant inference network with smaller fractions of false positives which stay below $5n$, which is anyway a large number of misclassifications. Nonetheless, we observe that, despite such large fractions of FP, the discovered true positives are statistically significant.

4.5. Unconditioned Transfer Entropy Network. We last tested whether conditioning to the past of all other variables gives better causality network retrievals than the unconditioned case. Here, transfer entropy, $T(\mathbf{Z_i} \rightarrow \mathbf{Z_j})$, is computed by using (3) with $\mathbf{W} = \emptyset$, the empty set. For the ridge case this unconditional transfer entropy depends only from the time series, $\mathbf{Z}_{i,t}, \{\mathbf{Z}_{i,t-1}, \ldots, \mathbf{Z}_{i,t-\tau}\}$ and $\{\mathbf{Z}_{j,t-1}, \ldots, \mathbf{Z}_{j,t-\tau}\}$ (with $\tau = 5$ in this case). Glasso and LoGo cases are instead hybrid because a conditional dependency has been already introduced in the sparse structure of the inverse covariance $\mathbf{J}$ (the inference network). Results are reported in Table 3 where we observe that these networks retrieve a larger quantity of

true positives than the ones constructed from conditional entropy. However, the fraction of false positive is also larger than the ones in Table 1 although it is smaller than what observed in the inference network in Table 2. Overall, these results indicate that conditioning is effective in discarding false positives.

4.6. Summary of All Results in a Single ROC Plot. In summary, we have investigated the networks associated with conditional transfer entropy, unconditional transfer entropy, and inference for three models under a range of different parameters. In the previous subsections we have provided some comparisons between the performances of the three models in different ranges of parameters. Let us here provide a summary of all results within a single ROC plot [23]. Figure 3 reports the ROC values, for each model and each parameter combination, x-axis is false positive rates (FP/m), and y-axis is true positive rates (TP/n). Each point is an average over 100 processes. Points above the diagonal line are associated with relatively well performing models with the upper left corner representing the point where models correctly discover all true causality links without any false positive. The plot reports with large symbols the cases for $\gamma = 0.1$ and validation at p value $p_v = 0.01$, which can be compared with the data reported in the tables. We note that, by construction, LoGo models are sparse (with a number of edges $\sim 3p$ [4]). This restrains the ROC results to the left-hand side of the plot. For this reason an expanded view of the figure is also proposed with the x-axis scaled. Note that this ROC curve is provided as a visual tool for intuitive comparison between models.

TABLE 3: *Unconditioned transfer entropy network: comparison between fraction of true positive (TP/n) and fraction of false positive (FP/n). Statistical validation of transfer entropy is at $p_v = 1\%$ p value. The table reports only the case for the parameter $\gamma = 0.1$.*

q	10	20	30	50	200	300	1000	20000
Ridge TP/n	0.02	0.39**	0.45**	0.51**	0.65**	0.69**	0.78**	0.92**
Ridge FP/n	0.07	1.06	0.95	0.85	0.93	0.99	1.20	1.73
Glasso TP/n	0.00	0.24**	0.35**	0.43**	0.57**	0.60**	0.67**	0.77**
Glasso FP/n	0.00	0.10	0.20	0.29	0.51	0.56	0.73	1.66
LoGo TP/n	0.11	0.34**	0.41**	0.47**	0.63**	0.66**	0.76**	0.89**
LoGo FP/n	0.02	0.16	0.25	0.34	0.59	0.66	0.87	1.49

**$P < 10^{-8}$.
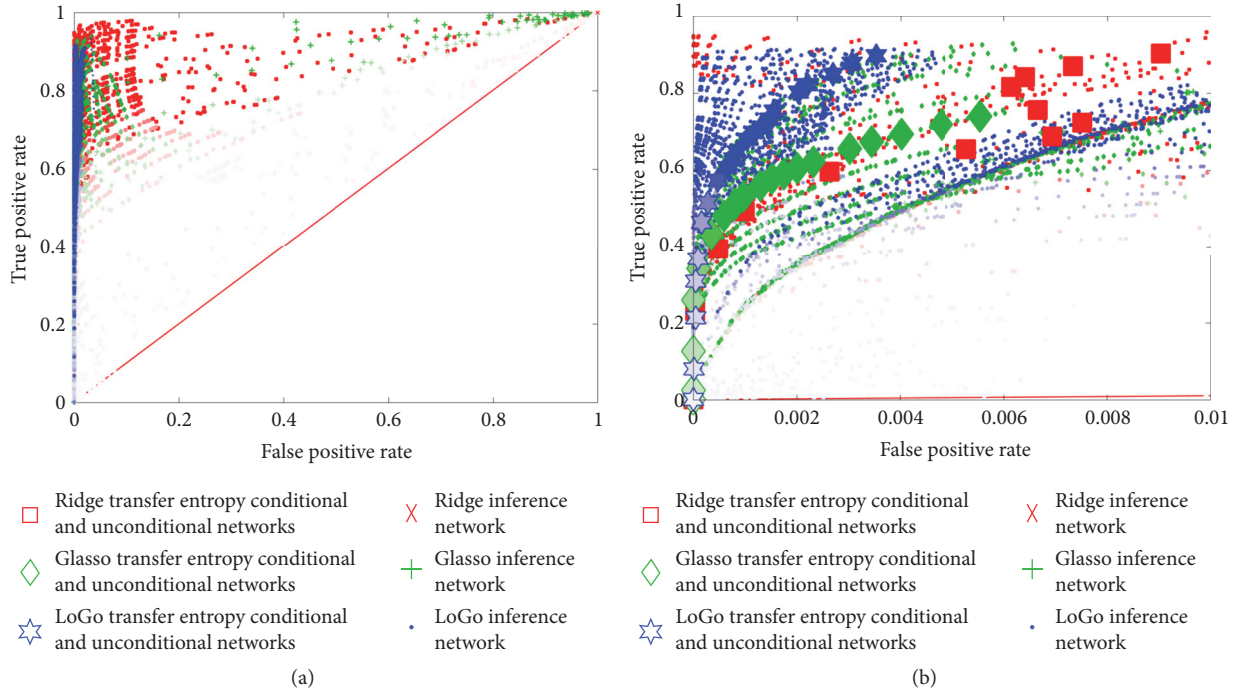
FIGURE 3: *ROC values, for each model and each parameter combination. x-axis false positive rates (FP/m); y-axis true positive rates (TP/n). (a) and (b) are the same with x-axis expanded on the low values only for (b) to better visualise the differences between the various models. Large symbols refer to $\gamma = 0.1$ and validation at p value $p_v = 0.01$. Color intensity is proportional to time series length. Inference network results are all outside the range of the plot. Reported values are averages over 100 processes.*

Overall from Tables 1, 2, and 3 and Figure 3 we conclude that all models obtain better results for longer time series and that conditional transfer entropy overperforms the unconditional counterparts (see, Tables 1 and 3 and the two separated ROC figures for conditional and unconditional transfer entropies reported in Figure 5 in Appendix D). In the range of short time series, when $q \leq p$, which is of interest for this paper, LoGo is the best performing model with better performances achieved for small $\gamma \lesssim 10^{-4}$ and validation with small p values $p_v \lesssim 10^{-4}$. LoGo is consistently the best performing model also for longer time series up to lengths of $q \sim 1000$. Instead, above $q = 2000$ ridge begins to provide better results. For long time series, at $q = 20,000$, the best performing model is ridge with parameters $\gamma = 10^{-5}$, p value $p_v = 5 \ 10^{-6}$. LoGo is also performing well when time series are long with best performance obtained at $q = 20,000$ for

parameters $\gamma = 10^{-10}$, p value $p_v = 5 \ 10^{-6}$. We note that LoGo instead performs poorly in the region of parameters with $\gamma \leq 0.1$ and $p_v \leq 0.01$ for short time series $q \leq p/2$.

5. Conclusions and Perspectives

In this paper we have undertaken the challenging task to explore models and parameter regions where the analytics of time series can retrieve significant fractions of true causality links from linear multivariate autoregressive process with known causality structure. Results demonstrate that sparse models with conditional transfer entropy are the ones who achieve best results with significant causality link retrievals already for very short time series even with $q \leq p/5 = 20$. This region is very critical and general considerations would suggest that no solutions can be discovered. Indeed, this result

is in apparent contradiction with a general analytical results in [24, 25] who find that no significant solutions should be retrieved for $q \leq N/2 = 150$. However, we notice that the problem we are addressing here is different from the one in [24, 25]. In this paper we have been considering an underlying sparse true causality structure and such a sparsity changes considerably the condition of the problem yielding to significant solutions even well below the theoretical limit from [24, 25] which is instead associated with nonsparse models.

Unexpectedly, we observed that the structure of the inference networks in the two sparse models, Glasso and LoGo, has excessive numbers of false positives yielding to rather poor performances. However, in these models false positive can be efficiently filtered out by imposing statistical significance of the transfer entropies.

Results are affected by the choice of the parameters and the fact that the models depend on various parameters (q, p, γ, p_v, P) make the navigation in this space quite complex. We observed that the choice of p values, p_v, for valid transfer entropies affects results. Within our setting we obtained best results with the smaller p values especially in the regions of short time series. We note that the regularizer parameter γ also plays an important role and best performances are obtained by combination of the two parameters γ and p_v. Not surprisingly, longer time series yield to better results. We observe that conditioning to all other variables or unconditioning is affecting the transfer entropy estimation with better performing causality network retrieval obtained for conditioned transfer entropies. However, qualitatively, results are comparable. Other intermediate cases, such as conditioning to past of all other variables only, have been explored again with qualitatively comparable results. It must be said that in the present system results are expected to be robust to different conditionings because the underlying network of the investigated processes is sparse. For denser inference structures, conditioning could affect more the results.

Consistently with the findings in [4] we find that LoGo outperforms the other methods. This is encouraging because the present settings of LoGo is using a simple class of information filtering networks, namely, the TMFG [7], obtained by retaining largest correlations. There are a number of alternative information filtering networks which should be explored. In particular, given the importance of statistical validation emerged from the present work, it would be interesting to explore statistical validation within the process of construction of the information filtering networks themselves.

In this paper we investigate a simple case with a linear autoregressive multivariate normal process analysed by means of linear models. Both LoGo and Glasso can be extended to the nonlinear case with LoGo being particularly suitable for nonparametric approaches as well [4].

There are alternative methods to extract causality networks from short time series, in particular Multispatial CCM [26, 27] appears to perform well for short time series. A comparison between different approaches and the application of these methods to real data will be extremely interesting. However this should be the object of future works.

Appendix

A. Conditional Transfer Entropy

Let us here briefly review two of the most commonly used information theoretic quantities that we use in this paper, namely, mutual information (quantifying dependency) and transfer entropy (quantifying causality) for the multivariate case [11–13].

A.1. Mutual Information. Let us first start from the simplest case of two random variables, $X \in \mathbb{R}^1$ and $Y \in \mathbb{R}^1$, where dependence can be quantified by the amount of shared information between the two variables, which is called mutual information: $I(X; Y) = H(X) + H(Y) - H(X, Y)$, where $H(X)$ is the entropy of variable X, $H(Y)$ is the entropy of variable Y, and $H(X, Y)$ is the joint entropy of variables X and Y [13]. Extending to the multivariate case, the shared information between a set of n random variables $\mathbf{X} = (X_1, \ldots, X_n)^T \in \mathbb{R}^n$ and another set of m random variables $\mathbf{Y} = (Y_1, \ldots, Y_m)^T \in \mathbb{R}^m$ is

$$I(\mathbf{X}; \mathbf{Y}) = H(\mathbf{X}) + H(\mathbf{Y}) - H(\mathbf{X}, \mathbf{Y}) \tag{A.1}$$

with $H(\mathbf{X})$, $H(\mathbf{Y})$ being the entropies, respectively, for the set of variables $\mathbf{X}$ and $\mathbf{Y}$ and $H(\mathbf{X}, \mathbf{Y})$ being their joint entropy. It must be stressed that this quantity is the mutual information between two sets of multivariate variables and it is not the multivariate mutual information between all variables $\{\mathbf{X}, \mathbf{Y}\}$ which instead measures the intersection of information between all variables. Mutual information in (A.1) can also be written as

$$I(\mathbf{X}; \mathbf{Y}) = H(\mathbf{Y}) - H(\mathbf{Y} \mid \mathbf{X}) = H(\mathbf{X}) - H(\mathbf{X} \mid \mathbf{Y}) \tag{A.2}$$

which makes use of the conditional entropy of $\mathbf{Y}$ given $\mathbf{X}$: $H(\mathbf{Y} \mid \mathbf{X}) = H(\mathbf{Y}, \mathbf{X}) - H(\mathbf{X}) = \mathbb{E}(H(\mathbf{Y}) \mid \mathbf{X})$.

Conditioning to a third set of variables $\mathbf{W}$ can also be applied to mutual information itself and its expression is a direct extension of (A.1) and it is called conditional mutual information:

$$I(\mathbf{X}; \mathbf{Y} \mid \mathbf{W}) = H(\mathbf{X} \mid \mathbf{W}) + H(\mathbf{Y} \mid \mathbf{W}) \\ - H(\mathbf{X}, \mathbf{Y} \mid \mathbf{W}). \tag{A.3}$$

Equations (A.1) and (A.3) coincide in the case of an empty set $\mathbf{W} = \emptyset$. Mutual information and conditional mutual information are symmetric measures with $I(\mathbf{X}; \mathbf{Y} \mid \mathbf{W}) = I(\mathbf{Y}; \mathbf{X} \mid \mathbf{W})$ always. Let us note that symmetry is unavoidable for information measures that quantify the simultaneous effect of a set of variables onto another. Indeed, in a simultaneous interaction cause and effect cannot be distinguished from the exchange of information and direction cannot be established. To quantify causality one must investigate the transmission of information not only between two sets of variables but also through time.

A.2. Conditional Transfer Entropy. Causality between two random variables, $X \in \mathbb{R}^1$ and $Y \in \mathbb{R}^1$, can be quantified by means of the so-called transfer entropy which quantifies

the amount of uncertainty on Y explained by the past of X given the past of Y. Let us consider a series of observations and denote with X_t being the random variable X at time t and with $X_{t-\tau}$ being the random variable at a previous time, τ lags before t. Using this notation, we can define transfer entropy from variable X to variable Y in terms of the following conditional mutual information: $T(X \rightarrow Y) = I(Y_t; X_{t-\tau} \mid Y_{t-\tau})$ [11, 13].

For the multivariate case, given two sets of random variables $\mathbf{X} \in \mathbb{R}^n$ and $\mathbf{Y} \in \mathbb{R}^m$, the transfer entropy is the conditional mutual information between the set of variables $\mathbf{Y}_t$ at time t and the past of the other set of variables, $\mathbf{X}_{t-\tau}$ conditioned to the past of the first variable $\mathbf{Y}_{t-\tau}$. That is, $T(\mathbf{X} \rightarrow \mathbf{Y}) = I(\mathbf{Y}_t; \mathbf{X}_{t-\tau} \mid \mathbf{Y}_{t-\tau})$ [13]. In general, the influence from the past can come from more than one lag and we can therefore extend the definition including different sets of lags for the two variables: $\tau_1, \ldots, \tau_k, \lambda_1, \ldots, \lambda_h$:

$$
\begin{aligned}
T(\mathbf{X} &\longrightarrow \mathbf{Y}) \\
&= I\left(\mathbf{Y}_t \left\{\mathbf{Y}_{t-\lambda_1} \cdots \mathbf{Y}_{t-\lambda_h}\right\}; \left\{\mathbf{X}_{t-\tau_1} \cdots \mathbf{X}_{t-\tau_k}\right\}\right) \\
&= H\left(\mathbf{Y}_t \left\{\mathbf{Y}_{t-\lambda_1} \cdots \mathbf{Y}_{t-\lambda_h}\right\}\right) \\
&\quad - H\left(\mathbf{Y}_t \left\{\mathbf{X}_{t-\tau_1} \cdots \mathbf{X}_{t-\tau_k}, \mathbf{Y}_{t-\lambda_1} \cdots \mathbf{Y}_{t-\lambda_h}\right\}\right);
\end{aligned}
\tag{A.4}
$$

a further generalization, which we use in this paper, includes conditioning to any other set of variables $\{\mathbf{W}_{t-\theta_1} \cdots \mathbf{W}_{t-\theta_g}\}$ lagged at $\theta_1, \ldots, \theta_g$:

$$
\begin{aligned}
T(\mathbf{X} \longrightarrow \mathbf{Y} \mid \mathbf{W}) &= I\left(\mathbf{Y}_t; \left\{\mathbf{X}_{t-\tau_1} \cdots \mathbf{X}_{t-\tau_k}\right\} \mid \right. \\
&\quad \left. \left\{\mathbf{Y}_{t-\lambda_1} \cdots \mathbf{Y}_{t-\lambda_h}, \mathbf{W}_{t-\theta_1} \cdots \mathbf{W}_{t-\theta_g}\right\}\right).
\end{aligned}
\tag{A.5}
$$

In this paper we simplify notation using $\mathbf{X}_t^{\text{lag}} = \{\mathbf{X}_{t-\tau_1} \cdots \mathbf{X}_{t-\tau_k}\}$, $\mathbf{Y}_t^{\text{lag}} = \{\mathbf{Y}_{t-\lambda_1} \cdots \mathbf{Y}_{t-\lambda_h}\}$, and $\mathbf{W}_t = \{\mathbf{W}_{t-\theta_1} \cdots \mathbf{W}_{t-\theta_g}\}$.

In the literature, there are several examples that use adaptations of (1) to compute causality and dependency measures [28]. A notable example is the directed information, introduced by Massey in [29], where τ spans all lags in a range between 0 and $s - 1$ and λ spans the lags from 1 to $s - 1$. The directed information is then defined as the sum over transfer entropies from $s = 1$ to present:

$$
\begin{aligned}
I\left(\{\mathbf{X}\}_1^t \longrightarrow \{\mathbf{Y}\}_1^t \mid \mathbf{W}\right) &= \sum_{s=1}^t I\left(\mathbf{Y}_s; \{\mathbf{X}\}_1^s \mid \{\mathbf{Y}\}_1^{s-1}, \mathbf{W}\right),
\end{aligned}
\tag{A.6}
$$

where we adopted the notations $\{\mathbf{X}\}_1^t = \{\mathbf{X}_1 \cdots \mathbf{X}_t\}$ and $\{\mathbf{Y}\}_1^t = \{\mathbf{Y}_1 \cdots \mathbf{Y}_t\}$. Interestingly, this definition includes the conditional synchronous mutual information contributions

between $\mathbf{X}_s$ and $\mathbf{Y}_s$. Following Kramer et al. [30, 31] we observe that for stationary processes

$$
\begin{aligned}
\lim_{t \to \infty} \frac{1}{t} I\left(\{\mathbf{X}\}_1^t \longrightarrow \{\mathbf{Y}\}_1^t\right) &= \lim_{t \to \infty} I\left(\{\mathbf{X}\}_1^t; \mathbf{Y}_t \mid \{\mathbf{Y}\}_1^{t-1}\right) \\
&= T\left(\{\mathbf{X}\}_1^{t-1} \longrightarrow \mathbf{Y}_t\right) + I\left(\{\mathbf{X}\}_1^t; \{\mathbf{Y}\}_1^t \mid \{\mathbf{X}\}_1^{t-1}\right),
\end{aligned}
\tag{A.7}
$$

with $T(\{\mathbf{X}\}_1^{t-1} \to \mathbf{Y}_t) = I(\mathbf{Y}_t; \{\mathbf{X}_1 \cdots \mathbf{X}_{t-1}\} \mid \{\mathbf{Y}_1 \cdots \mathbf{Y}_{t-1}\})$. This identity supports the intuition that the directed information accounts for the transfer entropy plus an instantaneous term.

B. Shannon-Gibbs Entropy

The general expression for the transfer entropy reported di in Section A, (1), is independent of the kind of entropy definition. In this paper we use Shannon entropy, which is defined as

$$
H(\mathbf{X}) = -\mathbb{E}\left[\log p(\mathbf{X})\right],
\tag{B.1}
$$

$$
H(\mathbf{Y}) = -\mathbb{E}\left[\log p(\mathbf{Y})\right],
\tag{B.2}
$$

where $p(\mathbf{X})$ and $p(\mathbf{Y})$ are the probability distribution function for the set of random variables $\mathbf{X}$ and $\mathbf{Y}$. Similarly, the joint Shannon entropy for the variables $\mathbf{X}$ and $\mathbf{Y}$ is defined as

$$
H(\mathbf{X}, \mathbf{Y}) = -\mathbb{E}\left[\log p(\mathbf{X}, \mathbf{Y})\right]
\tag{B.3}
$$

with $p(\mathbf{X}, \mathbf{Y})$ being the joint probability distribution function of $\mathbf{X}$ and $\mathbf{Y}$. This is the most common definition of entropy. It is a particularly meaningful and suitable entropy for linear modelling, as we focus in the paper.

B.1. Multivariate Normal Modelling. For multivariate normal variables the Shannon-Gibbs entropy is

$$
H(\mathbf{X}) = \frac{1}{2} \log\left(\det \mathbf{\Sigma}(\mathbf{X})\right) + \frac{n}{2} \log(2\pi e)
\tag{B.4}
$$

and its conditional counterpart is

$$
H(\mathbf{X} \mid \mathbf{W}) = \frac{1}{2} \log\left(\det \mathbf{\Sigma}(\mathbf{X} \mid \mathbf{W})\right) + \frac{n}{2} \log(2\pi e)
\tag{B.5}
$$

with $\mathbf{\Sigma}$ being the covariance matrix and $\det(\cdot)$ being the matrix determinant. In the paper we use these expressions to compute mutual information and conditional transfer entropy.

C. Computing Conditional Covariances for Subsets of Variables from the Inverse Covariance

Let us consider three sets of variables $\mathbf{Z}_1 \in \mathbb{R}^{p_1}$, $\mathbf{Z}_2 \in \mathbb{R}^{p_2}$, and $\mathbf{Z}_3 \in \mathbb{R}^{p_3}$ and the associated inverse covariance $\mathbf{J} \in \mathbb{R}^{(p_1+p_2+p_3) \times (p_1+p_2+p_3)}$ for $\{\mathbf{Z}_1, \mathbf{Z}_2, \mathbf{Z}_3\} \in \mathbb{R}^{(p_1+p_2+p_3)}$. The conditional covariance of $\mathbf{Z}_1$ given $\mathbf{Z}_2$ and $\mathbf{Z}_3$ is the inverse of the $p_1 \times p_1$ upper left part of $\mathbf{J}$ with indices in $V_1 = (1, \ldots, p_1)$ (see Figure 4):

$$
\mathbf{\Sigma}(\mathbf{Z}_1 \mid \mathbf{Z}_2, \mathbf{Z}_3) = (\mathbf{J}_{1,1})^{-1}.
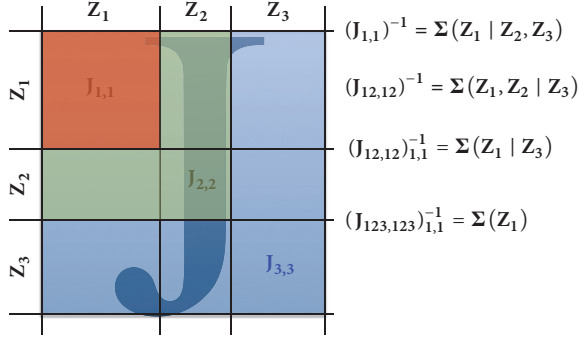\tag{C.1}
$$

FIGURE 4: The inverse of parts the inverse covariance $\mathbf{J}$ gives the covariance of the variables corresponding to that part conditioned to the other variables.

Instead, the conditional covariance of $\mathbf{Z}_3$ given $\mathbf{Z}_3$ is obtained by inverting the larger upper left part $\mathbf{J}_{12,12}$ with both indices in $\{V_1, V_2\}$ with $V_2 = (p_1 + 1, \ldots, p_1 + p_2)$, and then taking the inverse of the part with indices in V_1 which, using the Schur complement [13], is

$$\Sigma\left(\mathbf{Z}_1 \mid \mathbf{Z}_3\right) = \left(\mathbf{J}_{1,1} - \mathbf{J}_{1,2}\left(\mathbf{J}_{2,2}\right)^{-1}\mathbf{J}_{2,1}\right)^{-1}. \tag{C.2}$$

Figure 4 schematically illustrates these inversions and their relations with conditional covariances. Let us note that these conditional covariances can also be expressed directly in terms of subcovariances by using again the Schur complement:

$$\Sigma\left(\mathbf{Z}_1\mathbf{Z}_2, \mathbf{Z}_3\right) = \Sigma_{1,1} - \Sigma_{1,23}\left(\Sigma_{23,23}\right)^{-1}\Sigma_{23,1},$$
$$\Sigma\left(\mathbf{Z}_1 \mid \mathbf{Z}_3\right) = \Sigma_{1,1} - \Sigma_{1,3}\left(\Sigma_{3,3}\right)^{-1}\Sigma_{3,1}. \tag{C.3}$$

However, when p_3 (cardinality of V_3) is much larger than p_1 and p_2 (cardinalities of V_1 and V_2) then the equivalent expressions, (C.1) and (C.2), that use the inverse covariance involve matrices with much smaller dimensions. This can become computationally crucial when very large dimensionalities are involved. Furthermore, if the inverse covariance $\mathbf{J}$ is estimated by using a sparse modelling tool such as Glasso or LoGo [4, 14] (as we do in this paper), then computations in expressions (C.1) and (C.2) have to handle only a few nonzero elements providing great computational advantages over (C.3).

In the paper we make use of (C.1)-(C.2) to compute mutual information and conditional transfer entropy for the system of all variables and their lagged versions.

C.1. Mutual Information. Let us consider the mutual information between any two subsets $\mathbf{X} \in \mathbb{R}^n$ and $\mathbf{Y} \in \mathbb{R}^m$ of variables conditioned to all other variables, which we shall call $\mathbf{W} \in \mathbb{R}^{p-n-m}$. For these three sets of variables $\{\mathbf{X}, \mathbf{Y}, \mathbf{W}\} \in \mathbb{R}^p$ the conditional mutual information, $I(\mathbf{X}, \mathbf{Y}, \mathbf{W}) = H(\mathbf{X}, \mathbf{Y} \mid \mathbf{W}) - H(\mathbf{X} \mid \mathbf{Y}, \mathbf{W})$ (see (A.3)), can be expressed in terms of the conditional covariances by using (B.5):

$$I\left(\mathbf{X}; \mathbf{Y} \mid \mathbf{W}\right) = \frac{1}{2}\log\det\Sigma\left(\mathbf{X} \mid \mathbf{W}\right)$$
$$-\frac{1}{2}\log\det\Sigma\left(\mathbf{X} \mid \mathbf{Y}, \mathbf{W}\right). \tag{C.4}$$

Given the inverse covariance $\mathbf{J} \in \mathbb{R}^{p \times p}$, by using (C.1) and (C.2) and substituting

$$\mathbf{Z}_1 = \mathbf{X},$$
$$\mathbf{Z}_2 = \mathbf{Y}, \tag{C.5}$$
$$\mathbf{Z}_3 = \mathbf{W},$$

we can express the conditional mutual information, (C.4), directly in terms of the parts of $\mathbf{J}$:

$$I\left(\mathbf{X}; \mathbf{Y} \mid \mathbf{W}\right) = -\frac{1}{2}\log\det\left(\mathbf{J}_{1,1} - \mathbf{J}_{1,2}\left(\mathbf{J}_{2,2}\right)^{-1}\mathbf{J}_{2,1}\right)$$
$$+ \frac{1}{2}\log\left(\det\mathbf{J}_{1,1}\right). \tag{C.6}$$

Note that although this is not directly evident, (C.6) is symmetric by exchanging **1** and **2** (i.e., $\mathbf{X}$ and $\mathbf{Y}$).

C.2. Conditional Transfer Entropy. Conditional transfer entropy (see (1)) is conditional mutual information between lagged sets of variables and therefore it can be computed directly from (C.6). In this case we shall name

$$\mathbf{Z}_1 = \mathbf{Y}_t,$$
$$\mathbf{Z}_2 = \left\{\mathbf{X}_{t-\tau_1}\cdots\mathbf{X}_{t-\tau_k}\right\},$$
$$\mathbf{Z}_3 = \left\{\mathbf{Y}_{t-\lambda_1}\cdots\mathbf{Y}_{t-\lambda_h}, \mathbf{W}_{t-\theta_1}\cdots\mathbf{W}_{t-\theta_g}\right\},$$
$$T\left(\mathbf{X} \longrightarrow \mathbf{Y} \mid \mathbf{W}\right) \tag{C.7}$$
$$= -\frac{1}{2}\log\det\left(\mathbf{J}_{1,1} - \mathbf{J}_{1,2}\left(\mathbf{J}_{2,2}\right)^{-1}\mathbf{J}_{2,1}\right)$$
$$+ \frac{1}{2}\log\det\left(\mathbf{J}_{1,1}\right)$$

obtaining an expression which is formally identical to (C.6) but with indices **1** and **2** referring to the above sets of variables instead.

Note that index **3** does not appear in this expression. Information from variables **3** ($\mathbf{W}$) has been used to compute $\mathbf{J}$ but then only the subparts **1** and **2** are required to compute the conditional transfer entropy. The fact that these expressions for conditional mutual information and conditional transfer entropy involve only local parts (**1** and **2**) of the inverse covariance can become extremely useful when high-dimensional datasets are involved.

D. Comparison between Conditional and Unconditional Transfer Entropies

The two ROC plots for conditional and unconditional transfer entropies are displayed in Figure 5. Form the comparison it is evident that, for the process studied in this paper, conditional transfer entropy provides best results. This is in line with what observed in Tables 1, 3, 4, and 5.
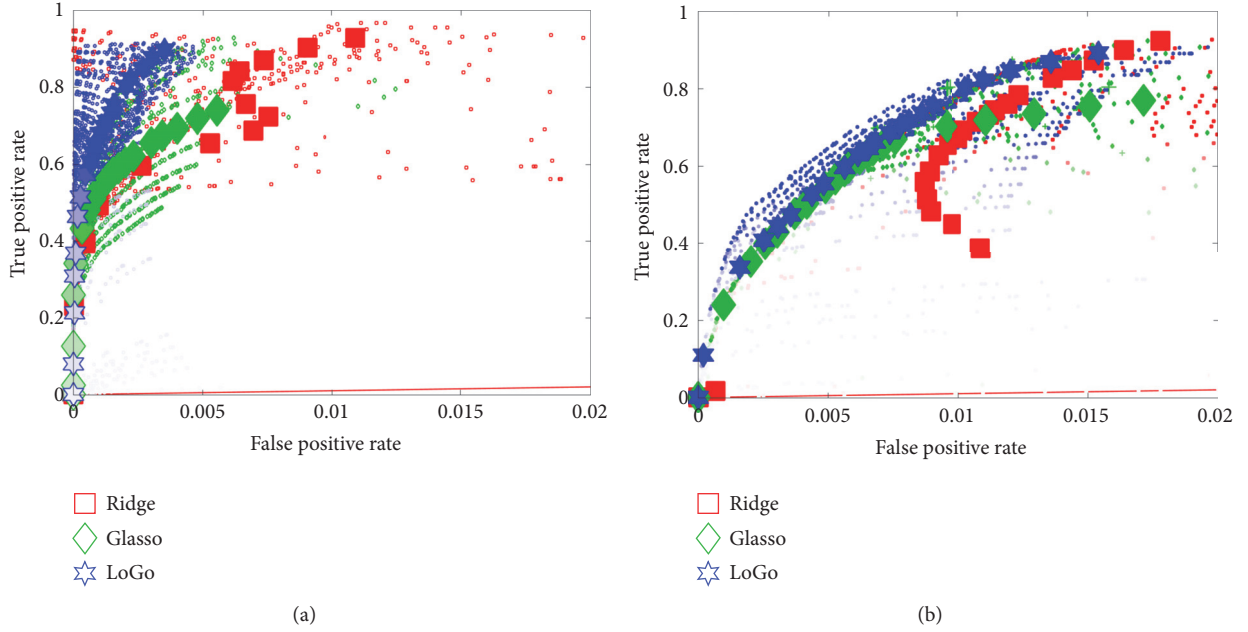
FIGURE 5: *ROC values, for conditional (a) and unconditional (b) transfer entropies. x-axis false positive rates (FP/m); y-axis true positive rates (TP/n). Large symbols refer to γ = 0.1 and validation at p value p_v = 0.01. Color intensity is proportional to time series length. Inference network results are all outside the range of the plot. Reported values are averages over 100 processes.*

TABLE 4: *Causality network validation with conditional transfer entropy validation at 1% Bonberroni adjusted p value. Fraction of true positive (TP/n) and fraction of false positive (FP/n), statistically validated, causality links for the three models, and different time series lengths. The table reports only the case for the parameter γ = 0.1.*

q	10	20	30	50	200	300	1000	20000
Ridge TP/n	0.00	0.00	0.00	0.00	0.00	0.30**	0.67**	0.89**
Ridge FP/n	0.00	0.00	0.00	0.00	0.00	0.01	0.18	0.75
Glasso TP/n	0.00	0.00	0.00	0.00	0.35**	0.43**	0.57**	0.71**
Glasso FP/n	0.00	0.00	0.00	0.00	0.01	0.03	0.13	0.45
LoGo TP/n	0.00	0.00	0.02	0.17**	0.50**	0.56**	0.69**	0.87**
LoGo FP/n	0.00	0.00	0.00	0.00	0.01	0.03	0.09	0.28

**$P < 10^{-8}$.

TABLE 5: *Causality network validation with unconditional transfer entropy validation at 1% Bonberroni adjusted p value. Fraction of true positive (TP/n) and fraction of false positive (FP/n), statistically validated, causality links for the three models, and different time-series lengths. The table reports only the case for the parameter γ = 0.1.*

q	10	20	30	50	200	300	1000	20000
Ridge TP/n	0.00	0.00	0.22**	0.36**	0.55**	0.59**	0.70**	0.88**
Ridge FP/n	0.00	0.00	0.09	0.21	0.47	0.55	0.77	1.32
Glasso TP/n	0.00	0.00	0.00	0.27**	0.48**	0.53**	0.62**	0.75**
Glasso FP/n	0.00	0.00	0.00	0.11	0.37	0.43	0.61	1.41
LoGo TP/n	0.00	0.00	0.22**	0.35**	0.53**	0.58**	0.69**	0.86**
LoGo FP/n	0.00	0.00	0.05	0.16	0.42	0.49	0.71	1.27

**$P < 10^{-8}$.

E. Causality Network Results for Transfer Entropy Validation with 1% Bonferroni Adjusted p Values

In Tables 4 and 5, true positive rates (TP/n) and fraction of false positives (FP/m) statistically validated and causality links with validation at 1% Bonberroni adjusted p value (i.e., $p_v \leq 10^{-6}$) are reported. These tables must be compared with Tables 1 and 3, in the main text where causality links are validated at $p_v = 1\%$ nonadjusted p value.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] A. M. Bruckstein, D. L. Donoho, and M. Elad, "From sparse solutions of systems of equations to sparse modeling of signals and images," *SIAM Review*, vol. 51, no. 1, pp. 34–81, 2009.

[2] S. Theodoridis, Y. Kopsinis, and K. Slavakis, "Sparsity-aware learning and compressed sensing: an overview," https://arxiv.org/abs/1211.5231.

[3] R. Tibshirani, "Regression shrinkage and selection via the lasso," *Journal of the Royal Statistical Society, Series B*, vol. 58, no. 1, pp. 267–288, 1996.

[4] W. Barfuss, G. P. Massara, T. Di Matteo, and T. Aste, "Parsimonious modeling with information filtering networks," *Physical Review E*, vol. 94, no. 6, Article ID 062306, 2016.

[5] T. Aste, T. Di Matteo, and S. T. Hyde, "Complex networks on hyperbolic surfaces," *Physica A: Statistical Mechanics and its Applications*, vol. 346, no. 1-2, pp. 20–26, 2005.

[6] M. Tumminello, T. Aste, T. di Matteo, and R. N. Mantegna, "A tool for filtering information in complex systems," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 102, no. 30, pp. 10421–10426, 2005.

[7] G. P. Massara, T. Di Matteo, and T. Aste, "Network filtering for big data: triangulated maximally filtered graph," *Journal of Complex Networks*, vol. 5, no. 2, article 161, 2017.

[8] A. Tikhonov, "Solution of incorrectly formulated problems and the regularization method," *Soviet Mathematics Doklady*, vol. 4, no. 4, pp. 1035–1038, 1984.

[9] A. E. Hoerl and R. W. Kennard, "Ridge regression: biased estimation for nonorthogonal problems," *Technometrics*, vol. 12, no. 1, pp. 55–67, 1970.

[10] A. Zaremba and T. Aste, "Measures of causality in complex datasets with application to financial data," *Entropy*, vol. 16, no. 4, pp. 2309–2349, 2014.

[11] T. Schreiber, "Measuring information transfer," *Physical Review Letters*, vol. 85, no. 2, pp. 461–464, 2000.

[12] C. E. Shannon, "A mathematical theory of communication," *ACM SIGMOBILE Mobile Computing and Communications Review*, vol. 5, no. 1, pp. 3–55, 2001.

[13] T. W. Anderson, *Multivariate Statistical Analysis*, Willey and Sons, New York, NY, USA, 1984.

[14] J. Friedman, T. Hastie, and R. Tibshirani, "Sparse inverse covariance estimation with the graphical lasso," *Biostatistics*, vol. 9, no. 3, pp. 432–441, 2008.

[15] M. H. Gruber, *Improving Efficiency by Shrinkage: The James–Stein and Ridge Regression Estimators*, vol. 156 of *CRC Press*, 1998.

[16] D. M. Witten and R. Tibshirani, "Covariance-regularized regression and classification for high dimensional problems," *Journal of the Royal Statistical Society. Series B. Statistical Methodology*, vol. 71, no. 3, pp. 615–636, 2009.

[17] N. Meinshausen and P. Bühlmann, "High-dimensional graphs and variable selection with the lasso," *The Annals of Statistics*, vol. 34, no. 3, pp. 1436–1462, 2006.

[18] J. Kruskal Jr., "On the shortest spanning subtree of a graph and the traveling salesman problem," *Proceedings of the American Mathematical Society*, vol. 7, no. 1, pp. 48–50, 1956.

[19] R. N. Mantegna, "Hierarchical structure in financial markets," *The European Physical Journal B—Condensed Matter and Complex Systems*, vol. 11, no. 1, pp. 193–197, 1999.

[20] J. D. Hamilton, *Time Series Analysis*, vol. 2, Princeton University Press, Princeton, NJ, USA, 1994.

[21] S. S. Wilks, "The large-sample distribution of the likelihood ratio for testing composite hypotheses," *The Annals of Mathematical Statistics*, vol. 9, no. 1, pp. 60–62, 1938.

[22] Q. H. Vuong, "Likelihood ratio tests for model selection and nonnested hypotheses," *Econometrica. Journal of the Econometric Society*, vol. 57, no. 2, pp. 307–333, 1989.

[23] J. A. Swets, *Signal Detection Theory and ROC Analysis in Psychology and Diagnostics: Collected Papers*, Psychology Press, 2014.

[24] I. Varga-Haszonits, F. Caccioli, and I. Kondor, "Replica approach to mean-variance portfolio optimization," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2016, no. 12, Article ID 123404, 2016.

[25] G. Papp, F. Caccioli, and I. Kondor, "Fluctuation-bias trade-off in portfolio optimization under expected shortfall with l_2 regularization," https://arxiv.org/abs/1602.08297.

[26] G. Sugihara, R. May, H. Ye et al., "Detecting causality in complex ecosystems," *Science*, vol. 338, no. 6106, pp. 496–500, 2012.

[27] A. T. Clark, H. Ye, F. Isbell et al., "Spatial convergent cross mapping to detect causal relationships from short time series," *Ecology*, vol. 96, no. 5, pp. 1174–1181, 2015.

[28] J. Pearl, *Causality*, Cambridge University Press, 2009.

[29] J. Massey, *Causality, Feedback And Directed Information*, Citeseer, 1999.

[30] G. Kramer, *Directed information for channels with feedback [PhD thesis]*, University of Manitoba, Winnipeg, Canada, 1998.

[31] P.-O. Amblard and O. J. Michel, "The relation between Granger causality and directed information theory: a review," *Entropy. An International and Interdisciplinary Journal of Entropy and Information Studies*, vol. 15, no. 1, pp. 113–143, 2013.

Research Article

A Novel Synchronization-Based Approach for Functional Connectivity Analysis

Angela Lombardi,[1] Sabina Tangaro,[2] Roberto Bellotti,[2,3] Alessandro Bertolino,[4,5] Giuseppe Blasi,[4,5] Giulio Pergola,[4] Paolo Taurisano,[4,6] and Cataldo Guaragnella[1]

[1]*Dipartimento di Ingegneria Elettrica e dell'Informazione, Politecnico di Bari, Via E. Orabona 4, 70125 Bari, Italy*
[2]*Istituto Nazionale di Fisica Nucleare, Sezione di Bari, Via E. Orabona 4, 70125 Bari, Italy*
[3]*Dipartimento Interateneo di Fisica "M. Merlin", Universitá degli Studi di Bari "A. Moro", Via E. Orabona 4, 70125 Bari, Italy*
[4]*Dipartimento di Scienze Mediche di Base, Neuroscienze e Organi di Senso, Universitá degli Studi di Bari "A. Moro", Piazza Giulio Cesare 11, 70124 Bari, Italy*
[5]*Azienda Ospedaliero-Universitaria Consorziale Policlinico, 70124 Bari, Italy*
[6]*IRCCS "Casa Sollievo della Sofferenza", 71013 San Giovanni Rotondo, Italy*

Correspondence should be addressed to Sabina Tangaro; sonia.tangaro@ba.infn.it

Complex network analysis has become a gold standard to investigate functional connectivity in the human brain. Popular approaches for quantifying functional coupling between fMRI time series are linear zero-lag correlation methods; however, they might reveal only partial aspects of the functional links between brain areas. In this work, we propose a novel approach for assessing functional coupling between fMRI time series and constructing functional brain networks. A phase space framework is used to map couples of signals exploiting their cross recurrence plots (CRPs) to compare the trajectories of the interacting systems. A synchronization metric is extracted from the CRP to assess the coupling behavior of the time series. Since the functional communities of a healthy population are expected to be highly consistent for the same task, we defined functional networks of task-related fMRI data of a cohort of healthy subjects and applied a modularity algorithm in order to determine the community structures of the networks. The within-group similarity of communities is evaluated to verify whether such new metric is robust enough against noise. The synchronization metric is also compared with Pearson's correlation coefficient and the detected communities seem to better reflect the functional brain organization during the specific task.

1. Introduction

The human brain, as many biological systems, can be seen as a complex network of interacting components whose integration leads to a hierarchical architecture of highly specialized modules [1]. A network formulation simplifies the analysis of a complex system by providing mathematical tools able to capture different aspects of its organization in a compact and straightforward manner. Graph theoretical methods have been extensively applied to many neuroimaging datasets in order to describe the topological properties of both functional and structural networks [2, 3].

In particular, over the past few years, there has been an increasing interest in inferring connectivity properties from fMRI data. Functional connectivity analysis aims at assessing the strength of functional coupling between the signal responses in distinct brain areas [4]. According to the complex network framework, the anatomical regions of interest are the nodes of the network, connected by edges resulting from the adopted interregional interaction metrics. Pairwise fMRI time series connections are usually estimated through zero-lag correlation metrics, leading to a weighted network whose links quantify the statistical similarity between pairs of regions. Different preprocessing techniques and strategies

are also applied in order to extract only relevant information from the functional network, for example, by considering only a range of weights or by applying several thresholds to filter out weak connections [3]. Functional connectivity studies have revealed interesting insights on normal functional brain organization such as property of small-worldness [5], modularity and presence of hub nodes [6], and the existence of critical alterations of low-frequency neural activity patterns in pathological conditions [7]. Among the proposed strategies, some techniques are more established than others, even if there is still no agreement on which ones are the most effective or appropriate.

A number of important questions regarding the identification of networks have to be addressed before considering any analysis technique. Recent studies have demonstrated that different edge definitions could affect the topological properties of brain networks obtaining variable findings [8, 9]. Thus, properties like time resolution of the physiological time series under investigation, the effect of the observational noise, and the presence of nonlinear effects should been taken into account for selecting measures for edge definition. The low temporal resolution of fMRI data limits the number of methods that can be used to assess the statistical interactions between the time series. Linear correlation metrics, including Pearson's correlation and partial correlation, have been used in simulation environment and resting state studies, showing good performances in estimating functional connections in both cases [9, 10]. On the other hand, nonlinear phenomena in the human brain have been explored at various scales, revealing complex coupling mechanisms in both resting state and task-based neural activity [11, 12]. Most of the functional connectivity studies are focused on configurations of intrinsic connectivity networks (ICNs) and therefore did not assess complex connectivity patterns that can arise in the presence of a cognitive task. Indeed, even if a steady intrinsic network architecture has been found at rest and across a large number of tasks and conditions, task-evoked changes of functional connectivity have been also documented, proving the existence of task-specific network configurations [13]. Exploring topological changes in functional networks when the neural activity is modulated by a cognitive task could improve the understanding of some important mechanisms of human cognition, for example, the dynamic balancing of specialization and integration of brain regions for supporting different cognitive loads [4] and the trade-off between connection cost and topological efficiency in information processing [14]. Assessing functional interactions during external tasks should require metrics that (i) are sensitive to nonlinear coupling between time series and (ii) are more robust with respect to noise.

In this work, we propose a novel approach for quantifying functional coupling between fMRI time series and constructing functional brain networks. We use a phase space framework to map pairs of signals in their reconstructed phase space, that is, a topological representation of their behavior under all possible initial conditions [15]. This method assumes that each signal represents a projection of a higher-dimensional dynamical system evolving in time, whose trajectories are embedded into a manifold, that is, a region of its phase space. Cross recurrence plots (CRPs) [16] are then employed to reduce the dimensionality of the phase space and compare the trajectories of the interacting systems. A synchronization metric is finally extracted from the CRP to assess the coupling behavior of the time series.

The proposed metric and Pearson's correlation coefficient are applied to the fMRI data of a cohort of healthy subjects acquired during performing a working memory task to construct weighted networks.

At macroscopic level, functional related brain regions exhibit similar BOLD responses. These groups of regions form dense communities that reflect the functional organization of the brain and whose properties can be linked to the topological features of the task-evoked network configuration [17, 18]. The analysis carried out in this work aims at investigating some properties of the modular structure of task-evoked functional networks obtained with Pearson's correlation metric and the proposed synchronization index in order to understand which index can better highlight the functional organization of distinct subsystems involved in the specific working memory task. Therefore, a modularity algorithm is used to determine the community structure of each functional network. The within-group similarity of communities is evaluated and exploited to verify whether the metrics are sufficiently robust against noise and effective in revealing correlation even in presence of external stimuli. The rationale underlying this choice is that community structure of a group of healthy subject is expected to be highly consistent in presence of the same task.

2. Materials

2.1. Subjects. We studied 50 healthy subjects (age: mean = 25, standard deviation SD = 6; 24 females) in the analysis. All of them were evaluated using the Non-Patient Structured Clinical Interview for DSM-IV [19] to exclude any psychiatric condition. Other exclusion criteria were a significant history of drug or alcohol abuse, active drug abuse in the previous year, experience of a head trauma with loss of consciousness, and any other significant medical condition. Socioeconomic status (Hollingshead Four Factor Index, [20]), handedness (Edinburgh Inventory) [21], and total IQ (WAIS-R [22]) were also measured (see Table 1). The present study was approved by the local ethics committee (Comitato Etico Locale Indipendente Azienda Ospedaliera Ospedale Policlinico Consorziale Bari). Written informed consent was obtained by all participants after a complete description of the procedures, in accordance with the Helsinki Declaration.

2.2. fMRI Task. Participants performed the N-Back working memory task, in which a sequence of stimuli is presented and the subject has to remember the stimulus from "N" steps earlier. The stimuli consisted of numbers (1–4) presented in random sequence and displayed at the points of a diamond-shaped box. The control condition (0-back) simply required the subjects to identify the current stimulus. In the working memory condition, the task required the collection of a stimulus seen two stimuli earlier (2-back). The task was organized in a block design, consisting of eight alternating

TABLE 1: Demographic data of the healthy cohort (mean ± standard deviation).

Demographic Data	
Age (years)	25 ± 6
Gender (M/F)	26/24
Handedness	0.60 ± 0.55
Socioeconomic status	39 ± 17
IQ	114 ± 4

0-back and 2-back conditions, each lasting 30 seconds. Each 30 sec. block includes 14 n-back trials with an interstimuli interval of 2000 ms. Each run lasted 4 minutes and 8 seconds, from which dummy scans were acquired and discarded, obtaining 120 volumes.

2.3. fMRI Data. Echo planar imaging blood oxygenation level dependent fMRI data were acquired on a GE Signa 3T scanner (GE Healthcare) equipped with a standard quadrature head coil. A gradient-echo planar imaging sequence (repetition time, 2000 ms; echo time, 30 ms; thickness, 4 mm; gap, 1 mm; flip angle, 90°; field of view, 24 cm; and matrix, 64× 64) was used to acquire images while the subjects performed the tasks.

Images were preprocessed using Statistical Parametric Mapping 8 software (SPM8; http://www.fil.ion.ucl.ac.uk/ spm). Images were realigned to the first volume in the time series to correct for head motion (<2 mm translation, <1° rotation), resampled to a 2 mm isotropic voxel size, spatially normalized into a standard stereotactic space (Montreal Neurological Institute template) using a 12-parameter nonlinear warping, and smoothed to minimize noise and residual differences in gyral anatomy with a Gaussian filter, set at 6 mm full-width at half-maximum.

3. Methods

3.1. Network Construction. The brain volume of each subject was divided into 246 nonoverlapping anatomical regions of interest (ROIs) according to the Brainnetome Atlas [23]. Thirty regions from the most ventral part of the brain not acquired during scans were discarded and are not included in the following analysis. For each of the 216 remaining ROIs, a single time series was extracted by averaging the fMRI time series over all the voxels within the ROI. The time series were high-pass filtered (cutoff frequency 1/128 s). For each subject, functional connectivity between all pairwise combinations of ROI time series was assessed:

(i) by calculating their Pearson's correlation coefficient;

(ii) by computing their CRP and then by calculating their synchronization (SYNC) index as described in the following subsection.

Finally, for each subject, we identified two undirected weighted networks, whose edges resulted from

(1) the signed pairwise Pearson's correlation coefficients;

(2) the SYNC indexes.

3.2. Synchronization Index. A state of a system is defined by the values of the variables that describe it at a given time. When such system evolves in time, the sequence of all its states forms a trajectory in the phase space, that is, a multidimensional space whose dimension depends on the number of the variables of the system. Starting from different initial conditions, a real physical dissipative system tends to evolve in similar ways, such that its trajectories converge in a region of the phase space called attractor which represents the steady-state behavior of the system [15].

In experimental contexts, where the time series $\{u_i\}_{i=1}^{N}$ obtained from the sampling of a single observable variable is available, it is possible to reconstruct the phase space of the system under investigation by means of Takens's Theorem [24]. Accordingly, a state in the reconstructed phase space is given by a m-dimensional time delay embedded vector obtained from time delayed versions of the output signals as

$$\vec{x}_i = (u_i, u_{i+\tau}, \ldots, u_{i+(m-1)\tau}), \tag{1}$$

where m is the embedding dimension and τ is the time delay.

Both parameters have to be properly selected to avoid redundancy in the phase space. The dimension m of the reconstructed phase space should be large enough to preserve the properties of the dynamical system ($m \geq 2D + 1$, where D is the correlation dimension of the original phase space). The correct time delay τ should be chosen by determining when the samples of the time series are independent enough to be useful as coordinates of the time delayed vectors. For the estimation of the embedded parameters m and τ several techniques have been proposed. As an example, the first local minimum of average mutual information algorithm [25] can be used to select the proper time delay. The minimum embedding dimension is usually estimated through the false nearest-neighbors (FNN) algorithm [26].

The trajectories of two distinct systems with the same embedded parameters can be compared in a CRP [16], a matrix whose entries include information on the degree of closeness of each state of the first system with each state of the second system. In detail, for two systems with trajectories, respectively, $\vec{x}_i$ ($i = 1, \ldots, N$) and $\vec{y}_j$ ($j = 1, \ldots, N$), the CRP is defined as

$$CR_{i,j}(\epsilon) = \Theta\left(\epsilon - \left\|\vec{x}_i - \vec{y}_j\right\|\right), \tag{2}$$

where Θ is the Heaviside function, ϵ is a threshold for closeness, N is the number of considered states for each system, and $\|\cdot\|$ is a norm function. A generic entry $CR_{i,j}$ in the resulting $N \times N$ array is set to one if the distance between the points $\vec{x}_i$ and $\vec{y}_j$ is smaller than the threshold ϵ or to zero elsewhere.

The value of the parameter ϵ must be estimated carefully, as it influences the creation of structures in the plot. The selection of an appropriate value for the threshold ϵ can be made by taking into account the influence of the observational noise that could affect the experimental measures and the minimum distance between the trajectories of the two

systems. In general, choosing ϵ equal to few percent of the maximum phase space diameter could ensure a sufficient number of structures in the cross recurrence plot [27], while the appearance of artifacts could be avoided by considering the signal-to-noise ratio for the underlying physical systems [28].

A CRP exhibits characteristic patterns that show local time relationships of the segments of the trajectories of the two interacting systems. Typical structures include single dots, diagonal lines, and vertical and horizontal lines. Diagonal lines occur when the evolution of the states is similar at different times and their lengths are related to the periods during which the two systems move in similar ways remaining close to each other [29]. A CRP can also exhibit the main diagonal known as line of synchronization (LOS). The presence of LOS implies the identity of the states of the two systems in the same time intervals, that is, the (i, i) states, so its structure can be analyzed to extract information about the synchronization of the two time series [30]. In particular, the presence of LOS suggests that the two time series are fully synchronized, while discontinuities appear when the two signals do not have the same frequency and the same phase. Hence, the *synchronization time* (SYNC) has been defined as a metric to quantify the mean period during which the two systems are synchronized in order to reflect the dynamical synchronization behavior of the series throughout the observation period. *SYNC* is proportional to the ratio of the sum of the lengths of the subsegments l_j along the LOS to the total number of samples N:

$$\text{SYNC} = \frac{1}{N_d} \frac{\sum_{j=1}^{N_d} l_j}{N}, \tag{3}$$

where N_d is the total number of subsegments.

For a visual reference, see Figure 1. In Figure 1(a), two fMRI unsynchronized time series are compared and in Figure 1(b) are shown two fully synchronized fMRI time series. It is worth noting that in the first case there are discontinuities of LOS, while in the second case the LOS is continuous. Their SYNC values are, respectively, 0.05 and 1.

3.3. Modularity Detection.
Several community detection methods have been proposed to find an optimum partition of the nodes into nonoverlapped communities, that is, clusters of nodes that are more densely connected to each other than to other nodes in the network [31–33]. All these methods aim at maximizing a modularity metric that evaluates the quality of a partition by comparing the density of connections within a community to that expected in a random network. Here, the Louvain algorithm [33] has been used to find communities of ROIs in the two functional networks obtaining two partitions for each subject. The Louvain method is divided into two phases that are repeated iteratively. The first step favors local optimizations of modularity, while during the second step the communities found in the first step define a new coarse-grained network to be evaluated. This algorithm was chosen because it is fast and seems to be less affected by the resolution

limit problem (i.e., the capability to detect modules smaller than a certain size) thanks to its multilevel nature. This method optimizes the modularity function defined as

$$Q = \frac{1}{2m} \sum_{i,j} \left[A_{i,j} - \frac{k_i k_j}{2m} \right] \delta \left(c_i, c_j \right), \tag{4}$$

where $A_{i,j}$ is the link between nodes i and j, k_i is the sum of the weights of the links attached to node i, c_i is the community assigned to the node i, m is the sum of all of the links of the networks, and δ is the δ function.

3.4. Statistical Analysis of Modularity.
A statistical framework was adopted in order to compare the partitions of all the subjects for each functional network [34].

The normalized mutual information (NMI) [35] was used to assess the similarity between a couple of community partitions. For two networks with partitions, respectively, A and B, it is defined as

$$\text{NMI} (A, B) = 2 \frac{I (A, B)}{[H (A) + H (B)]}, \tag{5}$$

where $I(A, B)$ is the mutual information between the two partitions and $H(A)$ and $H(B)$ are the entropy of A and B. This metric ranges between zero (if A and B are completely independent) and one (if A and B are identical).

The statistical relevance of the within-group community structure similarity was evaluated through a permutation test. First, a randomly rewired version of each functional network was generated preserving weights, density, and degree sequence, resulting in two groups of networks: the actual and its randomized matching network. Then, the NMI was calculated between all the possible pairs of network partitions within each group. A null distribution was generated by randomizing group labels 10000 times and by calculating the permuted within-group mean NMI at each permutation. Finally, a p value was assigned as the number of times that the permuted within-group mean-similarity was greater than the actual within-group similarity, divided by the number of permutations.

In order to inspect the consistency of node assignments to specific functional communities, we carried out further analyses on the networks. Since the labels of modules are arbitrarily assigned by the community detection algorithm at each iteration, it is necessary to match the partition values across the subjects for visualizing the group level community structure. This problem can be overcome by finding a template partition as a reference and by reassigning the labels of communities to match the template, while preserving the distinctions between different modules in each partition [34]. In this work, the partitions of each network for both metrics were matched to the most representative network partition of the group, that is, the median determined by pairwise NMI. Once the labels of partitions are reassigned, it is possible to assess the within-group consistency of each ROI in community membership by counting the number of occurrences with which a ROI appears with a particular label.
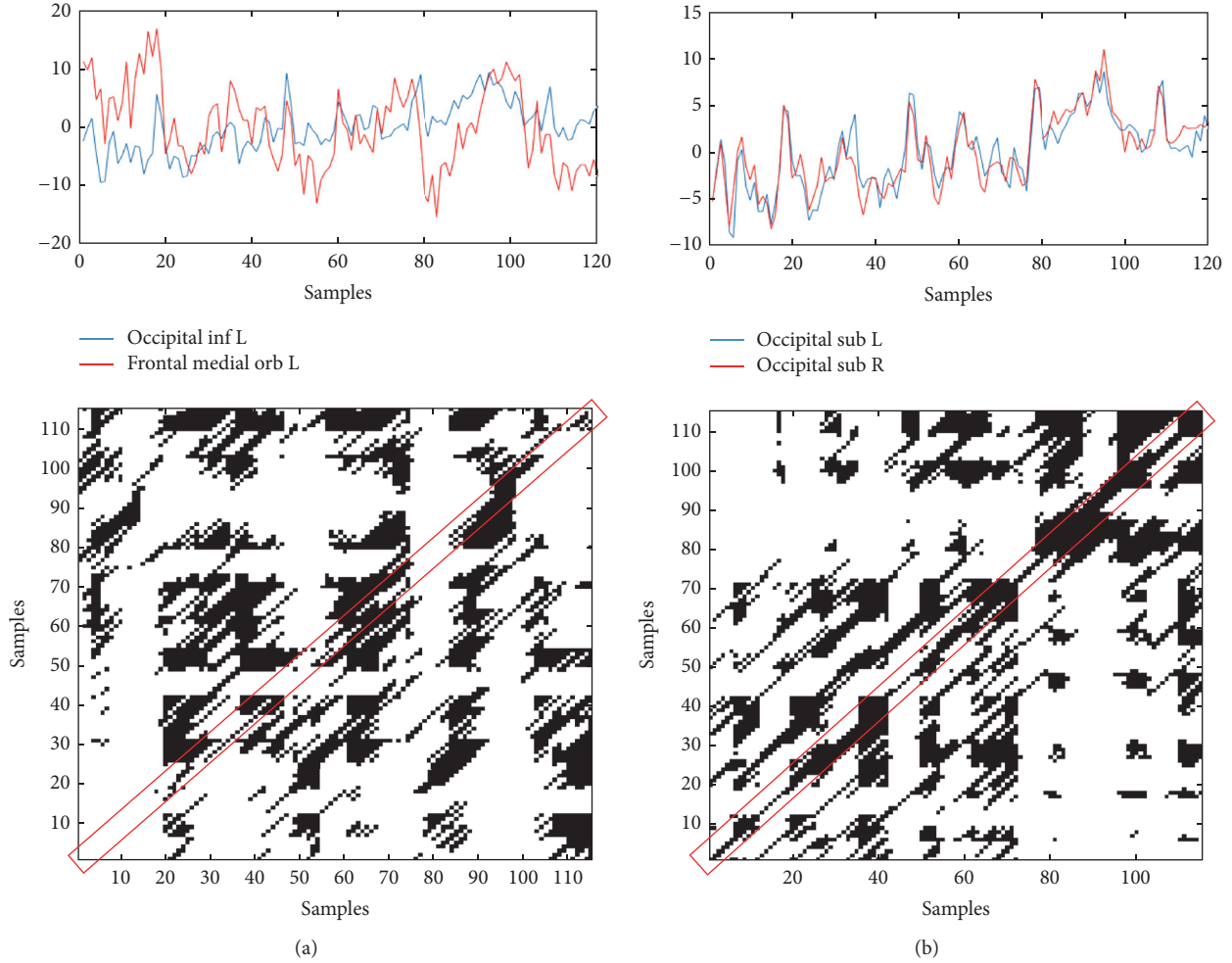
FIGURE 1: Pairs of fMRI time series and their CRPs for (a) occipital inferior L and frontal medial orbital L (SYNC = 0.05); (b) occipital superior L and occipital superior R (SYNC = 1).

4. Results

4.1. CRP Parameters. We randomly selected a subset of 5000 BOLD time series from the whole dataset and applied the FNN algorithm for estimating the embedded dimension and the first local minimum of the averaged mutual information for selecting the proper time delay. We obtained $m = 5.2 \pm 0.75$ and $\tau = 1.4 \pm 0.66$, so the embedded parameters were set to $m = 6$ and $\tau = 1$. Following the criteria reported in [27, 28], we identified the range [1.2–1.8] for the threshold ϵ. The analysis was carried out with the average value of the range, setting $\epsilon = 1.5$.

4.2. Statistical Analysis of Modularity. Permutation tests reveal significant differences of modularity structures between all the functional networks and their randomly rewired versions ($p = 0$ for both the couples), indicating different modular decompositions compared to the null models. However, as shown in Figure 2(a), Pearson's networks exhibit within-group NMI values much lower than those obtained by means of the SYNC metric (see Table 2

TABLE 2: Mean and median (interquartile range) quantities of NMI and Q distributions for the metrics synchronization and Pearson's correlation.

Distributions	NMI	Q
Synchronization	0.24; 0.23 (0.11)	0.17; 0.17 (0.04)
Pearson	0.15; 0.14 (0.15)	0.11; 0.11 (0.07)

for mean, median, and interquartile range quantities). The nonparametric Wilcoxon rank sum test confirmed significant differences between NMI values of SYNC and those of Pearson's metric ($p < 0.0001$, $\alpha = 0.05$). The ranges of NMI values of SYNC networks are also comparable to those found among control healthy subjects in resting studies at different threshold values of network density [34]. These results suggest that the functional networks constructed with the SYNC metric share more modularity structures than Pearson's networks and exhibit also a higher signal-to-noise ratio.

In addition, we evaluated the modularity index Q. This index ranges between 0 and 1 and measures the density
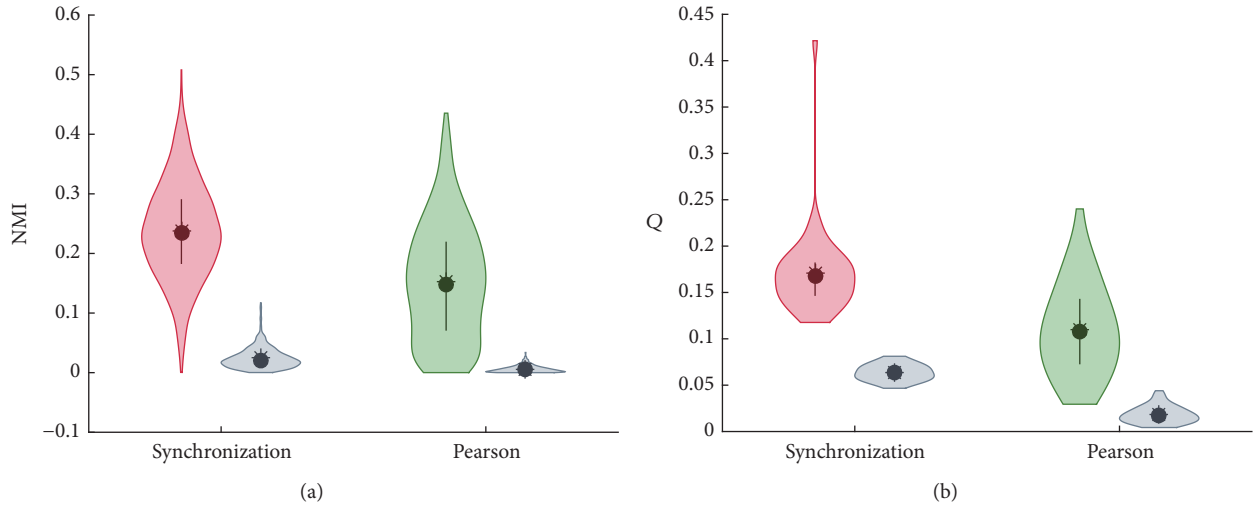
FIGURE 2: Violin plots of (a) within-group NMI distributions and (b) Q distributions for synchronization and Pearson's correlation connectivity metrics with their randomized matching networks (in gray).

of links inside communities as compared to links between communities. As shown in Figure 2(b), the two distributions are significantly different from their random versions (permutation tests: $p = 0$ for both pairings) and the modularity index of the networks obtained with the SYNC metric is higher than that of Pearson's networks (p value resulting from the nonparametric Wilcoxon rank sum test $p = 3.52 \cdot 10^{-8}$, $\alpha = 0.05$).

4.3. Comparison of Modular Partitions in SYNC and Pearson's Networks. Since we computed connectivity measures on a time series derived from a working memory task, we expected to find modules related to working memory performance involving the frontoparietal network [36] to motor activity related to the 0-back task [37] and to the default mode network, which is deactivated when performing the task [38]. Figure 3 shows the five modules detected by the Louvain algorithm at group level. The first module includes areas critical for visuospatial memory and closely resembles the classical frontoparietal network. In contrast, the second module includes more medial regions, with nodes belonging to both the anterior and the posterior default mode networks [39, 40]. The third module overlaps widely with the sensory-motor network, including pre- and postcentral nodes, but also areas of the temporal lobe involved in auditory perception. Interestingly, the fourth and fifth module map almost exclusively to subcortical regions, including the dorsal basal ganglia and the thalamus with the ventral striatum, respectively. These regions are involved in working memory performance [41, 42], but it is intriguing to notice that the technique here employed parsed the connectivity of cortical and subcortical regions based on the time series of activation, yielding anatomic information just based on functional activity patterns. Figure 4 shows the two communities identified at group level for Pearson's networks. The first module comprises most of the ROIs mapped in the first community

of the SYNC networks, while the rest of the ROIs are included in the second module.

The consistency of the assignment of brain regions to functional modules for the SYNC networks is shown in Figure 5. As can be seen, all the ROIs within the frontoparietal network are the most consistent among the subjects; in contrast, some nodes from the medial temporal lobes, insular gyrus, and globus pallidus are assigned less uniformly to the same community across the subjects. These findings are in line with the crucial involvement of the frontal parietal network in working memory processing [43]. As this map resembles closely an activity group map, these findings highlight that the connectivity assessment we developed is sensitive to the functional role of the modules identified. Overall, the network parsing obtained by the novel technique we reported reveals a pattern of coupling between brain regions consistent with known models of activation and deactivation during task performance. In Figure 6 is shown the within-group consistency of each ROI in community membership for Pearson's networks. Although the overall consistency seems generally higher due to the lower number of communities (two versus five), a direct comparison with the SYNC networks is possible only for the first module. The one-sided hypothesis Wilcoxon rank sum test confirmed a greater consistency of the ROIs within the first module for the SYNC matrices (median values of consistency: $m_{\text{sync}} = 85.5$, $m_{\text{pearson}} = 82$, $p = 0.0096$, $\alpha = 0.0$) proving a better identification of the frontoparietal network across the subjects in such matrices.

5. Discussion

In the current study, a modularity analysis is applied to networks defined with both the proposed SYNC index and Pearson's correlation coefficient in order to investigate the task-related functional organization of the brain. Modularity

Figure 3: The five group level functional communities detected in SYNC networks. In each row, a single community is shown in four brain views (left side, right side, top side, and bottom side).
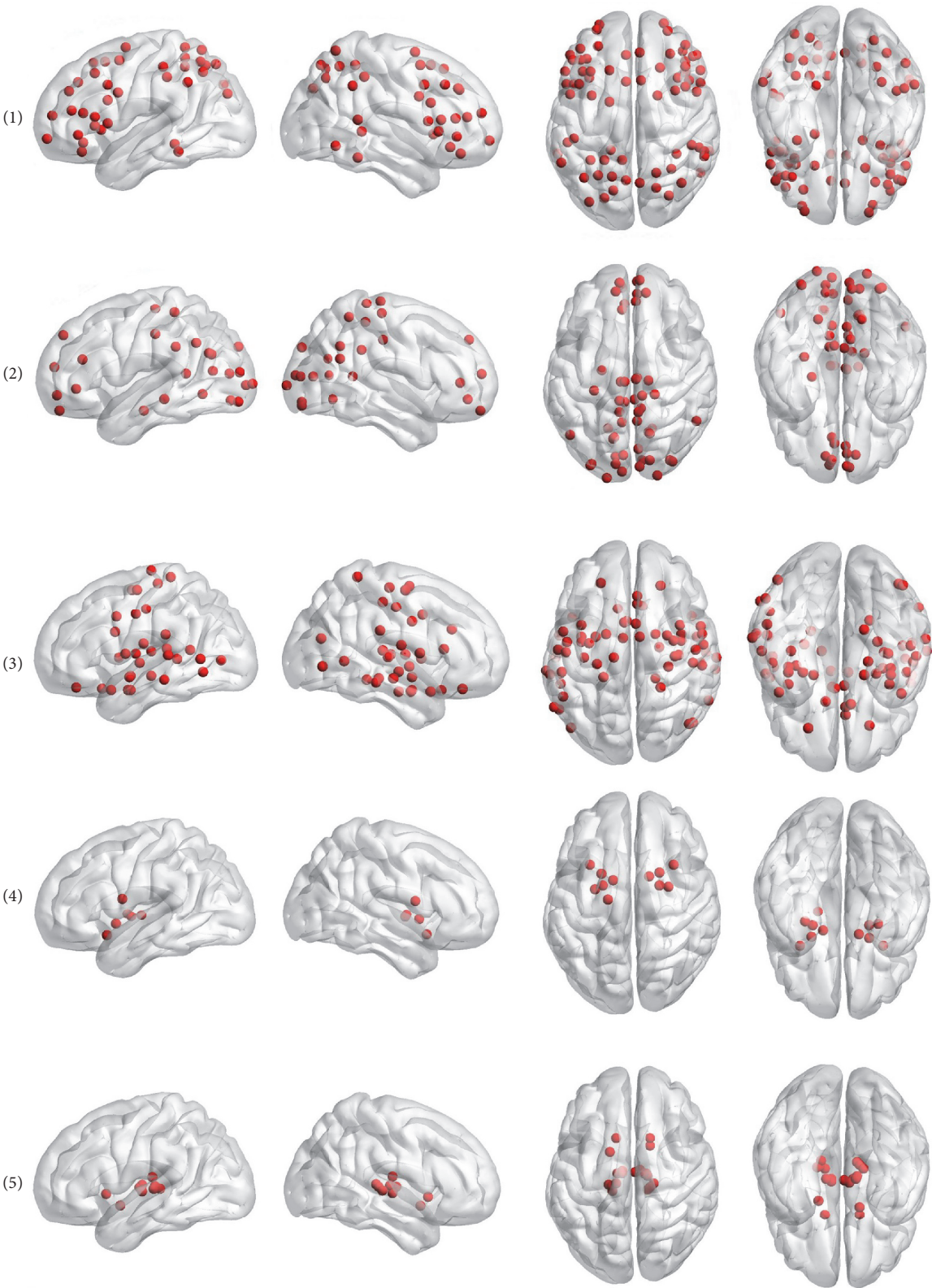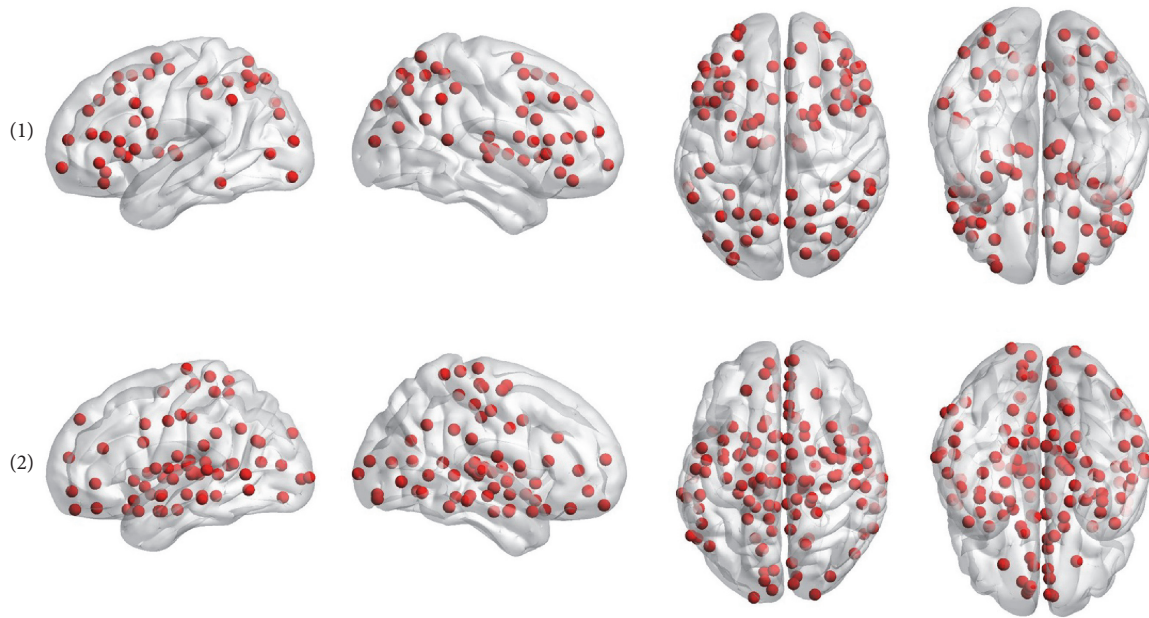
Figure 4: The two group level functional communities detected in Pearson's networks. In each row, a single community is shown in four brain views (left side, right side, top side, and bottom side).
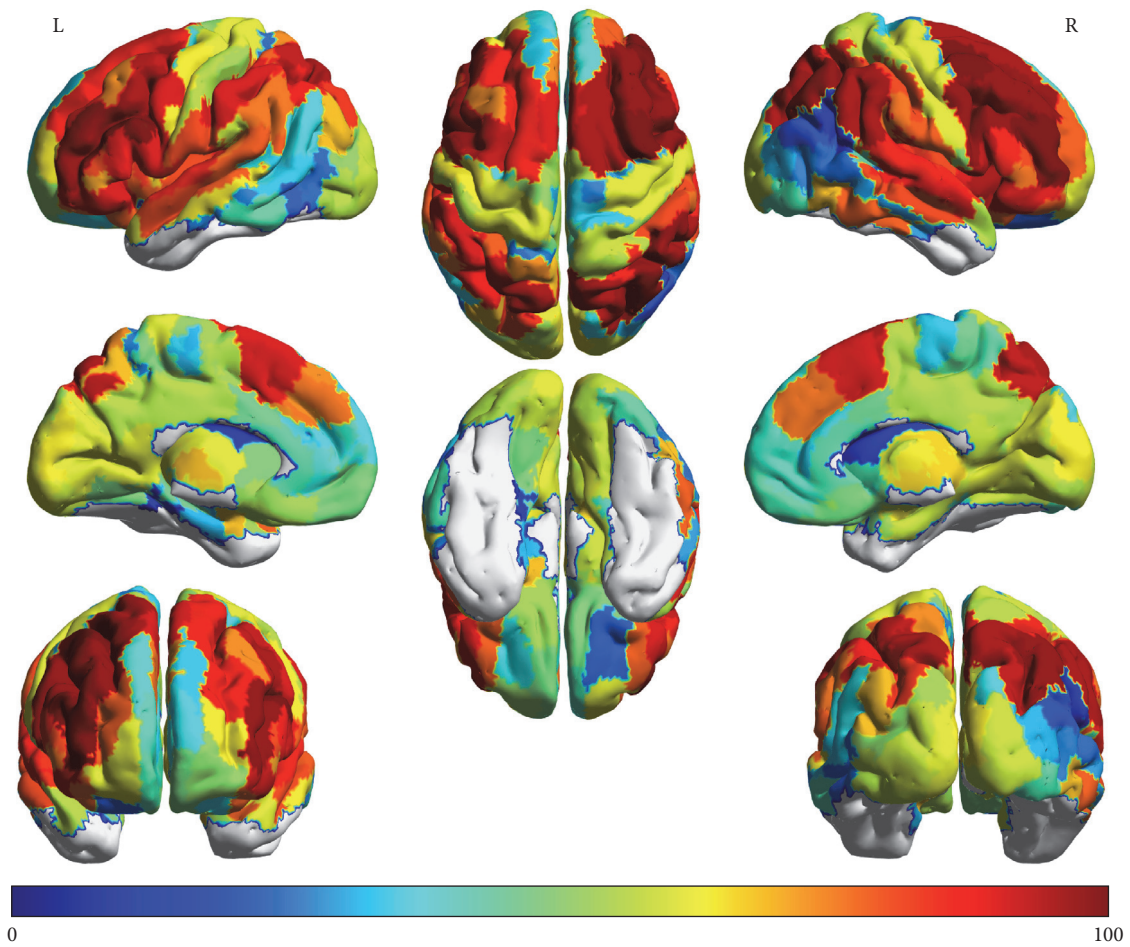

Figure 5: Consistency of the assignment of brain regions to modules measured as the frequency of occurrence of the node with a specific label (in percent) for SYNC networks.
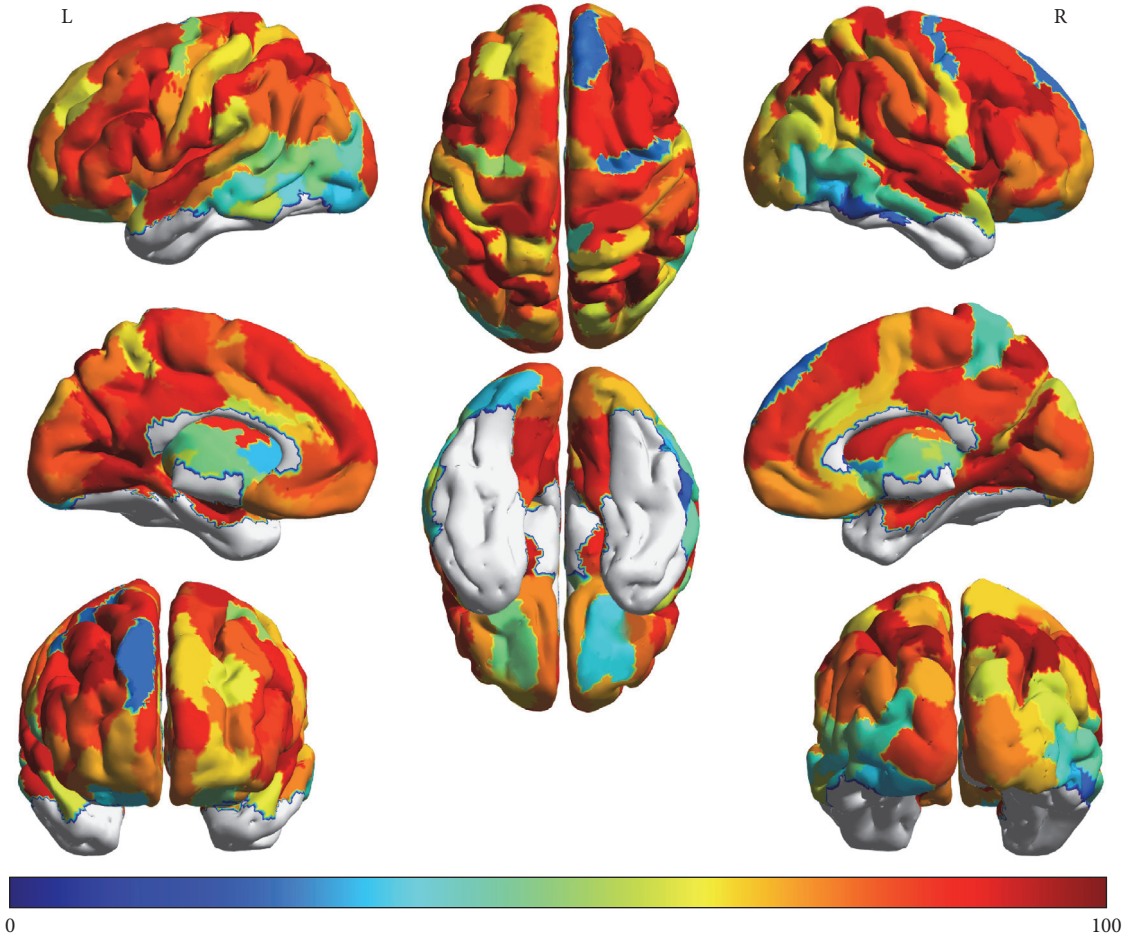
FIGURE 6: Consistency of the assignment of brain regions to modules measured as the frequency of occurrence of the node with a specific label (in percent) for Pearson's networks.

is implicitly related to significant self-regulating mechanisms of the human brain: efficient dense within-module processing and sparse fast integration among subsystems reduce noise propagation and latency [44]. Thus, this feature is strictly connected to critical functional organization between brain systems that are specialized to carry out different tasks: modularity is expected to be greater for optimal system organizations, while decreased modularity implies that there are less intramodular edges than intermodular edges [45]. A low level of modularity would not be compatible with a fast adaptation of the human brain in response to external stimuli. Indeed, lack of highly specialized modules may not allow a rapid execution of complex cognitive task [45, 46]. Consistently, a decreased modularity has been associated with brain disorders characterized by abnormal cognitive processing and has been found as a marker of abnormal brain network development [47–49]. Moreover, there is evidence that while the adaptation speed of the functional organization of the brain is not critical among healthy individuals that perform a specific task, modularity is stable across time, suggesting the existence of latent specific task-related modular configurations [17, 50]. The statistical analysis of

modularity reveals that a greater structure homogeneity and a higher number of functional communities activated during the working memory task seem to be better identified in SYNC networks, while Pearson's correlation does not reflect such features expected in a healthy population. In detail, the SYNC networks showed both higher NMI and Q values thus indicating that the extracted modular partitions are more similar to each other across the population and exhibit a clearer division into communities. Indeed the modularity index Q statistically quantifies the goodness of a hard partition as its value is related to the difference between the within-module interactions and the between-module interactions [51]. Furthermore, the consistency analysis in which the partition of each subject is compared with the median partition of the population points out two results: (i) both networks show at the group level a similar first community that resembles the frontoparietal network, but in the SYNC networks other modules that map to systems engaged during working memory performance are detected; (ii) the statistical comparison of the ROIs within the first module highlights a greater consistency of such task-related regions in SYNC networks. These findings suggest that a

problem of community resolution is evident in Pearson's networks, whereas all the regions not included in the first module are identified in a single community without distinction among sensory-motor network, default mode network, and subcortical areas and even the frontoparietal network is identified more weakly across the population.

In our framework, the same community detection algorithm was applied to both kinds of networks. Since the algorithm generates a node partition of a connectivity matrix, some properties of the index used to identify the network such as sensitivity to noise and to complex interaction mechanisms occurring among the brain regions could affect the degree of partition of the network into communities. Several brain connectivity metrics have been proposed as alternatives to Pearson's correlation coefficient. Coherence and partial coherence analysis were applied to fMRI data to extend linear metrics of zero-lag correlation. These spectral measures estimate the linear time-invariant relationship between time series by using phase and magnitude information for all the time lags [8]. Both coherence and partial coherence were proved effective in overcoming an important limitation of the zero-lag correlation, that is, its sensitivity to the shape of the regional hemodynamic response function that could result in spurious correlations of the underlying neural activity. In the last two decades, there has been a growing interest in developing new connectivity metrics sensitive to both linear and nonlinear interactions in human brain. In fact, the spatiotemporal nonlinearity was shown to be an important feature of the BOLD signal that should be considered to properly characterize the complex interactions between brain regions. In [52] a phase space multivariate approach was adopted to investigate the nonlinear properties of resting state fMRI data. The dynamics of the signals were reconstructed by using the time delay embedding of some principal components of the fMRI data and the correlation dimension and the spatiotemporal Lyapunov exponents were calculated to assess the nonlinear fractal property and the chaotic dynamic behavior of the signals. A surrogate data test confirmed an inherent deterministic nonlinear behavior in fMRI fluctuations. Other methods for exploring the dynamic behavior of physiological signals have been proposed. Recurrence plots and recurrence quantitative analysis of the structures therein contained were used to examine the recurrence properties of dynamic systems [29]. As an example, in [53] RQA was employed as a univariate data-driven technique to quantify recurrent patterns in fMRI data. This technique involves the projection of each time series in the phase space from which a recurrence plot is obtained. Several numerical descriptors are then used to quantify recurrent patterns in each time series. This method has been developed as an alternative to general linear model and probabilistic independent component analysis in activation studies. The underlying idea is that single-voxel signals become more regular in response to a stimulus, so RQA can detect the most active voxels without any model assumption. Recurrence plots and RQA were proved successful in analyzing very noisy and nonstationary signals. These methods afford a set of metrics able to capture comprehensively the dynamic behavior of a system in the phase space. Some studies confirmed their effectiveness also for the analysis of EEG and MEG data, particularly for detecting functional anomalies in several diseases [54–57]. Cross recurrence plots are bivariate extensions of the recurrence plots that consist in two-dimensional matrices showing the interactions of pairs of signals in the phase space. The proposed index, extracted from the CRP, represents an intuitively interpretable generalized dynamic synchronization metric that could be used to extend the set of known RQA measures.

6. Methodological Limitations

These results are promising with respect to the value of the novel technique we are proposing, even though they are not free of limitations. For example, since we used task-dependent time series, we do not know yet whether these results extend to resting state data, and this will be the object of future studies. We chose to examine task-driven functional connectivity as done in several other studies [17, 18, 51], by analyzing the modular architecture during working memory. In particular, we used both the presence of known task-related functional modules and their high consistence across a healthy cohort of subjects to evaluate the proposed synchronization metric. An advantage of the block-designed task we considered is that BOLD activity presents cyclostationary properties due to the ON-OFF periods of the task. Instead, spontaneous BOLD fluctuations are intrinsically dynamic over time and thus nonstationary [58]. For this reason, studying modularity with resting state data will require a modified dynamical framework to correctly identify stable ICNs for the declared purposes.

Another relevant issue concerns the modularity properties used to perform the comparison between the SYNC metric and Pearson's correlation index. Indeed, in our analysis we found both higher modularity and higher consistency of task-related communities in the SYNC matrices. These features are related to a greater homogeneity of the functional organization across the subjects in response to the same task and although they are compatible with behaviors expected in a healthy cohort, a more rigorous assessment of the sensitivity of the proposed synchronization metric should require further analysis. Future studies could employ alternative topological properties of SYNC networks and their correlation with task performance or behavioral data to uncover additional insights into the suitability of the SYNC index as a functional connectivity metric for fMRI time series.

Finally, our study has focused on an alternative method to define functional connectivity between pairs of BOLD time series. Generally, functional connectivity refers to a larger spectrum of neuroimaging techniques including EEG, MEG, and NIRS. As discussed above, recurrence plots have been used to explore dynamical properties of EEG and MEG, providing interesting features on complex phenomena in human brain. Although the SYNC metric is extracted from cross recurrence plots, a separate and accurate analysis may be needed to assert the validity of the index in a broader context and extend its use to more functional imaging techniques.

7. Conclusion

In this work, a new synchronization-based metric is proposed to assess functional connectivity in human brain. The metric is a generalized synchronization measure that takes into account both the amplitude and phase coupling between pairs of fMRI series. This method differs from the correlation measures used in the literature, as it is more sensitive to nonlinear coupling phenomena between time series and it is more robust against the physiological noise. In order to probe these latter two aspects, we performed a modularity analysis of task-related fMRI networks of a cohort of healthy subjects built with the new proposed metric. The aim was to verify whether the new metric was able to return networks whose functional modules were coherent with the actual organization of the brain regions during the task-based activity. We considered unthresholded complete connectivity matrices to test the effectiveness of the synchronization against noise and spurious correlations. Indeed unthresholded networks have lower signal-to-noise ratio as the most important links do not stand out among all the weights. By comparing the networks constructed by means of the proposed metric with those obtained through Pearson's coefficient, it seems that the synchronization metric better reflects the task-related network structure for number of detected communities, for the functional organization of the ROIs, and for greater consistency of communities across the subjects.

Conflicts of Interest

The authors declare that there are no conflicts of interest regarding the publication of this paper.

References

[1] O. Sporns and R. F. Betzel, "Modular brain networks," *Annual Review of Psychology*, vol. 67, pp. 613–640, 2016.

[2] E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," *Nature Reviews Neuroscience*, vol. 10, no. 4, pp. 186–198, 2009.

[3] M. Rubinov and O. Sporns, "Complex network measures of brain connectivity: Uses and interpretations," *NeuroImage*, vol. 52, no. 3, pp. 1059–1069, 2010.

[4] K. J. Friston, "Functional and effective connectivity: a review," *Brain Connectivity*, vol. 1, no. 1, pp. 13–36, 2011.

[5] D. S. Bassett and E. T. Bullmore, "Small-world brain networks," *The Neuroscientist*, vol. 12, no. 6, pp. 512–523, 2006.

[6] M. P. van den Heuvel and O. Sporns, "Network hubs in the human brain," *Trends in Cognitive Sciences*, vol. 17, no. 12, pp. 683–696, 2013.

[7] U. Braun, S. F. Muldoon, and D. S. Bassett, "On human brain networks in health and disease," eLS, 2015.

[8] F. T. Sun, L. M. Miller, and M. D'Esposito, "Measuring inter-regional functional connectivity using coherence and partial coherence analyses of fMRI data," *NeuroImage*, vol. 21, no. 2, pp. 647–658, 2004.

[9] X. Liang, J. Wang, C. Yan et al., "Effects of different correlation metrics and preprocessing factors on small-world brain functional networks: a resting-state functional MRI study," *PLoS ONE*, vol. 7, no. 3, Article ID e32766, 2012.

[10] S. M. Smith, K. L. Miller, G. Salimi-Khorshidi et al., "Network modelling methods for FMRI," *NeuroImage*, vol. 54, no. 2, pp. 875–891, 2011.

[11] G. Deco, V. K. Jirsa, P. A. Robinson, M. Breakspear, and K. Friston, "The dynamic brain: from spiking neurons to neural masses and cortical fields," *PLoS Computational Biology*, vol. 4, no. 8, Article ID e1000092, 2008.

[12] E. Pereda, R. Q. Quiroga, and J. Bhattacharya, "Nonlinear multivariate analysis of neurophysiological signals," *Progress in Neurobiology*, vol. 77, no. 1-2, pp. 1–37, 2005.

[13] M. W. Cole, D. S. Bassett, J. D. Power, T. S. Braver, and S. E. Petersen, "Intrinsic and task-evoked network architectures of the human brain," *Neuron*, vol. 83, no. 1, pp. 238–251, 2014.

[14] E. Bullmore and O. Sporns, "The economy of brain network organization," *Nature Reviews Neuroscience*, vol. 13, no. 5, pp. 336–349, 2012.

[15] G. L. Baker and J. P. Gollub, *Chaotic Dynamics*, Cambridge University Press, Cambridge, UK, 2nd edition, 1996.

[16] N. Marwan and J. Kurths, "Nonlinear analysis of bivariate data with cross recurrence plots," *Physics Letters A*, vol. 302, no. 5-6, pp. 299–307, 2002.

[17] D. S. Bassett, N. F. Wymbs, M. P. Rombach, M. A. Porter, P. J. Mucha, and S. T. Grafton, "Task-Based Core-Periphery Organization of Human Brain Dynamics," *PLoS Computational Biology*, vol. 9, no. 9, Article ID e1003171, 2013.

[18] Q. K. Telesford, M.-E. Lynall, J. Vettel, M. B. Miller, S. T. Grafton, and D. S. Bassett, "Detection of functional brain network reconfiguration during task-driven cognitive states," *NeuroImage*, vol. 142, pp. 198–210, 2016.

[19] D. DATA, *Structured Clinical Interview for DSM-IV Axis I Disorders*, American Psychiatric Press, Washington, DC, USA, 1997.

[20] A. B. Hollingshead, "Four factor index of social status," 1975.

[21] R. C. Oldfield, "The assessment and analysis of handedness: the Edinburgh inventory," *Neuropsychologia*, vol. 9, no. 1, pp. 97–113, 1971.

[22] A. Orsini and C. Laicardi, "Factor structure of the Italian version of the WAIS-R compared with the American standardization," *Perceptual and Motor Skills*, vol. 90, no. 4, pp. 1091–1100, 2000.

[23] L. Fan, H. Li, J. Zhuo et al., "The Human Brainnetome Atlas: A New Brain Atlas Based on Connectional Architecture," *Cerebral Cortex*, vol. 26, no. 8, pp. 3508–3526, 2016.

[24] F. Takens, "Detecting strange attractors in turbulence," in *Dynamical systems and Turbulence*, D. A. Rand and L. S. Young, Eds., vol. 898 of *Lecture Note in Mathematics*, pp. 366–381, Springer, Berlin, Germany, 1981.

[25] A. M. Fraser and H. L. Swinney, "Independent coordinates for strange attractors from mutual information," *Physical Review A: Atomic, Molecular and Optical Physics*, vol. 33, no. 2, pp. 1134–1140, 1986.

[26] M. B. Kennel, R. Brown, and H. D. I. Abarbanel, "Determining embedding dimension for phase-space reconstruction using a geometrical construction," *Physical Review A: Atomic, Molecular and Optical Physics*, vol. 45, no. 6, pp. 3403–3411, 1992.

[27] G. B. Mindlin and R. Gilmore, "Topological analysis and synthesis of chaotic time series," *Physica D: Nonlinear Phenomena*, vol. 58, no. 1-4, pp. 229–242, 1992.

[28] M. Thiel, M. C. Romano, J. Kurths, R. Meucci, E. Allaria, and F. T. Arecchi, "Influence of observational noise on the recurrence quantification analysis," *Physica D: Nonlinear Phenomena*, vol. 171, no. 3, pp. 138–152, 2002.

[29] N. Marwan, M. C. Romano, M. Thiel, and J. Kurths, "Recurrence plots for the analysis of complex systems," *Physics Reports*, vol. 438, no. 5-6, pp. 237–329, 2007.

[30] N. Marwan, M. Thiel, and N. R. Nowaczyk, "Cross recurrence plot based synchronization of time series," *Nonlinear Processes in Geophysics*, vol. 9, no. 3-4, pp. 325–331, 2002.

[31] M. E. J. Newman, "Modularity and community structure in networks," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 103, no. 23, pp. 8577–8582, 2006.

[32] M. Rosvall and C. T. Bergstrom, "Maps of random walks on complex networks reveal community structure," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 105, no. 4, pp. 1118–1123, 2008.

[33] V. D. Blondel, J. Guillaume, R. Lambiotte, and E. Lefebvre, "Fast unfolding of communities in large networks," *Journal of Statistical Mechanics: Theory and Experiment*, vol. 2008, no. 10, Article ID P10008, 2008.

[34] A. Alexander-Bloch, R. Lambiotte, B. Roberts, J. Giedd, N. Gogtay, and E. Bullmore, "The discovery of population differences in network community structure: New methods and applications to brain functional networks in schizophrenia," *NeuroImage*, vol. 59, no. 4, pp. 3889–3900, 2012.

[35] L. I. Kuncheva and S. T. Hadjitodorov, "Using diversity in cluster ensembles," in *Proceedings of the 2004 IEEE International Conference on Systems, Man and Cybernetics, SMC 2004*, pp. 1214–1219, nld, October 2004.

[36] R. Ptak, A. Schnider, and J. Fellrath, "The Dorsal Frontoparietal Network: A Core System for Emulated Action," *Trends in Cognitive Sciences*, vol. 21, no. 8, pp. 589–599, 2017.

[37] J. H. Callicott, N. F. Ramsey, K. Tallent et al., "Functional magnetic resonance imaging brain mapping in psychiatry: Methodological issues illustrated in a study of working memory in schizophrenia," *Neuropsychopharmacology*, vol. 18, no. 3, pp. 186–196, 1998.

[38] F. Sambataro, G. Blasi, L. Fazio et al., "Treatment with olanzapine is associated with modulation of the default mode network in patients with schizophrenia," *Neuropsychopharmacology*, vol. 35, no. 4, pp. 904–912, 2010.

[39] K. C. R. Fox, R. N. Spreng, M. Ellamil, J. R. Andrews-Hanna, and K. Christoff, "The wandering brain: Meta-analysis of functional neuroimaging studies of mind-wandering and related spontaneous thought processes," *NeuroImage*, vol. 111, pp. 611–621, 2015.

[40] P. Qin and G. Northoff, "How is our self related to midline regions and the default-mode network?" *NeuroImage*, vol. 57, no. 3, pp. 1221–1233, 2011.

[41] J. M. Scimeca and D. Badre, "Striatal Contributions to Declarative Memory Retrieval," *Neuron*, vol. 75, no. 3, pp. 380–392, 2012.

[42] J. Peräkylä, L. Sun, K. Lehtimäki et al., "Causal Evidence from Humans for the Role of Mediodorsal Nucleus of the Thalamus in Working Memory," *Cognitive Neuroscience*, pp. 1–13, 2017.

[43] J. H. Callicott, V. S. Mattay, A. Bertolino et al., "Physiological characteristics of capacity constraints in working memory as revealed by functional MRI," *Cerebral Cortex*, vol. 9, no. 1, pp. 20–26, 1999.

[44] D. S. Bassett, D. L. Greenfield, A. Meyer-Lindenberg, D. R. Weinberger, S. W. Moore, and E. T. Bullmore, "Efficient physical embedding of topologically complex information processing networks in brains and computer circuits," *PLoS Computational Biology*, vol. 6, no. 4, 2010.

[45] D. Meunier, R. Lambiotte, and E. T. Bullmore, "Modular and hierarchically modular organization of brain networks," *Frontiers in Neuroscience*, vol. 4, no. 200, 2010.

[46] D. A. Fair, A. L. Cohen, J. D. Power et al., "Functional brain networks develop from a "local to distributed" organization," *PLoS Computational Biology*, vol. 5, no. 5, Article ID e1000381, e1000381, 14 pages, 2009.

[47] A. F. Alexander-Bloch, N. Gogtay, D. Meunier et al., "Disrupted modularity and local connectivity of brain functional networks in childhood-onset schizophrenia," *Frontiers in Systems Neuroscience*, vol. 4, article 147, 2010.

[48] J. Wang, X. Zuo, Z. Dai et al., "Disrupted functional brain connectome in individuals at risk for Alzheimer's disease," *Biological Psychiatry*, vol. 73, no. 5, pp. 472–481, 2013.

[49] J. D. Rudie, J. A. Brown, D. Beck-Pancer et al., "Altered functional and structural brain network organization in autism," *NeuroImage: Clinical*, vol. 2, no. 1, pp. 79–94, 2013.

[50] D. S. Bassett, N. F. Wymbs, M. A. Porter, P. J. Mucha, J. M. Carlson, and S. T. Grafton, "Dynamic reconfiguration of human brain networks during learning," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 108, no. 18, pp. 7641–7646, 2011.

[51] U. Braun, A. Schäfer, H. Walter et al., "Dynamic reconfiguration of frontal brain networks during executive cognition in humans," *Proceedings of the National Acadamy of Sciences of the United States of America*, vol. 112, no. 37, pp. 11678–11683, 2015.

[52] X. Xie, Z. Cao, and X. Weng, "Spatiotemporal nonlinearity in resting-state fMRI of the human brain," *NeuroImage*, vol. 40, no. 4, pp. 1672–1685, 2008.

[53] M. Bianciardi, P. Sirabella, G. E. Hagberg, A. Giuliani, J. P. Zbilut, and A. Colosimo, "Model-free analysis of brain fMRI data by recurrence quantification," *NeuroImage*, vol. 37, no. 2, pp. 489–503, 2007.

[54] C. J. Stam, "Nonlinear dynamical analysis of EEG and MEG: review of an emerging field," *Clinical Neurophysiology*, vol. 116, no. 10, pp. 2266–2301, 2005.

[55] U. R. Acharya, S. V. Sree, S. Chattopadhyay, W. Yu, and P. C. A. Ang, "Application of recurrence quantification analysis for the automated identification of epileptic EEG signals," *International Journal of Neural Systems*, vol. 21, no. 3, pp. 199–211, 2011.

[56] D. E. McCarty, N. M. Punjabi, P. Y. Kim, C. Frilot, and A. A. Marino, "Recurrence analysis of the EEG during sleep accurately identifies subjects with mental health symptoms," *Psychiatry Research: Neuroimaging*, vol. 224, no. 3, pp. 335–340, 2014.

[57] L. T. Timothy, B. M. Krishna, and U. Nair, "Classification of mild cognitive impairment EEG using combined recurrence and cross recurrence quantification analysis," *International Journal of Psychophysiology*, vol. 120, pp. 86–95, 2017.

[58] D. T. Jones, P. Vemuri, M. C. Murphy et al., "Non-stationarity in the "resting brain's" modular architecture," *PLoS ONE*, vol. 7, no. 6, Article ID e39731, 2012.

Research Article

The Multiplex Dependency Structure of Financial Markets

Nicolò Musmeci,[1] Vincenzo Nicosia,[2] Tomaso Aste,[3,4] Tiziana Di Matteo,[1,3] and Vito Latora[2,5]

[1]*Department of Mathematics, King's College London, The Strand, London WC2R 2LS, UK*
[2]*School of Mathematical Sciences, Queen Mary University of London, Mile End Road, London E1 4NS, UK*
[3]*Department of Computer Science, University College London, Gower Street, London WC1E 6BT, UK*
[4]*Systemic Risk Centre, London School of Economics and Political Sciences, London WC2A 2AE, UK*
[5]*Dipartimento di Fisica ed Astronomia, Università di Catania and INFN, 95123 Catania, Italy*

Correspondence should be addressed to Vito Latora; v.latora@qmul.ac.uk

We propose here a multiplex network approach to investigate simultaneously different types of dependency in complex datasets. In particular, we consider multiplex networks made of four layers corresponding, respectively, to linear, nonlinear, tail, and partial correlations among a set of financial time series. We construct the sparse graph on each layer using a standard network filtering procedure, and we then analyse the structural properties of the obtained multiplex networks. The study of the time evolution of the multiplex constructed from financial data uncovers important changes in intrinsically multiplex properties of the network, and such changes are associated with periods of financial stress. We observe that some features are unique to the multiplex structure and would not be visible otherwise by the separate analysis of the single-layer networks corresponding to each dependency measure.

1. Introduction

In the last decade, network theory has been extensively applied to the analysis of financial markets. Financial markets and complex systems in general are comprised of many interacting elements, and understanding their dependency structure and its evolution with time is essential to capture the collective behaviour of these systems, to identify the emergence of critical states, and to mitigate systemic risk arising from the simultaneous movement of several factors. Network filtering is a powerful instrument to associate a sparse network to a high-dimensional dependency measure and the analysis of the structure of such a network can uncover important insights on the collective properties of the underlying system. Following the line first traced by the preliminary work of Mantegna [1], a set of time series associated with financial asset values is mapped into a sparse complex network whose nodes are the assets and whose weighted links represent the dependencies between the corresponding time series. Filtering correlation matrices has been proven to be very useful for the study and characterisation of the underlying

interdependency structure of complex datasets [1–5]. Indeed, sparsity allows filtering out noise, and sparse networks can then be analysed by using standard tools and indicators proposed in complex networks theory to investigate the multivariate properties of the dataset [6, 7]. Further, the filtered network can be used as a sparse inference structure to construct meaningful and computationally efficient predictive models [7, 8].

Complex systems are often characterised by nonlinear forms of dependency between the variables, which are hard to capture with a single measure and are hard to map into a single filtered network. A multiplex network approach, which considers the multilayer structure of a system in a consistent way, is thus a natural and powerful way to take into account simultaneously several distinct kinds of dependency. Dependencies among financial time series can be described by means of different measures, each one having its own advantages and drawbacks, and this has led to the study of different types of networks, namely, correlation networks, causality networks, and so on. The most common approach uses Pearson correlation coefficient to define the weight of a

link, because this is a quantity that can be easily and quickly computed. However, the Pearson coefficient measures the linear correlation between two time series [9], and this is quite a severe limitation, since nonlinearity has been shown to be an important feature of financial markets [10]. Other measures can provide equally informative pictures on assets relationships. For instance, the Kendall correlation coefficient takes into account monotonic nonlinearity [11, 12], while other measures, such as the Tail dependence, quantify dependence in extreme events. It is therefore important to describe quantitatively how these alternative descriptions are related but also differ from the Pearson correlation coefficient and also to monitor how these differences change in time, if at all.

In this work we exploit the power of a multiplex approach to analyse simultaneously different kinds of dependencies among financial time series. The theory of multiplex network is a recently introduced framework that allows describing real-world complex systems consisting of units connected by relationships of different kinds as networks with many layers, where the links at each layer represent a different type of interaction between the same set of nodes [13, 14]. A multiplex network approach, combined with network filtering, is the ideal framework to investigate the interplay between linear, nonlinear, and Tail dependencies, as it is specifically designed to take into account the peculiarity of the patterns of connections at each of the layers but also to describe the intricate relations between the different layers [15].

The idea of analysing multiple layers of interaction was introduced initially in the context of social networks, within the theory of frame analysis [16]. The importance of considering multiple types of human interactions has been more recently demonstrated in different social networks, from terrorist organizations [14] to online communities; in all these cases, multilayer analyses unveil a rich topological structure [17], outperforming single-layer analyses in terms of network modeling and prediction as well [18, 19]. In particular, multilayer community detection in social networks has been shown to be more effective than single-layer approaches [20]; similar results have been reported for community detection on the World Wide Web [21, 22] and citation networks [23]. For instance, in the context of electrical power grids, multilayer analyses have provided important insight into the role of synchronization in triggering cascading failures [24, 25]. Similarly, the analyses on transport networks have highlighted the importance of a multilayer approach to optimize the system against nodes failures, such as flights cancellation [26]. In the context of economic networks, multiplex analyses have been applied to study the World Trade Web [27]. Moreover, they have been extensively used in the context of systemic risk, where graphs are used to model interbank and credit networks [28, 29].

Here, we extend the multiplex approach to financial market time series, with the purpose of analysing the role of different measures of dependencies, namely, the Pearson, Kendall, Tail, and Partial correlation. In particular we consider the so-called Planar Maximally Filtered Graph (PMFG) [2–4, 7] as filtering procedure to each of the four layers. For each of the four unfiltered dependence matrices, the PMFG

filtering starts from the fully connected graph and uses a greedy procedure to obtain a planar graph that connects all the nodes and has the largest sum of weights [3, 4]. The PMFG is able to retain a higher number of links, and therefore a larger amount of information, than the Minimum Spanning Tree (MST) and can be seen as a generalization of the latter which is always contained as a proper subgraph [2]. The topological structures of MST and PMFG have been shown to provide meaningful economic and financial information [30–34] that can be exploited for risk monitoring [35–37] and asset allocation [38, 39]. The advantage of adopting a filtering procedure is not only in the reduction of noise and dimensionality but more importantly in the possibility of generating sparse networks, as sparsity is a requirement for most of the multiplex network measures that will be used in this paper [14]. Other kinds of filtering procedures, including thresholding based methods [35, 40], could have been considered. However, PMFG has the advantage to produce networks with fixed a priori $(3N - 6)$ number of links that make the comparison between layers and across time windows easier. It is worth mentioning that the filtering of the Partial correlation layer requires an adaptation of the PMFG algorithm to deal with asymmetric relations. We have followed the approach suggested in [41] that rules out double links between nodes. The obtained planar graph corresponding to Partial correlations has been then converted into an undirected graph and included in the multiplex.

2. Results

2.1. Multiplex Network of Financial Stocks. We have constructed a time-varying multiplex network with $M = 4$ layers and a varying number of nodes. Nodes represent stocks, selected from a dataset of $N_{tot} = 1004$ US stocks which have appeared at least once in S&P500 in the period between 03/01/1993 and 26/02/2015. The period under study has been divided into 200 rolling time windows, each of $\theta = 1000$ trading days. The network at time $T = 1, 2, \ldots, 200$ can be described by the adjacency matrix $a_{ij}^{\alpha}(T)$, with $i, j = 1, \ldots, N(T)$ and $\alpha = 1, 2, 3, 4$. The network at time window T has $N(T) < N_{tot}$ nodes, representing those stocks which were continuously traded in time window T. The links at each of the four layers are constructed by means of the PMFG procedure from Pearson, Kendall, Tail, and Partial dependencies. The reason for this choice is to provide a complete picture of the market dependency structure: Pearson layer accounts for linear dependency, Kendall layer for monotonic nonlinearity, and Tail dependency for correlation in the tails of returns distribution while Partial correlation detects direct asset-asset relationships which are not explained by the market (see Materials and Methods for details).

Figure 1(a) shows how the average link weight of each of the four dependency networks changes over time. We notice that the average edge weight is a meaningful proxy for the overall level of correlation in one of these dependency layers, since the distribution of edge weights within a layer is normally quite peaked around its mean. The curves shown in Figure 1(a) indicate an overall increase of the typical weights in the examined period 1993–2015 and reveals
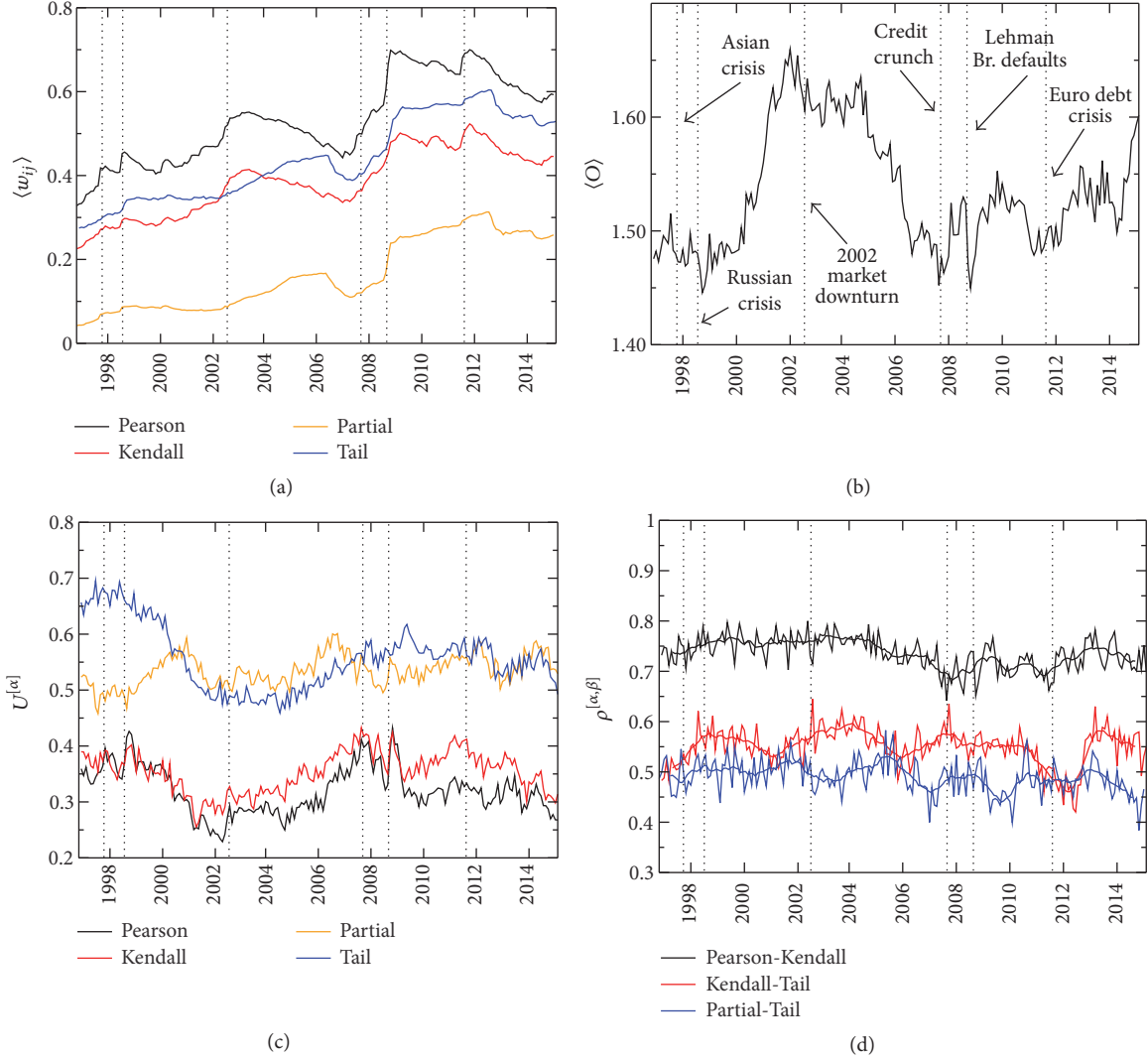
FIGURE 1: The multiplex nature of dependence among financial assets. The plots report the network analysis of a multiplex whose four layers are Planar Maximally Filtered Graphs (PMFGs) obtained from four classical dependence measures, namely, Pearson, Kendall, Tail, and Partial correlation, computed on rolling time windows of 23 trading days between 1993 and 2015. Each of the four layers provides different information on the dependency structure of a market. Although market events and trends have a somehow similar effect on the average dependence $\langle w_{ij} \rangle$ between nodes at the different layers (panel (a)), each layer has a distinct local structure. This is made evident by the plots of the average edge overlap $\langle O \rangle$ (panel (b)) and of the fraction $U^{[\alpha]}$ of edges unique to each layer, which confirm that an edge exists on average on less than two layers, and up to 70% of the edges of a layer are not present on any other layer. Moreover, the same node can have different degrees across the four layers, as indicated by the relatively low values of the pairwise interlayer degree correlation coefficient $\rho^{[\alpha,\beta]}$ reported in panel (d) for three pairs of layers over the whole observation interval.

a strongly correlated behaviour of the four curves (with linear correlation coefficients between the curves range in [0.91, 0.99]). In particular they all display a steep increase in correspondence with the 2007-2008 financial crisis, revealing how the market became more synchronized, regardless of the dependence measure used. This strong correlation in the temporal patterns of the four measures of dependence may lead to the wrong conclusion that the four networks carry very similar information about the structure of financial systems. Conversely, we shall see that even basic multiplex measures suggest otherwise. In Figure 1(b) we report the average edge overlap $\langle O \rangle$, that is, the average number of

layers of the financial multiplex network where a generic pair of nodes (i, j) is connected by an edge (see Materials and Methods for details). Since our multiplex network consists of four layer, $\langle O \rangle$ takes values in $[1, 4]$, and in particular we have $\langle O \rangle = 1$ when each edge is present only in one layer, while $\langle O \rangle = 4$ when the four networks are identical. The relatively low values of $\langle O \rangle$ observed in this case reveal the complementary role played by the different dependency indicators. It is interesting to note that the edge overlap $\langle O \rangle$ displays a quite dynamic pattern, and its variations seem to be related to the main financial crises highlighted by the vertical lines in Figure 1(b). Overall, what we observe is that periods

of financial turbulence are linked to widening differences among the four layers. Namely, the effect of nonlinearity in the cross-dependence increases, as well as correlation on the tails of returns: the dependence structure becomes richer and more complex during financial crisis. This might be related to the highly nonlinear interactions that characterise investors activities in turbulent periods and that make fat-tail and power-law distributions distinctive features of financial returns. Indeed, if returns were completely described by a multivariate normal distribution, the Pearson layer would be sufficient to quantify entirely the cross-dependence and its relation with the other layers would be trivial and would not change with time. Therefore any variation in the overlapping degree is a signature of increasing complexity in the market. In particular, the first event that triggers a sensible decrease in the average edge overlap is the Russian crisis in 1998, which corresponds to the overall global minimum of $\langle O \rangle$ in the considered interval. Then, $\langle O \rangle$ starts increasing towards the end of year 2000 and reaches its global maximum at the beginning of 2002, just before the market downturn of the same year. We observe a marked decrease in 2005, in correspondence with the second phase of the housing bubble, which culminates in the dip associated with the credit crunch at the end of 2007. A second, even steeper drop occurs during the Lehman Brothers default of 2008. After that, the signal appears more stable and weakly increasing, especially towards the end of 2014. Since each edge is present, on average, in less than two layers, each of the four layers effectively provides a partial perspective on the dependency structure of the market. This fact is made more evident by the results reported in Figure 1(c), where we show, for each layer $\alpha = 1, \ldots, 4$, the fraction of edges $U^{[\alpha]}$ that exist exclusively in that layer (see Materials and Methods for details). We notice that, at any point in time, from 30% to 70% of the edges of each of the four layers are unique to that layer, meaning that a large fraction of the dependence relations captured by a given measure are not captured by the other measures. For instance, despite the fact that Pearson and Kendall show similar behaviour in Figure 1(c), still between 30% and 40% of the edges on each of those layers exist only on that layer. This indicates that the Pearson and Kendall layers differ for at least 60% to 80% of their edges. In general, each of the four layers is contributing information that cannot be found in the other three layers. It was shown in a recent paper by some of the authors [36] that information filtering networks can be used to forecast volatility outbursts. The present results suggest that a multilayer approach could provide a further forecasting instrument for bear/bull markets. However, this requires further explorations. Interestingly, we observe an increase of $U^{[\alpha]}$ for all the layers since 2005, which indicates a build-up of nonlinearity and tail correlation in the years preceding the financial crisis: such dynamics might be related to early-risk warnings.

Another remarkable finding is that also the relative importance of a stock in the network, measured for instance by its centrality in terms of degree [39, 42], varies a lot across layers. This is confirmed by the degree correlation coefficient $\rho^{[\alpha,\beta]}$ for pairs of layers α and β. In general, high values of

$\rho^{[\alpha,\beta]}$ signal the presence of strong correlations between the degrees of the same node in the two layers (see Materials and Methods for details). Figure 1(d) shows $\rho^{[\alpha,\beta]}$ as a function of time for three pairs of dependence measures, namely, Pearson-Kendall, Kendall-Tail, and Tail-Partial. Notice that the degrees of the layers corresponding to Pearson and Kendall exhibit a relatively large correlation, which remains quite stable over the whole time interval. Conversely, the degrees of nodes in the Kendall and Tail layers are on average less correlated, and the corresponding values of $\rho^{[\alpha,\beta]}$ exhibit larger fluctuations. For example, in the tenth time window we find that General Electric stock (GE US) is a hub in Kendall layer with 71 connections, but it has only 16 connections in the Tail layer: therefore the relevance of this stock in the dependence structure depends sensitively on the layer. A similar pattern in observed in the interlayer correlation between the degrees of nodes in Partial and Tail. This might have important implications for portfolio allocation problems, since the asset centrality in the network is related to its risk in the portfolio.

The presence of temporal fluctuations in $\langle O \rangle$, in particular the fact that $\langle O \rangle$ reaches lower values during financial crises, together with the unique patterns of links at each layer, testified by high values of $U^{[\alpha]}$ and by relatively weak interlayer degree-degree correlations for some pairs of layers, confirms that an analysis of relations among stocks simply based on one dependence measure can neglect relevant information which can however be captured by other measures. As we will show below, a multiplex network approach which takes into account at the same time all the four dependence measures, but without aggregating them into a single-layer network, is able to provide a richer description of financial markets.

2.2. Multiedges and Node Multidegrees. As a first example of useful quantities that can be investigated in a multiplex network, we have computed the so-called multidegree $k_i^{\vec{m}}$ for each node i in the network, corresponding to different multiedges (see Materials and Methods) [43]. In particular, we have normalised the multidegree of node i dividing it by the corresponding node overlapping degree o_i, so that the resulting $k_i^{\vec{m}}/o_i$ is the fraction of multiplex edges of node i that exist only on a given subset of layers. In Figure 2 we report the average normalised multidegree of each of the 10 industry sectors of the Industry Classification Benchmark (ICB) classification. We focus on the edges existing exclusively in one of the four layers and on the combination of multiedges associated with edges existing in either of the Kendall, Tail, or Partial layer, but not in the Pearson layer. As shown in Figure 2, the multidegree exhibits strong variations in time and high heterogeneity across different industries. Industries such as Oil & Gas, Utilities, and Basic Materials show low values of normalised multidegree in all the four panels (Figures 2(a)–2(d)). Conversely, the edges of nodes corresponding to Industrials, Finance, Technology, Telecommunications, and Consumer Services tend to concentrate in one layer or in a small subset of layers only. For instance, we observe a relatively high concentration of edges at the Kendall layer for nodes corresponding to
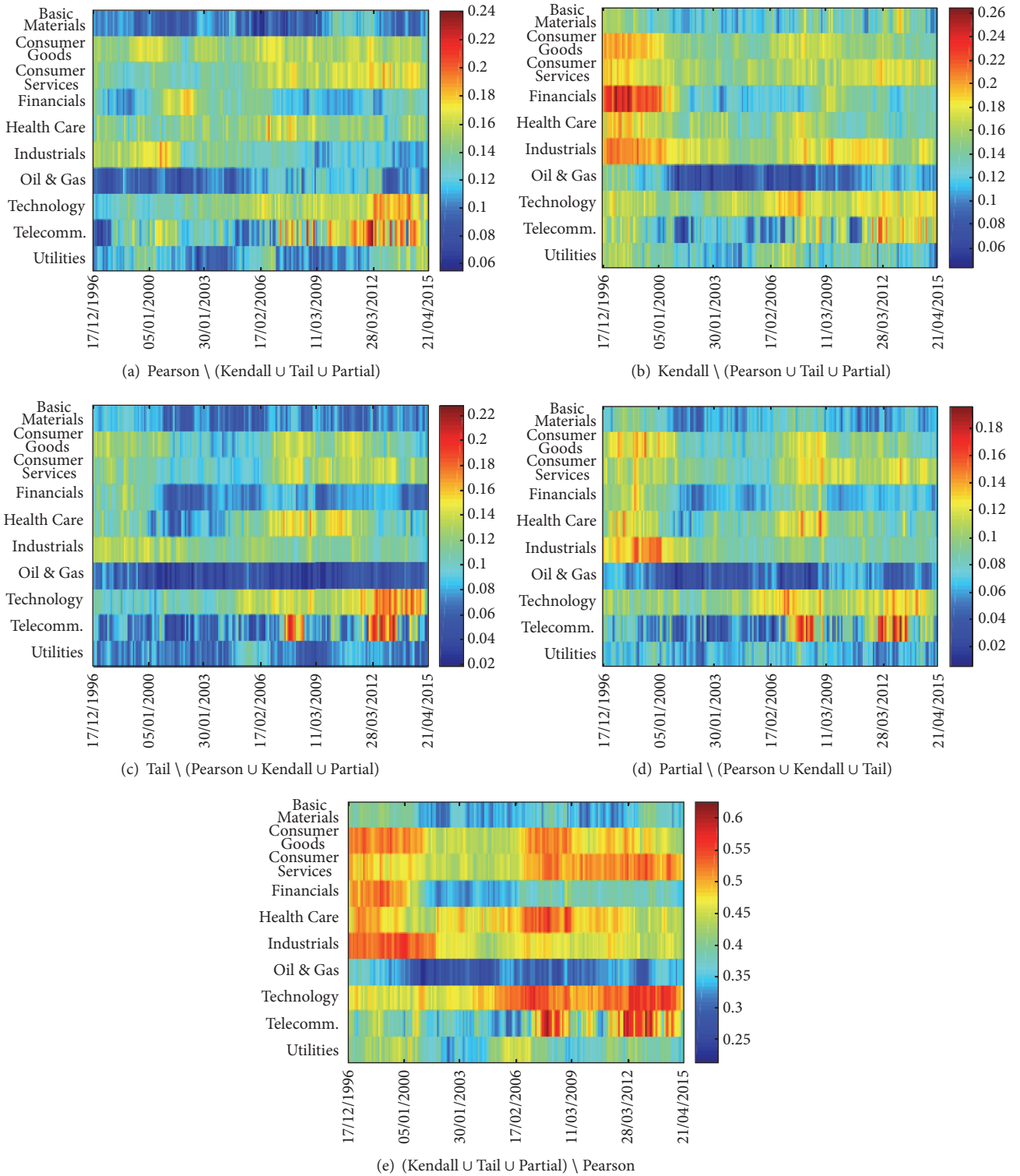
FIGURE 2: Multidegrees reveal the different role of industrial sectors during crises. The plots of the average multidegree of the nodes of the same industrial sector restricted to edges existing exclusively on the (a) Pearson, (b) Kendall, (c) Tail, and (d) Partial layers clearly show that some dependence measures can reveal structures which are unnoticed by other measures. In particular, the plot of the average multidegree associated with edges existing on at least one layer among Kendall, Tail, and Partial, but not on Pearson (panel (e)), reveals that Pearson correlation does not capture many important features such as the prominent role of Basic Materials, Financial, Consumer Goods, and Industrials during crises and the increasing importance of Technology and Consumer Services after the 2007-2008 crisis.

Finance, Industrials, and Consumer Goods stocks in the period preceding the dot-com bubble and the 2002 downturn, a feature not visible in the Pearson layer in Figure 2(a). This implies that for stocks belonging to those industries nonlinearity was a feature of their cross-dependency more important than for other stocks. Analogously, we notice a sudden increase of edges unique to the Tail layer for nodes in Consumer Goods, Consumer Services, and Health Care after the 2007-2008 crisis. Therefore, during this crisis period, the synchronization in the tail region has become a more relevant factor in their dependency structure than before the crisis: this has important implications for portfolio risk, as high Tail dependence can lead to substantial financial losses in case of large price movements. The presence of large heterogeneity and temporal variations in the relative role of different industrial sectors confirms the importance of using a multiplex network approach to analyse dependence among assets. Since industrial sectors have been often used for risk diversification, these findings point out that their use as a diversification benchmark might benefit remarkably from a continuous monitoring based on multiplex: an increase of edges in one layer for an industry can indicate the need of using the corresponding dependency measure for assessing the industry's risk and diversification potential. The fact that different industries display different degrees of nonlinearity and Tail dependence is not surprising after all, given that each industrial sector can be affected in a different way by new information: this industrial specific sensitivity might translate in different cross-dependency properties.

From this perspective it is particularly interesting to discuss the plot of multidegree restricted to edges that are present on either Kendall, Partial, or Tail layer but are not present in the Pearson layer as reported in Figure 2(e). Despite the fact that Pearson correlation coefficient is the most used measure to study dependencies, the plot reveals that until 2002 an analysis of the financial network based exclusively on Pearson correlations would have missed from 40% up to 60% of the edges of assets in sectors such as Basic Materials, Financial, Consumer Goods, and Industrials. The study of evolution with time in Figure 2(e) reveals that the relative role of such industrial sectors in Kendall, Tail, and Partial layers becomes relatively less important between the two crises in 2002 and in 2007, but then such sectors become central again during the 2007-2008 crisis and beyond. This prominent role is quite revealing but it would not had been evident from the analysis of the Pearson layer alone. Let us also note that the period following the 2007-2008 crisis is also characterised by a sensible and unprecedented increase of the normalised multidegree on Kendall, Partial, and Tail layers of stocks belonging to Technology and Telecommunications sectors, whose importance in the market dependence structure has been therefore somehow underestimated over the last ten years by the studies based exclusively on Pearson correlation.

2.3. Multiplex Cartography of Financial Systems. To better quantify the relative importance of specific nodes and groups of nodes we computed the overlapping degree and participation coefficient, respectively, measuring the total number of

edges of a node and how such edges are distributed across the layers (see Materials and Methods for details). We started by computing the average degree $k_I^{[\alpha]}$ at layer α of nodes belonging to each ICB industry sector I, defined as $k_I^{[\alpha]} = (1/N_I) \sum_{i \in I} k_i^{[\alpha]} \delta(c_i, I)$, where by c_i we denote the industry of node i and N_I is the number of nodes belonging to industry sector I. Figures 3(a)–3(d) show $k_I^{[\alpha]}$ as a function of time for each of the four layers.

Notice that nodes in the Financial sector exhibit a quite high average degree, no matter the dependence measure used, with a noticeable peak before the dot-com bubble in 2002. After that, the average degree of Financials has dropped sensibly, with the exception of the 2007-2008 crisis. Apart from the existence of similarities in the overall trend of Financials across the four layers, the analysis of the average degree suggests again the presence of high heterogeneity across sectors and over time.

In the Pearson layer, Basic Materials is the second most central industry throughout most of the observation interval, whereas Industrials and Oil & Gas acquired more connections in the period following the 2007-2008 crisis. The degree in the Kendall layer is distributed more homogeneously among the sectors than in the Pearson layer. Interestingly, the plot of degree on the Tail layer looks similar to that of the Pearson layer. Finally, in the Partial layer we observe the highest level of concentration of links in Finance (consistently to what was found in [41]) and, after the 2007-08 crisis, in Basic Materials.

We have also calculated for each industry I the average overlapping degree $o_I \equiv \langle o_i \rangle_{i \in I}$, where o_i is the overlapping degree of node i, which quantifies the overall importance of each industrial sector in the multiplex dependence network. The average overlapping degree of each industry is shown as a function of time in Figure 3(e). As we can see, o_I is able to highlight the prominent role played in the multiplex network by Financials, Basic Materials, Oil & Gas, and Industrials sectors, revealing also the presence of four different phases between 1997 and 2015. The first phase, during which Financials is the only prominent industry, covers the period between 1997 and 2000. The second phase, between 2000 and the 2007-08 crisis, is characterised by the emergence of Basic Materials as the second most central sector. In the third phase, between 2009 and 2014, Financials loses its importance in favour of Industrials, Oil & Gas, and Basic Materials (that becomes the most central one). Finally, in 2014 a new equilibrium starts to emerge, with Financials and Industrials gaining again a central role in the system.

The participation coefficient complements the information provided by the overlapping degree, quantifying how the edges of a node are distributed over the layers of the multiplex. In particular, the participation coefficient of node i is equal to 0 if i has edges in only one of the layers, while it is maximum and equal to 1 when the edges of node i are equally distributed across the layers (see Materials and Methods for details). In Figure 3(f) we report, as a function of time, the average participation coefficient P_I for each ICB industry I. Interestingly, the plot reveals that the increase of the overlapping degrees of Financials, Basic Materials, Industrials,
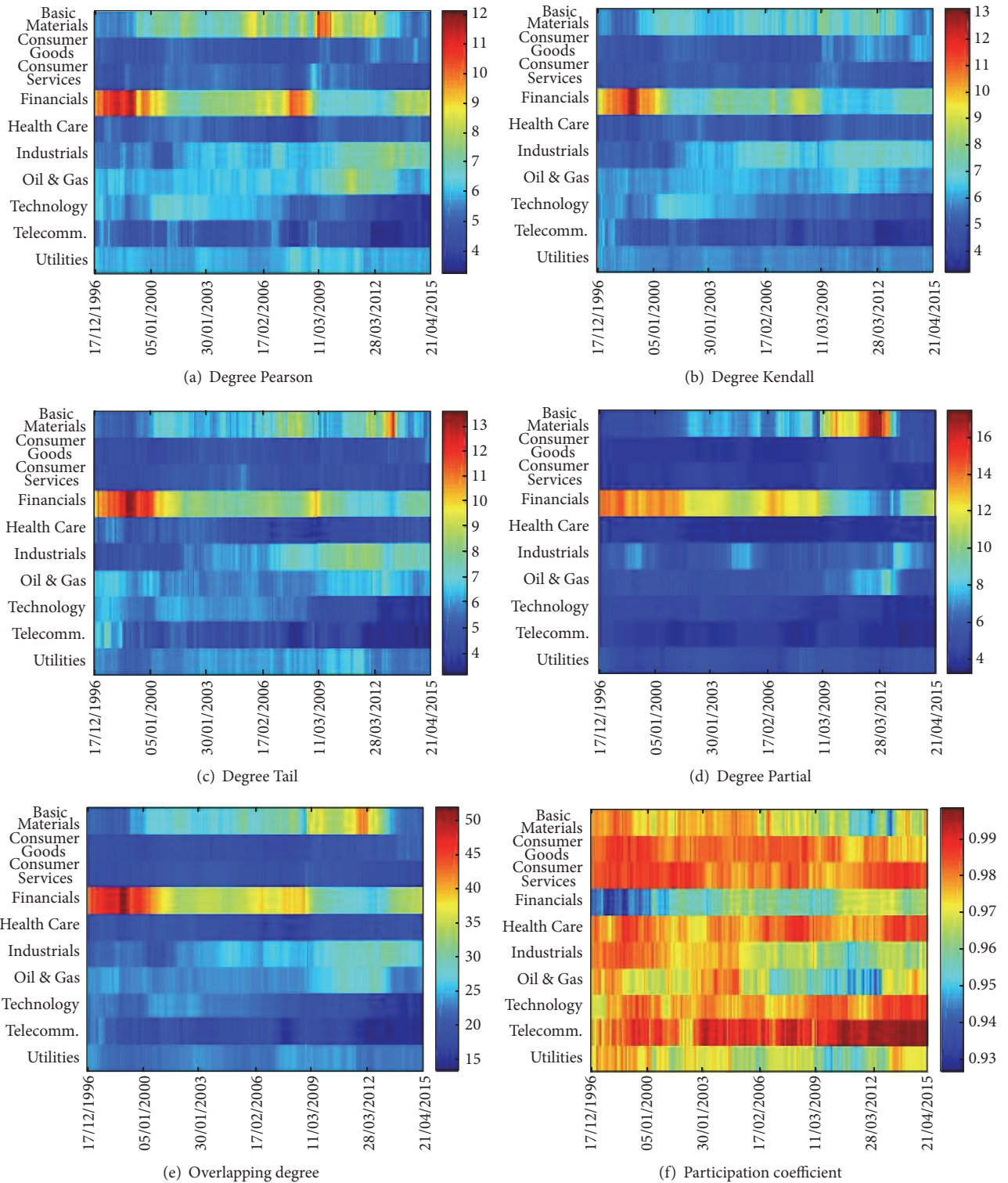
FIGURE 3: Average node degree as a proxy of the importance of an industry. The plots of average degree of the nodes belonging to the different industrial sectors restricted to the (a) Pearson, (b) Kendall, (c) Tail, and (d) Partial layers and of the average overlapping degree reported in panel (e) confirm the relative importance of Financials. However, the average participation coefficient (panel (f)) suggests that the dependence structure of some sectors such as Basic Materials, Industrials, and Oil & Gas has become more heterogeneous, that is, focusing only on a subset of the four layers, after the 2007-2008 crisis.

and Oil & Gas sectors shown in Figure 3(e) is normally accompanied by a substantial decrease of the corresponding participation coefficients. This indicates that those sectors accumulated degree on just one or two layers, confirming what we found in multidegree analysis. A somehow more detailed analysis of the temporal evolution of participation coefficient for each sector is reported in Section 4.5.

3. Discussion

By using filtered networks from different correlation measures we have demonstrated that a multiplex network approach can reveal features that would have otherwise been invisible to the analysis of each dependency measure in isolation. Although the layers produced, respectively, from Pearson, Kendall, Tail, and Partial correlations show a certain overall similarity, they exhibit distinct features that are associated with market changes. For instance, we observed that average edge overlap between the first three layers drops significantly during periods of market stress revealing that nonlinear effects are more relevant during crisis periods. The analysis of the average multidegree associated with edges not present on the Pearson layer, but existing on at least one of the three remaining layers, indicates that Pearson correlations alone can miss detecting some important features. We observed that the relative importance of nonlinearity and tails on market dependence structure, as measured by mean edge overlap between the last three layers, has dropped significantly in the first half of the 2000s and then risen steeply between 2005 and the 2007-08 crisis. Overall, financial crises trigger remarkable drops in the edge overlap, widening therefore the differences among the measures of dependence just when evaluation of risk becomes of the highest importance. Different industry sectors exhibit different structural overlaps. For instance, Financials, Industrials, and Consumer Goods show an increasing number of connections only on Kendall layer in the late 90s/early 2000s, at the edge of the dot-com bubble. After the 2007-08 crisis these industries tend to have many edges on the Kendall, Tail, and Partial which are not present on the Pearson layer. This observation questions whether we can rely on the Pearson estimator alone, when analysing correlations between stocks. A study of the overlapping degree and of the participation coefficient shows that asset centrality, an important feature for portfolio optimization [39, 40], changes considerably across layers with largest desynchronized changes occurring during periods of market distress. Overall our analysis indicates that different dependency measures provide complementary pieces of information about the structure and evolution of markets and that a multiplex network approach can be useful in capturing systemic properties that would otherwise go unnoticed.

4. Materials and Methods

4.1. Dataset. The original dataset consists of the daily prices of N_{tot} = 1004 US stocks traded in the period between 03/01/1993 and 26/02/2015. Each stock in the dataset has been included in S&P500 at least once in the period considered.
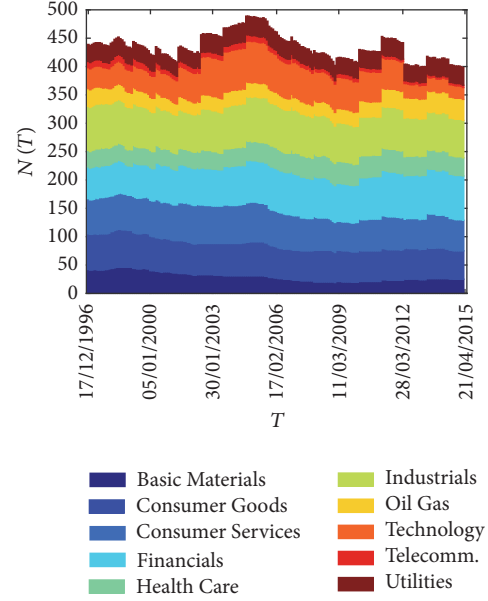

FIGURE 4: Number of stocks in each ICB industry in time. Number of stocks that are continuously traded in each time window together with their partition in terms of ICB industries.

Hence the stocks considered provide a representative picture of the US stock market over an extended time window of 22 years and cover all the 10 industries listed in the Industry Classification Benchmark (ICB) (Figure 4). It is important to notice that most of the stocks in this set are not traded over the entire period. This is a major difference with respect to the majority of the works on dynamic correlation-based networks, in which only stocks continuously traded over the period under study are considered, leading to a significant "survival bias." For each asset i we have calculated the series of log-returns, defined as $r_i(t) = \log(P_i(t)) - \log(P_i(t-1))$, where $P_i(t)$ is stock price at day t. The construction of the time-varying multiplex networks is based on log-returns and has been performed in moving time windows of θ = 1000 trading days (about 4 years), with a shift of dT = 23 trading days (about one month), adding up to 200 different multiplex networks, one for each time window. For each time window T, four different $N(T) \times N(T)$ dependence matrices have been computed, respectively, based on the four different estimators illustrated in Section 4.2. Since the number of active stocks changes with time, dependence matrices at different times can have different number of stocks $N(T)$, as shown in Figure 4. In the figure the ICB industry composition of our dataset in each time window is also shown, confirming that we have a representative sample of all market throughout the period. We have verified that the results we are discussing in the following are robust against change of θ and dT.

4.2. Dependence among Financial Time Series. We have considered four different measures of dependence between two time series $r_i(u)$ and $r_j(u)$, $i, j = 1, 2, \ldots, N$, $u = 1, 2, \ldots, \theta$, indicated in the following, respectively, as Pearson, Kendall, Tail, and Partial.

4.2.1. Pearson Dependence. It is a measure of linear dependence between two time series and is based on the evaluation of the Pearson correlation coefficient [44]. We have used the exponentially smoothed version of this estimator [45], in order to mitigate excessive sensitiveness to outliers in remote observations:

$$\rho_{ij}^{w}$$

$$= \frac{\sum_{u=1}^{\theta} w_u \left(r_i(u) - \overline{r_i}^w\right) \left(r_j(u) - \overline{r_j}^w\right)}{\sqrt{\sum_{u=1}^{\theta} w_u \left(r_i(u) - \overline{r_i}^w\right)^2} \sqrt{\sum_{u=1}^{\theta} w_u \left(r_j(u) - \overline{r_j}^w\right)^2}}, \quad (1)$$

with

$$w_u = w_0 \exp\left(\frac{u - \theta}{T^*}\right), \quad (2)$$

where T^* is the weight characteristic time ($T^* > 0$) that controls the rate at which past observations lose importance in the correlation and w_0 is a constant connected to the normalisation constraint $\sum_{u=1}^{\theta} w_u = 1$. We have chosen $T^* = \theta/3$ according to previously established criteria [45].

4.2.2. Kendall Dependence. It is a measure of dependence that takes into account the nonlinearity of a time series. It is based on the evaluation of the so-called Kendall's τ rank correlation coefficient, starting from the quantities $d_k(u, v) \equiv \text{sgn}(r_k(u) - r_k(v))$. The estimator counts the number of concordant pairs, that is, pairs of observations such that $d_i(u, v)$ and $d_j(u, v)$ have equal signs, minus the number of discordant pairs [11]. As for the case of the Pearson dependence, we have used the exponentially smoothed version of the estimator [45]:

$$\tau_{ij}^w = \sum_{u=1}^{\theta} \sum_{v=u+1}^{\theta} w_{u,v} d_i(u, v) d_j(u, v), \quad (3)$$

with

$$w_{u,v} = w_0 \exp\left(\frac{u - \theta}{T^*}\right) \exp\left(\frac{v - \theta}{T^*}\right), \quad (4)$$

where T^* is again the weight characteristic time.

4.2.3. Tail Dependence. It is a nonparametric estimator of tail copula that provides a measure of dependence focused on extreme events. It is based on the evaluation of the following estimator [46]:

$$C_{ij}(p_1, p_2) = \frac{\sum_{u=1}^{\theta} \mathbb{1}_{\{F^i(r_i(u)) < p_1 \land F^j(r_j(u)) < p_2\}}}{\sum_{u=1}^{\theta} \mathbb{1}_{\{F^i(r_i(u)) < p_1 \lor F^j(r_j(u)) < p_2\}}}, \quad (5)$$

where F^i and F^j are the empirical cumulative probabilities of returns $r_i(u)$ and $r_j(u)$, respectively, and p_1 and p_2 are two parameters representing the percentiles above which an observation is considered (lower) tail. We focus on lower tails since we are interested in risk management applications, where the attention is on losses. It can be shown that this is a consistent estimator of tail copula [46]. In this work we have chosen $p_1 = p_2 = 0.1$ (i.e., we consider tail every observation below the 10th percentile), as a trade-off between the need of statistic and the interest in extreme events.

4.2.4. Partial Dependence. It is a measure of dependence that quantifies to what extent each asset affects other assets correlation. The Partial correlation $\rho_{ik|j}$, or correlation influence, between assets i and k based on j, is the Pearson correlation between the residuals of $r_i(u)$ and $r_k(u)$ obtained after regression against $r_j(u)$ [47]. It can be written in terms of a Pearson correlation coefficient as follows [41]:

$$\rho_{ik|j} = \frac{\rho_{ik} - \rho_{ij}\rho_{kj}}{\sqrt{\left[1 - \rho_{ij}^2\right]\left[1 - \rho_{kj}^2\right]}}. \quad (6)$$

This measure represents the amount of correlation between i and k that is left once the influence of j is subtracted. Following [41], we define the correlation influence of j on the pair i, k as

$$d(i, k \mid j) = \rho_{ik} - \rho_{ik|j}. \quad (7)$$

$d(i, k \mid j)$ is large when a significant fraction of correlation between i and k is due to the influence of j. Finally, in order to translate this into a measure between i and j, the so-called Partial dependence, we average it over the index k:

$$d(i \mid j) = \langle d(i, k \mid j) \rangle_{k \neq i, j}. \quad (8)$$

$d(i \mid j)$ is the measure of influence of j on i based on Partial correlation. It is worth noting that, unlike the other measures of dependence, $d(i \mid j)$ provides a directed relation between assets (as in general $d(i \mid j) \neq d(j \mid i)$). In the rest of the paper we refer to this indicator as "Partial dependence," even though strictly speaking we are analysing the correlation influence based on Partial correlation.

4.3. Graph Filtering and the Construction of the Multiplex Network. For each of the 200 time windows we have then constructed a multiplex network with $M = 4$ layers obtained, respectively, by means of the four dependence indicators. In order to reduce the noise and the redundance contained in each dependence matrix we have applied the Planar Maximally Filtered Graph [2–4, 7]. It is worth mentioning that the filtering of the correlation influence layer requires an adaptation of the PMFG algorithm to deal with asymmetric relations. We have followed the approach suggested in [41] that rules out double links between nodes. The obtained planar graphs have been then converted into undirected graphs and included in the multiplex.

4.4. Multiplex Measures. Let us consider a weighted multiplex network $\mathcal{M}$ on N nodes, defined by the M-dimensional array of weighted adjacency matrices $\mathcal{W} = \{W^{[1]}, W^{[2]}, \ldots, W^{[M]}\}$, where $W^{[\alpha]} = \{w_{ij}^{[\alpha]}\}$ are the matrices of weights that determine the topology of the αth layer though the PMFG filtering. Here the weight $w_{ij}^{[\alpha]}$ represents the strength of the correlation between node i and node j on layer α, where the different layers are obtained through different correlation measures. In the following we will indicate by $W^{[\alpha]}$ the weighted adjacency matrix of the PMFG associated with layer α and by $A^{[\alpha]}$ the corresponding unweighted

adjacency matrix, where $a_{ij}^{[\alpha]} = 1$ if and only if $w_{ij}^{[\alpha]} \neq 0$. We denote by $K^{[\alpha]} = (1/2) \sum_{ij} a_{ij}^{[\alpha]}$ the number of edges on layer α and by $K = (1/2) \sum_{i,j} [1 - \prod_{\alpha} (1 - a_{ij}^{[\alpha]})]$ the number of pairs of nodes which are connected by at least one edge on at least one of the M layers. Notice that since the network at each layer is a PMFG, then we have $K^{[\alpha]} = 3(N - 2) \ \forall \alpha$ by construction.

We consider some basic quantities commonly used to characterise multiplex networks [14, 43]. The first one is the mean edge overlap, defined as the average number of layers on which an edge between two randomly chosen nodes i and j exists:

$$\langle O \rangle = \frac{1}{2K} \sum_{i,j} \sum_{\alpha} a_{ij}^{[\alpha]}. \tag{9}$$

Notice that $\langle O \rangle = 1$ only when all the M layers are identical; that is, $A^{[\alpha]} \equiv A^{[\beta]} \ \forall \alpha, \beta = 1, \ldots, M$, while $\langle O \rangle = 0$ if no edge is present in more than one layer, so that the average edge overlap is in fact a measure of how much similar the structures of the layers of a multiplex network are. A somehow dual quantity is the fraction of edges of layer α which do not exist on any other layer:

$$U^{[\alpha]} = \frac{1}{2K^{[\alpha]}} \sum_{i,j} a_{ij}^{[\alpha]} \prod_{\beta \neq \alpha} \left(1 - a_{ij}^{[\beta]} \right) \tag{10}$$

which quantifies how peculiar the structure of a given layer α is, since $U^{[\alpha]}$ is close to zero only when almost all the edges of layer α are also present on at least one of the other $M - 1$ layers.

More accurate information about the contribution of each node to a layer (or to a group of layers) can be obtained by the so-called multidegree of a node i. Let us consider the vector $\overrightarrow{m} = (m_1, m_2, \ldots, m_M)$, with M equal to the number of layers, where each m_α can take only two values $\{1, 0\}$. We say that a pair of nodes i, j has a multilink $\overrightarrow{m}$ if they are connected only on those layers α for which $m_\alpha = 1$ in $\overrightarrow{m}$ [43]. The information on the M adjacency matrices a_{ij}^α ($\alpha = 1, \ldots, M$) can then be aggregated in the multiadjacency matrix $A_{ij}^{\overrightarrow{m}}$, where $A_{ij}^{\overrightarrow{m}} = 1$ if and only if the pair i, j is connected by a multilink $\overrightarrow{m}$. Formally [13, 43]

$$A_{ij}^{\overrightarrow{m}} \equiv \prod_{\alpha=1}^{M} \left[a_{ij}^\alpha m_\alpha + \left(1 - a_{ij}^\alpha \right) \left(1 - m_\alpha \right) \right]. \tag{11}$$

From the multiadjacency matrix we can define the multidegree $\overrightarrow{m}$ of a node i, as the number of multilinks $\overrightarrow{m}$ connecting i:

$$k_i^{\overrightarrow{m}} = \sum_j A_{ij}^{\overrightarrow{m}}. \tag{12}$$

This measure allows us to calculate, for example, how many edges node i has on layer 1 only ($k_i^{\overrightarrow{m}}$ choosing $m_1 = 1, m_\alpha = 0, \ \forall \alpha \neq 1$), integrating the global information provided by $U^{[\alpha]}$.

The most basic measure to quantify the importance of single nodes on each layer is by means of the node degree $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$. However, since the same node i is normally present at all layers, we can introduce two quantities to characterise the role of node i in the multiplex [14], namely, the overlapping degree

$$o_i = \sum_{\alpha} k_i^{[\alpha]} \tag{13}$$

and the multiplex participation coefficient:

$$P_i = \frac{M}{M-1} \left[1 - \sum_{\alpha} \left(\frac{k_i^{[\alpha]}}{o_i} \right) \right]. \tag{14}$$

The overlapping degree is just the total number of edges incident on node i at any layer, so that node are classified as *hubs* if they have a relatively large value of o_i. The multiplex participation coefficient quantifies the dispersion of the edges incident on node i across the layers. In fact, $P_i = 0$ if the edges of i are concentrated on exactly one of the M layers (in this case i is a *focused* node), while $P_i = 1$ if the edges of i are uniformly distributed across the M layers, that is, when $k_i^{[\alpha]} = o_i/M \ \forall \alpha$ (in which case i is a *truly multiplex* node). The scatter plot of o_i and P_i is called *multiplex cartography* and has been used as a synthetic graphical representation of the overall heterogeneity of node roles observed in a multiplex.

In a multiplex network, it is important also to look at the presence and sign of interlayer degree correlations. This can be done by computing the interlayer degree correlation coefficient [15]:

$$\rho^{[\alpha,\beta]} = \frac{\sum_i \left(R_i^{[\alpha]} - \overline{R^{[\alpha]}} \right) \left(R_i^{[\beta]} - \overline{R^{[\beta]}} \right)}{\sqrt{\sum_i \left(R_i^{[\alpha]} - \overline{R^{[\alpha]}} \right)^2 \sum_j \left(R_j^{[\beta]} - \overline{R^{[\beta]}} \right)^2}}, \tag{15}$$

where $R_i^{[\alpha]}$ is the rank of node i according to its degree on layer α and $\overline{R^{[\alpha]}}$ is the average rank by degree on layer α. In general $\rho^{[\alpha,\beta]}$ takes values in $[-1, 1]$, where values close to $+1$ and -1, respectively, indicate the strong positive and negative correlations, while $\rho^{[\alpha,\beta]} \simeq 0$ if the degrees at the two layers are uncorrelated.

4.5. Time Evolution of the Average Participation Coefficient. In Figure 5 we plot the time evolution of the average participation coefficient P_I (x-axis) of stocks in the industrial sector I against the average overlapping degree o_I (y-axis). Each circle corresponds to one of the 200 time windows, while the size and colour of each circle represent different time windows. Each panel corresponds to one industrial sector I. The diagrams reveal that in the last 20 years the role of different sectors has changed radically and in different directions. For instance, stocks in the Financials sector evolved from a relatively large overlapping degree and a small participation coefficient in the late 1990s to a smaller number of edges, distributed more homogeneously across the layers, towards the end of the observation period. Conversely, Industrials
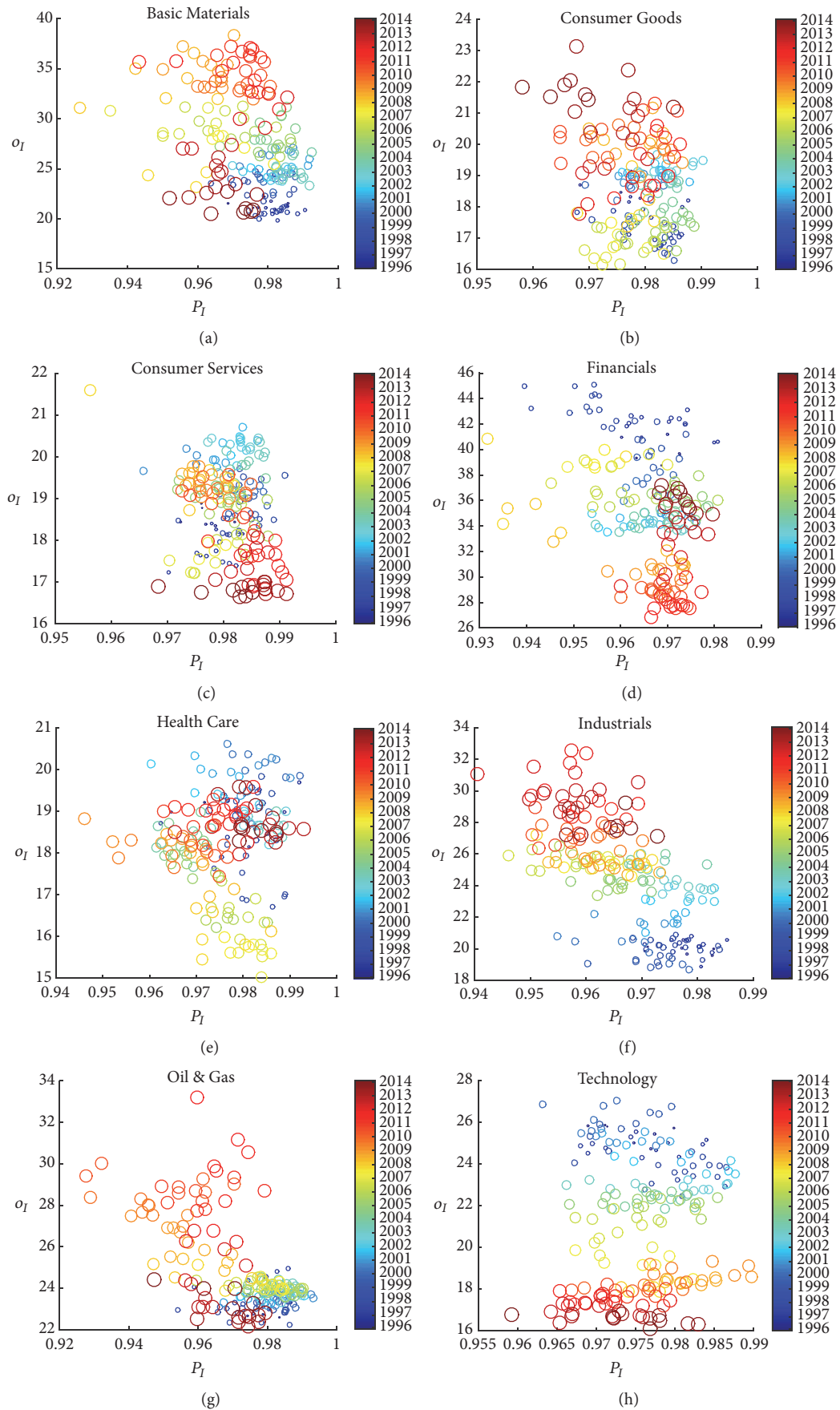
(a)

(b)

(c)

(d)

(e)
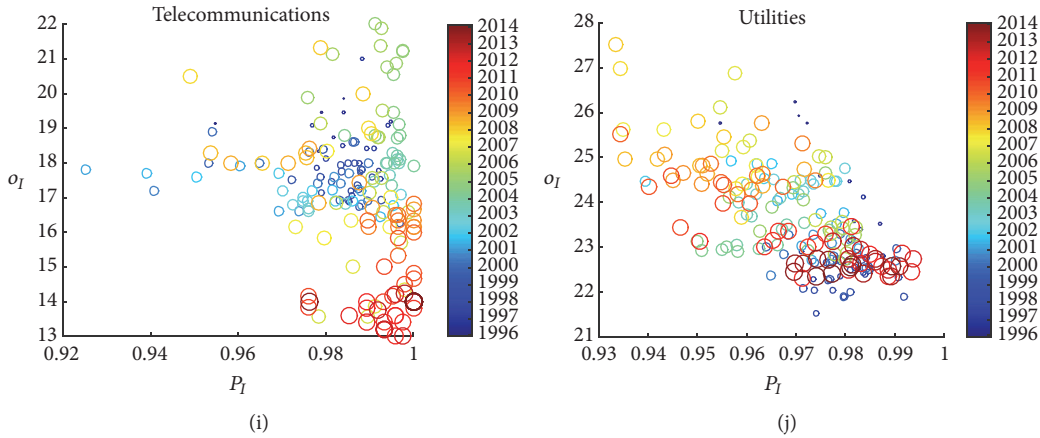
(f)

(g)

(h)

FIGURE 5: Continued.

FIGURE 5: Industries evolution in the overlapping degree/participation coefficient plane. Fixing an industry I, we have plotted for each time window a circle whose y coordinate is the average overlapping degree o_I and whose x coordinate is the average participation coefficient P_I. Points at different times are characterised by different sizes (small to large) and colours (legend on the right). In (a)–(j) we show the results, respectively, for Basic Materials, Consumer Goods, Consumer Services, Financials, Health Care, Industrials, Oil & Gas, Technology, Telecommunications, and Utilities.

stocks have acquired degree on some of the layers, resulting in a considerable decrease of participation coefficient. This is another indication of the importance of monitoring all the layers together, as an increase in the structural role of an industry (as measured by the overlapping degree) is typically due to only a subset of layers (as indicated by the corresponding decrease of the participation coefficient).

Conflicts of Interest

The authors declare no competing financial interests.

Authors' Contributions

Nicolò Musmeci and Vincenzo Nicosia contributed equally to this work. Nicolò Musmeci, Vincenzo Nicosia, Tomaso Aste, Tiziana Di Matteo, and Vito Latora devised the study, performed the experiments and simulations, analysed the results, wrote the paper, and approved the final draft.

Acknowledgments

References

[1] R. Mantegna, "Hierarchical structure in financial markets," *The European Physical Journal B*, vol. 11, no. 1, pp. 193–197, 1999.

[2] M. Tumminello, T. Aste, T. Di Matteo, and R. Mantegna, "A tool for filtering information in complex systems," *Proc. Natl. Acad. Sci*, vol. 102, pp. 10421–10426, 2005.

[3] T. Aste, T. Di Matteo, and S. T. Hyde, "Complex networks on hyperbolic surfaces," *Physica A 346*, 2005.

[4] T. Aste and T. Di Matteo, "Dynamical networks from correlations," *Physica A: Statistical Mechanics and its Applications*, vol. 370, no. 1, pp. 156–161, 2006.

[5] T. Aste, R. Grammatica, and T. Di Matteo, "Exploring complex networks via topological embedding on surfaces," *Phys. Rev. E*, vol. 86, Article ID 036109, 2012.

[6] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. W. Hwang, "Complex networks: Structure and dynamics," *Physics Reports*, vol. 424, no. 4-5, pp. 175–308, 2006.

[7] G. P. Massara, T. Di Matteo, and T. Aste, "Network filtering for big data: Triangulated maximally filtered graph," *J Complex Networks*, vol. 5, no. 2, 2017.

[8] W. Barfuss, G. P. Massara, T. Di Matteo, and T. Aste, "Parsimonious modeling with information filtering networks," *Covering Statistical, Nonlinear, Biological, and Soft Matter Physics*, vol. 94, no. 6, 2016.

[9] P. Embrechts, A. McNeil, and D. Straumann, "Correlation and dependency in risk management: properties and pitfalls," in *Risk Management: Value at Risk and Beyond*, M. Dempster et al., Ed., Cambridge University Press, Cambridge, UK, 2001.

[10] D. Sornette and J. Andersen, "A nonlinear super-exponential rational model of speculative financial bubbles," *Int. J. Mod. Phys. C*, vol. 13, pp. 171–188, 2002.

[11] M. G. Kendall, "A New Measure of Rank Correlation," *Biometrika*, vol. 30, no. 1/2, p. 81, 1938.

[12] G. Meissner, *Correlation Risk Modeling and Management*, Harvard University Press, 2014.

[13] S. Boccaletti et al., "The structure and dynamics of multilayer networks," *Phys Rep*, vol. 544, pp. 1–122, 2014.

[14] F. Battiston, V. Nicosia, and V. Latora, "Metrics for the analysis of multiplex networks," *Phys. Rev. E*, vol. 89, Article ID 032804, 2013.

[15] V. Nicosia and V. Latora, "Measuring and modelling correlations in multiplex networks," *Phys. Rev. E*, vol. 92, Article ID 032805, 2015.

[16] E. Goffman, *Frame Analysis: An Essay on the Organization of Experience*, Harvard University Press, 1974.

[17] M. Szell, R. Lambiotte, and S. Thurner, "Multirelational organization of large-scale social networks in an online world," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 107, no. 31, pp. 13636–13641, 2010.

[18] P. Klimek and S. Thurner, "Triadic closure dynamics drives scaling laws in social multiplex networks," *New Journal of Physics*, vol. 15, Article ID 063008, 2013.

[19] B. Corominas-Murtra, B. Fuchs, and S. Thurner, "Detection of the elite structure in a virtual multiplex social system by means of a generalised K-core," *PLoS ONE*, vol. 9, no. 12, Article ID e112606, 2014.

[20] L. Tang, X. Wang, and H. Liu, "Community detection via heterogeneous interaction analysis," *Data Min. Knowl. Discov*, vol. 25, pp. 1–33, 2012.

[21] T. G. Kolda, B. W. Bader, and J. P. Kenny, "Higher-order web link analysis using multilinear algebra," in *Proceedings of the 5th IEEE International Conference on Data Mining, ICDM 2005*, pp. 242–249, November 2005.

[22] G. A. Barnett, H. W. Park, K. Jiang, C. Tang, and I. F. Aguillo, "A multi-level network analysis of web- citations among the worlds universities," *Scientometrics*, vol. 99, p. 26, 2014.

[23] Z. Wu, W. Yin, J. Cao, G. Xu, and A. Cuzzocrea, "Community detection in multirelational social networks," in *Web Information Systems Engineering WISE 2013*, S. B. Heidelberg, Ed., vol. 8181 of *Lecture Notes in Computer Science*, pp. 43–56, 2014.

[24] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin, "Catastrophic cascade of failures in interdependent networks," *Nature*, vol. 464, no. 7291, pp. 1025–1028, 2010.

[25] C. D. Brummitt, R. M. DSouza, and E. A. Leicht, "Suppressing cascades of load in interdependent networks," *Proc. Natl. Acad. Sci. USA*, vol. 109, pp. E680–E689, 2012.

[26] A. Cardillo, M. Zanin, J. Gómez-Gardeñes, M. Romance, A. García del Amo, and S. Boccaletti, "Modeling the multi-layer nature of the european air transport network: Resilience and passengers re-scheduling under random failures," *Eur. Phys. J. Spec. Top*, vol. 215, pp. 23–33, 2013.

[27] M. Barigozzi, G. Fagiolo, and G. Mangioni, "Identifying the community structure of the international-trade multi-network," *Physica A: Statistical Mechanics and its Applications*, vol. 390, no. 11, pp. 2051–2066, 2011.

[28] M. Montagna and C. Kok, *Multi-layered interbank model for assessing systemic risk*, Kiel, 2013.

[29] R. Burkholz, M. V. Leduc, A. Garas, and F. Schweitzer, "Systemic risk in multiplex networks with asymmetric coupling and threshold feedback," *Physica D: Nonlinear Phenomena*, vol. 323-324, pp. 64–72, 2016.

[30] D. J. Fenn, M. A. Porter, P. J. Mucha et al., "Dynamical clustering of exchange rates," *Quantitative Finance*, vol. 12, no. 10, pp. 1493–1520, 2012.

[31] T. Di Matteo, F. Pozzi, and T. Aste, "The use of dynamical networks to detect the hierarchical organization of financial market sectors," *Eur. Phys. J. B*, vol. 73, pp. 3–11, 2010.

[32] T. Aste, W. Shaw, and T. Di Matteo, "Correlation structure and dynamics in volatile markets," *New J. Phys*, vol. 12, Article ID 085009, 2010.

[33] C. Borghesi, M. Marsili, S. MiccichΦ, and S. Miccichè, "Emergence of time-horizon invariant correlation structure in financial returns by subtraction of the market mode," *Phys. Rev. E*, vol. 76, Article ID 026104, 2007.

[34] N. Musmeci, T. Aste, and T. Di Matteo, "Relation between financial market structure and the real economy," *The PLOS ONE*, 2015.

[35] J.-P. Onnela, A. Chakraborti, K. Kaski, and J. Kertész, "Dynamic asset trees and Black Monday," *Physica A: Statistical Mechanics and its Applications*, vol. 324, no. 1-2, pp. 247–252, 2003.

[36] N. Musmeci, T. Aste, and T. Di Matteo, "Risk diversification: a study of persistence with a filtered correlation-network approach," *Journal of Network Theory in Finance*, vol. 1, pp. 1–22, 2015.

[37] N. Musmeci, T. Aste, and T. Di Matteo, "What does past correlation structure tell us about the future? an answer from network theory," *Portfolio Management*, 2016.

[38] V. Tola, F. Lillo, M. Gallegati, and R. N. Mantegna, "Cluster analysis for portfolio optimization," *Journal of Economic Dynamics and Control*, vol. 32, no. 1, pp. 235–258, 2008.

[39] F. Pozzi, T. Di Matteo, and T. Aste, "Spread of risk across financial markets: better to invest in the peripheries," *Scientific Reports*, vol. 3, 2013.

[40] H. Kaya, "Eccentricity in asset management," *The Journal of Network Theory in Finance*, vol. 1, no. 1, pp. 45–76, 2015.

[41] D. Y. Kenett, M. Tumminello, A. Madi, G. Gur-Gershgoren, R. N. Mantegna, and E. Ben-Jacob, "Dominating clasp of the financial sector revealed by partial correlation analysis of the stock market," *PLoS ONE*, vol. 5, no. 12, Article ID e15032, 2010.

[42] D. B. West, *Introduction to graph theory (Prentice-Hall*, Prentice-Hall, Englewood Cliffs, NJ, USA, 1996.

[43] G. Bianconi, "Statistical mechanics of multiplex networks: entropy and overlap," *Physical Review E—Statistical, Nonlinear, and Soft Matter Physics*, vol. 87, no. 6, Article ID 062806, 2013.

[44] K. Pearson, "Note on regression and inheritance in the case of two parents," *Proceedings of the Royal Society of London*, vol. 58, pp. 240–242, 1895.

[45] F. Pozzi, T. Di Matteo, and T. Aste, "Exponential smoothing weighted correlations," *Eur. Phys. J. B*, vol. 85, 2012.

[46] R. Schmidt and U. Stadmüller, "Nonparametric estimation of tail dependence," *Scandinavian Journal of Statistics*, vol. 33, pp. 307–335, 2006.

[47] K. Baba, R. Shibata, and M. Sibuya, "Partial correlation and conditional correlation as measures of conditional independence," *Australian and New Zealand Journal of Statistics*, vol. 46, no. 4, pp. 657–664, 2004.

Research Article

Evolutionary Network Games: Equilibria from Imitation and Best Response Dynamics

Giulio Cimini[1,2,3]

[1]*IMT School for Advanced Studies, 55100 Lucca, Italy*
[2]*Istituto dei Sistemi Complessi (ISC-CNR), 00185 Rome, Italy*
[3]*Universidad Carlos III, Leganés, 28911 Madrid, Spain*

Correspondence should be addressed to Giulio Cimini; giulio.cimini@imtlucca.it

We consider games of strategic substitutes and complements on networks and introduce two evolutionary dynamics in order to refine their multiplicity of equilibria. Within mean field, we find that for the best-shot game, taken as a representative example of strategic substitutes, replicator-like dynamics does not lead to Nash equilibria, whereas it leads to a unique equilibrium for complements, represented by a coordination game. On the other hand, when the dynamics becomes more cognitively demanding, predictions are always Nash equilibria: for the best-shot game we find a reduced set of equilibria with a definite value of the fraction of contributors, whereas, for the coordination game, symmetric equilibria arise only for low or high initial fractions of cooperators. We further extend our study by considering complex topologies through heterogeneous mean field and show that the nature of the selected equilibria does not change for the best-shot game. However, for coordination games, we reveal an important difference: on infinitely large scale-free networks, cooperative equilibria arise for any value of the incentive to cooperate. Our analytical results are confirmed by numerical simulations and open the question of whether there can be dynamics that consistently leads to stringent equilibria refinements for both classes of games.

1. Introduction

Strategic interactions among individuals located on a network, be it geographical, social, or of any other nature, are becoming increasingly relevant in many economic contexts. Decisions made by our neighbors on the network influence ours and are in turn influenced by their other neighbors to whom we may or may not be connected. Such a framework makes finding the best strategy a very complex problem, almost always plagued by a very large multiplicity of equilibria. Researchers are devoting much effort to this problem, and an increasing body of knowledge is being consolidated [1–3]. In this work we consider games of strategic substitutes and strategic complements on networks, as discussed in [4]. In this paper, Galeotti et al. obtained an important reduction in the number of game equilibria by going from a complete information setting to an incomplete one. They introduced incomplete information by assuming that each player is only aware of the number of neighbors he/she has, but not of their

identity nor of the number of neighbors they have in turn. We here aim at providing an alternative equilibrium refinement by looking at network games from an evolutionary viewpoint. In particular, we look for the set of equilibria which can be accessed according to two different dynamics for players' strategies and discuss the implications of such reduction. Furthermore, we go beyond the state-of-the-art mean field approach and consider the role of complex topologies with a heterogeneous mean field technique.

Our work belongs to the literature on strategic interactions in networks and its applications to economics [5–13]. In particular, one of the games we study is a discrete version of a public goods game proposed by Bramoullé and Kranton [14], who opened the way to the problem of equilibrium selection in this kind of games under complete information. Bramoullé further considered this problem [15] for the case of anticoordination games on networks, showing that network effects are much stronger than for coordination games. As already stated, our paper originates from Galeotti et al. [4],

for they considered one-shot games with strategic complements and substitutes and model equilibria resulting from incomplete information. Our approach is instead based on evolutionary selection of equilibria—pertaining to the large body of work emanating from the Nash programme [16–19]—and is thus complementary to theirs. In particular we focus on the analysis of two evolutionary dynamics (see Roca et al. [20] for a review of the literature) in two representative games and on how this dynamics leads to a refinement of the Nash equilibria or to other final states. The dynamics we consider are Proportional Imitation [21, 22], which does not lead in general to Nash equilibria, and best response [23, 24], which instead allows for convergence to Nash equilibria—an issue about which there are a number of interesting results in the case of a well-mixed population [25–27]. As we are working on a network setup, our specific perspective is close to that of Boncinelli and Pin [28]. They elaborate on the literature on stochastic stability [19, 29] (see [24, 30] for an early example of related dynamics on lattices) as a device that selects the equilibria that are more likely to be observed in the long run, in the presence of small errors occurring with a vanishing probability. They work from the observation [31] that different equilibria can be selected depending on assumptions on the relative likelihood of different types of errors. Thus, Boncinelli and Pin work with a best response dynamics and by means of a Markov Chain analysis find, counterintuitively, that when contributors are the most perturbed players, the selected equilibrium is the one with the highest contribution. The techniques we use here are based on differential equations and have a more dynamical character, and we do not incorporate the possibility of having special distributions of errors—although we do consider random mistakes. Particularly relevant to our work is the paper by López-Pintado [32] (see [33] for an extension to the case of directed networks) where a mean field dynamical approach involving a random subsample of players is proposed. Within this framework, the network is dynamic, as if at each period the network were generated randomly. Then a unique globally stable state of the dynamics is found, although the identities of free riders might change from one period to another. The difference with our work is that we do not deal with a time-dependent subsample of the population, but we use a global mean field approach (possibly depending on the connectivity of individuals) to describe the behavior of a static network.

In the remainder of this introduction we present the games we study and the dynamics we apply for equilibrium refinement in detail, discuss the implications of such a framework on the informational settings we are considering, and summarize our main contributions.

1.1. Framework

1.1.1. Games. We consider a finite set of agents I of cardinality n, linked together in a fixed, undirected, exogenous network. The links between agents reflect social interactions, and connected agents are said to be "neighbors." The network is defined through a $n \times n$ symmetric matrix G with null diagonal, where $G_{ij} = 1$ means that agents i and j are neighbors, while $G_{ij} = 0$ means that they are not. We indicate

with N_i the set of i's neighbors; that is, $N_i = \{j \in I : G_{ij} = 1\}$, where the number of such neighbors $|N_i| = k_i$ is the *degree* of the node.

Each player can take one of two actions $X = \{0, 1\}$, with $x_i \in X$ denoting i's action. Hence, only pure strategies are considered. In our context (particularly for the case of substitutes), action 1 may be interpreted as *cooperating* and action 0 as not doing so—or *defecting*. Thus, the two actions are labeled in the rest of the paper as C and D, respectively. There is a cost c, where $0 < c < 1$, for choosing action $x = 1$, while action $x = 0$ bears no cost.

In what follows we concentrate on two games, the best-shot game and a coordination game, as representative instances of strategic substitutes and strategic complements, respectively. We choose specific examples for the sake of being able to study analytically their dynamics. To define the payoffs we introduce the following notation: $x_{N_i} = \sum_{j \in N_i} x_j$ is the aggregate action in N_i and $y_i = x_i + x_{N_i}$.

(a) Strategic Substitutes: Best-Shot Game. This game was first considered by Bramoullé and Kranton [14] as a model of the local provision of a public good. As stated above, we consider the discrete version, where there are only two actions available, as in [4, 28]. The corresponding payoff function takes the form

$$\pi_i = \Theta_H (y_i - 1) - c x_i, \tag{1}$$

where $\Theta_H(\cdot)$ is the Heaviside step function $\Theta_H(x) = 1$ if $x \geq 0$ and $\Theta_H(x) = 0$ otherwise.

(b) Strategic Complements: Coordination Game. For our second example, we follow Galeotti et al. [4] and consider again a discrete version of the game, but now let the payoffs of any particular agent i be given by

$$\pi_i = \left(\alpha x_{N_i} - c \right) x_i. \tag{2}$$

Assuming that $c > \alpha > 0$, we are faced with a coordination game where, as discussed in [4], depending on the underlying network and the information conditions, there can generally be multiple equilibria.

1.1.2. Dynamics. Within the two games we have presented above, we now consider evolutionary dynamics for players' strategies. Starting at $t = 0$ with a certain fraction $\rho(0) = \sum_i x_i(0)/n$ of players randomly chosen to undertake action $x = 1$, at each round t of the game players collect their payoff $\pi(t)$ according to their neighbors' actions and the kind of game under consideration. Subsequently, a fraction q of players update their strategy. We consider two different mechanisms for strategy updating.

(a) Proportional Imitation (PI) [21, 22]. It represents a rule of imitative nature in which player i may copy the strategy of a selected counterpart j, which is chosen randomly among the k_i neighbors of i. The probability that i copies j's strategy

depends on the difference between the payoffs they obtained in the previous round of the game:

$$\mathscr{P}\left\{x_j(t) \longrightarrow x_i(t+1)\right\}$$

$$= \begin{cases} \dfrac{\left[\pi_j(t) - \pi_i(t)\right]}{\Phi} & \text{if } \pi_j(t) > \pi_i(t) \\ \epsilon & \text{otherwise,} \end{cases} \quad (3)$$

where Φ is a normalization constant that ensures $\mathscr{P}\{\cdot\} \in [0,1]$ and $0 \le \epsilon < 1$ allows for mistakes (i.e., copying an action that yielded less payoff in the previous round). Note that because of the imitation mechanism of PI, the configurations $x_i = 1 \ \forall i$ and $x_i = 0 \ \forall i$ are absorbing states: the system cannot escape from them and not even mistakes can reintroduce strategies, as they always involve imitation. On the other hand, it can be shown that PI is equivalent to the well-known replicator dynamics in the limit of an infinitely large, well-mixed population (equivalently, on a complete graph) [34, 35]. As was first put by Schlag [22], the assumption that agents play a random-matching game in a large population and learn the actual payoff of another randomly chosen agent, along with a rule of action that increases their expected payoff, leads to a probability of switching to the other agent's strategy that is proportional to the difference in payoffs. The corresponding aggregate dynamics is like the replicator dynamics. See also [36] for another interpretation of these dynamics in terms of learning.

(b) Best Response (BR). This rule was introduced in [23, 24] and has been widely used in the economics literature. BR describes players that are rational and choose their strategy (myopically) in order to maximize their payoff, assuming that their neighbors will again do what they did in the last round. This means that each player i, given the past actions of their partners $x_{N_i}(t)$, computes the payoffs that he/she would obtain by choosing action 1 (cooperating) or 0 (defecting) at time t, respectively, $\tilde{\pi}_C(t)$ and $\tilde{\pi}_D(t)$. Then actions are updated as follows:

$$\mathscr{P}\left\{x_i(t+1) = 1\right\} = \begin{cases} 1-\epsilon & \text{if } \tilde{\pi}_C(t) > \tilde{\pi}_D(t) \\ \epsilon & \text{if } \tilde{\pi}_C(t) < \tilde{\pi}_D(t); \end{cases}$$

$$\mathscr{P}\left\{x_i(t+1) = 0\right\} = \begin{cases} \epsilon & \text{if } \tilde{\pi}_C(t) > \tilde{\pi}_D(t) \\ 1-\epsilon & \text{if } \tilde{\pi}_C(t) < \tilde{\pi}_D(t) \end{cases} \quad (4)$$

and $x_i(t+1) = x_i(t)$ if $\tilde{\pi}_C(t) = \tilde{\pi}_D(t)$. Here again $0 \le \epsilon < 1$ represents the probability of making a mistake, with $\epsilon = 0$ indicating fully rational players.

The reason to study these two dynamics is because they may lead to different results as they represent very different evolutions of the players' strategies. In this respect, it is important to mention that, in the case $\epsilon = 0$, Nash equilibria are stable by definition under BR dynamics and, vice versa, any stationary state found by BR is necessarily a Nash equilibrium. On the contrary, with PI this is not always true: even in the absence of mistakes, players can change action

by copying better-performing neighbors, also if such change leads to a decreasing of their payoffs in the next round. Another difference between the two dynamics is the amount of cognitive capability they assume for the players: whereas PI refers to agents with very limited rationality, which imitate a randomly chosen neighbor on the only condition that he/she does better, BR requires agents with a much more developed analytic ability.

1.1.3. Analytical and Informational Settings. We study how the system evolves by either of these two dynamics, starting from an initial random distribution of strategies. In particular, we are interested in the global fraction of cooperators $\rho(t) = \sum_i x_i(t)/n$ and its possible stationary value ρ_s. We carry out our calculations in the framework of a homogeneous mean field (MF) approximation, which is most appropriate to study networks with homogeneous degree distribution $P(k)$ like Erdös-Rényi random graphs [37]. The basic assumption underlying this approach is that every player interacts with an "average player" that represents the actions of his/her neighbors. More formally, the MF approximation consists in assuming that when a player interacts with a neighbor of theirs, the action of such a neighbor is $x = 1$ with probability ρ (and $x = 0$ otherwise), independently of the particular pair of players considered [38]. Loosely speaking, this amounts to having a very incomplete information setting, in which all players know only how many other players they will engage with, and is reminiscent of that used by Galeotti et al. [4] for their refinement of equilibria. However, the analogy is not perfect and therefore, for the sake of accuracy, we do not dwell any further on the matter. In any case, MF represents our setup for most of the paper.

As an extension of the results obtained in the above context, we also study the case of highly heterogeneous networks, that is, networks with broad degree distribution $P(k)$, such as scale-free ones [39]. In these cases in fact there are a number of players with many neighbors ("hubs") and many players with only a few neighbors, and this heterogeneity may give rise to very different behaviors as compared to Erdös-Rényi systems. Analytically, this can be done by means of the heterogeneous mean field technique (HMF) [40] which generalizes, for the case of networks with arbitrary degree distribution, the equations describing the dynamical process by considering degree-block variables grouping nodes within the same degree. More formally, now when a player interacts with a neighbor of theirs, the action of such a neighbor is $x = 1$ with probability ρ_k (and $x = 0$ otherwise) if k is the neighbor's degree (ρ_k is the density of cooperators within players of degree k). By resorting to this second perspective we are able to gain insights on the effects of heterogeneity on the evolutionary dynamics of our games.

1.2. Our Contribution. Within this framework, our main contribution can be summarized as follows. In our basic setup of homogeneous networks (described by the mean field approximation): for the best-shot game, PI leads to a stationary state in which all players play $x_i = 0$, that is, to full defection, which is however non-Nash as any player surrounded by defectors would obtain higher payoff by choosing

cooperation (at odds with the standard version of the public goods game). This is the result also in the presence of mistakes, unless the probability of errors becomes large, in which case the stationary state is the opposite, $x_i = 1$, that is, full cooperation, also non-Nash. Hence, PI does not lead to any refinement of the Nash equilibrium structure. On the contrary, BR leads to Nash equilibria characterized by a finite fraction of cooperators ρ_s, whereas, in the case when players are affected by errors, this fraction coincides with the probability of making an error as the mean degree of the network goes to infinity. The picture is different for the coordination game. In this case, PI does lead to Nash equilibria, selecting the coordination in 0 below a threshold value of α and the opposite state otherwise. This threshold is found to depend on the initial fraction $\rho(0)$ of players choosing $x = 1$. Mistakes lead to the appearance of a new possibility, an intermediate value of the fraction of players choosing 1, and as before the initial value of this fraction governs which equilibrium is selected. BR gives similar results, albeit for the fact that a finite fraction of 1 actions can also be found even without mistakes, and with mistakes the equilibria are not full 0 or 1 but there is always a fraction of mistaken players. Finally, changing the analytical setting by proceeding to the heterogeneous mean field approach does not lead to any dramatic change in the structure of the equilibria for the best-shot game. Interestingly, things change significantly for coordination games—when played on infinitely large scale-free networks. In this case, which is the one where the heterogeneous mean field should make a difference, equilibria with nonvanishing cooperation are obtained for any value of the incentive to cooperate (represented by the parameter α).

The paper is organized in seven sections including this introduction. Section 2 presents our analysis and results for the best-shot game. Section 3 deals with the coordination game. In both cases, the analytical framework is that of the mean field technique. After an overall analysis of global welfare performed in Section 4, Section 5 presents the extensions of the results for both games within the heterogeneous mean field approach, including some background on the formalism itself. Finally, Section 6 contains an assessment of the validity of all these analytical findings in light of the results of recent numerical simulations of the system described above, and Section 7 concludes the paper summarizing our most important findings concerning the refinement of equilibria in network games and pointing to relevant open questions.

2. Best-Shot Game

2.1. Proportional Imitation.
We begin by considering the case of strategic substitutes when imitation of a neighbor is only possible if he/she has obtained better payoff than the focal player; that is, $\epsilon = 0$ in (3). In that case, the main result is the following.

Proposition 1. *Within the mean field formalism, under PI dynamics, when $\epsilon = 0$ the final state for the population is the absorbing state with a density of cooperators $\rho = 0$ (full defection) except if the initial state is full cooperation.*

Proof. Working in a mean field context means that individuals are well-mixed, that is, every player interacts with average players. In this case the differential equation for the density of cooperators ρ is

$$\frac{\dot{\rho}}{q} = (1 - \rho)\,\rho\mathscr{P}_{D \to C} - \rho\,(1 - \rho)\,\mathscr{P}_{C \to D}. \tag{5}$$

The first term is the probability $(1 - \rho)\rho$ of picking a defector with a neighboring cooperator, times the probability of imitation $\mathscr{P}_{D \to C}$. The second term is the probability $\rho(1 - \rho)$ of picking a cooperator with a neighboring defector, times the probability of imitation $\mathscr{P}_{C \to D}$. In the best-shot game a defector cannot copy a neighboring cooperator (who has lower payoff by construction), whereas, a cooperator eventually copies one of his/her neighboring defectors (who has higher payoff). Hence $\mathscr{P}_{D \to C} = 0$ and $\mathscr{P}_{C \to D}$ is equal to the payoff difference $1 - (1 - c) = c$. Since the normalization constant $\Phi = 1$ for strategic substitutes, (5) becomes

$$\frac{\dot{\rho}}{q} = -c\rho\,(1 - \rho). \tag{6}$$

The solution, for any initial condition $0 < \rho(0) < 1$, is

$$\rho(t) = \left\{ 1 + \left[\rho(0)^{-1} - 1\right] e^{cqt} \right\}^{-1}, \tag{7}$$

hence $\rho(t) \to 0$ for $t \to \infty$: the only stationary state is full defection unless $\rho(0) = 1$. $\qquad\square$

Remark 2. As discussed above, PI does not necessarily lead to Nash equilibria as asymptotic, stationary states. This is clear in this case. For any $\rho(0) < 1$ the population ends up in full defection, even if every individual player would be better off by switching to cooperation. This phenomenon is likely to arise from the substitutes or anticoordination character of the game: in a context in which it is best to do the opposite of the other players, imitation does not seem the best way for players to decide on their actions.

Proposition 3. *Within the mean field formalism, under PI dynamics, when $\epsilon \in (0, 1)$ the final state for the population is the absorbing state $\rho = 0$ (full defection) when $\epsilon < c$, $\rho = \rho(0)$ when $\epsilon = c$, and $\rho = 1$ when $\epsilon > c$. When the initial state is $\rho(0) = 0$ or $\rho(0) = 1$, it remains unchanged.*

Proof. Equation (5) is still valid, with $\mathscr{P}_{C \to D}$ unchanged, whereas, $\mathscr{P}_{D \to C} = \epsilon$. By introducing the effective cost $\tilde{c} = c - \epsilon$ we can rewrite (7) as

$$\rho(t) = \left\{ 1 + \left[\rho(0)^{-1} - 1\right] e^{\tilde{c}qt} \right\}^{-1}. \tag{8}$$

Hence $\rho(t) \to 0$ for $t \to \infty$ only for $\tilde{c} > 0$ ($\epsilon < c$) and instead for $\tilde{c} = 0$ ($\epsilon = c$) then $\rho(t) \equiv \rho(0)\ \forall t$, and for $\tilde{c} < 0$ ($\epsilon > c$) then $\rho(t) \to 1$ for $t \to \infty$ (cooperation is favored now). $\quad\square$

Remark 4. As before, PI does not drive the population to a Nash equilibrium, independently of the probability of making a mistake. However, mistakes do introduce a bias towards cooperation and thus a new scenario: when their probability exceeds the cost of cooperating, the whole population ends up cooperating.

2.2. Best Response. We now turn to the case of the best response dynamics, which (at least for $\epsilon = 0$) is guaranteed to drive the system towards Nash equilibria. In this scenario, we have not been able to find a rigorous proof of our main result, but we can make some approximations in the equation that support it. As we will see, our main conclusion is that, within the mean field formalism under BR dynamics, when $\epsilon = 0$ the final state for the population is a mixed state $\rho = \rho_s$, $0 < \rho_s < 1$, for any initial condition.

Indeed, for BR dynamics without mistakes, the homogeneous mean field equation for $\dot\rho$ is

$$\frac{\dot\rho}{q} = -\rho Q\left[\pi_C < \pi_D\right] + (1 - \rho) Q\left[\pi_C > \pi_D\right], \quad (9)$$

where the first term is the probability of picking a cooperator who would do better by defecting, and the second term is the probability of picking a defector who would do better by cooperating. This far, no approximation has been made; however, these two probabilities cannot be exactly computed and we need to estimate them.

To evaluate the two probabilities, we can recall that $\pi_C = 1 - c$ always, whereas $\pi_D = 0$ when none of the neighbors cooperates and $\pi_D = 1$ otherwise. Therefore, for an average player of degree k we have that $Q_k[\pi_C > \pi_D] = (1 - \rho)^k$. Consistently with the mean field framework we are working on, as a rough approximation, we can assume that every player has degree $\bar k$ (the average degree of the network), so that $Q[\pi_C > \pi_D] = 1 - Q[\pi_C < \pi_D] = (1 - \rho)^{\bar k}$. Thus, we have

$$\frac{\dot\rho}{q} = (1 - \rho)^{\bar k} - \rho. \quad (10)$$

To go beyond this simple estimation, we can work out a better approximation by integrating $Q_k[\pi_C > \pi_D]$ over the probability distribution of players' degrees $P(k)$. For Erdös-Rényi random graphs, in the limit of large populations ($n \to \infty$), it is $P(k) \simeq \bar k^k e^{-\bar k}/k!$. This leads to $Q[\pi_C > \pi_D] = e^{-\bar k \rho}$ and, subsequently,

$$\frac{\dot\rho}{q} = e^{-\bar k \rho} - \rho. \quad (11)$$

Remark 5. The precise asymptotic value for the density of cooperators, ρ_s, depends on the approximation considered above. However, at least for networks that are not too inhomogeneous, the approximations turn out to be very good, and therefore the corresponding picture for the evolution of the population is quite accurate. It is interesting to note that, whatever its exact value, in both cases ρ_s is such that the right-hand sides of (10) and (11) vanish and, furthermore, ρ_s is an attractor of the dynamics, because $d(\dot\rho)/d\rho|_{\rho_s} < 0$.

How is the above result modified by mistakes? When $\epsilon \in (0, 1)$, (9) becomes

$$\frac{\dot\rho}{q} = Q\left[\pi_C < \pi_D\right] \left\{-\rho(1 - \epsilon) + (1 - \rho)\epsilon\right\}$$

$$+ Q\left[\pi_C > \pi_D\right] \left\{(1 - \rho)(1 - \epsilon) - \rho\epsilon\right\}$$

$$= Q\left[\pi_C < \pi_D\right] (-\rho + \epsilon)$$

$$+ Q\left[\pi_C > \pi_D\right] (1 - \rho - \epsilon), \quad (12)$$

where the first term accounts for cooperators rightfully switching to defection and defectors wrongly selecting cooperation, while the second term accounts for defectors correctly choosing cooperation and cooperators mistaken to defection. Proceeding as before, and again in the limit $n \to \infty$, we approximate $Q[\pi_C > \pi_D] = e^{-\bar k \rho}$, thus arriving at

$$\frac{\dot\rho}{q} = (1 - 2\epsilon) e^{-\bar k \rho} - (\rho - \epsilon) \quad (13)$$

from which it is possible to find the attractor of the dynamics ρ_s. Such attractor in turn exists if

$$\epsilon < \frac{\left(1 + \bar k^{-1} e^{\bar k \rho_s}\right)}{2}, \quad (14)$$

a threshold that is bounded below by 1/2, which would be tantamount to players choosing their action at random. Therefore, all reasonable values for the probability of errors allow for equilibria.

Remark 6. To gain some insight on the cooperation levels arising from BR dynamics in the Nash equilibria, we have numerically solved (13). The values for ρ_s are plotted in Figure 1 for different values of ϵ, as a function of $\bar k$. We observe that the larger the $\bar k$, the lower the cooperation level. The intuition behind such result is that the more the connections that every player has, the lower the need to play 1 to ensure obtaining a positive payoff. It could then be thought that this conclusion is reminiscent of the equilibria found for best-shot games in [4], which are nonincreasing in the degree. However, this is not the case, as in our work we are considering an iterated game that can perfectly lead to high degree nodes having to cooperate. Note also that this approach leads to a definite value for the density of cooperators in the Nash equilibrium, but there can be many action profiles for the player compatible with that value, so multiplicity of equilibria is reduced but not suppressed.

Remark 7. From Figure 1 it is also apparent that as the likelihood of mistakes increases, the density of cooperators at equilibrium increases. Note that for very large values of the connectivity $\bar k$, (13) has solution $\rho(t) = \rho(0) e^{-qt} + \epsilon$, and thus $\rho_s \equiv \epsilon$, in agreement with the fact that when a player has many neighbors he/she can assume that a fraction ϵ of them will cooperate, thus turning defection into his/her BR.

3. Coordination Game

We now turn to the case of strategic complements, exemplified by our coordination game. As above, we start from the case without mistakes, and we subsequently see how they affect the results.
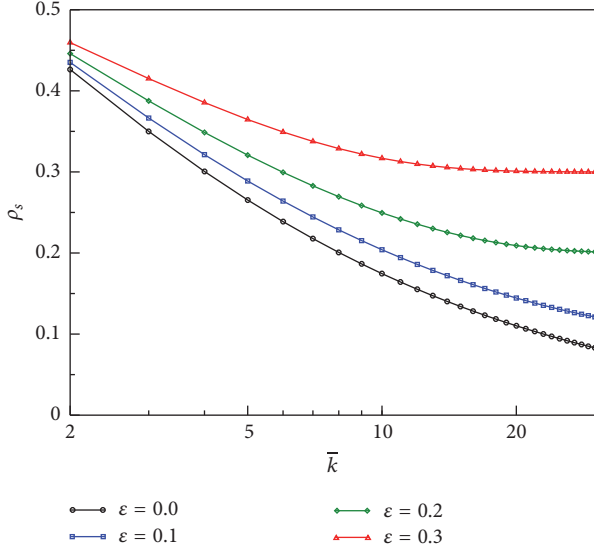
FIGURE 1: Best-shot game under BR dynamics in the mean field framework. Shown are the asymptotic cooperation values ρ_s versus the average degree $\bar{k}$ for different values of the probability of making a mistake ϵ. Values are obtained by numerically solving (13).

3.1. Proportional Imitation

Proposition 8. *Within the mean field formalism, under PI dynamics, when $\epsilon = 0$ the final state for the population is the absorbing state with a density of cooperators $\rho = 0$ (full defection) when $\alpha < \alpha_c \equiv c/[\bar{k}\rho(0)]$, and the absorbing state with $\rho = 1$ when $\alpha > \alpha_c$. In the case $\alpha = \alpha_c$ both outcomes are possible.*

Proof. Still within our homogeneous mean field context, the differential equation for the density of cooperators ρ is again (5). As we are in the case in which $\epsilon = 0$, we have that $\mathcal{P}_{D \to C} = (\pi_C - \pi_D)Q[\pi_C > \pi_D]/\Phi$ and $\mathcal{P}_{C \to D} = (\pi_D - \pi_C)Q[\pi_C < \pi_D]/\Phi = -(\pi_C - \pi_D)(1 - Q[\pi_C > \pi_D])/\Phi$, where for strategic complements $\Phi = \alpha k_{\max}$. Given that $\pi_D = 0$ and that, consistently with our MF framework, $\pi_C = \alpha \bar{k}\rho - c$, we find

$$\frac{\dot{\rho}}{q} = \frac{\rho(1 - \rho)(\alpha\bar{k}\rho - c)}{\Phi} = \frac{c\rho(1 - \rho)(\rho/\rho_c - 1)}{\Phi}, \quad (15)$$

where we have introduced the values $\rho_c = \rho(0)[\alpha_c/\alpha]$ and $\alpha_c \equiv c/[\bar{k}\rho(0)]$.

It is easy to see that $\rho = \rho_c$ is an unstable equilibrium, as $\dot{\rho} < 0$ for $\rho < \rho_c$ and $\dot{\rho} > 0$ for $\rho > \rho_c$. Therefore, we have two different cases: when $\alpha > \alpha_c$ then $\rho_c < \rho(0)$ and the final state is full cooperation ($\rho = 1$), whereas when $\alpha < \alpha_c$ then $\rho_c > \rho(0)$ and the outcome is full defection ($\rho = 0$). When $\alpha \equiv \alpha_c$ then $\rho_c \equiv \rho(0)$, so both outcomes are in principle possible. □

Remark 9. The same (but opposite) intuition we discussed in Remark 2 about the outcome of PI on substitute games suggests that imitation is indeed a good procedure to choose

actions in a coordination setup. In fact, contrary to the case of the best-shot game, in the coordination game PI does lead to Nash equilibria, and indeed it makes a very precise prediction: a unique equilibrium that depends on the initial density. Turning around the condition for the separatrix, we have $\rho(0) < c/(\bar{k}\alpha)$; that is, when few people cooperate initially then evolution leads to everybody defecting, and vice versa. In any event, having a unique equilibrium (except exactly at the separatrix) is a remarkable achievement.

Remark 10. In a system where players may have different degrees, while full defection is always a Nash equilibrium for the coordination game, full cooperation becomes a Nash equilibrium only when $\alpha > c/k_{\min}$, where $k_{\min}$ is the smallest degree in the network—which means that only networks with $k_{\min} > c/\alpha > 1$ feature a fully cooperative Nash equilibrium.

When $\epsilon \in (0, 1)$, the problem becomes much more involved and we have not been able to prove rigorously our main result. In fact, now we have $\mathcal{P}_{D \to C} = (\pi_C - \pi_D)Q[\pi_C > \pi_D]/\Phi + \epsilon Q[\pi_C < \pi_D]$ and $\mathcal{P}_{C \to D} = (\pi_D - \pi_C)Q[\pi_C < \pi_D]/\Phi + \epsilon Q[\pi_C > \pi_D]$. Equation (15) thus becomes

$$\frac{\dot{\rho}}{q} = \rho(1 - \rho)\left\{\frac{c(\rho/\rho_c - 1)}{\Phi} + \epsilon(1 - 2Q[\rho > \rho_c])\right\}, \quad (16)$$

where we have used $Q[\pi_C > 0] \simeq Q[\alpha\bar{k}\rho > c] = Q[\rho > \rho_c]$. We then have three different cases which we can treat approximately:

(i) When $\rho \simeq \rho_c$, then $Q[\rho > \rho_c] \simeq 1/2$ and (16) reduces to (15); that is, we would recover the result for the case with no mistakes.

(ii) When $\rho \gg \rho_c$, then $Q[\rho > \rho_c] \simeq 1$ and (16) can be rewritten as

$$\frac{\dot{\rho}}{q} = \rho(1 - \rho)\left(\frac{c}{\Phi} + \epsilon\right)\left(\frac{\rho}{\rho_+} - 1\right) \quad (17)$$

with $\rho_+ = \rho_c(1 + \Phi\epsilon/c) > \rho_s$. This value ρ_+ leads to an unstable equilibrium; in particular, $\dot{\rho} < 0$ for $\rho < \rho_+$ so that $\rho \to \rho_c$ and hence (16) holds.

(iii) Finally when $\rho \ll \rho_c$, then $Q[\rho > \rho_c] \simeq 0$ and (16) can be rewritten as

$$\frac{\dot{\rho}}{q} = \rho(1 - \rho)\left(\frac{c}{\Phi} - \epsilon\right)\left(\frac{\rho}{\rho_-} - 1\right) \quad (18)$$

with $\rho_- = \rho_c(1 - \Phi\epsilon/c) < \rho_s$. As before, ρ_- gives an unstable equilibrium, because $\dot{\rho} < 0$ for $\rho > \rho_-$ so that again $\rho \to \rho_c$ where (16) holds.

Remark 11. In summary, the region $\rho_- < \rho < \rho_+$ becomes a finite basin of attraction for the dynamics. Note that when $\epsilon > c/(\alpha_c k_{\max})$, then $\rho_+ = \rho(0)$ has no solution and ρ_c becomes the attractor in the whole α space. Our analysis thus shows that, for a range of initial densities of cooperators, there is a dynamical equilibrium characterized by an intermediate value of ρ, which is neither full defection nor full cooperation. Instead, for small enough or large enough values of $\rho(0)$, the

system evolves towards the fully defective or fully cooperative Nash equilibrium, respectively.

Remark 12. The intuition behind the result above could be that mistakes can take a number of people away from the equilibrium, be it full defection or full cooperation, and that this takes place in a range of initial conditions that grows with the likelihood of mistakes.

3.2. Best Response. Considering now the case of BR dynamics, the case of the coordination game is no different from that of the best-shot game and we cannot find rigorous proofs for our results, although we believe that we can substantiate them on firm grounds. To proceed, for this case, (9) becomes

$$\frac{\dot{\rho}}{q} = -\rho + Q\left[\pi_C > 0\right], \tag{19}$$

where we have taken into account that $\pi_D = 0$ and $Q[\pi_C < \pi_D] = 1 - Q[\pi_C > \pi_D]$. Assuming that every node has degree $\bar{k}$, that is, a regular random network, it is clear that there must be at least $[c/\alpha] + 1$ neighboring cooperators in order to have $\pi_C > \pi_D$. Thus

$$Q\left[\pi_C > \pi_D\right] = Q\left[\pi_C > 0\right]$$

$$= \sum_{l=[c/\alpha]+1}^{\bar{k}} \binom{\bar{k}}{l} \rho^l (1-\rho)^{\bar{k}-l}, \tag{20}$$

$$\frac{\dot{\rho}}{q} = -\rho + \sum_{l=[c/\alpha]+1}^{\bar{k}} \binom{\bar{k}}{l} \rho^l (1-\rho)^{\bar{k}-l}.$$

Once again, the difficulty is to show that $\rho_c = \rho(0)(\alpha_c/\alpha)$ is the unstable equilibrium. However, we can follow the same approach used with PI and write $Q[\pi_C > 0] \simeq Q[\alpha\bar{k}\rho > c] = Q[\rho > \rho_c]$; that is, we approximate $Q[\pi_C > 0]$ as a Heaviside step function with threshold in ρ_c. We then again have three different cases as follows:

(i) If $\rho \simeq \rho_c$, then $Q[\rho > \rho_c] \simeq 1/2$: we have $\dot{\rho}/q = -\rho + 1/2$ and the attractor becomes $\rho \equiv 1/2$.

(ii) If $\rho \gg \rho_c$, then $Q[\rho > \rho_c] \simeq 1$: we have $\dot{\rho}/q = -\rho + 1$ and a stable equilibrium at $\rho \equiv 1$.

(iii) Finally if $\rho \ll \rho_c$, then $Q[\rho > \rho_c] \simeq 0$: we have $\dot{\rho}/q = -\rho$ and a stable equilibrium at $\rho \equiv 0$.

Remark 13. As one can see, even without mistakes, BR equilibria with intermediate values of the density of cooperators can be obtained in a range of initial densities. Compared to the situation with PI, in which we only found the absorbing states as equilibria, this points to the fact that more rational players would eventually converge to equilibria with higher payoffs. It is interesting to note that such equilibria could be related to those found by Galeotti et al. [4] in the sense that not everybody in the network chooses the same action; however, we cannot make a more specific connection as we cannot detect which players choose which action—see, however, Section 5.2.2.

A similar approach allows some insight on the situation $\epsilon > 0$. We start again from (12), which now reduces to

$$\frac{\dot{\rho}}{q} = -(\rho - \epsilon) + Q\left[\pi_C > 0\right](1 - 2\epsilon). \tag{21}$$

Approximating as before $Q[\pi_C > 0] \simeq Q[\rho > \rho_c]$ we again have the same three different cases.

(i) If $\rho \simeq \rho_c$, then the attractor $\rho \equiv 1/2$ is unaffected by the particular value of ϵ.

(ii) If $\rho \gg \rho_c$, then the stable equilibrium lies at $\rho \equiv 1 - \epsilon$.

(iii) If $\rho \ll \rho_c$, then the stable equilibrium is at $\rho \equiv \epsilon$.

Remark 14. Adding mistakes to BR does not change dramatically the results, as it did occur with PI. The only relevant change is that equilibria for low or high densities of cooperators are never homogeneous, as there is a percentage of the population that chooses the wrong action. Other than that, in this case the situation is basically the same with a range of densities converging to an intermediate amount of cooperators.

4. Analysis of Global Welfare

Having found the equilibria selected by different evolutionary dynamics, it is interesting to inspect their corresponding welfare (measured in terms of average payoffs). We can again resort to the mean field approximation to approach this problem.

Best-Shot Game. In this case the payoff of player i is given by (1): $\pi_i = \Theta_H(y_i - 1) - cx_i$. Within the mean field approximation, for a generic player i with degree k_i, we can approximate the theta function as $\Theta_H(y_i - 1) \simeq \rho + (1 - \rho)[1 - (1 - \rho)^{k_i}]$, where the first term is the contribution given by player i cooperating ($x_i = 1$), whereas the second term is the contribution of player i defecting ($x_i = 0$) and at least one of i's neighbors cooperating ($x_j = 1$ for at least one $j \in N_i$). It follows easily that

$$\langle \pi \rangle = \sum_k P(k) \left\{ \rho + (1 - \rho)\left[1 - (1 - \rho)^k\right] - c\rho \right\}. \tag{22}$$

If $P(k) = \delta(k - \bar{k})$ (where $\delta(\cdot)$ stands for the Dirac delta function), then $\langle \pi \rangle = 1 - c\rho - (1 - \rho)^{\bar{k}+1}$, whereas, if $P(k) = \bar{k}^k e^{-\bar{k}}/k!$, then $\langle \pi \rangle = 1 - c\rho - (1 - \rho)e^{-\rho\bar{k}}$. We recall that in the simple case where players do not make mistakes ($\epsilon = 0$), PI leads to a stationary cooperation level $\rho \equiv 0$, which corresponds to $\langle \pi \rangle = 0$. On the other hand, with BR the stationary value of ρ_s is given by (10) or (11), both leading to $\langle \pi_s \rangle = 1 - c\rho_s - \rho_s(1 - \rho_s)$. As long as $\rho_s < c$, it is $\langle \pi_s \rangle > 1 - c$ (the payoff of full cooperation). We thus see that under BR players are indeed able to self-organize into states with high values of welfare in a nontrivial manner: defectors are not too many and are placed on the network to allow any of them to be connected to at least one cooperator (and thus to get the payoff $\pi = 1$); this, together with cooperators having $\pi = 1 - c$

by construction, results in a state of higher welfare than full cooperation.

Coordination Game. Now player i's payoff is given by (2): $\pi_i = (\alpha x_{N_i} - c) \, x_i$. Again within the mean field framework we approximate the term x_{N_i} as ρk_i, and we immediately obtain

$$\langle \pi \rangle = \sum_k P(k) \, \rho \left\{ \alpha \rho k - c \right\} = \rho \left(\alpha \rho \overline{k} - c \right). \qquad (23)$$

$\langle \pi \rangle$ is thus a convex function of ρ, which (considering that $0 \leq \rho \leq 1$) attains its maximum value at $\rho = 0$ when $\alpha < \alpha_\pi := c/\overline{k}$, and at $\rho = 1$ for $\alpha > \alpha_\pi$. Recalling that, in the simple case $\epsilon = 0$, with both PI and BR there are two different stationary regimes ($\rho \to 0$ for $\alpha \ll \alpha_c = c/[\rho(0)\overline{k}]$ and $\rho \to 1$ for $\alpha \gg \alpha_c$), we immediately see that for $\alpha > \alpha_c > \alpha_\pi$ the stationary state $\rho = 1$ maximizes welfare, and the same happens for $\alpha < \alpha_\pi$ with $\rho = 0$. However, in the intermediate region $\alpha_\pi < \alpha < \alpha_c$, the stationary state is $\rho = 0$ but payoffs are not optimal.

5. Extension: Higher Heterogeneity of the Network

In the two previous sections we have confined ourselves to the case in which the only information about the network we use is the mean degree, that is, how many neighbors players do interact with on average. However, in many cases, we may consider information on details of the network, such as the degree distribution, and this is relevant as most networks of a given nature (e.g., social) are usually more complex and heterogeneous than Erdös-Rényi random graphs. The heterogeneous mean field (HMF) [40] technique is a very common theoretical tool [41] to deal with the intrinsic heterogeneity of networks. It is the natural generalization of the usual mean field (homogeneous mixing) approach to networks characterized by a broad distribution of the connectivity. The fundamental assumption underlying HMF is that the dynamical state of a vertex depends only on its degree k. In other words, all vertices having the same number of connections have exactly the same dynamical properties. HMF theory can be interpreted also as assuming that the dynamical process takes place on an annealed network [41], that is, a network where connections are completely reshuffled at each time step, with the sole constraints that both the degree distribution $P(k)$ and the conditional probability $P(k \mid k')$ (i.e., the probability that a node of degree k' has a neighbor of degree k, thus encoding topological correlations) remain constant.

Note that in the following HMF calculations we always assume that our network is uncorrelated; that is, $P(k' \mid k) = k'P(k')/\overline{k}$. This is consistent with our minimal informational setting, meaning that it represents the most natural assumption we can make.

5.1. Best-Shot Game

5.1.1. Proportional Imitation. In this framework, considering more complex network topologies does not change the results

we found before, and we again find a final state that is not a Nash equilibrium, namely, full defection.

Proposition 15. *In the HMF setting, under PI dynamics, when $\epsilon = 0$ the final state for the population is the absorbing state with a density of cooperators $\rho = 0$ (full defection) except if the initial state is full cooperation.*

Proof. The HMF technique proceeds by building the k-block variables: we denote by ρ_k the density of cooperators among players of degree k. The differential equation for the density of cooperators ρ_k is

$$\frac{\dot{\rho}_k}{q} = (1 - \rho_k) \sum_{k'} \rho_{k'} P\left(k' \mid k\right) \mathscr{P}_{D \to C}^{kk'}$$
$$- \rho_k \sum_{k'} \left(1 - \rho_{k'}\right) P\left(k' \mid k\right) \mathscr{P}_{C \to D}^{kk'}. \qquad (24)$$

The first term is the probability of picking a defector of degree k with a neighboring cooperator of degree k' times the probability of imitation (all summed over k'), whereas the second term is the probability of picking a cooperator of degree k with a neighboring defector of degree k' times the probability of imitation (again, all summed over k'). For the best-shot game when $\epsilon = 0$, we have

$$\mathscr{P}_{C \to D}^{kk'} = c$$
$$\mathscr{P}_{D \to C}^{kk'} = 0 \qquad (25)$$
$$\forall k, k'.$$

We now introduce these values in (24) and, using the uncorrelated network assumption, we arrive at

$$\frac{\dot{\rho}_k}{q} = -c\rho_k \sum_{k'} \left(1 - \rho_{k'}\right) \frac{k' P\left(k'\right)}{\overline{k}} = -c \left(1 - \Theta\right) \rho_k, \qquad (26)$$

where we have introduced the probability to find a cooperator following a randomly chosen link:

$$\Theta := \sum_{k'} \frac{k' P\left(k'\right) \rho_{k'}}{\overline{k}}. \qquad (27)$$

The corresponding differential equation for Θ reads

$$\dot{\Theta} = \sum_k \frac{k P(k) \dot{\rho}_k}{\overline{k}} = -qc\Theta \left(1 - \Theta\right), \qquad (28)$$

and its solution has the same form of (7):

$$\Theta(t) = \left\{ 1 + \left[\Theta(0)^{-1} - 1 \right] e^{cqt} \right\}^{-1} \qquad (29)$$

with $\Theta(0) \equiv \rho(0)$ as $\rho_k(0) = \rho(0) \; \forall k$. Hence $\Theta(t) \to 0$ for $t \to \infty$ which implies $\rho_k(t) \to 0$ for $t \to \infty$ and $\forall k$. □

Remark 16. For the best-shot game with PI, the particular form of the degree distribution does not change anything. The

outcome of evolution still is full defection, thus indicating that the failure to find a Nash equilibrium arises from the (bounded rational) dynamics and not from the underlying population structure. Again, this suggests that imitation is not a good procedure for the players to decide in this kind of games.

Proposition 17. *In the HMF setting, under PI dynamics, when $\epsilon \in (0, 1)$ the final state for the population is the absorbing state $\rho = 0$ (full defection) when $\epsilon < c$, $\rho = \rho(0)$ when $\epsilon = c$, and $\rho = 1$ when $\epsilon > c$. When the initial state is $\rho(0) = 0$ or $\rho(0) = 1$, it remains unchanged.*

Proof. Equation (24) is still valid, but now $\mathscr{P}_{C \to D}^{kk'} = c$ and $\mathscr{P}_{D \to C}^{kk'} = \epsilon \; \forall k, k'$. Again, using the uncorrelated network assumption, and introducing the effective cost $\widetilde{c} = c - \epsilon$, we arrive at

$$\frac{\dot{\rho}_k}{q} = -c(1 - \Theta)\rho_k + \epsilon\Theta(1 - \rho_k),$$
$$\dot{\Theta} = -q\widetilde{c}\Theta(1 - \Theta), \tag{30}$$

and at the end at a solution of the same form of (29):

$$\Theta(t) = \left\{1 + \left[\Theta(0)^{-1} - 1\right]e^{\widetilde{c}qt}\right\}^{-1} \tag{31}$$

with $\Theta(0) \equiv \rho(0)$. Hence $\Theta(t) \to 0$ for $t \to \infty$ (which implies $\rho_k(t) \to 0$) only for $\widetilde{c} > 0$ ($\epsilon < c$) and instead for $\widetilde{c} = 0$ ($\epsilon = c$) then $\Theta(t) \equiv \Theta(0)$ ($\rho(t) \equiv \rho(0)$) $\forall t$, and for $\widetilde{c} < 0$ ($\epsilon > c$) then $\Theta(t) \to 1$ for $t \to \infty$ (which implies $\rho_k(t) \to 1$). □

5.1.2. Best Response. Always within the deterministic scenario with $\epsilon = 0$, for the case of best response dynamics the differential equation for each of the k-block variables ρ_k has the same form as (9) above, where now to evaluate $Q_k[\pi_C > \pi_D]$ we have to consider the particular values of neighbors' degrees. As before, we consider the uncorrelated network case and introduce the variable Θ from (27). We thus have

$$Q_k[\pi_C > \pi_D] = \left[\sum_{k'}\left(1 - \rho_k'\right)P\left(k' \mid k\right)\right]^k$$
$$= (1 - \Theta)^k,$$
$$\frac{\dot{\rho}_k}{q} = -\rho_k Q_k[\pi_C < \pi_D] \tag{32}$$
$$+ (1 - \rho_k)Q_k[\pi_C > \pi_D]$$
$$= (1 - \Theta)^k - \rho_k.$$

The differential equation for Θ is thus

$$\frac{\dot{\Theta}}{q} = -\Theta + \sum_k \frac{(1 - \Theta)^k kP(k)}{\overline{k}} \tag{33}$$

whose solution depends on the form of degree distribution $P(k)$. Nevertheless, the critical value Θ_s such that the right-hand side of (33) equals zero is also in this case the attractor of the dynamics.
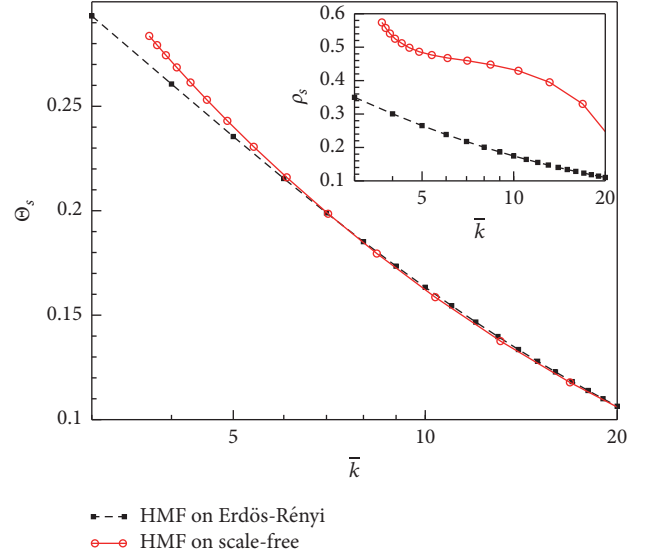

FIGURE 2: Asymptotic value of the cooperator density for the best-shot game with BR dynamics for Erdös-Rényi and scale-free random graphs (with $k_{\min} = 3$ and varying γ): Θ_s (main panel) and ρ_s (inset) versus the average degree $\overline{k}$, obtained by numerically solving (33).

Remark 18. In order to assess the effect of degree heterogeneity, we have plotted in Figure 2 the numerical solution for two random graphs, an Erdös-Rényi graph with a homogeneous degree distribution and a scale-free graph with a much more heterogeneous distribution $P(k) = (\gamma - 1)k_{\min}^{(\gamma-1)}/k^{\gamma}$. In both cases, the networks are uncorrelated so our framework applies. As we can see from the plot, the results are not very different, and they become more similar as the average degree increases. This is related on one hand to the particular form of Nash equilibria for strategic substitutes, where cooperators are generally the nodes with low degree and on the other hand to the fact that the main difference between a homogeneous and a scale-free $P(k)$ lies in the tail of the distribution. In this sense, the nodes with the highest degrees (that can make a difference) do not contribute to Θ_s and thus their effects on the system are negligible.

If we allow for the possibility of mistakes, the starting point of the analysis is—for each of the k-block variables ρ_k—the differential equation given by (9). Recalling that $Q_k[\pi_C > \pi_D] = (1 - \Theta)^k$, we easily arrive at

$$\frac{\dot{\rho}_k}{q} = \epsilon - \rho_k + (1 - 2\epsilon)(1 - \Theta)^k$$
$$\frac{\dot{\Theta}}{q} = \epsilon - \Theta + (1 - 2\epsilon)\sum_k \frac{(1 - \Theta)^k kP(k)}{\overline{k}}. \tag{34}$$

A sufficient condition for the existence of a dynamical attractor Θ_s is $\epsilon < 1/2$: also, in heterogeneous networks, all reasonable values for the probability of errors allow for the existence of stable equilibria.

5.2. Coordination Game. Unfortunately, for the coordination game, working in the HMF framework is much more

complicated, and we have been able to gain only qualitative but important insights on the system's features. For the sake of clarity, we illustrate only the deterministic case in which no mistakes are made ($\epsilon = 0$).

5.2.1. Proportional Imitation.
The average payoffs of cooperating and defecting for players with degree k are

$$\pi_D^k = 0 \quad \forall k;$$

$$\pi_C^k = \alpha k \left[\sum_{k'} P\left(k' \mid k\right) \rho_{k'} \right] - c = \alpha k \Theta - c, \tag{35}$$

where Θ is the same as defined in (27).

We then use our starting point for the HMF formalism, (24), where now the probabilities of imitation are

$$\mathscr{P}_{D \to C}^{kk'} = \frac{\left(\pi_C^{k'} - \pi_D^k\right) Q\left[\pi_C^{k'} > \pi_D^k\right]}{\Phi}$$

$$= \frac{\left(\alpha k' \Theta - c\right) Q_{k'}\left[\pi_C > 0\right]}{\Phi}$$

$$\mathscr{P}_{C \to D}^{kk'} = \frac{\left(\pi_D^{k'} - \pi_C^k\right) Q\left[\pi_D^{k'} > \pi_C^k\right]}{\Phi}$$

$$= -\frac{\left(\alpha k \Theta - c\right) \left\{1 - Q_k\left[\pi_C > 0\right]\right\}}{\Phi}. \tag{36}$$

Once again within the assumption of an uncorrelated network, we find

$$\frac{\Phi \dot{\rho}_k}{q} = \left(1 - \rho_k\right) \sum_{k'} \rho_{k'} \frac{k' P\left(k'\right)}{\overline{k}} \left(\alpha k' \Theta - c\right)$$

$$\cdot Q_{k'}\left[\pi_C > 0\right] + \rho_k \sum_{k'} \left(1 - \rho_{k'}\right) \frac{k' P\left(k'\right)}{\overline{k}} \left(\alpha k \Theta - c\right) \tag{37}$$

$$\cdot \left\{1 - Q_k\left(\pi_C > 0\right)\right\}.$$

In the second term we can carry out the sum over k', which yields $\sum_{k'}(1 - \rho_{k'})k' P(k')/\overline{k} = 1 - \Theta$. We are now ready to write the differential equation for Θ:

$$\frac{\Phi \overline{k} \dot{\Theta}}{q} = \sum_k \frac{k P(k) \dot{\rho}_k}{q} = \sum_k k P(k) \left(1 - \rho_k\right) \sum_{k'} \rho_{k'}$$

$$\cdot \frac{k' P\left(k'\right)}{\overline{k}} \left(\alpha k' \Theta - c\right) Q_{k'}\left[\pi_C > 0\right] + \sum_k k P(k) \tag{38}$$

$$\cdot \rho_k \left(1 - \Theta\right) \left(\alpha k \Theta - c\right) \left\{1 - Q_k\left[\pi_C > 0\right]\right\}.$$

Carrying out the summation over k in the first term (which results again in $1 - \Theta$), and relabeling k' as k, we are left with

$$\frac{\Phi \overline{k} \dot{\Theta}}{q} = \sum_k k P(k) \rho_k \left(1 - \Theta\right) \left(\alpha k \Theta - c\right) Q_k\left[\pi_C > 0\right]$$

$$+ \sum_k k P(k) \rho_k \left(1 - \Theta\right) \left(\alpha k \Theta - c\right) \left\{1 - Q_k\left[\pi_C > 0\right]\right\}$$

$$= \sum_k k P(k) \rho_k \left(1 - \Theta\right) \left(\alpha k \Theta - c\right) = \left(1 - \Theta\right)$$

$$\cdot \Theta \left[\alpha \sum_k k^2 P(k) \rho_k - c \overline{k} \right]. \tag{39}$$

Finally, introducing the new variable

$$\Theta_2 := \sum_{k'} \frac{\left(k'\right)^2 P\left(k'\right) \rho_{k'}}{\overline{k}} \tag{40}$$

we arrive at

$$\frac{\Phi \dot{\Theta}}{q} = \left(1 - \Theta\right) \Theta \left(\alpha \Theta_2 - c\right). \tag{41}$$

Remark 19. While is it difficult to solve (41) in a self-consistent way, qualitative insights can be gained by defining $\alpha_c(t) = c/\Theta_2(t)$, and by rewriting (41) as $\dot{\Theta}(t) \simeq \alpha/\alpha_c(t) - 1$ (the term $(1 - \Theta)\Theta \geq 0$ always and thus can be discarded). Now, starting from the beginning of the game at $t = 0$, the initial conditions $\{\rho_k(0)\}$ univocally determine the value of $\Theta_2(0)$ and thus of $\alpha_c(0)$. For $\alpha < \alpha_c(0)$, $\dot{\Theta}(0) < 0$ and ρ decreases. Because of this, in the next time step $t = 1$ we have on average that $\Theta_2(1) < \Theta_2(0)$, meaning $\alpha_c(1) > \alpha_c(0) > \alpha$: $\dot{\Theta}(1) < 0$ again and ρ keeps decreasing. By iterating such a reasoning, we conclude that in this case the stable equilibrium is $\Theta = 0$. Symmetrically, for $\alpha > \alpha_c(0)$, the attractor becomes $\Theta = 1$, and the transition between the two regimes lies at $\alpha \equiv \alpha_c(0)$. Note that Θ_2 is basically the second momentum of the degree distribution, where each degree k is weighted with the density ρ_k. Recalling that $\overline{k^2}$ may diverge for highly heterogeneous networks (e.g., it diverges for scale-free networks with $\gamma < 3$), and that for the coordination game cooperation is more favorable for players with many neighbors (hence $\rho_k \simeq 1$ for high k), we immediately see that in these cases Θ_2 diverges as well (as the divergence is given by nodes with high degree). Thus, while at the transition point the product $\alpha_c \Theta_2$ remains finite (and equal to c), $\alpha_c \to 0$ to compensate for the divergence of Θ_2 (Figure 3). We can conclude that, in networks with broad $P(k)$ and in the limit $n \to \infty$, cooperation emerges also when the incentive to cooperate (α) vanishes. This is likely to be related to the fact that as the system size goes to infinity, so does the number of neighbors of the largest degree nodes. This drives hubs to cooperate, thus triggering a nonzero level of global cooperation. However, if the network is homogeneous, neither $\overline{k^2}$ nor Θ_2 diverge, so that α_c remains finite and the fully defective state appears also in the limit $n \to \infty$.

5.2.2. Best Response.
For BR dynamics, we would have to begin again from the fact that the differential equation for each of the k-block variables ρ_k has the same form of (19). We would then need to evaluate $Q_k[\pi_C > 0] = \sum_{l=[c/\alpha]+1}^k \binom{k}{l} [\sum_{k'} \rho_{k'} P(k' \mid k)]^l [1 - \sum_{k'} \rho_{k'} P(k' \mid k)]^{k-l} = \sum_{l=[c/\alpha]+1}^k \binom{k}{l} \Theta^l (1-\Theta)^{k-l}$. As in the homogeneous case, such
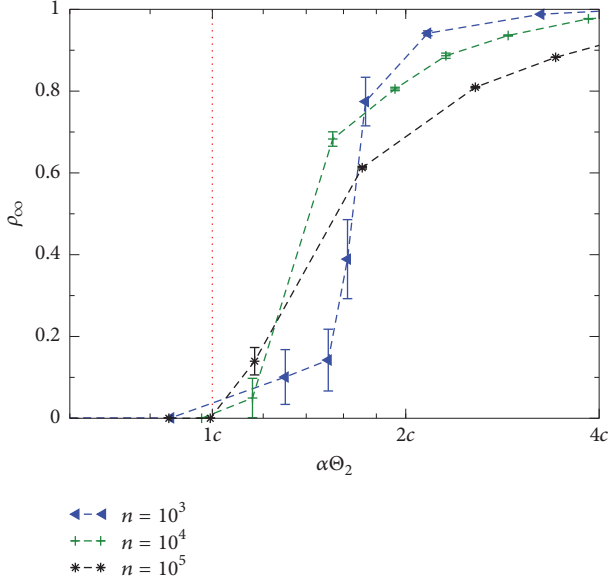
FIGURE 3: Coordination game with PI dynamics for scale-free networks (with $\gamma = 2.5$, $k_{\min} = 3$, and $k_{\max} = \sqrt{n}$): stationary cooperation levels ρ_∞ versus $\alpha\Theta_2$ for various system sizes n. The vertical solid line identifies the critical value of $\alpha_c\Theta_2 = c$.

expression is difficult to treat analytically. Alternatively, we can perform the approximation of setting $Q_k[\pi_C > 0] = Q[\alpha k\Theta > c]$; that is, we approximate $Q_k[\pi_C > 0]$ with a Heaviside step function with threshold in $\Theta = c/(\alpha k)$. This leads to

$$\frac{\dot{\rho}_k}{q} = -\rho_k \quad \text{for } k < \frac{c}{(\alpha\Theta)} \tag{42}$$

$$\frac{\dot{\rho}_k}{q} = -\rho_k + 1 \quad \text{for } k > \frac{c}{(\alpha\Theta)} \tag{43}$$

$$\frac{\dot{\Theta}}{q} = -\Theta + \sum_{k > c/(\alpha\Theta)} \frac{kP(k)}{\bar{k}} \tag{44}$$

and to the following self-consistent equation for the equilibrium Θ_s:

$$\Theta_s = \sum_{k > c/(\alpha\Theta_s)} \frac{kP(k)}{\bar{k}} \tag{45}$$

whose solution strongly depends on the form of degree distribution $P(k)$. Indeed, if the network is highly heterogeneous (e.g., a scale-free network with $2 < \gamma < 3$), it can be shown that Θ_s is a stable equilibrium whose dependence on α is of the form $\Theta_s \sim \alpha^{(\gamma-2)/(3-\gamma)}$; that is, there exists a nonvanishing cooperation level Θ_s no matter how small the value of α. However, if the network is more homogeneous (e.g., $\gamma > 3$), Θ_s becomes unstable and for $\alpha \to 0$ the system always falls in the fully defective Nash equilibria. Another important characterization of such system comes from considering (42) and (43): we have $\rho_k(t) \to 0$ when $k < c/(\alpha\Theta_s)$ and $\rho_k(t) \to 1$ for $k > c/(\alpha\Theta_s)$. In this sense, we find a qualitative agreement

between the features of our equilibria and those found by Galeotti et al. [4], in which players' actions show a monotonic, nondecreasing dependence on their degrees.

6. Comparison with Numerical Simulations

Before discussing and summarizing our results, one question that arises naturally is whether, given that mean field approaches are approximations in so far as they assume interactions with a typical individual (or classes of typical individuals), our results are accurate descriptions of the real dynamics of the system. Therefore, in this section, we present a brief comparison of the analytical results we obtained above with those arising from a complete program of numerical simulations of the system recently carried out by us, whose details can be found in [42] (along with many additional findings on issues that cannot be analytically studied). In this comparison, we focus on the scenario in which mistakes are not allowed ($\epsilon = 0$) as it, being deterministic, allows for a meaningful comparison of theory and simulations without extra effects arising perhaps from poor sampling.

Concerning the best-shot game, numerical simulations fully confirm our analytical results. With PI, the dynamical evolution is in perfect agreement with that predicted by both MF and HMF theory—which indeed coincide when (as in our case) $\rho_k(t = 0)$ does not depend on k. Simulations and analytics agree well also when the dynamics is BR: The final state of the system is, for any initial condition, a Nash equilibrium with cooperators ratio ρ_s (which decreases for increasing network connectivity). Yet, the ρ_s solution of $\dot{\rho} = 0$ from (13) slightly underestimates the one found in simulations—probably because of the approximation made in computing the probabilities Q of (9). Notwithstanding this minor quantitative disagreement, we can safely confirm the validity of our analytical results.

On the other hand, the agreement between theory and simulations is also good for coordination games with PI dynamics. On homogeneous networks, numerical simulations show an abrupt transition from full defection to full cooperation as α crosses a critical value α_T. The MF theory is thus able to qualitatively predict the behavior of the system; furthermore, while α_T is somewhat smaller than the α_c predicted by the theory, simulations also show that $\alpha_T \to \alpha_c$ in the infinite network size, which implies that for reasonably large systems our analytical predictions are accurately fulfilled. Finally, simulations cannot find other Nash equilibrium (with intermediate cooperation levels) than full defection, again as predicted by the MF calculations. On heterogeneous networks instead, simulations show a smooth crossover between full defection and full cooperation, and the point at which the transition starts (α_T) tends to zero as the system size grows. Therefore, the most important prediction of HMF theory, namely, that the fully defective state disappears in the large size limit (a phenomenon not captured by the simple MF approach), is fully confirmed by simulations. Finally, concerning BR dynamics for coordination games, we have a similar scenario: in homogeneous networks, simulations allow finding a sharp transition at α_T from full defection to full cooperation, featuring many

nontrivial Nash equilibria (all characterized by intermediate cooperation levels) in the transient region. This behavior, together with $\alpha_T \to \alpha_c$ in the infinite network size, agrees well with the approximate theoretical results. Heterogeneous networks instead feature a continuous transition, and it appears from numerical simulations that—in the infinite network size—a Nash equilibrium with nonvanishing cooperation level exists no matter how small the value of α, exactly as predicted by the HMF calculations.

We can conclude that the set of analytical results we are presenting in this paper provides, in spite of its approximate character, a very good description of the evolutionary equilibria of our two prototypical games, particularly so when considering the more accurate HMF approach.

7. Conclusion

In this paper, we have presented two evolutionary approaches to two paradigmatic games on networks, namely, the best-shot game and the coordination game as representatives, respectively, of the wider classes of strategic substitutes and complements. As we have seen, using the MF approximation we have been able to prove a number of rigorous results and otherwise to get good insights on the outcome of the evolution. Importantly, numerical simulations support all our conclusions and confirm the validity of our analytical approach to the problem.

Proceeding in order of increasing cognitive demand, we first summarize what we have learned about PI dynamics, equivalent to replicator dynamics in a well-mixed population. For the case of the best-shot game, this dynamics has proven unable to refine the set of Nash equilibria, as it always leads to outcomes that are not Nash. On the other hand, the asymptotic states obtained for the coordination game are Nash equilibria and constitute indeed a drastic refinement, selecting a restricted set of values for the average cooperation. We believe that the difference between these results arises from the fact that PI is imitative dynamics and in a context such as the best-shot game, in which equilibria are not symmetric, this leads to players imitating others who are playing "correctly" in their own context but whose action is not appropriate for the imitator. In the coordination game, where the equilibria should be symmetric, this is not a problem and we find equilibria characterized by a homogeneous action. Note that imitation is quite difficult to justify for rational players (as humans are supposed to act), because it assumes bounded rationality or lack of information leaving players no choice but copying others' strategies [22]. Indeed, imitation is much more apt to model contexts as biological evolution, where payoffs are interpreted as reproductive successes within natural selection [43]. Under this interpretation, in the best-shot game, for instance, it is clear that a cooperator surrounded by defectors would die out and be replaced by the offspring of one of its neighboring defectors.

When going to a more demanding evolutionary rule, BR does lead by construction to Nash equilibria—when players are fully rational and do not make mistakes. We are then able to obtain predictions on the average level of cooperation for the best-shot game but still many possible equilibria are compatible with that value. Predictions are less specific for the coordination game, due to the fact that—in an intermediate range of initial conditions—different equilibria with finite densities of cooperators are found. The general picture remains the same in terms of finding full defection or full cooperation for low or high initial cooperation, but the intermediate region is much more difficult to study.

Besides, we have probed into the issue of degree heterogeneity by considering more complex network topologies. Generally speaking, the results do not change much, at least qualitatively, for any of the dynamics applied to the best-shot game. The coordination game is more difficult to deal with in this context but we were able to show that when the number of connections is very heterogeneous, cooperation may be obtained even if the incentive for cooperation vanishes. This vanishing of the transition point is reminiscent of what occurs for other processes on scale-free networks, such as percolation of epidemic spreading [41]. Interestingly, our results are in contrast with [15], in the sense that—for our dynamical approach—coordination games are more affected by the network (and are henceforth more difficult to tackle) than anticoordination ones.

Finally, a comment is in order about the generality of our results. We believe that the insight on how PI dynamics drives the two types of games studied here should be applicable in general; that is, PI should lead to dramatic reductions of the set of equilibria for strategic complements, but is likely to be off and produce spurious results for strategic substitutes, due to imitation of inappropriate choices of action. On the other hand, BR must produce Nash equilibria, as already stated, leading to significant refinements for strategic substitutes but to only moderate ones for strategic complements. This conclusion hints that different types of dynamics should be considered when refining the equilibria of the two types of games, and raises the question of whether a consistently better refinement could be found with only one dynamics. In addition, our findings also hint to the possible little relevance of the particular network considered on the ability of the dynamics to cut down the number of equilibria. In this respect, it is important to clarify that while our results should apply to a wide class of networks going from homogeneous to extremely heterogeneous, networks with correlations, clustering, or other nontrivial structural properties might behave differently. These are relevant questions for network games that we hope will attract more research in the near future.

Abbreviations

PI: Proportional imitation
BR: Best response
MF: Homogeneous mean field
HMF: Heterogeneous mean field.

Conflicts of Interest

The author declares that there are no conflicts of interest regarding the publication of this paper.

Acknowledgments

References

[1] S. Goyal, *Connections: An Introduction to the Economics of Networks*, Princeton University Press, Princeton, NJ, USA, 2007.

[2] F. Vega-Redondo, *Complex Social Networks*, Econometric Society Monographs, Cambridge University Press, Cambridge, UK, 2007.

[3] M. O. Jackson, *Social and Economic Networks*, Princeton University Press, Princeton, NJ, USA, 2008.

[4] A. Galeotti, S. Goyal, M. O. Jackson, F. Vega-Redondo, and L. Yariv, "Network games," *Review of Economic Studies*, vol. 77, no. 1, pp. 218–244, 2010.

[5] G.-M. Angeletos and A. Pavan, "Efficient use of information and social value of information," *Econometrica: Journal of the Econometric Society*, vol. 75, no. 4, pp. 1103–1142, 2007.

[6] C. Ballester, A. Calvó-Armengol, and Y. Zenou, "Who's who in networks. Wanted: the key player," *Econometrica: Journal of the Econometric Society*, vol. 74, no. 5, pp. 1403–1417, 2006.

[7] C. Ballester and A. Calvó-Armengol, "Interactions with hidden complementarities," *Regional Science and Urban Economics*, vol. 40, no. 6, pp. 397–406, 2010.

[8] D. Bergemann and S. Morris, "Robust implementation in direct mechanisms," *Review of Economic Studies*, vol. 76, no. 4, pp. 1175–1204, 2009.

[9] Y. Bramoullé, R. Kranton, and M. D'Amours, "Strategic interaction and networks," *The American Economic Review*, vol. 104, no. 3, pp. 898–930, 2014.

[10] A. Calvó-Armengol, E. Patacchini, and Y. Zenou, "Peer effects and social networks in education," *Review of Economic Studies*, vol. 76, no. 4, pp. 1239–1267, 2009.

[11] E. L. Glaeser and J. A. Scheinkman, "Non-market interactions," in *Advances in Economics and Econometrics: Theory and Applications*, vol. 1, Cambridge University Press, Cambridge, UK, 2003.

[12] S. Goyal and J. L. Moraga-Gonzalez, "R&D networks," *Rand Journal of Economics*, vol. 32, pp. 686–707, 2001.

[13] X. Vives, *Oligopoly Pricing: Old Ideas and New Tools*, MIT Press, Cambridge, UK, 1999.

[14] Y. Bramoullé and R. Kranton, "Public goods in networks," *Journal of Economic Theory*, vol. 135, no. 1, pp. 478–494, 2007.

[15] Y. Bramoullé, "Anti-coordination and social interactions," *Games and Economic Behavior*, vol. 58, no. 1, pp. 30–49, 2007.

[16] J. F. Nash, *Non-cooperative games [Ph.D. thesis]*, Princeton University, Princeton, NJ, USA, 1950.

[17] A. Mas-Colell, "Bargaining games," in *Cooperation: Game-Theoretic Approaches*, pp. 69–90, Springer, Berlin, Germany, 1997.

[18] D. Fudenberg and D. K. Levine, "Learning in games," *European Economic Review*, vol. 42, no. 3–5, pp. 631–639, 1998.

[19] H. P. Young, *Individual Strategy and Social Structure: An Evolutionary Theory of Institutions*, vol. 93 of *Princeton University Press*, Princeton, NJ, USA, 1999.

[20] C. P. Roca, J. A. Cuesta, and A. Sánchez, "Evolutionary game theory: temporal and spatial effects beyond replicator dynamics," *Physics of Life Reviews*, vol. 6, no. 4, pp. 208–249, 2009.

[21] D. Helbing, "Interrelations between stochastic equations for systems with pair interactions," *Physica A. Statistical and Theoretical Physics*, vol. 181, no. 1-2, pp. 29–52, 1992.

[22] K. H. Schlag, "Why imitate, and if so, how? A boundedly rational approach to multi-armed bandits," *Journal of Economic Theory*, vol. 78, no. 1, pp. 130–156, 1998.

[23] A. Matsui, "Best response dynamics and socially stable strategies," *Journal of Economic Theory*, vol. 57, no. 2, pp. 343–362, 1992.

[24] L. E. Blume, "The statistical mechanics of strategic interaction," *Games and Economic Behavior*, vol. 5, no. 3, pp. 387–424, 1993.

[25] S. Hart and A. Mas-Colell, "A simple adaptive procedure leading to correlated equilibrium," *Econometrica: Journal of the Econometric Society*, vol. 68, no. 5, pp. 1127–1150, 2000.

[26] S. Hart and A. Mas-Colell, "Uncoupled dynamics do not lead to Nash equilibrium," *American Economic Review*, vol. 93, no. 5, pp. 1830–1836, 2003.

[27] S. Hart, "Commentary: Nash equilibrium and dynamics," *Games and Economic Behavior*, vol. 71, no. 1, pp. 6–8, 2011.

[28] L. Boncinelli and P. Pin, "Stochastic stability in best shot network games," *Games and Economic Behavior*, vol. 75, no. 2, pp. 538–554, 2012.

[29] L. E. Blume, "How noise matters," *Games and Economic Behavior*, vol. 44, no. 2, pp. 251–271, 2003.

[30] L. E. Blume, "The statistical mechanics of best-response strategy revision," *Games and Economic Behavior*, vol. 11, no. 2, pp. 111–145, 1995.

[31] J. Bergin and B. L. Lipman, "Evolution with state-dependent mutations," *Econometrica: Journal of the Econometric Society*, vol. 64, no. 4, pp. 943–956, 1996.

[32] D. López-Pintado, "The spread of free-riding behavior in a social network," *Eastern Economic Journal*, vol. 34, no. 4, pp. 464–479, 2008.

[33] D. López-Pintado, "Public goods in directed networks," *Economics Letters*, vol. 121, no. 2, pp. 160–162, 2013.

[34] H. Gintis, *Game Theory Evolving*, Princeton University Press, Princeton, NJ, USA, 2009.

[35] D. K. Levine and W. Pesendorfer, "The evolution of cooperation through imitation," *Games and Economic Behavior*, vol. 58, no. 2, pp. 293–315, 2007.

[36] T. Börgers and R. Sarin, "Learning through reinforcement and replicator dynamics," *Journal of Economic Theory*, vol. 77, no. 1, pp. 1–14, 1997.

[37] SP. Erdös and A. Rényi, "On the evolution of random graphs," *Publications of the Mathematical Institute of the Hungarian Academy of Sciences*, vol. 5, pp. 17–61, 1960.

[38] J.-M. Lasry and P.-L. Lions, "Mean field games," *Japanese Journal of Mathematics*, vol. 2, no. 1, pp. 229–260, 2007.

[39] A.-L. Barabási and R. Albert, "Emergence of scaling in random networks," *Science*, vol. 286, no. 5439, pp. 509–512, 1999.

[40] R. Pastor-Satorras and A. Vespignani, "Epidemic spreading in scale-free networks," *Physical Review Letters*, vol. 86, no. 14, pp. 3200–3203, 2001.

[41] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, "Critical phenomena in complex networks," *Reviews of Modern Physics*, vol. 80, no. 4, pp. 1275–1335, 2008.

[42] G. Cimini, C. Castellano, and A. Sánchez, "Dynamics to equilibrium in network games: Individual behavior and global response," *PLoS ONE*, vol. 10, no. 3, Article ID e0120343, 2015.

[43] J. Maynard Smith and E. Szathmary, *The Major Transitions in Evolution*, Freeman, Oxford, UK, 1995.

Research Article

Predicting the Rise of EU Right-Wing Populism in Response to Unbalanced Immigration

Boris Podobnik,[1,2,3,4,5] **Marko Jusup,**[6] **Dejan Kovac,**[7,8,9] **and H. E. Stanley**[1]

[1]*Boston University, Boston, MA 02215, USA*
[2]*Faculty of Information Studies, SI-8000 Novo Mesto, Slovenia*
[3]*University of Rijeka, 51000 Rijeka, Croatia*
[4]*Zagreb School of Economics and Management, 10 000 Zagreb, Croatia*
[5]*Luxembourg School of Business, 2453 Luxembourg, Luxembourg*
[6]*Center of Mathematics for Social Creativity, Hokkaido University, Sapporo 060-0812, Japan*
[7]*CERGE-EI, Praha 1, 11000 Nové Město, Czech Republic*
[8]*Woodrow Wilson School of Public and International Affairs, Princeton University, Princeton, NJ 08544, USA*
[9]*Adriatic Economic Association, 10000 Zagreb, Croatia*

Correspondence should be addressed to Boris Podobnik; bp@phy.hr

Among the central tenets of globalization is the free migration of labor. Although much has been written about the benefits of globalization, little is known about its limitations and how antiglobalist sentiment can be strongly affected by high levels of immigration. Analyzing poll data from a group of EU countries affected by the recent migrant crisis, we find that over the last three years the percentage of right-wing (RW) populist voters in a given country depends on the prevalence of immigrants in this country's population and the total immigration inflow into the entire EU. The latter is likely due to the perception that the EU functions as a supranational state in which a lack of inner borders means that "someone else's problem" can easily become "my problem." We find that the increase in the percentage of RW voters substantially surpasses the percentage of immigration inflow, implying that if this process continues, ongoing democratic processes will cause RW populism to prevail and globalization to rapidly decrease. We locate tipping points between the fraction of immigrants and the rise of RW populism, and we model our empirical findings using a complex network framework in which the success of globalization rests on a balance between immigration and immigrant integration.

1. Introduction

An important goal of globalization is to allow both capital and labor to move freely across national borders [1–4]. Although one EU country that lacks a sufficient labor force can draw from another EU country that has an overabundance of labor, this economic consideration neglects how movement of people can affect public opinion and alter the outcome of subsequent elections. A volatile situation can arise when either the native majority or the migrant minority sense that their national, ethnic, or religious identity is being threatened. Thus the unprecedented inflow of immigrants into the EU during the recent migrant crisis allows fresh insights into

the relationship between immigration and the popular vote and—by extension—the factors that directly affect the success of further globalization.

Although a large body of literature is dedicated to the analysis of how migration affects the global economy [1–3, 5–10] and right-wing (RW) populism [11–17], much less is known about the limitations of globalization [4, 18], especially how large-scale migrations sway the popular vote and what the economic consequences may be. For example, analyzing national elections in 16 European countries from 1981 to 1998 Swank and Betz reported that the welfare state directly depresses RW populism [9]. Smith reported that RW populist parties benefit from higher levels of crime by linking crime

with higher levels of immigration [16]. Borjas reported that the integration of immigrants into the US was slow and that it took four generations for the earnings of immigrants to equal the earnings of natives, not one or two as commonly believed [19].

Although the recent migrant crisis in the EU was caused by political turmoil and armed confrontation, not globalization, some of the central tenets of globalization have nevertheless been tested. The surge in the population of immigrants has been matched by a surge in voters supporting RW populist parties. The growing number of RW voters across the EU suggests that tolerance towards immigrants is conditional [20] and that members of the general public previously not identifying with RW populism have become supporters. If globalization is to succeed, we must understand this "change of heart." When immigration is more rapid than integration over a prolonged period of time [20], RW populist movements can win elections and this can cause globalization to decline. Apart from concerns about globalization, expected global climate change strongly indicates possible massive displacements in the global population [21].

Why immigration in the past decades did not stir as much RW populism across the EU as the most recent inflow did is perhaps because the EU societies are approaching a tipping point characterized by drastic political and economic upheaval. The term tipping point is used (i) to denote a possible change during which RW populism becomes the ruling political option and (ii) to emphasize a potentially sudden (nonlinear) aspect of such a change that may not even be visible in the available data. It may not be globalization and immigration themselves that are the root of the problem, but rather the high speed at which these processes are occurring right now. Herein, we find that during the last three years the rise of RW populism in a set of EU countries has been substantially more rapid than the immigrant inflow into the EU. We report that the percentage of RW voters in a country significantly depends on the percentage of immigrants within this country and the total immigration inflow into the entire EU. We also find that the occurrence of violent incidents is unrelated to the rate of immigrant inflow and has little effect on the rising percentage of RW voters. We identify tipping points connecting the percentage of immigrants and the popularity of RW populist parties. Globalization is meant to be a tolerant form of democracy in which cooperation between nations supersedes individual national interests, but there are circumstances under which growing RW populist movements can overthrow this tolerant form of democracy. We analyze whether the interconnectedness between countries characteristic of globalization can facilitate cascades (i.e., domino effects) in which growing populist movements within one country trigger similar movements in other countries.

2. Results

It seems reasonable that the causes of populist behavior are, at least partly, nation-dependent. Generally, when a society is relatively far from a tipping point, the nation's interests likely span a multidimensional space. For example, according to Eurobarometer 65 published in 2006, the main concerns of European citizens were unemployment (49%), crime (24%), economic health (23%), immigration (14%), and terrorism (10%). In a concurrent survey in the UK, however, 38% of the respondents listed race and immigration as the top issue. Because RW populism in the UK eventually led to a win over the BREXIT vote, previous surveys suggest an intriguing possibility that, as a society approaches a tipping point, its multidimensional space starts to shrink, finally reducing to a one-dimensional space in which a single issue dominates the current affairs. Inspired by the developments in the UK, we analyze the rise of RW populism in the EU, wherein the recent migrant crisis suggests that the Union is approaching a tipping point.

RW populism often embraces intolerance, which is a widespread social phenomenon that produces conflicts and generates segregation [22–28]. Intolerance combined with radicalization is the main cause of violence and terrorism [29–33]. However, RW populism often shares certain values, for example, antiglobalization, protectionism, and Euroscepticism, with left wing (LW) populism. In principle, people not satisfied by the current government can swing between RW and LW populism. Such a cyclic dynamics may arise as the prevailing issues (e.g., economic versus ethnic) change at a faster-than-generational time scale. Here we focus only on RW populism and the fraction of RW populist voters in response to unbalanced immigration as the driving factor.

For each country affected by the recent migrant crisis, we calculate the percentage of immigrants from September 2013 to September 2016 by combining the official value for 2013 with the number of visa applicants recorded monthly [34, 35]. We then collect the available election poll data and election results for the same range of months [36]. Figure 1(a) shows that in June 2016 in this set of EU countries there is a rising trend in the percentage of RW populist supporters in response to the increasing percentage of immigrants in the general population. We calculated the best linear fit that can be interpreted as the cumulative probability function of a uniform distribution. From this fit we estimate that when the fraction of immigrants reaches approximately 30%, RW populism attains the majority. The slope value of 1.80 is highly statistically significant (t-statistic = 2.62). Besides the linear fit, we relate the fractions of immigrants and RW votes using the cumulative exponential probability function, thus finding that as the percentage of immigrants approaches approximately 22%, the percentage of RW populist voters exceeds 50%, which is again the threshold at which in democratic societies a party can take over the government. The considerable scattering of the data suggests that tolerance levels may differ among countries, and the lower the percentage of immigrants needed to trigger high levels of RW populism in a country, the lower the level of tolerance in that country. Figure 1(b) shows, for example, that in Austria the 50% threshold is reached even when the percentage of immigrants is below 20%.

The current fraction of immigrants in the general population is not the only factor that affects voter sentiment. Figure 2 uses the data for Austria and Germany for 2013–2016 and shows that the growing percentage of RW populist supporters
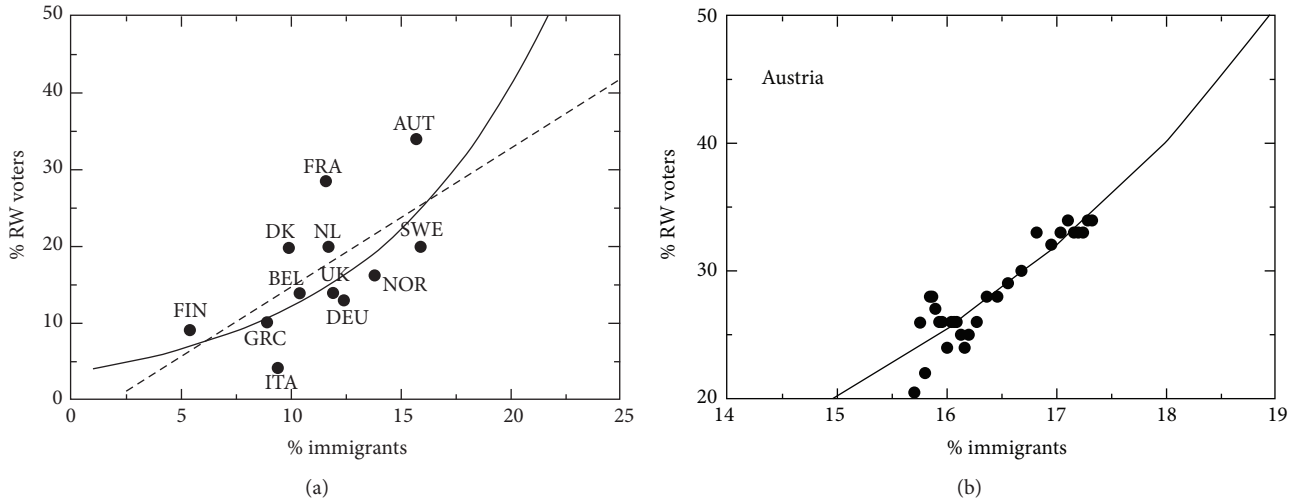
(a)

(b)

FIGURE 1: *Immigration affects the support for right-wing populism I.* (a) Among the EU countries involved in the recent migrant crisis, support for RW populism is generally higher in those countries that accepted a larger number of immigrants relative to the country's population size. Shown is June 2016. Seeing democracy as the majority rule principle, we presume that RW populism becomes a dominant political option when the percentage of RW voters exceeds 50%. Judging based on a cumulative exponential function that fits the data reasonably well—$y = 4.43 \exp(0.097x)$—RW populism in the examined EU countries may take over if the percentage of immigrants in the total population approaches 22%. Shown is also a linear fit $y = -3.67 + 1.80x$ roughly giving that the RW populism reaches the majority for the percentage of immigrants equal to approximately 30%. Coefficients of determination (R^2) for the two models are 0.3866 and 0.4077, respectively. Akaike weights indicate that, given the dataset at hand, the exponential model is better with probability 44.3%, while the linear model is better with the remaining probability 55.7%. (b) Similar as in the other EU countries, Austrian data reveal that the increase in the percentage of immigrants is accompanied with an increase in the percentage of RW populist voters. Here too a cumulative exponential function fits the data well. This function predicts the rise of RW populism in Austria when the percentage of immigrants is slightly below the 20% mark.

is responding to the inflow of immigrants. For example, in Austria the far-right party won 20.5% of the popular vote in 2013 in response to the percentage of immigrants living in Austria at the time, but in the second half of 2015 the inflow of immigrants increased sharply [34, 35] and a local election in Vienna saw the percentage of RW votes jump to 33%. This sudden change suggests the presence of a phase transition, tipping point, or critical point [37, 38]; that is, the nearer the percentage of immigrants comes to the tipping point, the more quickly voters turn to extreme political alternatives. Figures 2(b) and 2(c) show a qualitatively similar phenomenon occurring in Germany.

In an attempt to probe deeper into the internal dynamics of RW populism in the EU as a function of the inflow of immigrants, next we analyze how the immigration rate affects the rise in RW populist voters. In Figure 3, we qualitatively represent the annualized increase in RW votes by taking the differences between the popularity of RW populist parties in September 2016 and September 2013 and subsequently annualizing these differences by dividing them with the number of years in the specified period. Surprisingly, for a group of countries in which the annualized increase in the percentage of RW voters exceeded 2%, Figure 3 shows that this increase is practically independent of the inflow of immigrants. Why would countries with a relatively high and a relatively low inflow of immigrants exhibit about the same increase in the percentage of RW voters? This result

may be a consequence of the EU's political organization. Because the EU functions practically as a supranational state with no internal borders, if one country decides to accept immigrants, this decision may have repercussions for all the other member states. The increase in the percentage of RW populist voters may therefore more systematically depend on the total inflow of immigrants into the entire EU, expressed here as a percentage of the total EU population, than the inflow in any individual country.

Anecdotal evidence to this effect can be seen in the case of Sweden and Norway. Sweden was among the countries hit hard by the recent migrant crisis, yet Norway had approximately the same annualized increase in the percentage of RW voters. A similar occurrence happened in Germany and Poland. Germany experienced a high inflow of immigrants, and in Poland 53% of the population wanted their government to refuse asylum seekers from the Middle East and North Africa, and only 33% thought the opposite. If Poland has already transitioned from the tolerant mode of democracy associated with globalization to a mode dominated by RW populism, then the fraction of immigrants at which the Polish population is pushed beyond the tipping point is much lower than in western EU countries. Poland and Hungary both share decades of socialist experience and are both among the toughest opponents of immigration into the EU. Both strongly oppose EU quotas designed to evenly spread the shock of the migrant crisis.
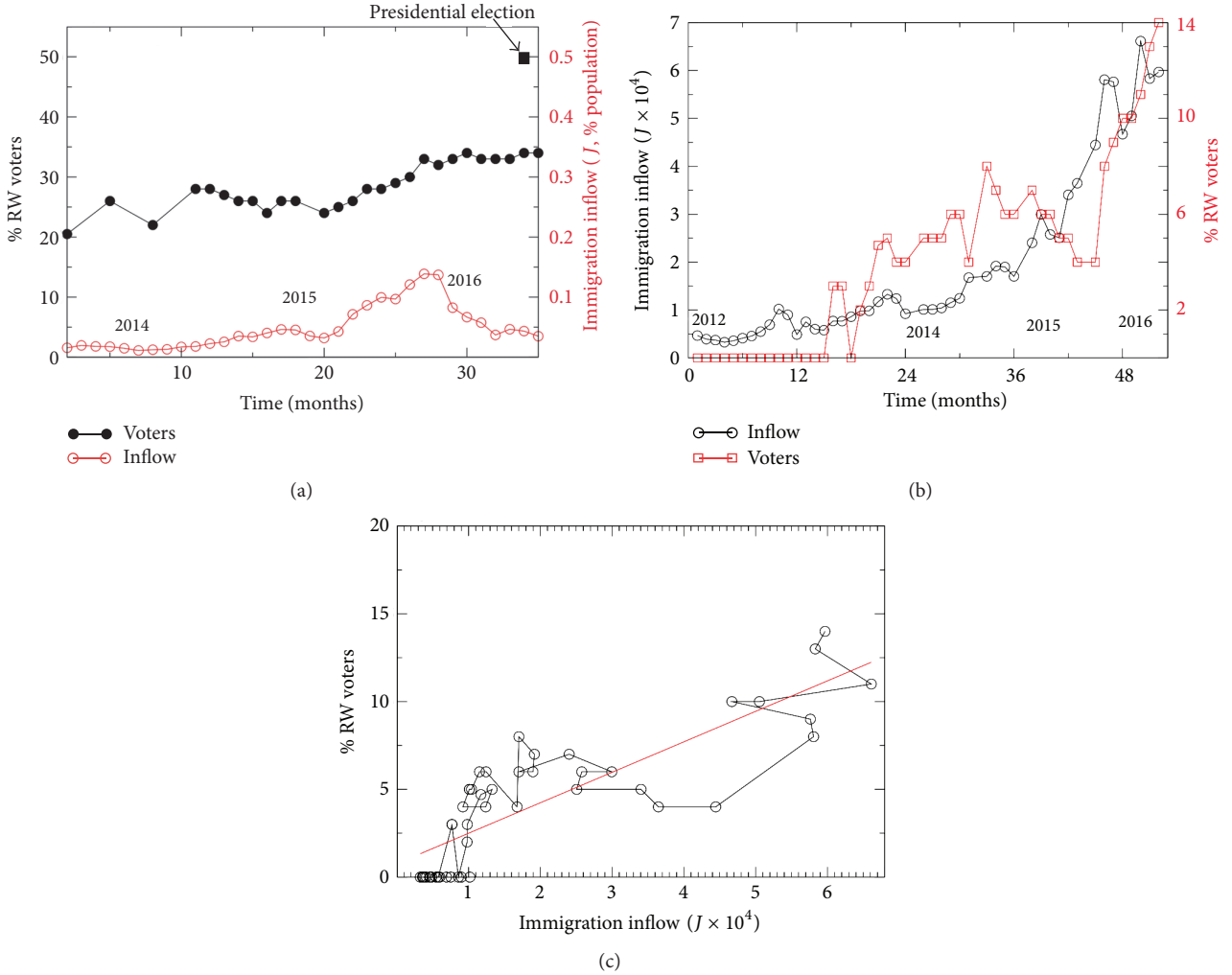
(a)

(b)

(c)

FIGURE 2: *Immigration affects the support for right-wing populism II*. (a) An unprecedented inflow of immigrants into Austria coincided with a steady increase in the fraction of RW populist voters. A solitary black dot represents the results of Austrian presidential election in May 2016 in which an RW populist candidate secured almost 50% of votes. This election shows that even after the record immigrant inflow at the end of 2015 had subsided, a decreasing trend in the number of immigrants that enter Austria did not automatically translate into lower support for the RW populist political option, that is, RW populism seems to be more than just a craze. (b) In Germany, the increasing inflow of immigrants (monthly data [34, 35]) rather clearly coincided with the increasing support for an RW populist party. (c) A significant regression emerges when the German case is presented as a scatter plot between the inflow of immigrants and the percentage of far-right voters.

Figure 3 indicates that the interplay of factors influencing the rising popularity of RW populism is more complex than a simple bivariate regression, and thus we turn to econometric analysis and multivariate regression. Using the results of the simple regression in Figure 1, we assume that the fraction of RW voters (response variable, RW_t) in a given country is determined by the fraction of immigrants (IM_t^L) living in the country. We further assume that the fraction of RW voters depends on the overall inflow of immigrants into the EU (IM_t^{EU}) calculated relative to the total EU population. This variable represents an "immigration shock" in the model. To also take into account the possibility that violent incidents involving immigrants could affect the popular vote, we include the total number of injuries (I) and casualties

(D) involving immigrants recorded across the EU [39] in the model. Additionally, we take into account the unemployment rate (U) that might also affect the popular vote. Finally, we add a variable $M_{it} = (1 - \text{MIPEX}/100)$ in which MIPEX is the migrant integration policy index [40], a proxy for the integration rate—the larger the MIPEX, the better the integration.

We perform econometric analysis using a pooled time-series cross-section (TSCS) method that combines the cross-sectional data on multiple countries. Here the number of countries is $N = 10$, entirely consisting of the so-called old democracies: Germany, France, Austria, Netherlands, Sweden, Norway, Denmark, Finland, Greece, and Italy. Because for each country there are T observations along the temporal dimension, the entire dataset has $N \times T = 370$ observations.
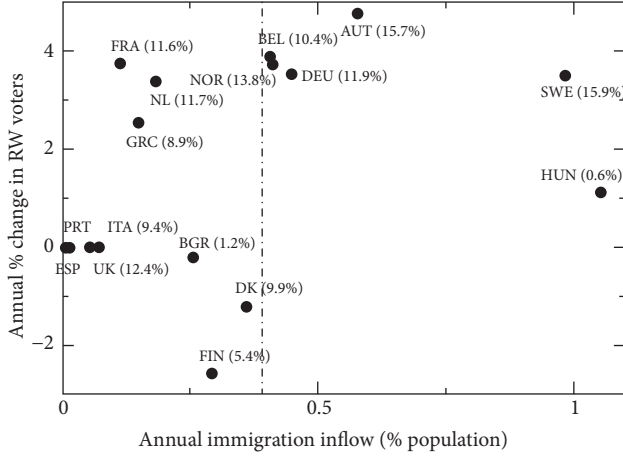
FIGURE 3: *Immigrant inflows and the popularity of right-wing populist movements—a nonlinear threshold.* Shown is the annualized immigrant inflow into a given country (horizontal axis) as a percentage of that country's population, as well as the corresponding percentage change in RW populist votes (vertical axis). In parentheses are the fractions of immigrants in the total population of the corresponding country. For a group of countries in which the annualized increase in the percentage of RW voters exceeded 2%, this increase is virtually independent of the inflow of immigrants. Such a result may reflect the EU's political organization, that is, the lack of internal borders whereby if one country decides to accept immigrants, the decision may have repercussions for all the other member states. We also observe a threshold indicated by a dashed line at which the immigrant inflow into a given country is sufficiently high to invariably provoke an increase in the percentage of RW populist voters. In the model construction, this threshold suggests $\alpha = 0.004$ on an annual basis.

We have an extra index $i = 1, 2, \ldots, N$ that refers to a cross-sectional unit giving

$$
\begin{aligned}
\mathrm{RW}_{it} = \beta_0 &+ \beta_{\mathrm{IM}}^L \mathrm{IM}_{it}^L + \beta_{\mathrm{IM}}^{\mathrm{EU}} \mathrm{IM}_{it}^{\mathrm{EU}} + \beta_D^{\mathrm{ter}} D_{it} + \beta_I^{\mathrm{ter}} I_{it} \\
&+ \beta_U U_{it} + \beta_M M_{it} + e_{it},
\end{aligned}
\tag{1}
$$

where e_t is the random error.

Table 1 shows the results of the TSCS regression model, indicating that the fraction of immigrants in the general population and the immigration inflow into the entire EU are the significant explanatory variables. In addition, the coefficient β_{IM}^L is not only significant but also higher than unity. Perhaps surprisingly, the response variable is not significantly affected by the number of injuries and casualties in violent incidents involving immigrants. We also show that unemployment insignificantly affects the popular vote.

The aforementioned survey data suggest that not every EU nation is equally tolerant to immigrants. We believe that a proxy for this tolerance can be a fraction of RW votes just after the Second World War (RW_0) when the fractions of immigrants were considerably smaller than nowadays. Note that by far the largest fraction of RW votes was recorded in Austria, 11.67%. Table 2 show the results of the TSCS regression model when RW_0 is included. We find that this

TABLE 1: Pooled time series cross-section (TSCS) analysis with random-effects GLS regression as defined in (1). Test statistics: Wald $\chi^2(6) = 147.87$ and Prob $> \chi^2 = 0.000$.

	Coeff.	Std. err.	z	$P > |z|$
β_{IM}^L	2.17	0.527	4.12	0.000
$\beta_{\mathrm{IM}}^{\mathrm{EU}}$	431.1	53.2	8.10	0.000
β_{ter}^D	$-3.7e-04$	$4.9e-04$	-0.78	0.437
β_{ter}^I	$2.1e-04$	$1.6e-04$	1.28	0.202
β_M	0.208	0.212	0.98	0.325
β_U	-0.167	0.348	-0.48	0.630
β_0	-0.298	0.103	-2.89	0.004

TABLE 2: Pooled time series cross-section (TSCS) analysis with random-effects GLS regression as defined in (1). Test statistics: Wald $\chi^2(6) = 151.54$ and Prob $> \chi^2 = 0.000$.

	Coeff.	Std. err.	z	$P > |z|$
β_{IM}^L	1.32	0.445	2.96	0.003
$\beta_{\mathrm{IM}}^{\mathrm{EU}}$	455.9	53.5	8.52	0.000
β_D^{ter}	$-4.2e-04$	$4.9e-04$	-0.86	0.392
β_I^{ter}	$2.2e-04$	$1.6e-04$	1.38	0.167
β_M	-0.019	0.2	-0.09	0.926
β_U	-0.015	0.289	-0.05	0.959
RW_0	0.904	0.562	1.61	0.108
β_0	-0.139	0.093	-1.50	0.133

TABLE 3: Pooled time series cross-section (TSCS) analysis with random-effects GLS regression. Test statistics: Wald $\chi^2(5) = 142.4$ and Prob $> \chi^2 = 0.000$.

	Coeff.	Std. err.	t stat.	$P > t$
β_{IM}^L	2.17	0.466	4.65	0.000
$\beta_{\mathrm{IM}}^{\mathrm{EU}}$	467.2	50.1	9.33	0.000
β_0	-0.235	0.056	-4.15	0.000

new regression insignificantly contributes to modern RW populism.

Table 3 shows the results of the TSCS regression model without MIPEX, unemployment, and violent incidents. Interestingly, with the total inflow of immigrants into the EU of 100,000 on a monthly basis, coefficient values in Table 3 suggest that around 30% of immigrants in the total population of a country are sufficient to cause larger than 50% support for RW populist parties, which corresponds to the result obtained by a linear fit in Figure 1.

The analyzed data do not indicate whether the rise of RW populism is a transient phenomenon or a longer-term change in political orientation. One factor is the persistence of voter memory. In the October 2015 local election in Vienna the RW populist party won 33% of the popular vote. During the presidential election a few months later the migrant crisis had reached a peak and the RW populist movement candidate secured almost 50% of the votes, narrowly losing to a leftist rival. These results have been contested and a new election in December 2016 brought around 48% to the RW populist candidate, indicating that this rise in RW populism is not

short-term. This phenomenon has been described in the literature. Betz [41] describes how a substantial increase in the number of refugees and illegal immigrants in European countries during the 1980s provoked a wave of radical RW populism. Following these events in the early 1990s there remained between 11% and 14% of Europeans who felt the presence of other nationalities, races, and religions to be unsettling [41].

3. Model

Human interactions are often heterogeneous and prone to abrupt nonlinear responses. Because this characterizes the rise of the RW populist party and its RW candidate in the Austrian elections, such linear approaches as the regression in (1) yield only partial results. A useful intuition is gained by thinking about the election system in a democratic country as a random walker. Even in a bipolar political system, not every left government is equally left, and similarly not every right government is equally right. Therefore, as a result of election, the random walker can move either left or right, with the size of this move depending on the standard deviation. Without limitations such that democracy and the election process possess an unlimited tolerance for every political option, after a long enough time the random walker is bound to end up either extremely right or extremely left. Both of these limits exhibit ideological rigidity likely to substantially reduce the level of democracy and tolerance—in agreement with Popper: "unlimited tolerance must lead to the disappearance of tolerance" [42]. Thus, RW populism can be considered as just one of the two random walk limits. However, the random walker intuition does not explain why a society would move left or right nor does it provide a microscopic interpretation of the societal processes at the individual agent level.

To mechanistically characterize the rise of RW populism and account for the existence of tipping points in social dynamics, we use a complex network approach [20, 43–45]. Complex network science is able to emulate heterogeneity in human interactions and goes beyond capturing the dynamics near tipping points. Heterogeneity is important when considering immigration and integration issues, because immigrants in order to sustain their identity live in "hubs" that make them more difficult to integrate than immigrants mixed into the native population.

We construct our model by setting a constant number of native "insider" agents and arranging them in an Erdős-Rényi random network of business and personal contacts. Immigrant "outsider" agents are then added to the network. Each insider notices the percentage of outsiders in their neighborhood and based on this percentage decides whether to be supportive of globalization or RW populism. Insider agents get information from their neighborhood and the interaction is local, and this data is essential in understanding tipping points [46], but there are also other relevant, nonlocal interactions. There is furthermore the factor that information can be misperceived or misinterpreted. In the following, we formalize these concepts with three assumptions.

Assumption i (media and economic influences). At each one-month period of time t, insider agents are influenced by media at a probability rate p and remain influenced for a period τ. We assume that this influence transforms insiders into RW populist supporters. Although media can affect insiders in both directions, we focus on the growth rate and disregard negative values of p. The effect of the media is global, and hearing that immigration is occurring can transform some insiders into RW populist supporters irrespective of their local situation. For example, most likely as the effect of media coverage of immigration to both the EU and the UK, during the BREXIT referendum, most UK districts with low immigration voted mainly for Leave [47]. The media, alongside a lower level of tolerance to immigration, can be an important reason why, in ex-socialist countries, the large fraction of voters oppose receiving even a low overall fraction of immigrants. The probability rate p can also reflect such economic factors as unemployment. Thus we use equation [43] $p^* = 1 - \exp(-p\tau)$ to calculate the probability that a randomly chosen insider agent is being influenced by the media.

Assumption ii (local influence of outsiders). In local elections in Greece in November 2010, although the far-right Golden Dawn party received only 5.3% of the vote, in some neighborhoods of Athens with large immigrant communities the party won nearly 20% [48]. This suggests that contacts between insiders and outsiders do matter. Within our network model, we maintain a constant number of N insiders and then add $I(0)$ outsiders. We then increase the number of outsiders with an inflow J_t at each moment t (representing one month). We randomly place newly arriving outsiders between insider agents both of whom initially have an average number of connections (i.e., a degree) k. The total number of outsiders $I(t)$ is obtained by summing the monthly J values according to $I(t) = I(0) + \sum_{s=1}^{t} J_s$. At any given moment the fraction of outsiders will equal $f_I(t) = I(t)/(N + I(t))$. To account for the effect of contacts between insiders and outsiders, we assume that any insider agent i with k_i total connections turns to RW populism at a rate p' when this agent is surrounded by at least $m_i = f_I' k_i$ outsiders [43, 46, 49], where $0 < f_I' < 1$ is a constant model parameter quantifying the toleration of insiders. This assumption merits a few additional comments.

First, the probability that randomly chosen insider agent i with k_i connections is surrounded by m_i outsiders and therefore prone to RW populism is $p_1(k_i, m_i, f_I) \equiv \sum_{j=m_i}^{k_i} f_I^j (1 - f_I)^{k_i-j} \binom{k_i}{k_i-j}$. In this formula, f_I is the true current state of the network. However the information may be biased and cause insiders to think there are more outsiders than is actually the case. If the bias is Δf_I, then $p_1(k_i, m_i, f_I + \Delta f_I) > p(k_i, m_i, f_I)$. This increased probability $p(k_i, m_i, f_I + \Delta f_I)$ implies that the tolerance parameter f_I' must decrease by amount $\Delta f_I'$, which we estimate using condition $p(k_i, m_i - \Delta f_I' k_i, f_I) = p(k_i, m_i, f_I + \Delta f_I)$. An implicit assumption here is that all insider agents are equally tolerant to immigrants because the tolerance parameter f_I' is defined as a global network property rather than an individual agent property.

An alternative would be to assume a distribution of tolerance levels, in which case f_I' would represent the mean [20].

We now expand assumption (ii) with extension (A): when the immigration inflow J is below some threshold J' the society becomes more tolerant. When $J < J'$ at a given time moment, the tolerance parameter f_I' increases by an amount $\Delta f_I' = \delta > 0$; there is a balance between immigration and immigrant integration—outsiders are successfully integrated—and insiders are able to acclimate to the changes in their society. Figure 2(b) shows that the J' value for Germany is approximately 10,000 people per month. Being able to accurately determine the maximum J' value is highly relevant to the success of globalization. According to our model, immigration and integration can be the inevitable consequences of globalization when $J < J'$. If this is not the case, globalization will be threatened by the rise of RW populism, the democratic system will enter an intolerant mode, and cooperation between nations will be downgraded on the list of political priorities.

Empirical evidence suggests that we add an opposite extension (B): as the inflow of outsiders J crosses some threshold J'' (which does not need to be equal to J') society becomes less tolerant. Mathematically, extension (B) indicates that when $J > J''$, the tolerance parameter f_I' decreases by

$$\Delta f_I' = -\gamma J. \qquad (2)$$

Using the econometric models in (1), we decrease the tolerance parameter proportional to inflow J, where γ is a proportionality coefficient expressing the sensitivity of insiders to high levels of outsider inflow. Figure 3(b) shows threshold J'' in terms of total population; that is, $J'' = \alpha N$, where α is another proportionality coefficient. Figure 3(b) estimates the α value when all of the EU countries are hit by the migrant crisis. The dashed line is annual inflow below which countries have a mixed response to immigration and above which support for RW populism increases. Because $\alpha = J''/N$ and, using Figure 3(b), $12J''/N \approx 0.004$, we obtain

$$\alpha \approx 0.00033. \qquad (3)$$

Extensions (A) and (B) are opposite limiting cases, one in which immigration is slow and the other in which immigration is rapid. Apart from the empirical evidence that these extensions are needed, brain science, for example, offers a physiological interpretation: political attitudes have a counterpart in brain structure [50–52]. If outsiders increase at a rate and in a manner perceived as controllable by insiders, the prefrontal cortex of the human brain responsible for decision making and for moderating social behavior acclimates to the new circumstances, but if the insiders perceive the outsiders to be invaders, the prefrontal cortex is supplanted by the amygdala, which induces a fighting reaction, and tolerance is suppressed.

Although with assumptions (i) and (ii) our model accounts for the processes that affect individual insider opinion, the expansion of RW populism can become extremely rapid when insiders are influenced by their peers. This well-documented phenomenon in human interactions is further accelerated when social media is added. Thus the spread of RW populism can be highly nonlinear, much like the spread of a highly contagious disease. We include this nonlinear collective spreading mechanism in our third model assumption.

Assumption iii (mutual insider contagion). At any given moment t, an insider agent i with K_i connections to other insiders turns to RW populism at rate p'' if at time $t - 1$ this agent has at least $M_i = K_i/2$ RW populist supporters in their neighborhood. Note that, for simplicity, factor 1/2 plays a role analogous to the tolerance parameter in assumption (ii). Because of connections between insider agents, when RW populism emerges anywhere in the network the populist movement is able to spread like a contagion. This collective spreading indicates that insider agents can become RW populist supporters even when there are no outsiders in the immediate neighborhood. Thus some EU countries with almost no immigration are opposed to accepting even a small group of immigrants. This may have affected the outcome of the recent US presidential election in which the winning candidate was often ridiculed in the mainstream media—an attitude represented in our model by assumption (i).

The model here assumes imitative interactions, without testing the validity of such an assumption. However, this assumption could be tested by the method of Agliari et al. [53] if the required data were available. The authors use the tools from statistical mechanics to determine from data the nature of interactions in social customs such as local and mixed marriages in Italy and neighboring European countries. The fraction of actual marriages satisfying a given condition scales differently with the fraction of all possible couples that satisfy the same condition depending on the nature of interactions between agents. If the interactions are independent (one-body model), the scaling is linear, whereas if the interactions are imitative (two-body model), the scaling is square-rooted.

We now turn to the simulation results and their implication. Figure 4 shows how in a network of 5000 agents the fraction of RW populist supporters increases when there is a constant inflow of outsiders; here $J = 2$ per month. This inflow is annually approximately 0.5% of the total population, which is slightly more than the threshold value implied in Figure 3(b). Simulations with $J = 2$ per month are designed to emulate the rapid limit in extension (B) described above. After approximately 37 years of rapid globalization, the network reaches a tipping point and abruptly shifts to a mode dominated by RW populism. RW populism dominates when more than 50% of the network is made up of RW populist supporters (i.e., $P > 0.5$, where P denotes the fraction of RW populists). The threshold is 50% because in a democracy the majority rules.

Figure 4 also shows how the simulated network under assumptions (i)–(iii) responds to shocks (red curve). At this stage, extension (B) does not yet operate. The constant annual inflow of outsiders, $J = 2$, is supplemented by two events at times t_1 and t_2, when $J = 200$. The state of the network, characterized by the proportion of RW populists (P), exhibits a much stronger response at t_2 than at t_1, although the shock inflow ($J = 200$) is the same. This occurs because at t_2 the
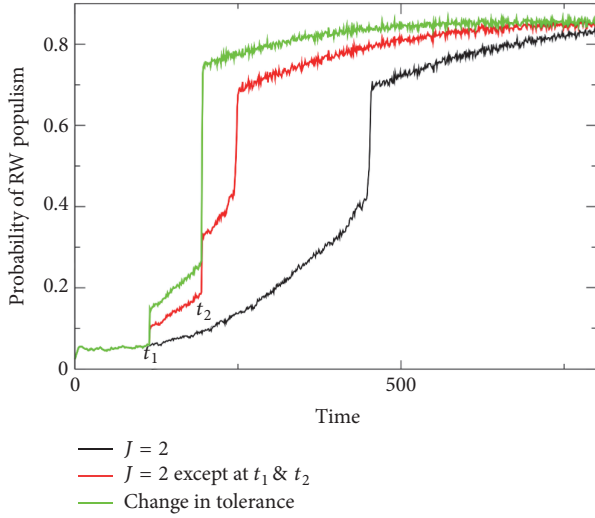
Figure 4: *Nonlinearity, a tipping point, and the rise of right-wing populism*. Using a network of N = 5,000 agents, each with an average of 15 connections, we examine the effect of a constant inflow of outsiders at rate J = 2 at each time step. In this setup, the total number of outsiders at any moment in time is $I(t)$ = $\sum_i J_i$ = $\langle J \rangle t$. As the fraction of outsiders, f_I = $I/(N + I)$, approaches the tolerance parameter, $f_I' $ = 0.15, the presence of a tipping point causes the fraction of RW populist supporters to start increasing nonlinearly and eventually undergo a sudden jump (i.e., a discontinuous change) at about 37 years (450 months) into the simulation (black curve). The sudden jump happens much earlier if the inflow of outsiders experiences shocks at times t_1 and t_2 at which J = 200 outsiders enter the network. In particular, as the network approaches the tipping point, the effect of exactly the same shock becomes disproportionately higher (red curve). In this case, however, the tolerance parameter is still kept constant. Finally, we also examine the case in which shocks at times t_1 and t_2 affect the tolerance parameter, where responsiveness is controlled by parameter γ = 0.0001. Here, the second shock at t_2 is sufficient to instantly tip the network into RW populism (green curve). Other parameters are p = 0.007, τ = 15, p' = 0.5, p'' = 0.5, and α = 0.001.

system is closer to the tipping point and consequently more unstable than at t_1. Because in real-world data the value of J can be biased due to estimation errors or misinterpreted information, our results suggest that approaching the tipping point can be concurrent with strong nonlinear effects such that even a small shock can trigger a transition to a mode dominated by RW populism. This can be even more explosive (in terms of P) if extension (B) is allowed to operate, that is, if the tolerance parameter f_I' changes with J.

Figure 4 shows a third simulation (green curve) in which the dynamics operate under assumptions (i)–(iii) with extension (B). The tolerance parameter f_I thus changes with J as prescribed by (2). Because of the decreasing tolerance, the second shock at t_2 can now push the system beyond the tipping point, and thus the dominance of RW populism occurs earlier than in the two previous simulations.

From the simulations in Figure 4 alone, it is unclear how much the local influence of outsiders (assumption (ii)) contributes to the rise of RW populism relative to the insider
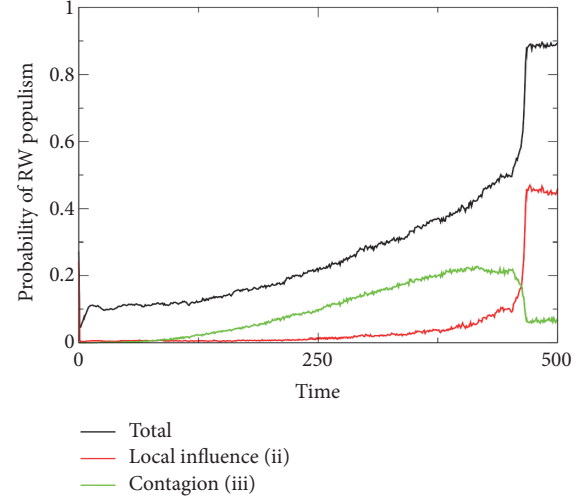

Figure 5: *Breakdown of the causes of right-wing populism*. Figure 4 shows that the probability of RW populism, P, suddenly increases as society approaches a tipping point but remains silent on the underlying causes. Here we discern between the contributions of local outsider influence (assumption (ii)) and mutual insider contagion (assumption (iii)). Far from the tipping point, P mainly responds to local outsider influence (ii). By contrast, as the network approaches its tipping point, mutual insider contagion (iii) takes over and accelerates the transition to RW populist dominance. Parameter values are N = 5,000 with an average degree of 15, J = 2, p = 0.007, τ = 15, p' = 0.7, and p'' = 0.8.

contagion (assumption (iii)). Figure 5 shows these contributions. After the initial transients fade, the local influence of outsiders drives the increase in RW populists in the network. The contribution of mutual insider contagion is relatively small until the system approaches a tipping point. Near the tipping point, contagion spreads rapidly and overtakes the local outsider influence as the main contributor to the rise of RW populism. Thereafter the RW populist movement can sustain itself without support from the outside.

In the regime of moderate immigration inflows ($J' < J < J''$), we can examine the dynamics of our complex network using the mean-field theory (MFT) analytic technique. When the number of agent connections does not deviate greatly from the network average, the probability P that a randomly chosen insider agent i is an RW populist supporter due to any of the processes underlying assumptions (i)–(iii) is

$$P = p^* + p' p_1 (k_i, m_i, f_I) + p'' p_1 (K_i, M_i, P)$$
$$- p^* p' p_1 (k_i, m_i, f_I) - p^* p'' p_1 (K_i, M_i, P) \quad (4)$$
$$- p' p_1 (k_i, m_i, f_I) p'' p_1 (K_i, M_i, P),$$

where the last three terms avoid double counting in accordance with the probability theory formula $P(A \cup B \cup C)$ = $P(A) + P(B) + P(C) - P(A)P(B) - P(A)P(C) - P(B)P(C)$ for three mutually independent events A, B, and C that cannot occur simultaneously. In the MFT approximation, we can drop index i because no single agent is markedly different from the collective average. Previously we set M = $K/2$
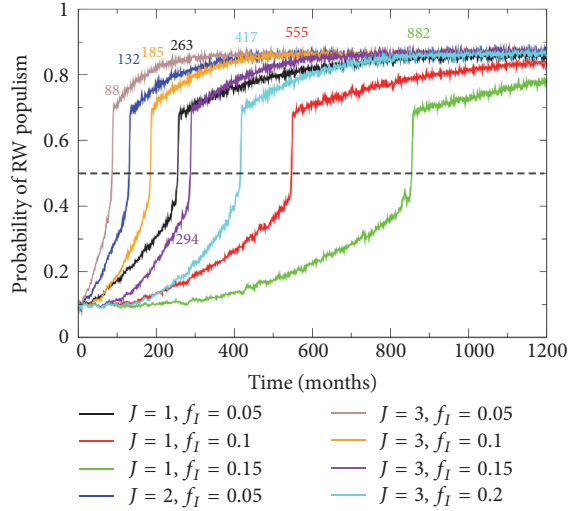
FIGURE 6: *Predicting the timing of RW populism*. We find that for a broad range of outsider inflows (J) and tolerance parameter values (f_I'), (5) predicts the moment at which $P > 0.5$ in a manner that agrees favorably with the simulation results. Except for $\gamma = 0$, other parameters are the same as in Figure 4.

for simplicity, but the peer pressure measured by the value of the proportionality factor between M and K can differ between countries or regions. In addition, parameters p' and p'' are constants only in theory. The real social dynamics are such that these parameters may change in response to rumors, political manipulations, or outside shocks. If parameters p' and p'' substantially increase, the value of P also increases, thus further improving the prospects for RW populism dominance. In our framework, due to democracy's majority rule principle, when P approaches 0.5 the nonlinear processes embedded in assumptions (ii) and (iii) cause a sudden transition to RW populist mode.

Because a theoretical model is more useful if it has predictive power [54, 55], we show how a network of agents under assumptions (i)–(iii) leads to a simple formula for the timing at which RW populism starts to dominate,

$$t_{th} = \frac{N f_I'}{J(1 - f_I')} - \frac{I(0)}{J}. \quad (5)$$

Gaining this result involves three steps. First, if the immigration inflow is constant, then the number of outsiders in the network after t time steps is $I(0) + Jt$. Second, the total population size thus equals $N + I(0) + Jt$. Finally, (5) follows if the current fraction of outsiders $(I(0) + Jt)/(N + I(0) + Jt)$ is equated with the critical parameter f_I'. Figure 6 shows that, for a number of immigration inflow-tolerance parameter pairs (J, f_I'), the simulated timing of the shift to RW populist mode (i.e., $P > 0.5$) fits theoretical predictions. In conjunction with the empirical data on tolerance towards immigrants in the EU countries, the formula in (5) could be used to provide an estimate of when a given country might be approaching a possible tipping point to RW populism dominance.

Numerical simulations allow us to examine not only isolated single networks, but also the interdependence between two or more networks. Note that there is a potential for a cascade effect when an RW populist movement in one network causes the rise of RW populist movements in other networks. This cascade effect is growing in relevance because expanding globalization is causing countries to become increasingly similar to each another, and this similarity increases in such supranational organizations as the EU in which borders between nation states are rapidly fading.

To examine how interdependence affects the rise of RW populism, we set up two random, economically equivalent ER networks (equal p in the model) with different tolerance levels towards outsiders (different f_I' in the model). To the usual intraconnections within each network we add interconnecting agents that link with their counterparts in the other network. Apart from this addition, we retain the same model assumptions as previously held.

To examine the effects of interconnectedness, we first run numerical simulations of two independent networks without interconnections [see Figure 7(a)]. As expected, when the inflow level of outsiders is the same ($J = 2$) the network with a higher tolerance parameter ($f_{I,1}' = 0.4$) reaches the tipping point much later than the network with a lower tolerance parameter ($f_{I,2}' = 0.2$). When the networks are interconnected and the more tolerant network experiences an increased inflow ($J_2 = 4$) and a shock ($J = 500$) at time $t_1 = 500$ its susceptibility to RW populism increases and it also affects the other network and shortens the time of transition to RW populism [see Figure 7(b)]. It would appear that countries do not want to be the first to cross the line, but in an interconnected world being second is easier.

4. Discussion and Conclusion

Why some countries (e.g., the ex-socialist EU countries) strongly oppose receiving immigrants while others (e.g., the USA) have a long history of receiving immigrants is an important topic in the social sciences and, more recently, a major issue for the EU. Perhaps receiving immigrants has been an ongoing pattern in the USA because there is no single dominant ethnicity and thus a clear distinction is made between USA national identity and the country of origin of its citizens. In addition, because USA citizens are people from all over the world, there is no single dominant religious, ethnic, or cultural group that can organize and threaten the established social order. In France we find the opposite. There is a large group of immigrants with different language and religion from the French majority, whose presence can instill fear among the majority. Fear exacerbated by an inflow rate of immigrants that exceeds the rate of their integration can lead to a volatile situation, one that is often resolved in one of two ways. Either there is an immigrant uprising as exemplified by the Visigoth immigrants and their ex-Roman commander Alaric who plundered Rome in 410 or the majority population suppresses the inflow of immigrants. Currently this second option is often accomplished by supporting populist right-wing parties.
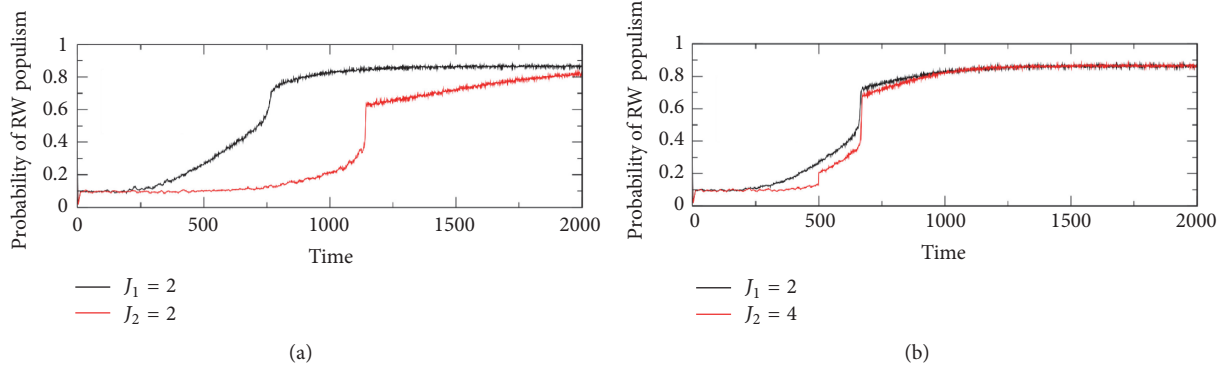
Although globalization was conceived to allow capital and labor to move freely across national borders, real-world globalization is affected by a multitude of noneconomic factors such as ethnicity, culture, and religion. With so many factors at play, globalization is an enormously complex process and often noneconomic factors do not align with purely economic factors. This misalignment can lead to frustrations that feed populist movements. By opposing the collaboration that comes with globalization, populist movements act as a feedback mechanism that pushes the general population towards becoming more ideologically rigid, and this, in turn, further accelerates the populism [56]. A strengthened populist movement can also trigger tectonic shifts in world affairs as exemplified by BREXIT in the UK and the recent 2016 US presidential election.

Problems in a globalized world rarely confine themselves to one place. Interdependence makes the developed countries more alike and synchronizes their social dynamics. This synchronization can cause political shifts in one country to spill over into other countries, and this is what enables the rise of RW populism to spread across large regions of the world. After BREXIT and the 2016 US presidential election, political elites should not expect to continue business as usual. Being the first to adopt a major political shift with a high probability of negative economic consequences is difficult, but once that line has been crossed the interdependence of globalization makes cascades (domino effects) an active possibility. No one wants to be first, but many are ready to be second.

The tipping point that brings on the rise of RW populism may be reached more quickly when voters face a binary choice, for example, the "yes" or "no" choice in the UK BREXIT referendum and the two major-party candidates in the recent US presidential election. In both cases the populist option secured a narrow victory. As mentioned above, the polls indicate that the populist party in Austria can expect to receive 34% of the votes, but the populist party candidate in the presidential race can expect close to 51%. A simple way to understand these percentages is to assume that political attitudes of voters are approximately symmetrical in their distribution across the political spectrum from left to right. Consequently, if leftist voters comprise ψ_L% of the total population, then RW populist voters will maintain a similar presence; that is, ψ_R% $\approx \psi_L$%. Facing this binary choice, the centrist voters have no one representing their views and thus are likely to vote evenly between the two available options. An implication is that when ψ_L% $\approx \psi_R$% ≈ 33%, even a slight (statistically significant) imbalance in favor of ψ_R% over ψ_L% can tip the society towards RW populism. In Austria ψ_R% ≈ 36% seems to be sufficient to make the RW populist candidate a front runner in the presidential race.

The final outcome of the battle between the conflicting factors surrounding globalization will almost surely have tremendous economic implications. If a country approaches its tipping point, how will the ensuing volatility affect its long-term credit rating? If a domino effect cascades across a large region of the EU or the entire EU, how will that affect the Euro and the common banking system? Without a proper resolution of the migrant crisis, what will be the impact on systemic risk? In an attempt to shed light on some of the factors and underlying processes that affect the success of globalization, we offer an empirically backed theoretical model of the rise of RW populism in response to unsustainable immigration inflows.

Our model emphasizes the need for controlled globalization in which immigration inflows into a society are balanced with the ability of the society to integrate the immigrants. This ability is arguably improved when immigrants mix with the native population, which is a principle practiced in Singapore where immigrant hubs are discouraged, and tenants in government-built housing (comprising 88% of all housing) must be of mixed ethnic origin [20]. Because tolerance towards immigrants is conditional, when immigration inflows overshadow integration rates the society can be tipped into RW populism, an intolerant mode of functioning. The rise of RW populism can occur because elections are by their nature stochastic, and they resemble a mathematical random walker. Left to its own devices, a random walker will eventually hit an absorbing barrier. Here this barrier,

of course, is a metaphor for the demise of globalization, ironically at the very hands of the progressive system (i.e., democracy) that made globalization possible in the first place.

Disclosure

Conflicts of Interest

The authors declare that they have no conflicts of interest.

Acknowledgments

References

[1] H. Grubel and A. Scott, "The international flow of human capital," *American Economic Review*, vol. 56, pp. 268–274, 1966.

[2] G. C. Hufbauer and K. Suominen, *Globalization at Risk: Challenges to Finance and Trade*, Yale University Press, 2010.

[3] F. Docquier and H. Rapoport, "Globalization, Brain Drain, and Development," *Journal of Economic Literature*, vol. 50, pp. 681–730, 2012.

[4] D. Rodrik, *The Globalization Paradox: Democracy and the Future of the World Economy*, W. W. Norton & Company, New York, USA, 2011.

[5] B. R. Chiswick, "The Effect of Americanization on the Earnings of Foreign-born Men," *Journal of Political Economy*, vol. 86, no. 5, pp. 897–921, 1978.

[6] G. J. Borjas, *Heaven's Door: Immigration Policy and the American Economy*, Princeton University Press, 1999.

[7] G. J. Borjas, "The economics of immigration," *Journal of Economic Literature*, vol. 32, pp. 1667–1717, 1994.

[8] G. J. Borjas, "Assimilation, Changes in Cohort Quality, and the Earnings of Immigrants," *Journal of Labor Economics*, vol. 3, no. 4, pp. 463–489, 1985.

[9] D. Swank and H. G. Betz, "Globalization, the welfare state and right-wing populism in Western Europe," *Socio-Economic Review*, vol. 1, no. 2, pp. 215–245, 2003.

[10] J. Gibson and D. McKenzie, "Eight Questions about Brain Drains," *Journal of Economic Perspectives*, vol. 25, no. 3, pp. 107–128, 2011.

[11] H. Betz, "Politics of Resentment: Right-Wing Radicalism in West Germany," *Comparative Politics*, vol. 23, no. 1, pp. 45–60, 1990.

[12] R. Knight, "Haider, the freedom party and the extreme right in Austria," *Parliamentary Affairs*, vol. 45, no. 3, pp. 285–299, 1992.

[13] P. Fysh and J. Wolfreys, "Le pen, the National Front and the extreme right in France," *Parliamentary Affairs*, vol. 45, no. 3, pp. 309–326, 1992.

[14] G. Voerman and P. Lucardie, "The extreme right in the Netherlands: The centrists and their radical rivals," *European Journal of Political Research*, vol. 22, no. 1, pp. 35–54, 1992.

[15] W. D. Chapin, "Explaining the electoral success of the new right: The German case," *West European Politics*, vol. 20, no. 2, pp. 53–72, 1997.

[16] J. M. Smith, "Does crime pay? Issue ownership, political opportunity, and the populist right in Western Europe," *Comparative Political Studies*, vol. 43, no. 11, pp. 1471–1498, 2010.

[17] B. Bonikowski and P. DiMaggio, "Varieties of American Popular Nationalism," *American Sociological Review*, vol. 81, no. 5, pp. 949–980, 2016.

[18] D. Rodrik, "How far will international economic integration go?" *Journal of Economic Perspectives*, vol. 14, no. 1, pp. 177–186, 2000.

[19] B. Davis, "Despite his heritage, prominent economist backs immigration cut," *The Wall Street Journal*, 1996, https://www.hks.harvard.edu/fs/gborjas/publications/popular/WSJ042696.pdf.

[20] B. Podobnik, M. Jusup, Z. Wang, and H. E. Stanley, "How fear of future outcomes affects social dynamics," *Journal of Statistical Physics*, vol. 167, no. 3-4, pp. 1007–1019, 2017.

[21] R. Reuveny, "Climate change-induced migration and violent conflict," *Political Geography*, vol. 26, no. 6, pp. 656–673, 2007.

[22] T. C. Schelling, "Dynamic models of segregation," *Journal of Mathematical Sociology*, vol. 1, pp. 143–186, 1971.

[23] R. Axelrod, "The dissemination of culture: a model with local convergence and global polarization," *Journal of Conflict Resolution*, vol. 41, no. 2, pp. 203–226, 1997.

[24] M. McPherson, L. Smith-Lovin, and J. M. Cook, "Birds of a Feather: Homophily in Social Networks," *Annual Review of Sociology*, vol. 27, pp. 415–444, 2001.

[25] T. Antal, P. L. Krapivsky, and S. Redner, "Dynamics of social balance on networks," *Physical Review E*, vol. 72, Article ID 036121, 2005.

[26] S. Marvel, J. Jon Kleinbergb, D. Robert, R. D. Kleinberg, and S. Strogatz, "Continuous-Time Model of Structural Balance," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 108, no. 5, pp. 1771–1776, 2011.

[27] C. Gracia-Lazaro, L. M. Flora, and Y. Moreno, "Selective Advantage of Tolerant Cultural Traits in the Axelrod-Schelling Model," *Physical Review E*, vol. 83, Article ID 056103, 2011.

[28] F. Aguiar and A. Parravano, "Tolerating the Intolerant: Homophily, Intolerance, and Segregation in Social Balanced Networks," *Journal of Conflict Resolution*, vol. 59, no. 1, pp. 29–50, 2015.

[29] M. Lim, R. Metzler, and Y. Bar-Yam, "Global pattern formation and ethnic/cultural violence," *Science*, vol. 317, no. 5844, pp. 1540–1544, 2007.

[30] R. M. Dancygier, *Immigration and Conflict in Europe*, Studies in Comparative Politics, Cambridge University Press, 2010.

[31] R. M. Dancygier and D. D. Laitin, "Immigration into Europe: Economic Discrimination, Violence and Public Policy," *Annual Review of Political Science*, vol. 17, pp. 43–64, 2014.

[32] S. Galam and M. A. Javarone, "Modeling Radicalization Phenomena in Heterogeneous Populations," *PLoS ONE*, vol. 11, Article ID e0155407, 2016.

[33] R. J. Sampson, S. W. Raudenbush, and F. Earls, "Neighborhoods and violent crime: a multilevel study of collective efficacy," *Science*, vol. 277, no. 5328, pp. 918–924, 1997.

[34] http://www.unhcr.org/.

[35] http://www.lucify.com/the-flow-towards-europe/.

[36] http://www.electograph.com/.

[37] C. Castellano, M. Marsilli, and A. Vespignani, "Nonequilibrium Phase Transition in a Model for Social Influence," *Physical Review Letters*, vol. 85, p. 3536, 2000.

[38] D. Card, A. Mas, and J. Rothstein, "Tipping and the Dynamics of Segragation," *The Quarterly Journal of Economics*, vol. 123, no. 1, pp. 177–218, 2008.

[39] https://en.wikipedia.org/wiki/List_of_non-state_terrorist_incidents.

[40] http://www.mipex.eu/what-is-mipex.

[41] H. Betz, "The New Politics of Resentment: Radical Right-Wing Populist Parties in Western Europe," *Comparative Politics*, vol. 25, no. 4, pp. 413–427, 1993.

[42] K. Popper, *The Open Society and Its Enemies: The Spell of Plato*, vol. 1, Routledge, UK, 1945.

[43] A. Majdanzic, B. Podobnik, S. V. Buldyrev, Y. Kenett, S. Havlin, and H. E. Stanley, "Spontaneous recovery in dynamical networks," *Nature Physics*, vol. 10, pp. 34–38, 2014.

[44] B. Podobnik, V. Vukovic, and H. E. Stanley, "Does the wage gap between private and public sectors encourage political corruption?" *PLoS ONE*, vol. 10, no. 10, Article ID e0141211, 2015.

[45] J.-H. Lee, M. Jusup, B. Podobnik, and Y. Iwasa, "Agent-based mapping of credit risk for sustainable microfinance," *PLoS ONE*, vol. 10, no. 5, Article ID e0126447, 2015.

[46] D. J. Watts, "A simple model of global cascades on random networks," *Proceedings of the National Academy of Sciences of the United States of America*, vol. 99, no. 9, pp. 5766–5771, 2002.

[47] http://theconversation.com/hard-evidence-how-areas-with-low-immigration-voted-mainly-for-brexit-62138.

[48] http://www.dw.com/en/golden-dawn-trial-begins-in-greece/a-18393111.

[49] B. Podobnik, D. Horvatic, T. Lipic, M. Perc, J. M. Buldu, and H. E. Stanley, "The Cost of Attack in Competing Networks," *Journal of The Royal Society Interface*, vol. 12, Article ID 20150770, 2015.

[50] R. Kanai, T. Feilden, C. Firth, and G. Rees, "Political Orientations Are Correlated with Brain Structure in Young Adults," *Current Biology*, vol. 21, no. 8, pp. 677–680, 2011.

[51] G. Zamboni, M. Gozzi, F. Krueger, J.-R. Duhamel, A. Sirigu, and J. Grafman, "Individualism, conservatism, and radicalism as criteria for processing political beliefs: a parametric fMRI study," *Social Neuroscience*, vol. 4, no. 5, pp. 367–383, 2009.

[52] D. R. Oxley, K. B. Smith, J. R. Alford et al., "Political attitudes vary with physiological traits," *Science*, vol. 321, no. 5896, pp. 1667–1670, 2008.

[53] E. Agliari, A. Barra, A. Galluzzi, M. A. Javarone, A. Pizzoferrato, and D. Tantari, "Emerging heterogeneities in Italian customs and comparison with nearby countries," *PLoS ONE*, vol. 10, no. 12, Article ID e0144643, 2015.

[54] S. Galam, "The Trump phenomenon: an explanation from sociophysics," *International Journal of Modern Physics B*, vol. 31, no. 10, 17 pages, 2017.

[55] R. Kennedy, S. Wojcik, and D. Lazer, "Improving election prediction internationally," *Science*, vol. 355, no. 6324, pp. 515–520, 2017.

[56] M. Jusup, T. Matsuo, and Y. Iwasa, "Barriers to Cooperation Aid Ideological Rigidity and Threaten Societal Collapse," *PLoS Computational Biology*, vol. 10, no. 5, Article ID e1003618, 2014.