UNIVERSIDAD DE ZARAGOZA
Facultad de Ciencias
Departamento de Física de la Materia Condensada

Tesis Doctoral

BEYOND SIMPLE COMPLEX-NETWORKS: COEVOLUTION, MULTIPLEXITY AND TIME-VARYING INTERACTIONS

Alessio Vincenzo Cardillo

Memoria presentada para acceder al grado de doctor, bajo la dirección del Doctor Jesús Gómez-Gardeñes y el Doctor Sandro Meloni.

Beyond simple complex-networks: coevolution, multiplexity and time-varying interactions

Colección de Estudios de Física
Vol. 117

Esta colección recoge las tesis presentadas en el Departamento de Física de la Materia Condensada de la Universidad de Zaragoza desde su constitución en 1987.

Colección de Estudios de Física Vol. 117

Beyond simple complex-networks: coevolution, multiplexity and time-varying interactions

Alessio Vincenzo Cardillo

Prensas de la Universidad
Universidad Zaragoza
1542

D. Jesús Gómez-Gardeñes, Investigador Ramón y Cajal en el Departamento de Física de la Materia Condensada de la Universidad de Zaragoza
y
D. Sandro Meloni, investigador Juan De La Cierva del Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) de la Universidad de Zaragoza

HACEN CONSTAR

que la presente memoria de Tesis Doctoral presentada por Alessio Vincenzo Cardillo y titulada "*Beyond simple complex-networks: coevolution, multiplexity and time-varying interactions*" ha sido realizada bajo su dirección en el Departamento de Física de la Materia Condensada y en el Instituto de Biocomputación y Física de Sistemas Complejos (BIFI). El trabajo recogido en esta memoria se corresponde con lo planteado en el proyecto de tesis doctoral aprobado en su día por el órgano responsable del programa de doctorado. Asimismo, los directores declaran que la contribución del candidato ha sido altamente satisfactoria y decisiva para la consecución de todas las fases de la investigación aquí presentada. Incluyendo el modelado teórico, el diseño de las herramientas computacionales y el análisis de datos en cada uno de los trabajos presentados en esta memoria. Por último, autorizan la presentación de dicha tesis en el formato de compendio de publicaciones.

Y para que conste, cumplimiento de la legislación vigente, informan favorablemente sobre la referida Tesis Doctoral y autorizan su presentación para su admisión a trámite.

Zaragoza, a 23 de Mayo de 2014
Los directores de la Tesis

Fdo: Jesús Gómez-Gardeñes　　　　Fdo: Sandro Meloni

D. Jesús Gómez-Gardeñes, Investigador Ramón y Cajal en el Departamento de Física de la Materia Condensada de la Universidad de Zaragoza
y
D. Sandro Meloni, investigador Juan De La Cierva del Instituto de Biocomputación y Física de Sistemas Complejos (BIFI) de la Universidad de Zaragoza

CERTIFICAN QUE:

los siguentes co-autores de los artículos incluidos en el trabajo de tesis doctoral titulado: "Beyond simple complex-networks: coevolution, multiplexity and time-varying interactions", presentado por D. Alessio Vincenzo Cardillo son doctores.

- Yamir Moreno,
- Giovanni Petri,
- Vincenzo Nicosia,
- Roberta Sinatra,
- Vito Latora,
- Fernando Naranjo,
- Massimiliano Zanin,
- Miguel Romance,
- David Papo,
- Francisco del Pozo,
- Stefano Boccaletti,
- Alejando J. García del Amo,
- Lucia Valentina Gambuzza,
- Alessandro Fiasconaro,
- Luigi Fortuna,
- Mattia Frasca.

Zaragoza, a 23 de Mayo de 2014
Los directores de la Tesis

Fdo: Jesús Gómez-Gardeñes

Fdo: Sandro Meloni

Aqui va la hoja de creditos.... asi que habra que sustituir esta pagina en el pdf final por la que nos den

All'Italia, terra di: navigatori, poeti e … scienziati emigrati.

Alla Sicilia, terra di: storie, colori, profumi e sapori unici al mondo.

Contents

Prólogo

Cuando se trata de entender el funcionamiento de un *sistema complejo* conviene enfocar su descripción hacia las interacciones entre sus constituyentes elementales, en lugar de tratar de modelar al detalle la dinámica de cada uno de ellos. Este enfoque se justifica en que nos interesa capturar el comportamiento colectivo del mismo (sus propiedades *macroscópicas*). Por esta razón, las herramientas computacionales y el enfoque teórico de la física estadística son herramientas particularmente adecuadas para estudiar estos sistemas y sobre este paradigma han proliferado en las últimas décadas multitud de modelos físicos sencillos que permiten extraer información sobre una gran variedad de fenómenos colectivos que aparecen más allá del contexto de la física tradicional, "invadiendo" campos y disciplinas como la biológia, la sociología, la epidemiología, etc [1, 2].

A lo largo de la última década, el estudio de la física estadística de sistemas complejos ha avanzado enormemente, entre otras razones, gracias a la incorporación de las herramientas proporcionadas por la *ciencia de las redes complejas*, una nueva disciplina que tiene su origen en la teoría (matemática) de grafos fundada a mitad del siglo XVIII por Leonard Euler. La gran abundancia de datos sobre los patrones de interacción en múltiples sistemas complejos reales (desde las redes de transporte e Internet hasta el cerebro humano) y el hecho de que estos patrones sobre quién interactúa con quién no se puedan explicar como un resultados del azar, ha demandado una generalización de los modelos propuestos por la teoría de grafos y su incorporación al modelado de los procesos dinámicos que gobiernan el funcionamiento de los sistemas complejos. En este sentido, la ciencia de redes complejas nos ha proporcionado una herramienta suficientemente general para abordar problemas de diferentes disciplinas permitiendo codificar los elementos de un sistema y la relaciones entre ellos mediante un grafo donde los *elementos* son los nodos de la red y las *interacciones* sus enlaces [3, 4].

A pesar de los éxitos conseguidos durante la última década, el avance en la captación de nuevos datos sobre las interacciones en sistemas naturales, sociales y tecnológicos, ha obligado recientemente a una reformulación de los modelos de redes usados hasta la fecha. Asimismo, fenómenos tan importantes, como por ejemplo el gran apagón ocurrido en Italia en el año 2007, han puesto de manifiesto deficiencias en modelado de sistemas y redes complejas, como por ejemplo la interdependencia de dos o más redes formando metaredes, la coexistencia de diferentes tipos de interacciones entre los mismos constituyentes de una red (redes multicapa), el carácter temporal de las interacciones (redes tempo-variantes), etc. Estos nuevos tipos de formulaciones presentan propiedades que pueden ser muy diferentes de las encontradas en las redes constituyentes (para las redes inter-dependientes o las redes multicapa) o en las red acumulada (caso de las redes tempo-variantes). El salto conceptual al que estamos asistiendo en el campo de la ciencia de redes obliga también a reformular los modelos dinámicos y volver a estudiar los procesos colectivos asociados. Algunos de estos fenómenos como por ejemplo la emergencia de la cooperación [5], la difusión/control de epidemias [6–8], el estudio de la resistencia que un sistema manifiesta cuando se enfrenta a un ataque o un fallo [9–11] han sidos estudiados y ampliados recientemente a los nuevos (y adaptados) marcos conceptuales de la ciencia de redes [12–14].

En esta línea, el trabajo desarrollado a lo largo de esta tesis doctoral, tiene como uno de sus objetivos principales el estudio y desarrollo de modelos que permitan observar nuevos fenómenos y comportamientos colectivos derivados de las nuevas formas y modos de interacción mencionadas anteriormente. Asimismo, y en consonancia con el devenir de la ciencia de redes, otro objetivo fundamental de esta tesis ha sido la manipulación y análisis de grandes cantidades de datos procedentes de sistemas complejos reales con el fin de poder comparar empíricamente los resultados previstos por los modelos. En resumen, a lo largo de todo la etapa doctoral, se ha querido explorar tanto las fronteras, como regiones ya previamente exploradas en busca de nuevos escenarios que pongan de manifiesto la importancia de incorporar las características fundamentales de interacción entre los elementos de un sistema complejo. Los problemas investigados pueden ser agrupados bajo uno (o más) de estos temas:

- Teoría evolutiva de juegos.
- Sincronización.
- Difusión de epidemias.
- Codificación de interacciones de diferente naturaleza mediante estructuras multicapas.
- Tratamiento de interacciones que varían en el tiempo a través de redes con estructura tempo-variante.

Esta memoria de tesis doctoral pretende ser un resumen de los principales resultados logrados durante todo el recorrido de la etapa doctoral. Por ello, esta memoria de tesis ha sido redactada como *compendio de publicaciones*. Sin embargo, se ha querido dotar a la memoria de un capitulo inicial donde se introducen los conceptos básicos utilizados en los trabajos presentados para facilitar la comprensión del lector. Los capítulos sucesivos a la introducción general están constituidos por los artículos, colocados de forma que los tres ingredientes básicos de nuestros sistemas (*coevolución, multiplexidad y interacciones tempo-variantes*) se visiten en orden (casi) secuencial. Finalmente, esta memoria acaba con las conclusiones de la tesis, donde se resumen los resultados principales logrados en cada artículo y se delinean las posibles direcciones futuras.

Listado de las publicaciones incluidas en la memoria

Los trabajos incluidos en la siguente memoria de tesis son:

1. Cardillo, A., Meloni, S., Gómez-Gardeñes, J., & Moreno, Y. (2012). *Velocity-enhanced cooperation of moving agents playing public goods games.* Physical Review E, **85**(6), 067101. doi:10.1103/PhysRevE.85.067101

2. Cardillo, A., Petri, G., Nicosia, V., Sinatra, R., Gómez-Gardeñes, J., & Latora, V. (2013) *Evolutionary dynamics of time-resolved social interactions.* Los Alamos e-print archive, arXiv:1302.0558 [physics.soc-ph]. *Submitted for publication*

3. Cardillo, A., Reyes-Suárez, C., Naranjo, F., & Gómez-Gardeñes, J. (2013). *Evolutionary vaccination dilemma in complex networks.* Physical Review E, **88**(3), 032803. doi:10.1103/PhysRevE.88.032803

4. Cardillo, A., Gómez-Gardeñes, J., Zanin, M., Romance, M., Papo, D., del Pozo, F., & Boccaletti, S. (2013). *Emergence of network features from multiplexity.* Scientific Reports, **3**, 1344. doi:10.1038/srep01344

5. Cardillo, A., Zanin, M., Gómez-Gardeñes, J., Romance, M., García del Amo, A. J., & Boccaletti, S. (2013). *Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures.* The European Physical Journal Special Topics, **215**(1), 23–33. doi:10.1140/epjst/e2013-01712-8

6. Gambuzza, L. V., Cardillo, A., Fiasconaro, A., Fortuna, L., Gómez-Gardeñes, J., & Frasca, M. (2013). *Analysis of remote synchronization in complex networks.* Chaos: An Interdisciplinary Journal of Nonlinear Science, **23**(4), 043103. doi:10.1063/1.4824312

Con respecto a las revistas donde han sido publicados los articulos, las informaciones sobre factor de impacto, area tematica y cuartil son (datos procedentes desde el Journal of Citation Reports año 2012):

REVISTA	FACTOR DE IMPACTO	ÁREA	RANK	CUARTIL
Physical Review E	2.313	Physics, fluids & plasmas	10	Q2
		Physics, mathematical	6	Q1
Chaos	2.188	Mathematics, applied	8	Q1
		Physics, mathematical	7	Q1
European Physical Journal–Special Topics	1.796	Physics, multidisciplinary	23	Q2
Scientific Reports	2.927	Multidisciplinary sciences	8	Q1

Chapter 1

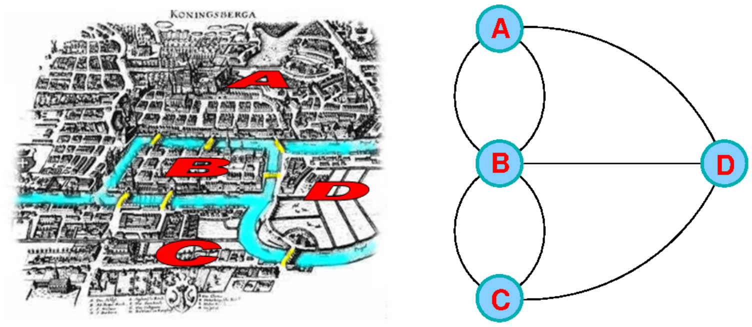
Introduction & objectives

1.1. Going beyond complex networks

If we look back at the history of complex systems science, and complex networks in particular, a lot of water has flowed beneath the bridge since 1736 when the mathematician Leonard Euler published the solution to the Könisberg bridges problem (formulated as finding a round trip that traversed each of the Könisberg's seven bridges exactly once; cf. Fig 1.1(a)). Since then, many steps forward have been made [3, 4, 15–19], and the use of graph theory to study those systems whose properties depend more on the interactions among their elements rather than on the microscopic composition of the element itself, has allowed to find the answer to a lot of questions in many different disciplines spanning from social science to applied mathematics. Graphs (or networks) are thus a general, yet powerful, means of representing patterns of connections or interactions between the parts of the system.

Across the years, the science of complex networks has benefited enormously from the wide range of viewpoints brought to it by practitioners from so many different disciplines. At the same time, it has also suffered of this interdisciplinary character since knowledge about networks is dispersed across the scientific community and researchers in one area often do not have ready access to discoveries made in another.

In the last fifteen years, complex networks have been used with success to tackle many problems allowing to find an answer to questions such as: why epidemics spread very rapidly across the world? Which are the interaction schemes among people in the era of social networks and smartphones? To find the answer to questions like these, a multidisciplinary approach is needed. As time passed, the models presented became more and more refined in order to



(a) (b)

Figure 1.1: A schematic view of the Könisberg bridges: (a) map view and (b) the corresponding graph.

capture as many real-system characteristics as possible. However, this must not be seen as just a mere *complication* of the models. The fact is that, in order to solve more complex problems, one cannot use anymore those simple models (or rely on just one kind of approach) developed so far, but rather use two, or more, interdependent approaches at the same time. As an example, in the forecast of Epidemics, one has to use concepts belonging to applied mathematics, computer science, and epidemiology of course. In the light of that, the last few years have seen the flourishing of a new generation of models in all the fields where complex networks have been previously used. Another reason behind the urge for a new generation of models is related with the appearance of a huge amount of data coming from the opportunity to crawl many technological infrastructure like: mobile phones [20], GPS tracks [21] or temporal trade data among countries [22] just to cite a few. The models developed so far are unable to account for the richness of phenomena observed in such systems, and the *big data challenge* cannot be tackled unless we perform the step forward that we have just talked about.

In the last five years, we have witnessed the appearance of more sophisticate coevolutive models to understand phenomena that depends on many factors at the same time; the inclusion of time dimension in the description of interactions among the nodes of a network [15]; the expansion to multiple kind of interactions and interconnected sub-systems in the description of realistic systems [12, 23–26]. These, and other, new *perspectives* have allowed and will allow us to shed light on many open issues and will consolidate the role of complex systems science as a powerful, and multidisciplinary, way to understand the world in which we live in.

1.2. Objectives

1.2.1. Main objectives

The doctorate of the candidate began during one of the most tumultuous moments experienced by complex network science over the past ten years. Thus, at its beginning, more than one line of research were available to choose. From one side, it would have been possible to follow the route of continuing the study of coevolutive dynamical processes acting on top of complex topologies. On the other side, instead, the appearance of new formalisms (multilevel, time-varying) like those introduced above, offered very intriguing perspectives in terms of new problems to tackle/solve. For this reason, it has been decided to "*ride the wave*" generated by the appearance of such new formalisms, while continuing, at the same time, the exploration of the "*previously known landscape*" of dynamics on top of complex network topologies. Lastly, an aspect to do not underestimate has been that related to the analysis/manipulation of big amounts of data. Such aspect will play in the near future a prominent role in the empirical comparison aimed at validating the existing models as well as the new ones. Summing up, the main objectives pursued during the whole duration of the doctorate are:

Objective 1
Extension of known models used to study complex systems to achieve a more realistic description of them.

Objective 2
Retrieval, analysis, and management of big amount of data originated by real complex systems.

Objective 3
Study emerging properties and phenomena in systems described using the new frameworks of time-varying and multilevel interactions.

1.2.2. Objectives of each paper and contribution of the candidate

Keeping in mind the trail indicated by the main objectives listed above, it is possible to go deeper and comment more in detail the objectives pursued by each one of the papers, connecting them with the general ones of the whole doctorate. At the same time, we also highlight which aspects of the research work (design of the models, analysis of the results, preparation of the manuscript, and presentation of the results at conferences/public events) involved the active contribution of the candidate. So, in conclusion, we have that:

1. In Chap. 3, we investigate the effects that a coevolutionary approach based on evolutionary game theory and time-varying interactions, the latter generated by the motion of the agents, may have on the emergence of cooperation (objectives 1 & 3). The contributions given by the candidate concern the design of the agent-based simulations, the analysis of their outcome/results, the preparation of part of the manuscript, and the presentation of the results at an international conference named "*Net-Works 2011*", held in El Escorial (Madrid), Spain, in October 2011.

2. In Chap. 4, we follow the path shown in the previous point, and continue the study about the effects that time-varying interactions have on the emergence of cooperation. However, in this case, we further extended the description used above including the use of real time-varying social interactions (objectives 1, 2 & 3) to infer the role that correlations exert on the survival of cooperation. The candidate contributed providing the topological characterization of the dataset, performing the agent-based simulation of the evolutionary dynamic, preparing some parts of the manuscript (especially those concerning the description of the data used and the evolution of topological properties over time), and displaying the results obtained, as an oral contribution, at the one of the main international conferences of network science named "*NetSci 2013*" held in Copenhagen, Denmark, in June 2013.

3. In Chap. 5, through the use of a coevolutionary dynamic (obtained putting in cascade two different dynamical processes), we study the phenomenon of voluntary vaccination as an active method to prevent the spread of diseases. In particular, the decision of whether getting vaccinated or not has been modeled as a social dilemma using the formalism of evolutionary games (objective 1). The contributions of the candidate have involved: the design of the model, the realization of the simulations, the analysis of the results produced, and the preparation of the manuscript. The candidate has also presented the results at international conferences both as oral ("*Complenet 2014*", held in Bologna, Italy, in March 2014) and poster ("*NetSci 2013*", held in Copenhagen, Denmark, in June 2013) contributions.

4. In Chap. 6, we investigate the emergence of topological properties when the system possesses different type of interactions and each one of them is encoded as a different layer. To gauge the emergence of such features when the layers are projected onto a single one, we make use of real data

belonging to the European airline network (objectives 3 & 2). The candidate has contributed to the design of the model, the preparation of the dataset, its topological characterization, the analysis of the results, and the writing of the whole manuscript. Also, with the aim of fostering the public availability of big data, the candidate prepared a public repository where the data (cleaned and filtered) are freely available for the whole scientific community.

5. In Chap. 7, we were interested in observing the effects that the use of data belonging to a real system made of several layers has on the dynamical outcome of a re-scheduling procedure whose behavior has been previously studied in systems with just one kind of interaction among their elements (objectives 1, 2 & 3). The contribution of the candidate involved all the phases of the work, from the design of the model up to the writing of the manuscript.

6. In Chap. 8, we studied the emergence of a collective phenomenon like synchronization. In particular, we were interested in the characterization of the onset of a novel kind of synchronization, named remote synchronization, in a population of amplitude oscillators whose interactions are arranged like a complex network (objectives 1 & 3). The contributions of the candidate have focused on the design of the simulations, the analysis of the results.

After having described the aims, and the contributions of the candidate, of the work done during the doctorate, we are ready to pass to the main body of the manuscript. However, before doing so, it is worth spending some words to illustrate the structure of this thesis to let the reader create some sort of mental map of it. The structure of this thesis is the following: Chapter 2 we provide the fundamental knowledge needed to an unexperienced reader to fully understand all the concepts used and results found in the following chapters. In particular, in Sec. 2.1 we introduce some basic mathematical notions of graph theory and some structural measures. In Sec. 2.2, instead, we present some general models of complex networks including multiplex (2.2.4) and time-varying (2.2.5) ones. Then, in Sec. 2.3, an overview on some dynamical processes that have been implemented on top of complex networks topologies is made. In particular: spreading phenomena, synchronization, and evolutionary game theory are illustrated in sections 2.3.1, 2.3.2, and 2.3.3 respectively as they have been the main dynamical frameworks used in the works of this thesis. After introducing the basic concepts used within the papers collected here, we are ready to understand the results contained into them and displayed in chapters from 3 to 8. Finally, in Chap. 9 we draw some general conclusions and give a few future perspectives.

Chapter 2

Methods

2.1. Introduction to graph theory

Many real-world situations can be conveniently described by means of a diagram consisting of a set of points together with lines joining certain pairs of points. For example, the points could represent people, with lines joining pairs of friends; or the points might be communication centres, with lines representing communication links. Notice that in such diagrams one is mainly interested in whether two given points are joined by a line; the manner in which they are joined is sometimes not physical. A mathematical abstraction of situations of this type gives rise to the concept of **graph**.

Graphs are so named because they can be represented graphically, and it is this graphical representation which helps us understand many of their properties. Each vertex is indicated by a point, and each edge by a line joining the points representing its ends. There is no single correct way to draw a graph; the relative positions of points representing vertices and the shapes of lines representing edges usually have no significance.

However, we often draw a diagram of a graph and refer to it as the graph itself; in the same spirit, we call its points **vertices** and its lines **edges**. Most of the definitions and concepts in graph theory are suggested by this graphical representation. The ends of an edge are said to be *incident* with the nodes, and vice versa. Two vertices which are incident with a common edge are *adjacent*, as are two edges which are incident with a common vertex, and two distinct adjacent vertices are **neighbours**.

However, sometimes drawing a graph is completely useless, especially when the number of vertices and edges is very large. In fact, the big number of connections and points could transform a brief clear sketch into a complete

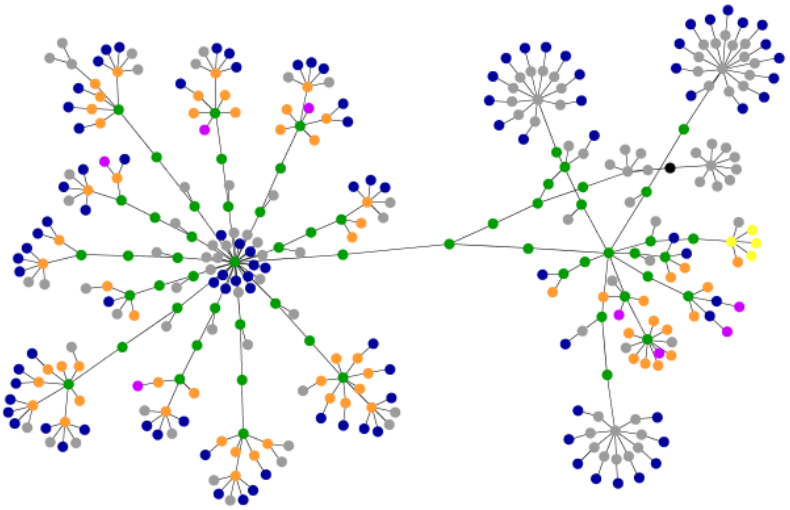


Figure 2.1: An example of a graphical representation of a graph.

mess. In addition, since positions of vertices are usually meaningless, the "properties" of such systems cannot be deduced only by a visual analysis. This facts seems to limit the range of validity of graph theory as it has been described until now. Fortunately for us, mathematics provides a rigorous formalism under which all the properties of a graph can be expressed. So, for a proper comprehension of the results, it is necessary to introduce this common language providing some definitions and notations.

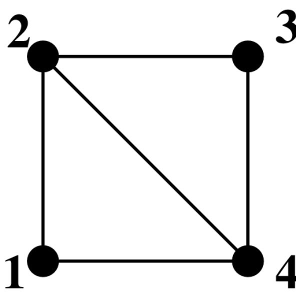
2.1.1. Basic notions

Let us start with the construction of the mathematical language used to describe graphs and in particular their structural properties. *Graph theory* [4, 27–30] is the natural framework for the exact mathematical treatment of complex networks and, formally, a complex network can be represented as a graph. An **undirected** (*directed*) **graph** $G = (\mathcal{N}, \mathcal{K})$ is a mathematical object which consists of two sets: $\mathcal{N}$ and $\mathcal{K}$, such that $\mathcal{N} \neq \emptyset$ and $\mathcal{K}$ is a set of unordered (ordered) pairs of elements of $\mathcal{N}$. The elements of $\mathcal{N} \equiv \{n_1, n_2, \ldots, n_N\}$ are the **nodes** (or *vertices*, or *points*) of the graph G, while the elements of $\mathcal{K} \equiv \{l_1, l_2, \ldots, l_K\}$ are its **links** (or *edges*, or *lines*). The number of elements in $\mathcal{N}$ and $\mathcal{K}$ are denoted by N and K, respectively. Then, later, a graph will be indicated as $G(N, K) = (\mathcal{N}, \mathcal{K})$, or simply $G(N, K)$ or G_N, whenever it is necessary to emphasize the number of nodes and links in the graph.

A powerful way to represent a graph is through the **adjacency** (or *connectivity*) **matrix** $\mathcal{A}$, a $N \times N$ square matrix whose entry $a_{i,j}$ $(i, j = 1, \dots, N)$ is equal to:

$$a_{i,j} = \begin{cases} 1, & \text{if the edge between nodes } i \text{ and } j \text{ exists,} \\ 0, & \text{otherwise.} \end{cases}$$

The simplest kind of network one can think of is that whose connections are undirected and have the same intensity. In such case we will talk about **undirected** and **unweighted** network as those networks whose adjacency matrix is *symmetric* and whose nonzero elements assume all the same value.



$$\mathcal{A} = \begin{pmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{pmatrix}$$

Figure 2.2: A graph and its corresponding adjacency matrix $\mathcal{A}$.

The adjacency matrix $\mathcal{A}$ fully characterizes the topological properties of the graph under study. In particular, as we will discover soon, its properties determine the particular type of network under consideration. In graph theory, two extremal cases of networks are the **complete graph** that is a graph in which each pair of vertices are adjacent so it has $\frac{N(N-1)}{2}$ edges corresponding to an adjacency matrix whose elements are all (except those on the diagonal) equal to one; and the **empty graph** that is a graph with no connections (i.e. where all the elements of $\mathcal{A}$ are equal to zero).

However, many real networks display a large heterogeneity in both the intensity and direction of connections. Examples are the electrical resistance in resistors networks, passengers flow in airline networks, and the presence of weak ties and hierarchies between individuals in social networks to name a few [31–36]. These systems can be better described in terms of **weighted networks**, i.e. networks in which each link carries a numerical value measuring the strength of the connection. In this sense a new set $\mathcal{W} \equiv \{w_1, w_2, \dots, w_K\}$ must be considered. In the light of such assumption, a graph G is defined as: $G^W = \{\mathcal{N}, \mathcal{K}, \mathcal{W}\}$.

When weighted networks are considered, it is useful to represent them redefining the adjacency matrix. In these cases it is useful to consider a **weight matrix** $\mathcal{W}$, a $N \times N$ square matrix whose entry w_{ij} $(i, j = 1, \ldots, N)$ is equal to the link weight. Directed networks, instead, can be represented in terms of asymmetric adjacency/weight matrices, thus accounting for the possibility that not every connection may be reciprocated.

After this brief introduction about the macro categories upon which complex networks may be classified, and the mathematical tools used to represent them, it is now time to define those measures that will help us in the characterization of several topological properties. For the sake of simplicity, whenever not indicated, we will consider unweighted undirected networks.

Node degree, strength and their distributions

The **degree** (or *connectivity*) k_i of a node i is the number of edges incident with it. It is defined in terms of the adjacency matrix $\mathcal{A}$ elements as:

$$k_i = \sum_{j \in \mathcal{N}} a_{ij} \,. \tag{2.1}$$

In the case of weighted graphs, the counterpart of the degree is played by the so-called node **strength**, s_i, which is the sum of the edges weights incident with the node i. In terms of weight matrix $\mathcal{W}$ we could express it as:

$$s_i = \sum_{j \in \mathcal{N}} w_{ij} \,. \tag{2.2}$$

The most basic topological characterization of a graph G can be obtained in terms of its **degree distribution** $P(k)$, defined as the probability that a randomly chosen node has degree k or, equivalently, as the fraction of nodes in the graph having degree k (analogously, one can define a *strength distribution* $P(s)$). Information on how the degree is distributed among the nodes of an undirected network can be obtained either by a plot of $P(k)$, or by the calculation of the moments of the distribution. The n-moment of $P(k)$ is defined as:

$$\langle k^n \rangle = \sum_{k} k^n \, P(k) \,. \tag{2.3}$$

The first moment $\langle k \rangle$ is the **average degree** of G. The second moment measures the fluctuations of the connectivity distribution around the mean $\langle k \rangle$. The degree distribution completely determines the statistical properties of uncorrelated networks as shown by Newman et al. in [4, 37].

When analyzing real networks [38], it may happen that the data have rather strong intrinsic noise due to the finiteness of the sampling. Therefore, when the size of the system is small, it is sometimes advisable to measure the **cumulative** degree distribution $P_{>}(k)$, defined as:

$$P_{>}(k) = \sum_{k'=k}^{\infty} P(k') \,. \tag{2.4}$$

Indeed, by using $P_{>}(k)$, the statistical fluctuations generally present in the tails are smoothed while preserving all other indicators. Some examples of various degree distribution belonging to several real networks can be seen in Fig. 2.3.

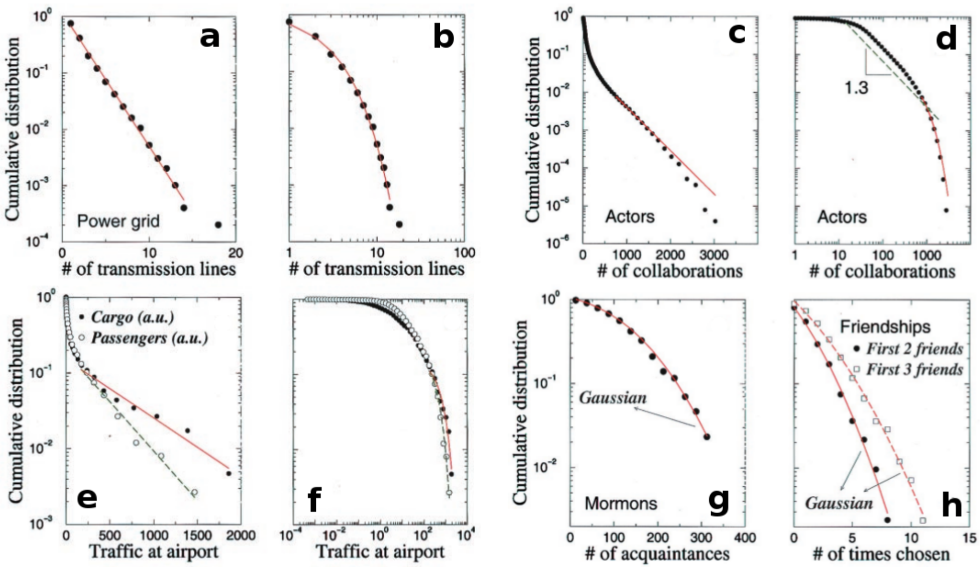


Figure 2.3: Shapes of cumulated degree distributions belonging to various real-world networks. Panels a,b,e, and f correspond to technological and economic networks, while panels c,d,g, and h to social ones (from Amaral *et al.* [38]).

Degree correlations and rich-club phenomenon

A large number of real networks are *correlated* in the sense that the probability that a node of degree k is connected to another node of degree, k', depends on k. In these cases, it is necessary to introduce the ***conditional probability*** $P(k'|k)$, defined as the probability that a link from a node of degree k points to a node of degree k'. $P(k'|k)$ satisfies the degree detailed balance condition $kP(k'|k)P(k) = k'P(k|k')P(k')$ [39]. For uncorrelated graphs, the balance condition gives $P(k'|k)P(k) = k'P(k')/\langle k \rangle$.

Correlated graphs are classified as **assortative** if the average degree of nearest neighbour $k_{nn} = \sum_{k'} k' P\left(k'|k\right)$ is an increasing function of k, whereas they are referred to as **disassortative** when k_{nn} is a decreasing function of k [40]. In other words, in assortative networks the nodes tend to connect to their connectivity peers, while in disassortative networks nodes with a lower degree are more likely connected with highly connected ones.

Another interesting quantity that is able to capture the presence of degree correlations is the so-called *rich-club phenomenon*, which refers to the tendency of nodes with high centrality, the dominant elements of the system, to form tightly interconnected communities, and it is one of the crucial properties accounting for the formation of dominant communities in both computer and social sciences [24, 41, 42]

Essentially, nodes with a large number of links, usually referred to as *rich nodes*, are much more likely to form tight and well-interconnected subgraphs (clubs) than low-degree nodes. A first quantitative definition of the rich-club phenomenon is given by the **rich-club coefficient** ϕ, introduced in the context of the Internet in [41]. Denoting by $E_{>k}$ the number of edges among the $N_{>k}$ nodes having degree higher than a given value k, the rich-club coefficient is expressed as:

$$\phi(k) = \frac{2E_{>k}}{N_{>k}\,(N_{>k}-1)}\,, \tag{2.5}$$

where $N_{>k}(N_{>k}-1)/2$ is nothing less than the maximum possible number of edges among the $N_{>k}$ nodes. Therefore, $\phi(k)$ measures the fraction of edges connecting those nodes with respect to the maximum number of edges there might be. It is worth mentioning that, the rich-club phenomenon is not necessarily associated with assortative mixing. In other words, the rich-club phenomenon and the mixing properties express different features that are not trivially related to each other. Also, it is worth stressing that a simple inspection of the $\phi(k)$ trend is potentially misleading in the discrimination of the rich-club phenomenon since an increasing behavior (for high values of k) does not confirm the presence of the rich-club. Nevertheless, the behaviour of the rich-club coefficient as a function of the degree k is a probe for the topological correlations in a complex network, and it yields important information about its underlying architecture.

As we have seen so far, the rich-club coefficient $\phi(k)$ is a quantity to deal with care. To reduce the risks that may lead to wrong conclusions, and to allow the comparison between different kind of networks, an alternative measure has been proposed. In particular, Colizza *et al.* have proposed to use a normalized coefficient rather than simply Eq. (2.5). More in detail, $\phi(k)$ is divided by the same quantity calculated on a network obtained using the appropriate null model providing a suitable normalization $\phi'(k)$. Such null models have to ensure that the obtained network has the same degree distribution of the original one but with any eventual degree correlation washed out.

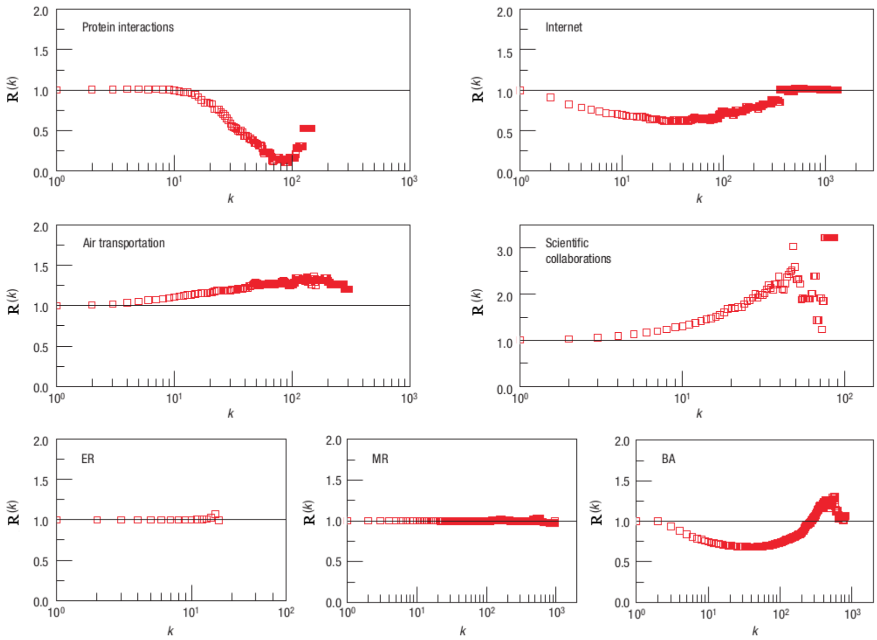


Figure 2.4: Normalized rich-club coefficient R as a function of degree k for various real and synthetic networks. We see how Internet (at the autonomous system level) and Protein interactions networks do not display rich-club effect, while Air transportation and Scientific collaboration do. Concerning the synthetic networks models (ER, MR and BA), only BA ones possess a rich-club (from [42]).

From the discussion above, a possible choice for the normalization of the rich-club coefficient is provided by the ratio $R(k) = \phi(k)/\phi'(k)$, where $\phi'(k)$ is the *link abundance* on a network with the same degree sequence of the original but with connections randomly shuffled (i.e. no degree correlations) that can be calculated analytically. A ratio larger than one is the actual evidence for the presence of a rich-club phenomenon leading to an increase in the interconnectivity of high-degree nodes in a more pronounced way than in the random case. In contrast, a ratio $R(k) < 1$ is a signature of an opposite organizing principle that leads to a lack of interconnectivity among high-degree nodes.

$$R(k) = \frac{\phi(k)}{\phi'(k)} = \frac{2E_{>k}}{N_{>k}(N_{>k}-1)} \frac{N_{>k}(N_{>k}-1)}{2E'_{>k}} = \frac{E_{>k}}{E'_{>k}}, \tag{2.6}$$

where $E'_{>k}$ is the number of nodes with degree greater or equal to k in such network. An illustrative example displaying the behavior of R as a function of k for various real and synthetic networks is displayed in Fig. 2.4.

Shortest path lengths and diameter

Graphs are often used to model systems in which goods move through them. Consider, for example, the case of transportation or communication systems, but social systems may, as well, have "things" moving through them (gossip news or viruses to cite some). For such reason, it is very important to know which route is the best one, to ensure a fast and safe delivery, as well as the size and the robustness of the system in which we are moving through. In this section we focus on quantities that can provide an answer to such kind of problems.

As previously stated, the knowledge of the shortest path between two points in the system is a crucial issue in the context of transport and communication within a network. Suppose, for example, that one needs to send a data packet from one computer called i to another named j through the Internet: the *geodesic* or **shortest path** is the shortest walk connecting two nodes (in an unweighted network the minimum number of hops needed to reach node j from i). Geodesics provide optimal pathway, since one would achieve a fast transfer and save of system resources. For such a reason, shortest paths have also played an important role in the characterization of the internal structure of a graph [24]. It is useful to represent all the shortest path lengths of a graph G as a matrix, $\mathcal{D}$, in which the entry d_{ij} corresponds to the length of the geodesic from node i to node j. The maximum value of d_{ij} is thus called the **diameter** of the graph. In the light of the previous definition, in a spatial (weighted) graph, the shortest path length is defined as the smallest sum of the edge lengths throughout all the possible paths in the graph connecting, say, a node i to another one named j [30, 35, 43]. In this way both the topology and the geography of the system are taken into account.

A measure of the typical separation between two nodes in the graph is given by the **average shortest path length**, L, (a global property) also known as the **characteristic path length**, defined as the mean of geodesic lengths over all couples of nodes [3, 4, 24].

$$L = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} d_{ij}. \tag{2.7}$$

A problem arising from the above definition is that L diverges if there are disconnected components in the graph. One possibility to avoid such a divergence is to limit the summation in Eq. (2.7) only to pairs of nodes belonging to the largest connected component, or **giant component** [9]. An interesting alternative, called *efficiency*, is described below.

Efficiency

The global structural properties of a graph can be evaluated by the analysis of the shortest paths between all pairs of nodes. However, as we have mentioned above, there exist a divergence issue with Eq. (2.7). We already spoke about one possibility to avoid such inconsistence. Another possible way to solve this problem is to consider the harmonic mean of the geodesic lengths, and to define the so-called **efficiency**, E, of a graph [35, 43]. In this way we are still able to quantify how easy is the communication between the nodes.

$$E = \frac{1}{N(N-1)} \sum_{i,j \in \mathcal{N}, i \neq j} \frac{1}{d_{ij}} . \tag{2.8}$$

Where d_{ij} is an element of the geodesic matrix $\mathcal{D}$. Such quantity is an indicator of the traffic capacity of a network, and avoids the divergence in Eq. (2.7), since any pairs of nodes belonging to disconnected component of the graph yields a contribution equal to zero to the summation in Eq. (2.8) (if two nodes have no paths connecting them then we set $d_{ij} = \infty$). In particular, the extreme boundaries of efficiency are represented by the complete graph and the empty one having $E_{\text{complete}} = 1$ and $E_{\text{empty}} = 0$ respectively.

Clustering

Many real networks, even belonging to different contexts, share common topological properties like being *small-world* or *scale-free* for example (we will comment later on some of these properties). Among the various topological features that networks may share, one of particular relevance is that of displaying characteristic patterns of interconnections among the nodes. Those patterns can be viewed as the fundamental "building blocks" of the network structure, and often play a crucial role in the development of special functions associated with the system under consideration. These fundamental sub-units of networks have attracted a lot of attention from the scientific community over the years and have been studied in many different contexts [44–47].

A **motif** $\tilde{G}$ is a pattern of interconnections occurring either in a undirected or in a directed graph G at a number of times significantly higher than in randomized version of the graph itself. As a pattern of interconnections, a motif $\tilde{G}$ is usually meant as a connected (directed or not) n-node graph which is a *subgraph* of G. The concept of motifs was originally studied by Alon and coworkers [44, 48–51], who studied small n motifs in biological networks but, of course, also other kind of networks display such features [47, 52, 53].

Among the various motifs, and cycles in particular, available those of length three (*i.e.* the triangles) play a special role. In fact, such interconnection patterns have caught the attention of scientists quite before than when the role played by other motifs have come up to light. Clustering, also known as transitivity, is a typical property of acquaintance networks, where two individuals with a common friend are likely to know each other [24]. This can be quantified by defining the *transitivity* T of the graph as the relative number of triples, i.e. the fraction of connected nodes which also form triangles. An alternative possibility is to use the graph's **clustering coefficient** C, a measure introduced by Watts and Strogatz [54] defined as follows. A quantity c_i (local clustering of node i) is introduced, expressing how likely $a_{jm} = 1$ for two neighbors j and m of node i. Its value is obtained counting the number of edges (denoted by e_i) in the subgraph G_i of node i neighbours (thus, a local property). The local clustering coefficient is defined as the ratio between e_i and $k_i(k_i - 1)/2$, that is the maximum possible number of edges in G_i, so:

$$c_i = \frac{2e_i}{k_i(k_i - 1)} = \frac{\sum_{j,m} a_{ij}\, a_{jm}\, a_{mi}}{k_i(k_i - 1)} \,. \tag{2.9}$$

The clustering coefficient of the graph C is the average of c_i over all nodes in G:

$$C \equiv \langle c \rangle = \frac{1}{N} \sum_{j \in \mathcal{N}} c_i \,. \tag{2.10}$$

By definition, $0 \leqslant c_i \leqslant 1$, and $0 \leqslant C \leqslant 1$. In some cases, it may be useful to consider $c(k)$, the clustering coefficient of a connectivity class k, which is defined as the average of c_i over all nodes with degree k.

2.2. Network models

The topological measures exposed so far are completely general and applicable to any kind of graph. The characterization/study of real systems mappable as complex networks, by means of such measures, constitutes one of the main aims of network science. However, the need for theoretical models able to reproduce the salient features of such real systems (in order to unveil the underlying mechanisms responsible for their presence) constituted a key aspect in the early stages of network science. For this reason, along the years, a plethora of network models have been proposed each of them trying to focus on this or that particular aspect, but always with the scope of providing useful tools to the scientific community. The main distinction can be made between *equilibrium* models in which the system is already generated with its final number of nodes; and *out of equilibrium* (or growth) models where the system undergoes an evolution over "time" until it reaches a dynamical equilibrium.

The systematic study of random graphs was initiated by Erdős and Rényi in 1959 with the original purpose of studying, by means of probabilistic methods, the properties of graphs as a function of the increasing number of random connections [55]. The term *random graph* refers to the disordered nature of the arrangement of links between different nodes. Some paradigmatic examples of equilibrium models are those of Erdős-Rényi, De Solla Price, and Molloy-Reed [55–58]. In 1998, instead, Watts and Strogatz developed a model of random network capable of reproducing at the same time the finite clustering C of lattices, and the small average path length L typical of random graphs. They named graphs obtained with this model **small-world**, in analogy with the small-world phenomenon [59].

Soon after, instead, Barabási and Albert, [60], developed a non equilibrium model for growing networks able to reproduce the scale-free property (*i.e.* the fact that the degree distribution is a power law of the kind $P(k) \propto k^{-\gamma}$, with $2 < \gamma < 3$) observed in many real systems [61]. Since then, we have witnessed to a continuous increase over the years of the number of models proposed to account for all the properties displayed by the real networks including the scale-free property. This flurry of activity is responsible for the appearance, in the last few years, of some new interesting frameworks that have allowed network science to make another step forward towards a better comprehension of complex systems, paving the way to the understanding of phenomena that were previously not completely (or partly) understood. Two typical examples of these new approaches are: the *time-varying networks* [15], and the *multilayer networks* [23]. In the following sections we will illustrate the fundamental properties of some of those models that will be used in the papers displayed in chapters from 3 to 8.

2.2.1. Erdős Rényi random graphs

In their first article [62], Erdős and Rényi proposed a model to generate random graphs with N nodes and K links, that we will henceforth call **Erdős and Rényi random graphs** (ER) and denote as $G^{ER}_{N,K}$. Starting with N disconnected nodes, ER random graphs are generated by connecting couples of randomly selected nodes, avoiding multiple connections, until the number of edges equals K [62]. We emphasize that a given graph is only one outcome of the many possible realizations, an element of the statistical ensemble of all possible combinations of connections. For the complete description of $G^{ER}_{N,K}$ one would need to describe the entire statistical ensemble of possible realizations, that is, in the matrix representation, the ensemble of adjacency matrices [63]. An alternative model for ER random graphs consists in connecting each couple of nodes with a probability $p \in [0, 1]$. This procedure defines a different ensemble, denoted as $G^{ER}_{N,p}$ and contains graphs with different number of links: graphs with K links will appear in the ensemble with a probability

$p^K(1-p)^{N(N-1)/2-K}$.

The two models have a strong analogy, respectively, with the canonical and grand canonical ensembles in statistical mechanics [64], and coincide in the limit of large N. Notice that the limit $N \to \infty$ is taken at fixed $\langle k \rangle$, which corresponds to fixing $2K/N$ in the first model and $p(N-1)$ in the second one. Although the first model seems to be more pertinent to applications, analytical calculations are easier and usually are performed in the second model.

Let us comment now about some of the structural properties of ER graphs. In our discussion we consider ER networks obtained using the $G^{ER}_{N,p}$ method. In this case, the structural properties of ER random graphs vary as a function of p showing, in the case of the size of the giant component, a dramatic change at a critical probability $p_c = \dfrac{1}{N}$, corresponding to a critical average degree $\langle k \rangle_c = 1$. The transition at p_c has the typical features of a second order phase transition. Concerning the degree distribution, $P(k)$, the probability that a node i has $k = k_i$ edges is given by the binomial distribution. However, for large N and fixed $\langle k \rangle$, the degree distribution is well approximated by a Poisson distribution:

$$P(k) = e^{-\langle k \rangle} \frac{\langle k \rangle^k}{k!} \,. \tag{2.11}$$

The shortest path length L, like the diameter, scales like: $L \sim \ln N / \ln\langle k \rangle$. The clustering coefficient of $G^{ER}_{N,p}$ is equal to $C = p = k/N$. Hence, in the limit of large system size, ER random graphs have a vanishing C and finite (but small) average path length L.

2.2.2. Random geometric graphs

Usually, most of the paradigmatic examples of networks one could think of are *relational* networks, that is, graphs in which distances do not have physical meaning and are just dimensionless quantities measured in terms of edge hops. However, in many cases the physical space in which networks are embedded and the actual distances between nodes are important, such as in rail and road networks, ad hoc communication networks, and other geographical and transportation networks (for a quite exhaustive bibliography on spatial networks we suggest to look at the report of Barthélemy [16]). The **random geometric graph** (RGG) is a standard spatial network model that plays a role for spatial networks similar to the one played by the Erdős-Rényi random graph for relational ones [65, 66]. The construction process of a RGG with N nodes and radius R, on a d-dimensional metric space, can be summarized as follows:

1. The N nodes are placed on the Euclidean space $\Omega \in \mathbb{R}^d$ with uniform distribution.

2. An edge is created for every pair of nodes whose Euclidean distance is $r < R$.

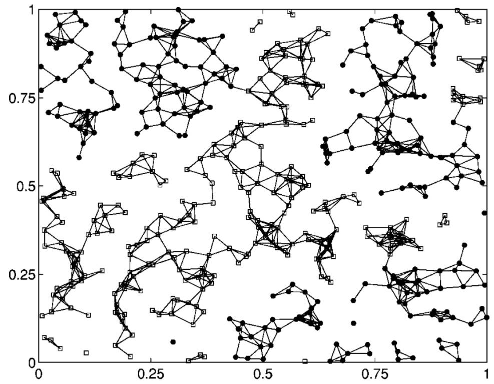


Figure 2.5: A 2D random geometric graph with $N = 500$, $\langle k \rangle = 5$ and no periodic boundary conditions. (from [65]).

The degree distribution of a RGG is given by a Poissonian as well. However, due to the spatial embedding, other topological estimators behave differently from those of ER graph. Also, the use of periodic or open boundary conditions affect severely such behavior. The relation between space and topological quantities stems from the intersection properties of d-dimensional hyperspheres (of radius r) centered on each node with those of the other nodes. Under such assumptions, for example, it is possible to determine that the average degree is equal to:

$$\langle k \rangle = Np = N\, 2^d\, V = N\, 2^d \frac{\pi^{d/2}\, r^d}{\Gamma\left(\dfrac{d+2}{2}\right)}, \tag{2.12}$$

where N is the number of nodes, p plays the role of the connection probability of $G^{ER}_{N,p}$ model, Γ is the gamma function, and V is the volume of the d-dimensional hypersphere of radius r. The clustering coefficient C, instead, is given by:

$$C \sim 3\sqrt{\frac{2}{\pi\, d}} \left(\frac{3}{4}\right)^{\dfrac{d+1}{2}}. \tag{2.13}$$

It is worth noting that the clustering coefficient does not depend on the size of the system thus, in contrast with ER random graphs, it does not vanish in the thermodynamic limit. The spatial embedding of RGG deeply affects the average path length L. As one could guess, the absence (for values of R smaller than the size of the system) of long range links draws L to behave like in a regular lattice. On the other hand, if we consider periodic boundary conditions the scenario changes and we are able to recover the results observed for ER networks. Concerning the second-order features and in particular, the degree correlation functions some results can be found in the paper of Antonioni *et al.* [67].

2.2.3. Barabási-Albert scale-free networks

As we have commented before, many real networks display degree distributions behaving like $P(k) \propto k^{-\gamma}$ thus having any characteristic scale for the degree k and for such reason named **scale-free networks**. The Barabási–Albert (BA) model is a non equilibrium model of network growth inspired by the formation of the World Wide Web, and is based on two basic ingredients: *growth* and *preferential attachment*.

More precisely, an undirected graph $G^{BA}_{N,K}$ is constructed as follows. Starting with m_0 nodes, at each time step $t = 1, 2, 3, \ldots, N - m_0$ a new node j with $m \leq m_0$ links is added to the network. The preferential attachment prescribes that the probability that a link will connect j to an existing node i is linearly proportional to the actual degree of i:

$$\prod_{j \to i} = \frac{k_i}{\sum_l k_l} \,. \tag{2.14}$$

In the limit $t \to \infty$, the model produces a degree distribution $P(k) \sim k^{-\gamma}$, with an exponent $\gamma = 3$. The case of a growing network with a constant attachment probability $\prod_{j \to i} = 1/(m0 + t - 1)$ produces, instead, a degree distribution $P(k) = e/m \exp(-k/m)$. This implies that the preferential attachment is an essential ingredient of the model. Also, since every new node has m links, the network at time t will have $N = m_0 + t$ nodes and $K = mt$ links, corresponding to an average degree $\langle k \rangle = 2m$ for large times.

Analyzing the behavior of other topological measures, we can observe that, in the BA model, the average distance L is smaller than an equivalent ER random graph, and increases logarithmically with N. Analytical results predict a double logarithmic correction to the logarithmic dependence $L \sim \log N / \log(\log N)$. The clustering coefficient vanishes with the system size as $C \sim N^{-0.75}$. This is a slower decay than that observed for random graphs, $C \sim N^{-1}$, but it is still different from the behavior in small-world models and real networks, where C is a constant independent of N.

Apart from the BA model, other models able to generate power law degree distribution have been proposed over the years. Among them, are worth to mention: the De Solla Price model, Molloy-Reed (MR) *configurational model*, the *atractiveness model* of Dorogovtsev *et al.* , and the so called *fitness model* of Caldarelli *et al.* just to cite a few [56, 57, 68, 69].

2.2.4. Multiplex networks

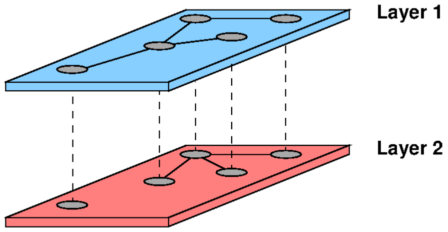
Recently, due also to the possibility of harvesting huge amount of data, there have been increasingly intense efforts to investigate networks with multiple types of connections or "*multilayer*" networks [70]. Such systems were examined decades ago in disciplines like sociology and engineering, albeit the explosive attempt to develop frameworks to study multilayer complex systems, and to generalize a large body of familiar tools from network science, is a recent phenomenon. Social networks are, indeed, a good example of multilevel (multilayer) networks because they intrinsically possess a rich variety of hierarchies, various kind of nodes and interactions, may have coarse-grained structures, can be bipartite, and so on. For such reasons, together with transportation networks, they can be considered a paradigmatic example of multilayer networks.

Within the efforts made to study such systems, a considerable amount of interest has been devoted to the characterization and modeling of multiplex networks, with the aim of creating a consistent mathematical framework to study, understand, and reproduce the structure of these systems [12, 23, 25, 26, 71–77]. In particular, in the last four years, two main tracks have been followed: from one side, people have worked to formulate a fully consistent mathematical framework to extend all the measures developed for single-layer networks to the multilayer ones [23, 76–78]; on the other side, there has been other people more interested in the study of well known dynamics on top of multilevel networks to see whether the multilayer structure produces new phenomena [13, 14, 71, 79, 80]. As an example, for some types of cascading-failure processes, a multilayer system can exhibit a "first-order" (*i.e.*, discontinuous) phase transition instead of the "second-order" (*i.e.*, continuous) phase transitions that are typical for single layer systems [12].

Before continuing with the mathematical formulation of multilevel networks, we must clarify an important aspect about terminology. With **multilayer network** one means a network made of two or more distinct layers where a node in a layer may not be present in the others, and where there are *intra-layer* connections among nodes of the same layer, and *inter-layer* connections among nodes belonging to different layers. For **multiplex network**, instead, we mean a particular kind of multilayer where each node is represented (as the same node) in every layer of the system, and where the inter-layer connections are only among the representatives of the same node. In this thesis we consider only multiplex networks. However, all the formalism that is going to be introduced can be easily extended to the case of multilayer networks and the interested reader could find a very good introduction to such formalism in [23, 78].

Having settled the terminology issue, we are now ready to introduce the mathematical framework used to describe multiplex networks. Three different formalisms have been proposed: the *adjacency tensor* [23], the *supra-adjacency matrix* [14, 23], and the *array of adjacency matrix* [77]. The first formalism is the most "elegant" but less intuitive, the second is nothing but a projection of the former with the advantage of being more useful for practical applications, the third is designed to work in the case where the inter-layer connections are less important than the intra-layer ones. In the following, we will introduce only the last two.

Considering a simple multiplex composed of just two layers (like that of Fig. 2.6), its **supra-adjacency matrix** $\mathcal{M}$ is a block matrix made of nothing else than the adjacency matrices of each layer $\mathcal{A}_i \mid i = 1, 2$ in the diagonal blocks, and a diagonal matrix $\mathcal{B}$ (whose elements are all the same) elsewhere. The diagonal blocks account for the intra-layer structure, while the off-diagonal blocks account for the inter-layer connections. Having built the supra-adjacency matrix with such hierarchical constraints, it is trivial to retrieve all the topological estimators we have seen so far. The interested reader could refer to the review of Kivelä *et al.* [78].



$$\mathcal{M} = \left(\begin{array}{c|c} \mathcal{A}_1 & \mathbb{I} \\ \hline \mathbb{I} & \mathcal{A}_2 \end{array}\right) \tag{2.15}$$

Figure 2.6: A multilayer graph (left) and its corresponding supra-adjacency matrix $\mathcal{M}$ (right). In particular, in this multilayer, all the links between layers have weight equal to one (thus $\mathcal{B} = \mathbb{I}$).

The supra-adjacency matrix can be obtained from the tensorial representation of the multiplex. Starting from the adjacency tensor of the multilayer, one can reduce its rank (i.e., order) by constraining the space of possible multilayer networks or by "flattening" the tensor [14, 81, 82].

The alternative formalism developed by Battiston *et al.* [77] has the advantage of being very intuitive and allowing a straightforward extension of all the topological estimators developed for simple networks. It also contemplates the possibility of projecting the layers into a single one as if we were looking at the multilevel system from above. We consider first a system composed of N nodes and M unweighted layers. We can associate to each layer α, $\alpha = 1, \ldots, M$, an adjacency matrix $\mathcal{A}^{[\alpha]} = \{a_{ij}^{[\alpha]}\}$, where $a_{ij}^{[\alpha]} = 1$ if node i and node j are connected through a link on layer α, so that each of the M layers is an unweighted network. Such a multiplex system is completely specified by the vector of the adjacency matrices of the M layers

$$\vec{\mathcal{A}} = \left\{ \mathcal{A}^{[1]}, \ldots, \mathcal{A}^{[M]} \right\} . \tag{2.16}$$

We define the degree of a node i on a given layer α as $k_i^{[\alpha]} = \sum_j a_{ij}^{[\alpha]}$, from which follows that $0 \leq k_i^{[\alpha]} \leq N - 1 \ \forall i, \forall \alpha$. Consequently, the degree of node i in a multiplex network is the vector

$$\vec{k_i} = \left(k_i^{[1]}, \ldots, k_i^{[M]} \right), \quad i = 1, \ldots, N . \tag{2.17}$$

As one could guess, $\sum_i k_i^{[\alpha]} = 2K^{[\alpha]}$, where $K^{[\alpha]}$ is the total number of links on layer α. Vectorial variables, such as $\vec{\mathcal{A}}$ and $\vec{k_i}$, are necessary to properly store all the richness of multiplex networks. However, it is also useful to define aggregated adjacency matrices (in which we disregard the fact that the links belongs to different layers) to be used as a term of comparison. We will show that aggregated matrices and the corresponding aggregated measures, with which one may be tempted to analyze the multi-layer structure, have limited potential and often fail in detecting the key structural features of a multiplex network. We define the aggregated topological adjacency matrix $\mathcal{A}' = \{a'_{ij}\}$ of a multiplex network, where

$$a'_{ij} = \begin{cases} 1 & \text{if } \ \exists \, \alpha : a_{ij}^{\alpha} = 1 \\ 0 & \text{otherwise} \end{cases} \tag{2.18}$$

This is the adjacency matrix of the unweighted network obtained from the multi-layer structure joining all pairs of nodes i and j which are connected by an edge in at least one layer of the multiplex network, and neglecting the possible existence of multiple ties between a pair of nodes and the nature of each tie as well. For the degree of node i on the aggregated topological network, we have

$$k'_i = \sum_j a'_{ij}. \tag{2.19}$$

Summing k'_i over all elements of the system, we obtain

$$\sum_i k'_i = 2K' , \tag{2.20}$$

where K' is the total number of links (also called the *size*) of the aggregated topological network. Matrix $\mathcal{A}'$ describes a single-layer binary network which can be studied using the well-established set of measures defined for single-layer networks. This representation turns out to be very simplistic and often insufficient to unveil the key features of multi-layer systems but proving it is out of the scope of this thesis. The interested reader could find it in the paper of Battiston *et al.* [77].

2.2.5. Time varying networks

Until recently, in most network studies, the time dimension has been projected out by aggregating the contacts between vertices to (sometimes weighted) edges, even in cases when detailed information on the temporal sequences of contacts or interactions would have been available.

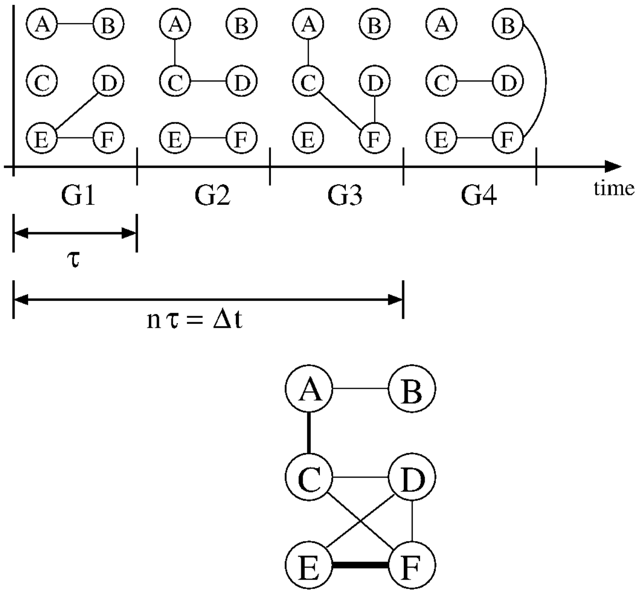


Figure 2.7: A time-varying graph represented as a series of graph $G_1, G_2, \ldots$, each one corresponding to an instant of time τ. If we project a certain number of snapshots, corresponding to a time window $\Delta t = n\,\tau$, onto a single static (weighted) graph we will obtain the graph displayed at the bottom.

On one hand, one will always lose information when projecting a temporal network structure to a static graph (see Fig. 2.7 for an illustration). On the other hand, in some cases, this loss of information is probably too insignificant to make up for the more complicated analysis and modeling needed for the temporal graph approach. So, we are legitimated to ask: *when are temporal networks a suitable framework for analysis and modeling*? A special case of the requirement that a system should have temporal structure for it to suit a temporal-network framework, relates to time scales [83]. If the dynamical system on the network is too fast compared to the dynamics of the contacts, or when edges are active, then there is no need to model the system as a temporal network. One example is the Internet where the data packets travel much faster than the topology changes. So, the reason behind the neglection of temporal structure of complex networks until some years ago may be also due to the fact that in recent times we are experiencing an acceleration of the time scales at which infrastructures evolve compared to the dynamical phenomenon acting on them. A typical example of such acceleration is given by the proximity patterns of humans – data on who is close to whom at what time – that are important both for understanding the spread of airborne pathogens and word-of-mouth spreading of information. Nowadays, in fact, trips across the world are much faster than, say, twenty years ago. Another fundamental aspect, that exerts deep consequences on the evolution of dynamical processes is that related with *causality*. The *arrow of time*, in fact, is one of the most important concepts in physics and can matter a lot, and as we shall see below, the timings of connections and their correlations do have effects that go beyond what can be captured by static networks.

Temporal networks can be divided into two (rough and overlapping) classes corresponding to the two types of representations. In the first representation, (Fig. 2.8(a)) there is a set of N vertices, $\mathcal{N}$, interacting with each other at certain times, and the durations of the interactions are negligible. In this case, the system can be represented by a **contact sequence**, *i.e.* a set of $\mathcal{C}$ contacts, triples (i, j, t) where $i, j \in \mathcal{N}$ and t denotes time. Equivalently, one can represent the system by $\mathcal{N}$, a set of K edges (pairs of vertices) $\mathcal{K}$, and, for $e \in \mathcal{E}$, a non-empty set of times of contacts $T_e = \{t_1, \ldots, t_n\}$. In

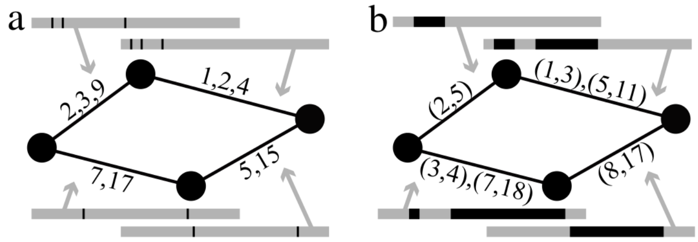


Figure 2.8: Contact sequences and interval graphs. This figure illustrates the two fundamental temporal network representations, namely: contact sequences (a) and interval graphs (b). The times of the contacts are states next to the edges. (from [15]).

the second class of temporal networks we discuss, **interval graphs**, where edges are not active over a set of times but rather over a set of intervals $T_e = \{(t_1, t'_1), \ldots, (t_n, t'_n)\}$, where the parentheses indicate the periods of activity. The static graph with an edge between i and j if and only if there is a contact between i and j is called the (time) **aggregated graph**.

$$a(i,j,t) = \begin{cases} 1 & \text{if } i \text{ and } j \text{ are connected at time } t \\ 0 & \text{otherwise} \end{cases} \tag{2.21}$$

An approach, similar to that used by Battiston *et al.* for multiplex networks [77], can also be formulated for time-varying networks. Also, it is straightforward to extend all the topological features seen so far to the time-varying graphs. The interested reader could find a complete treatment of those quantities in [15] (and references therein).

In most cases, the contacts between the same node pair in time-varying systems tend to be clustered in time, *i.e.*, they show persistence over time [84, 85]. For instance, people tend to engage in relations for continuous intervals of time. Hence, a given link has a higher probability to appear in graph G_t if it was already present in graph G_{t-1} (see Fig. 2.8 snapshots G_2 and G_3). To quantify this effect, following Refs. [86, 87], one can compute C, the average topological overlap of the neighbor set of a node i between two successive graphs in the sequence:

$$C^i = \frac{1}{T-1} \sum_{t=1}^{T-1} \frac{\sum_j a_{ij}^t a_{ij}^{t+1}}{\sqrt{k_i^t k_i^{t+1}}}, \tag{2.22}$$

where a_{ij}^t are the elements the adjacency matrix of the time-varying graph at snapshot t, $k_i^t = \sum_j a_{ij}^t$ is the degree of node i at snapshot t, and T is the duration of the whole observation interval. Tang *et al.* [87], name this metric the **temporal-correlation coefficient** of graph G but an alternative, and more intuitive, name is *temporal clustering.* Notice that C^i values fall in $[0,1]$ interval. In general, a higher value of C^i is obtained when the interactions of node i persist longer in time, while C^i tends to zero if the interactions of i are highly volatile. Also, it is worth mentioning that, following the ideas exposed in [88] for simple graphs, it is possible to extend the concept of temporal clustering to the case of weighted time-varying networks.

As anticipated above, the inclusion of the time dimension in the description of the systems has dramatic consequences on them because it severely affects two fundamental concepts: *path* and *distance.* Paths, in fact, must necessarily be constrained to sequences of link activations that follow one another in time. Thus, in a temporal graph, paths are usually defined as *sequences of contacts with non-decreasing times that connect sets of vertices.* Kempe *et al.* [89] and Holme *et al.* [90] call such paths **time-respecting**. A striking difference between static and temporal networks is that the paths *are not transitive.* The existence of time-respecting paths from i to j and j to k does not imply that there is a path from i to k. The set of vertices that can be reached by time-respecting paths from vertex i is called the **set of influence** of i. This is important e.g. for disease spreading, as it represents the set of vertices that can eventually be infected if i is the source of infection. One can also define the **source set** of i as the set of vertices that can reach i through time-respecting paths within the observation window. This set consists of all vertices that may have been the source of an infection to node i. Finally, two vertices i and j of a temporal network are defined to be **strongly connected** if there is a directed, time-respecting path connecting i to j and vice versa, while they are **weakly connected** if there are undirected time-respecting paths from i to j and j to i, *i.e.* the directions of the contacts are not taken into account.

To every time-respecting path is associated a *duration*, measured as the time difference between the last and first contacts on the path; note that some authors have called it the **temporal path length** [91]. Analogously to the shortest paths that define the geodesic distance, one can find the fastest time-respecting path(s) between two nodes; the shortest time within which i can reach j is called their **latency** or **temporal distance** [91]. As the concepts of temporal duration and link-wise distance have been used interchangeably in the literature. In the following, we will reserve the word "distance" for measuring numbers of links, and "duration" and "latency" for measuring times. Concerning the *diameter*, one option is to define it as the longest average latency albeit one would then have to choose how to deal with infinite latencies.

Concerning the identification of those meso-structures that play an important role in the system, the **reachability graphs**, or *path graphs*, or *associated influence digraph* [92, 93]. In such a case one puts a directed edge from ver-

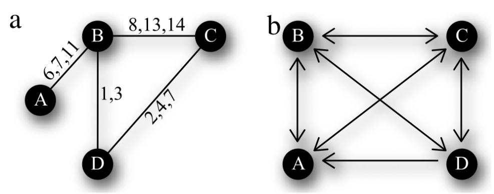


Figure 2.9: Reachability graphs. Panel (a) shows a contact sequence and (b) shows the corresponding graph (from [15]).

tices A to B if there is a time-respecting path from A to B (see Fig. 2.9). Such a graph, thus, shows which vertices can possibly affect which others. The average degree k of a reachability graph is thus the average worst-case outbreak size minus one. In other words, for any contact structure that supports a pandemic, the reachability graph will be dense ($k \sim N$).

Temporal activity and activity driven model

Temporal networks are often used to map time-varying interactions among agents. So the information encoded in the contact sequence may be used as a proxy to study the **temporal activity** of the system. Especially, for temporal networks of human communication, it has been discovered that the timings among adjacent events are often bursty, periodic, and deviate from the more uniform times expected from a memoryless, random Poisson process [94–97]. In fact, if one displays the scaled inter-contact time distributions and compare it against similar distributions for uniformly random contact times, broad tails are typically observed.

Goh and Barabási [98] use as their starting point the coefficient of variation, defined as the ratio of the standard deviation of the inter-contact times to their mean, $\sigma_\mathcal{T}/m_\mathcal{T}$. For a Poissonian contact sequence, $\sigma_\mathcal{T}/m_\mathcal{T} = 1$. Using this quantity, the **burstiness** of a sequence is then defined as

$$B = \frac{\dfrac{\sigma_\mathcal{T}}{m_\mathcal{T}} - 1}{\dfrac{\sigma_\mathcal{T}}{m_\mathcal{T}} + 1} = \frac{\sigma_\mathcal{T} - m_\mathcal{T}}{\sigma_\mathcal{T} + m_\mathcal{T}}; \tag{2.23}$$

Burstiness allows us to measure this non Poissonian behavior of contact sequence, but the question now becomes: *are we able to build a model of time-varying network that is able to mimic the burstiness patterns displayed by real systems*?

The **activity-driven model**, introduced by Perra *et al.* in [99], is a simple model to generate time-varying graphs starting from the empirical observation of the activity of each node, in terms of number of contacts established per unit time. Given a characteristic time-window Δt, one measures the activity potential x_i of each agent i, defined as the total number of interactions (edges) established by i in a time-window of length Δt divided by the total number of interactions established on average by all agents in the same time interval. Then, each agent is assigned an activity $\chi_i = \eta x_i$, which is the probability per unit time to create a new connection or contact with any another agent j. The coefficient η is a rescaling factor, whose value is appropriately set in order to ensure that the total number of active nodes per unit time in the system is equal to $\eta\langle x\rangle N$, where N is the total number of agents. Notice that η effectively determines the average number of connections in a temporal snapshot whose length corresponds to the resolution of the original data set.

The model works as follows. At each time t the graph G_t starts with N disconnected nodes. Then, each node i becomes active with probability $\chi_i\Delta t$

and connects to m other randomly selected nodes. At the following time-step, all the connections in G_t are deleted, and a new snapshot is sampled.

Notice that time-varying graphs constructed through the activity-driven model preserve the average degree of nodes in each snapshot, but impose that connections have, on average, a duration equal to Δt, effectively washing out any temporal correlation among edges. Thus, activity-driven model can be thought as an equivalent of ER model for time-varying networks with no temporal correlations. Also, activity-driven generated networks have been used to study the behavior of dynamical models like spreading and random walk on top of time-varying networks [100].

2.3. Dynamical processes

Up to now, we have commented on topological properties and generative models of complex networks. The concept of time, plays a role in the growth process of the system and the term *evolution* of the system refers, indeed, to such growing. However, this is just one side of the coin. The other side is represented by the study of **dynamical processes** acting on top of a complex topology. In such context, the word evolution assumes a new meaning because it refers to the change of the dynamical state of the system along time. The behavior of a variety of processes, spanning from synchronization up to diffusion, has been studied over the years and a lot of groundbreaking results have been found. However, in this thesis we will focus on just three of them, namely: **spreading of infections** (Sec. 2.3.1), **evolution of cooperation** (Sec. 2.3.3), and **synchronization** (Sec. 2.3.2). In the following sections, after introducing the basic concepts and ideas of these processes, we will provide some of main results that have been obtained combining those processes on systems where the pattern of interactions among the elements is described by a complex network. The aim of this section is act as a sky jump, providing the reader with the fundamental tools to understand results obtained going beyond simple models.

2.3.1. Spreading of infections

The study of how a disease spreads in a population is a fundamental topic in medical research. Since the 20^{th} century it attracted a lot of attention from mathematicians and, nowadays the mathematical modeling of infectious diseases is a key concept in epidemiology. Physicists and engineers entered the field when the similarities between the spreading of a disease and a percolation process were enlightened [101]. However, the history of disease spread mathematical modeling is much longer and is rooted in the attempts of John Graunt to provide demographical methods to monitor the diffusion of bubonic plague in 1662. We had to wait until 1927 to see the first compartmental models of epidemic spreading thanks to the work of Kermack and McKendrick. Then, another leap lead us to 2001 when, for the first time, an epidemic spreading model acting on top of an heterogeneous network, by Pastor-Satorras and Vespignani, leading to results that have revolutionized the entire field of mathematical epidemiology [102].

A fundamental building block in modeling infectious diseases is represented by the so-called **compartmental models**, in which the population is divided into groups each of them representing a possible state of the disease [6, 103]. In the simplest case, population can be divided into two groups: **susceptible** (S) healthy people that can catch the disease if in contact with infected individuals; and **infected** (I) people that currently have the disease and can transmit it to the others. Within this framework, adding additional states like the **recovered** (R) one, describing people that have been infected and are now healthy again (or died), it is possible to model a variety of different diseases.

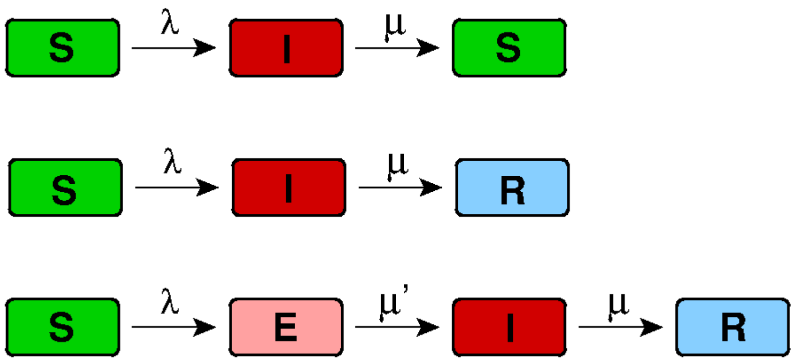


Figure 2.10: Examples of compartmental models. From top to bottom: the SIS, SIR, and SEIR models. Each compartment accounts for a possible state in which an agent can be, namely: Susceptible (S), Exposed (E), Infected (I), and Recovered (R). The transitions between compartments are controlled by rates λ, μ, and μ'.

One of the simplest epidemiological models describes diseases that can get caught only once and end up in a immunization or death of the infected is the *Susceptible, Infected, Recovered* (SIR). The model is based on two parameters, the **transmission rate** λ, and the **recovery rate** μ. At the beginning of the spreading process, an initial seed I_0 of infected individuals is inserted in the population. Then, at each time step, a susceptible individual l if in contact with an infected j is infected with probability λ. At the same time step, an infected individual m becomes recovered with a probability μ. The process continues its evolution until all the agents are recovered or until there are not any infected agent. The transitions explained above are summarized in the scheme below.

$$S(l) + I(j) \xrightarrow{\lambda} I(l) + I(j) \tag{2.24}$$

$$I(m) \xrightarrow{\mu} R(m)\,. \tag{2.25}$$

It is also possible to model diseases that do not provide immunization to their survivors, such as the common cold. These diseases are well described in terms of the *Susceptible, Infected, Susceptible* (SIS) model in which an infected subject returns to the S state at rate μ. In the simplest case, the spreading of the disease is considered much faster than the mean lifespan of an individual, so birth and death rates are not taken into account. Nevertheless, it is possible to imagine more complicated scenarios in which births and deaths are considered or, alternatively, other classes such as **exposed** (latent) (E) – *i.e.* individuals that are infected but are not infectious–. The two presented models (SIR and SIS), although very similar, lead to a totally different behavior. In the SIS two steady states are allowed: one with $I = 0$ in which the disease is absorbed by the system and *no real outbreak* takes place, or an *endemic state* $I > 0$, in which the infected population reaches a macroscopic stationary size and the disease propagates indefinitely. The SIR model prescribes that, in the final state, the number of infected individuals is always zero giving rise to two possible scenarios: the disease did not produce an outbreak and the final recovered population size is close to that of the initial seed, or, the disease propagate to a finite fraction of the population.

Even though the dynamical behavior of the two models is very different, in both cases the two parameters λ and μ (or, more correctly, their ratio $\sigma = \lambda/\mu$) control the appearance (or not) of an epidemic outbreak. In particular, we are interested in predicting the critical point σ_c at which the epidemic transition from the absorbing phase (*i.e.* no finite outbreak) to the endemic phase (*i.e.* finite fraction of infected, or recovered) occurs. To get some initial insight on the value of the critical point and the nature of the epidemic transition, it is possible to consider a simple scenario named **homogeneous mixing**. In this approximation, both the SIR and SIS models are considered within the hypothesis [6], that the contacts between individuals are chosen randomly among the entire population. Although this strong approximation does not consider any geographical or local detail, it permits to represent the system as a set of ordinary differential equations for the densities of individuals belonging to each class. For the SIS model, the equation is:

$$\frac{d\,\rho(t)}{dt} = \; -\mu\,\rho(t) + \lambda\;(1 - \rho(t))\,\bar{k}\,\rho(t)\,, \tag{2.26}$$

where $\bar{k}$ is the number of contacts in the unit time (that is fixed for all individuals) and $\rho(t)$ is the *fraction* of infected individuals at time t. Note that the normalization condition $s(t)+\rho(t) = 1\,\forall t$ must hold thus implying that the fraction of susceptible individual at time t, $s(t)$, must be given by $s(t) = 1 - \rho(t)$. Equation (2.26) can be explained in the following way: susceptibles become infected at a rate that is proportional to the infection probability λ times the densities of infected, susceptibles, and the number of contacts per unit time; infected, instead, decay into susceptible state at rate μ. It is worth to notice that λ and μ are fixed and equal for all the contacts. The Eq. (2.26) can be solved analytically and predicts the presence of a non-zero epidemic threshold λ_c (assuming $\mu = 1$) in order for the outbreak of the disease to appear. In particular, by considering the so- called **epidemic incidence** (indicated as: $i_\infty = \lim_{t\to\infty} i(t)$); if $\lambda > \lambda_c$, the value of i_∞ assumes a finite nonzero value. Otherwise, i_∞ is infinitesimally small in the very large population limit.

In order to calculate the critical threshold, λ_c, one has to solve the linear stability analysis of the so-called *healthy state* (*i.e.* $s^\star(t) = 1$) of the system. In particular, starting from Eq. (2.26) we have:

$$\frac{d\,\rho(t)}{dt} = 0 \;\Leftrightarrow\; -\mu_c\,\rho^\star(t) + \lambda_c\,\left(1-\rho^\star(t)\right)\bar{k}\,\rho^\star(t) = 0\,,$$
$$\mu_c\,\rho^\star(t) = \lambda_c\,\left(1-\rho^\star(t)\right)\bar{k}\,\rho^\star(t)\,.$$

The last equation has two solutions: a "trivial" one corresponding to the healthy state $\rho^\star = 0$, and another one corresponding to an endemic state with a small (but finite) fraction of infected individuals $\rho^\star = \varepsilon$ with $\varepsilon \ll 1$. In the latter case, the equation above becomes:

$$\mu_c\,\varepsilon = \lambda_c\,(1-\varepsilon)\,\bar{k}\,\varepsilon\,.$$

Since $\varepsilon \ll 1$, we have that $(1-\varepsilon) \simeq 1$ leading to:

$$\mu_c = \lambda_c\,\bar{k}\,.$$

Without loss of generality, we can set $\mu = 1$ and obtain:

$$\lambda_c = \frac{1}{\bar{k}}\,.$$

Epidemics on complex networks

Pastor-Satorras and Vespignani [7, 104], analysed the effects of network connections on the rate and diffusion patterns of a disease. Using a mean-field analysis, they solved a modified version of SIS model on heterogeneous graphs with a generic degree distribution $P(k)$, and a finite average connectivity $\langle k\rangle$.

This approach is called **heterogeneous mean-field** (HMF). The HMF approach 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 class k. Let us consider again the SIS model where the quantities $s_k(t)$ and $\rho_k(t)$ are the densities of susceptible and infected nodes in the degree class k at time t. We can also write the normalization condition as:

$$s_k(t) + \rho_k(t) = 1 \quad \forall t\,, \tag{2.27}$$

and express the global values of the epidemic incidence averaging over the various connectivity classes: $i_\infty = \lim_{t\to\infty} i(t)$, with $i(t) = \sum_k P(k)\, i_k(t)$. At the mean-field level, these densities satisfy the same differential equation as in Eq. (2.26), but differentiated by connectivity classes:

$$\frac{d\,\rho_k(t)}{dt} = -\mu\rho_k(t) + \lambda\, k\, \left(1 - \rho_k(t)\right)\, \Theta_k(t)\,, \tag{2.28}$$

where $\Theta_k(t)$ represents the probability that any given link points to an infected node [7, 104].

$$\Theta_k(t) = \sum_{k'} P\left(k'|k\right)\rho_k(t) = \frac{\sum_k kP(k)\rho_k(t)}{\langle k\rangle}\,, \tag{2.29}$$

because we have considered a random network in which the conditional probability does not depend on the originating node. Solving the system (2.28), and setting again $\mu = 1$, it is possible to obtain the epidemic threshold:

$$\lambda_c = \frac{\langle k\rangle}{\langle k^2\rangle}\,. \tag{2.30}$$

This latter result has profound implications in highly heterogeneous network. In fact, for graphs in which $\langle k^2\rangle < \infty$ (as ER graphs), the threshold has a finite value and a standard phase transition is observed. Instead, for graphs with highly fluctuating degree distributions $\langle k^2\rangle$ can assume high values and in some cases like scale-free networks with $2 < \gamma \le 3$ can diverge, leading to a vanishing epidemic threshold for the $N \to \infty$ limit.

The latter result spurred a lot of research related on the analysis and contention of epidemic outbreaks in networked populations. For instance, in the recent years, a large body of works have addressed the design of effective immunization strategies, the study of **epidemic metapopulation models** [105], coevolutionary model based on the combination of motion of agents and spreading dynamics [8, 106], up to the design of agent based models capable of making “disease forecast” [107]. Nevertheless, disease spreading still have a lot of question without an answer and more efforts are needed in order to find them.

2.3.2. Synchronization

Most of you have probably attended to a concert or a show. Usually, if the performer is not that bad, at the end of the show, the audience start to clap its hands to applaud. Surprisingly, the claps of the hands that initially are erratic will gradually get more and more organized until becoming completely synchronous. **Synchronization** is the coordination of events to operate a system in unison. The first results on synchronization were obtained by Huygens showing that only when two systems interact among them they can get synchronized [108]. Nevertheless, the works that shed light on the synchronizability of systems composed by many interacting agents are those of Winfree and Strogatz being inspired by biological systems [109, 110]. In particular, Winfree discovered that a population of non-identical oscillators can exhibit a remarkable cooperative phenomenon. When the variance of the distribution of *natural* (intrinsic) frequencies of these oscillator is large, the oscillators run incoherently, each one near its own natural frequency. This behavior remains when reducing the variance until a certain threshold is crossed. Below the threshold the oscillators begin to synchronize spontaneously.

Within the framework of non-linear oscillators, one of the most studied models is that of **Kuramoto** [111, 112]. This model, presented in 1975, has allowed the study of synchronization of limit cycle phase oscillator populations weakly coupled among them. The Kuramoto model approach to synchronization was a breakthrough for the understanding of synchronization in large populations of oscillators. An interesting introduction on the Kuramoto model can be found in [113]. In a nutshell, given a population of N oscillators, the angular velocity of an oscillator i, $\dot{\theta}_i$ obeys to the following law:

$$\dot{\theta}_i = \omega_i + \frac{K}{N} \sum_{j=1}^{N} \sin\left(\theta_j - \theta_i\right) . \tag{2.31}$$

Where, ω_i is the natural frequency of oscillator i that is drawn from a distribution, $g(\omega)$, of frequencies. K, instead, is the *coupling constant* which is supposed to be constant. The synchronization of the system is measured through the **order parameter** r, defined as:

$$r\, e^{i\psi} = \frac{1}{N} \sum_{j=1}^{N} e^{i\theta_j} . \tag{2.32}$$

To visualize the dynamics of the phases, it is convenient to imagine a swarm of points running around the unit circle in the complex plane. The complex order parameter is a macroscopic quantity that can be interpreted as the collective rhythm produced by the whole population. It corresponds to the centroid of the phases. The radius $r(t)$ measures the phase coherence, and $\psi(t)$ is the average phase. The values $r \simeq 1$ and $r \simeq 0$ describe the limits in which all oscillators are either phase locked or move incoherently, respectively. Another interesting property of Kuramoto model is that it exist a critical value of the coupling K_c, above which the long-term behavior of the system corresponds to the phase locked state, and below which we observe incoherence. Such critical value can be derived analytically (for the complete derivation, the interested reader could look at [19]) and it is equal to:

$$K_c = \frac{2}{\pi\, g(0)}\,; \tag{2.33}$$

where $g(0)$ is the distribution of the natural frequencies ω calculated in $\omega = 0$.

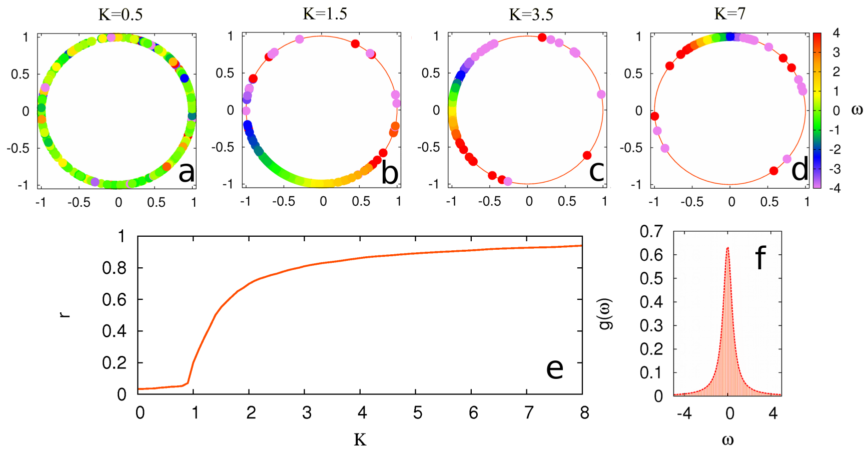


Figure 2.11: (panels a to d) Geometric interpretation of the order parameter r. The phases θ_j are plotted on the unit circle as a function of the coupling K, and their centroid is given by the complex number $r\,e^{i\psi}$. (panel e) Order parameter r as a function of the coupling K. (panel f) Distribution of the natural frequencies $g(\omega)$. From [114].

Kuramoto model on complex networks

So far we have commented on coupled oscillators where each element "feels" the presence of all the other ones. Of course, we know that such assumption is no longer valid in real systems and that the pattern of interaction among elements is better described in terms of a complex network. Thus, a question arise: what are the effects of considering an underlying topology on the onset of synchronization? Various attempts have been made through the years and we have now a better comprehension of the phenomena taking place when an ensemble of oscillator goes into the synchronized state.

Studies on synchronization in complex topologies where each node is considered to be a Kuramoto oscillator, were first reported for WS networks [115] and BA graphs [116]. These works are mainly numerical explorations of the onset of synchronization, being their main goal being the characterization of the critical coupling beyond which groups of nodes beating coherently first appear. The Kuramoto model on a generic complex network is defined as:

$$\dot{\theta}_i = \omega_i + \lambda \sum_{j=1}^{N} a_{ij} \, \sin\left(\theta_j - \theta_i\right) \, , \tag{2.34}$$

where a_{ij} are the elements of the adjacency matrix, and the coupling has been modified from K/N into λ. In analogy with what we have seen in the mean-field case, it is possible to obtain an equivalent of the order parameter r of Eq. (2.32) also in the networked case.

As well as in the mean field case, there exists a critical value of the coupling λ_c and it follows that such value is related to both K_c and the largest eigenvalue v_{max} of the adjacency matrix, yielding:

$$\lambda_c = \frac{K_c}{v_{max}} = K_c \frac{\langle k \rangle}{\langle k^2 \rangle} = \frac{2}{\pi \, g(0)} \, \frac{\langle k \rangle}{\langle k^2 \rangle} \, ; \tag{2.35}$$

The use of mean field approximation in Eq. (2.35) produces, as a surprising result, that the critical coupling λ_c in complex networks is nothing else but the one corresponding to the all-to-all topology K_c re-scaled by the ratio between the first two moments of the degree distribution, regardless of the many differences between the patterns of interconnections [117, 118]. It is worth mentioning that Eq. (2.35) has the same functional form for the critical points of other dynamical processes such as percolation and epidemic spreading processes [119]. It would imply that the critical properties of many dynamical processes on complex networks are essentially determined by the topology of the graph.

Before concluding the discussion about the Kuramoto model on complex networks, it is worth making few comments on the way the system reach the synchronized state. Concerning the region where we are neither close to the onset of synchronization nor at complete synchronization? How is the latter state attained when different topologies are considered?

Gómez-Gardeñes *et al.* have seen in [120] that the onset of synchronization first occurs for SF networks. As the network substrate becomes more homogeneous, the critical point λ_c shifts to larger values and the system seems to be less synchronizable. On the other hand, they also showed that the route to complete synchronization, $r = 1$, is sharper for homogeneous networks.

From a microscopic analysis, it turns out that for homogeneous topologies, many small clusters of synchronized pairs of oscillators are spread over the graph and merge together to form a giant synchronized cluster when the effective coupling is increased. On the contrary, in heterogeneous graphs, a central core containing the hubs first comes up driving the evolution of synchronization patterns by absorbing small clusters [120].

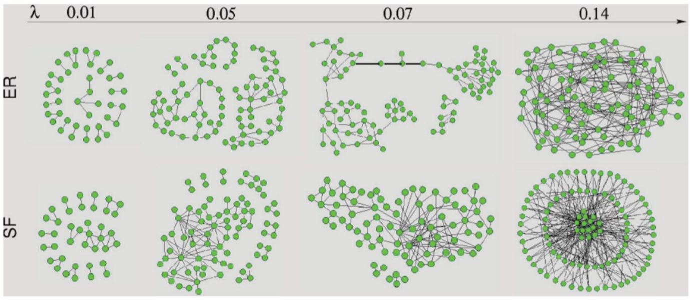


Figure 2.12: Synchronized clusters for several values of λ for two different topologies (ER and SF). The evolution of local synchronization patterns is always agglomerative, however, it follows two different routes. In the ER case, the growth of the GC proceeds by aggregation of small clusters of synchronized nodes, while for the SF network the central core groups the smaller clusters around it. (from [120]).

The Stuart–Landau model on complex networks

Another limit-cycle oscillator is the one described by the **Stuart-Landau model** (SL) [121]. We introduce directly the SL on complex networks and comment some of the properties of this model. Let us consider a network of N coupled Stuart-Landau oscillators [121]. Each node i is characterized by two variables, $(x_i, y_i)^T$, whose dynamical evolution follows:

$$\begin{aligned} \dot{x}_i &= \alpha x_i - x_i^3 - x_i y_i^2 - \omega_i y_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\left(x_j - x_i\right) \\ \dot{y}_i &= \omega_i x_i + \alpha y_i - x_i^2 y_i - y_i^3 + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\left(y_j - y_i\right) \ , \end{aligned} \tag{2.36}$$

where $\sqrt{\alpha}$ and ω_i are respectively the amplitude and the (natural) frequency of oscillator i when de-coupled from the rest of the system. The second term on the right accounts for the coupling of the dynamics of node i with its k_i neighbors. The strength of the coupling is controlled by λ ($\lambda = 0$ in the uncoupled limit) while $\mathcal{A} = \{a_{ij}\}$ represents the adjacency matrix of the network.

To study the synchronization properties of system (2.36), we work with the phase variable of each oscillator, defined as $\theta_i = \tan^{-1}(y_i/x_i)$. Re-writing Eqs.(2.36) into polar coordinates (ρ, θ) we obtain:

$$\begin{aligned} \dot{\rho}_i &= \alpha\rho_i - \rho_i^3 + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\left[\rho_j \cos(\theta_j - \theta_i) - \rho_i\right] \\ \dot{\theta}_i &= \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\,\frac{\rho_j}{\rho_i}\,\sin(\theta_j - \theta_i)\ . \end{aligned} \tag{2.37}$$

It is possible to demonstrate that, under certain circumstances, the Stuart-Landau oscillator behave like a Kuramoto one. In particular, starting from Eq. (2.37), we impose the following change of variable $\rho_i = \sqrt{\alpha}\,R_i$. Under such change, Eq. (2.37) become:

$$\begin{aligned} \sqrt{\alpha}\,\dot{R}_i &= \alpha\sqrt{\alpha}\,R_i - \sqrt{\alpha^3}\,R_i^3 + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\left[\sqrt{\alpha}\,R_j \cos(\theta_j - \theta_i) - \sqrt{\alpha}\,R_i\right] \\ \dot{\theta}_i &= \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\,\frac{\sqrt{\alpha}\,R_j}{\sqrt{\alpha}\,R_i}\,\sin(\theta_j - \theta_i)\ . \end{aligned} \tag{2.38}$$

Rescaling the time in the first equation of the above system setting $dT = \alpha\,dt$, we have:

$$\begin{aligned} \frac{dR_i}{dT} &= R_i - R_i^3 + \frac{\lambda}{\alpha\,k_i}\sum_{j=1}^{N} a_{ij}\left[R_j \cos(\theta_j - \theta_i) - R_i\right] \\ \dot{\theta}_i &= \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\,\frac{R_j}{R_i}\,\sin(\theta_j - \theta_i)\ . \end{aligned} \tag{2.39}$$

If we put ourselves in the limit for $\alpha \to \infty$, we observe how the equation above become:

$$\begin{aligned} \frac{dR_i}{dT} &= 0 \\ \dot{\theta}_i &= \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}\,\frac{R_j}{R_i}\,\sin(\theta_j - \theta_i)\ . \end{aligned} \tag{2.40}$$

because, in this limit, $R \to 1$. So, in conclusion, we are left with a Kuramoto oscillator where the coupling is re-scaled by k_i (as it was studied in [122]).

2.3.3. Evolutionary game theory

If natural selection among individuals favors the survival of the fittest, why would one individual help another at a cost to itself? Charles Darwin himself noted the difficulty of explaining why a worker bee would labor for the good of the colony, because its efforts do not lead to its own reproduction. The social insects are "one special difficulty, which first appeared to me insuperable, and actually fatal to my theory," he wrote in his book entitled *On the Origin of Species* [123]. So pervasive is cooperation that Martin Nowak of Harvard University ranks it *third pillar of evolution*, alongside of mutation and natural selection. "Natural selection and mutation describe how things change at the same level of organization", he explains. "But natural selection and mutation alone would not explain how you get from the world of bacteria 3 billion years ago to what you have now" [124]. Evolution is based on a fierce competition between individuals and should therefore reward only selfish behavior. Every gene, every cell, and every organism should be designed to promote its own evolutionary success at the expense of its competitors. Yet, we observe cooperation on many levels. That puzzle has inspired biologists, mathematicians, even economists to come up with ways to explain how cooperation can arise and thrive.

A **cooperator** is someone who pays a cost, c, for another individual to receive a benefit, b. A **defector** has no cost and does not deal out benefits. Cost and benefit are measured in terms of fitness. Over the years, many mechanisms have been proposed to explain the emergence of cooperation [125]. It is worth commenting about some of them. Assuming that cooperation can emerge only if the cost over benefit c/b ratio is suitable, some of most important mechanisms allowing the emergence of cooperation are: **kin selection**, **direct reciprocity**, **group selection** and, **network reciprocity**.

The mechanism named kin selection was proposed by J. B. S. Haldane and can be resumed by the following statement: "*Would I lay down my life to save my brother? No, but I would to save two brothers or eight cousins*", arguing that cooperation could emerge only among *members belonging to the same genetic pool.* To help explain our cooperative nature, in the 1970s, Robert Trivers came up with the idea of direct reciprocity or, alternatively, *reciprocal altruism* which work somehow like: "*You scratch my back, and I will scratch yours*". Kevin Foster, and others, have become convinced that *competition among groups* (*i.e.* group selection) can promote cooperation. In other words, evolutionary forces can act on several levels, with natural selection pushing to make individuals less cooperative being countered by competition at the level of the group, because groups with greater cooperation among members tend to survive better. An example of that is the so-called **green-beard gene effect**, which consist in the fact that a gene that enables an individual to recognize –as one could recognize a green beard– and cooperate with others who carry that same gene, promote the formation of groups where the collective efforts are directed towards the preservation of such feature.

All the mechanisms commented above are based on a well-mixed populations, where everybody interacts equally likely with everybody else. This approximation is used by all standard approaches to evolutionary game dynamics [5, 126, 127]. Of course, such assumption is not realistic. Spatial structures or social networks imply that some individuals interact more often than others. One approach of capturing this effect is evolutionary graph theory [52], which allows us to study how spatial structure affects evolutionary and ecological dynamics [128]. The individuals of a population occupy the vertices of the graph. In this setting, cooperators can prevail by forming network clusters, where they help each other. The resulting *network reciprocity* is a generalization of "spatial reciprocity" [129]. Let us get more insight on this last mechanism for the promotion of cooperation, introducing first some notions on evolutionary games played one against another or in a group. After this, we will comment about some results of evolutionary games on complex networks, *i.e.* on **evolutionary graph theory**.

Pairwise games

We shall model interactions among individuals in terms of two-people (or **pairwise**) games in which both players can either cooperate or defect when interacting with each other. After each round of the game the agents accumulate a **payoff** according to the strategy chosen by both players. Mutual cooperation leads to the **reward**, R, whereas mutual defection leads to the **punishment**, P. The other two possibilities occur when one player cooperates and the other defects, for which we have S (**sucker's** payoff) and T (**temptation**) for the cooperator and the defector, respectively. Provided that mutual cooperation is always preferred over mutual defection, three dilemmas arise naturally [130], depending on the relative ordering of these four payoffs: The **Snowdrift game**, for which $T > R > S > P$; the **Stag-Hunt game**, for which $R > T > P > S$; and the **Prisoner's Dilemma game**, for which $T > R > P > S$. For all dilemmas, mutual cooperation is also preferred over unilateral cooperation S. Tension becomes apparent when the preferred choices of each player lead to individual actions resulting in mutual defection, despite the fact that mutual cooperation is more beneficial.

Indeed, tension will arise when players prefer unilateral defection to mutual cooperation ($T > R$), when players prefer mutual defection to unilateral cooperation ($S < P$), or when both situations arise, which is precisely what happens in the Snowdrift game, the Stag-Hunt game and the Prisoner's Dilemma game, respectively. Formally, these dilemmas span a four-dimensional parameter space. Without loss of generality, normalizing mutual cooperation R to 1 and mutual defection P to 0, we are left with two parameters, T and S. The behavior of all dilemmas, summarized in Tab. 2.1, can be explored considering the following ranges for the parameters: $0 \leq T \leq 2$ and $-1 \leq S \leq 1$.

	Abbreviation	**Payoff order**	**Optimal Strategy**
Snowdrift	SG	$T > R > S > P$	Players prefer unilateral defection to mutual cooperation
Stag Hunt	SH	$R > T > P > S$	Players prefer mutual defection to unilateral cooperation
Prisoner's Dilemma	PD	$T > R > P > S$	Both tensions above are incorporated in this dilemma

Table 2.1: Characteristics of the three social dilemmas considered in this thesis and mappable as a 2×2 games. For each dilemma we consider its abbreviation, the payoff order and which strategy is the optimal one. (from [131]).

Groupal interaction: the public goods game

Group interactions are indeed inseparably linked with our increasingly interconnected world, and thus lie at the interface of many different fields of research. N-player interactions are almost as fundamental as pairwise ones. More importantly, group interactions, in general, cannot be reduced to the corresponding sum of pairwise interactions [132, 133]. It has been recently noted [134] that the oversimplifying restriction of pairwise social interactions has dominated the interpretation of many biological data that would probably be much better interpreted in terms of group interactions instead. Also, it has been shown that increasing the group size does not necessarily lead to mean-field behaviour [135], as is traditionally observed instead for games governed by pairwise interactions [136].

Let us now introduce a game suited for group interactions named: **public goods game** (PGG) [137]. Consider a group of N individuals where a strategy, s, being it either cooperate ($s = 1$) or defect ($s = 0$), is assigned to each one of them. Cooperators contribute a fixed amount c (*i.e.* the *cost*) to the common pool while defectors contribute nothing. Finally, the sum of all contributions in the group is multiplied by a **synergy** (or *enhancing*) **factor** r, and the resulting public goods are distributed equally among all the group members. Under such setup, the payoff of an agent i is given by:

$$P_i(C) = r\,\frac{\sum_j^N s_j\, c}{N} - c\,,$$
$$P_i(D) = r\,\frac{\sum_j^N s_j\, c}{N}\,,$$

thus, in general:

$$P_i = r\,\frac{\sum_j^N s_j\, c}{N} - c\, s_i = r\,\frac{N_C\, c}{N} - c\, s_i = r x_C\, c - c\, s_i\,, \tag{2.41}$$

where N_C and x_C are the total number and the fraction of cooperators in the system respectively. If the system is divided into groups of players, each player plays a number of games equal to the number of group she is involved into. For such reason, it results that the total payoff is equal to the sum of the payoffs obtained in each group. As shown by Hardin, the most convenient strategy in the PGG is defection leading to the so-called **tragedy of the commons** effect [138].

Evolutionary games and complex networks

Pairwise games Evolutionary game theory has been extended to account for the possibility that every agent does not interact with every other but rather through an interaction pattern described by means of a complex network giving rise to a discipline known as **evolutionary graph theory**. Both pairwise [139, 140], and groupal games [141] games have been studied on top of complex networks, and the general conclusion is that *complex networks enhance the emergence and survival of cooperation.* We comment briefly on such results (pairwise first and groupal after), the interested reader could look at some reviews (and references therein) [18, 142, 143].

Looking at Fig. 2.13, and taking the results for well mixed populations (*i.e.* leftmost panel) as a reference, it is clear that network reciprocity enhance cooperation in all dilemmas. Such promotion is stronger in BA and configurational scale-free networks. What are the reasons behind such enhancement?

Gómez-Gardeñes *et al.* [144] have found that, in order to survive invasion from defectors, cooperators tend to coalesce and form clusters. If we observe the strategies adopted by the agents along the time, we can basically distinguish two kind of behaviours: agents that never change their strategy (either cooperating or defecting), and agents whose strategy switch one round after

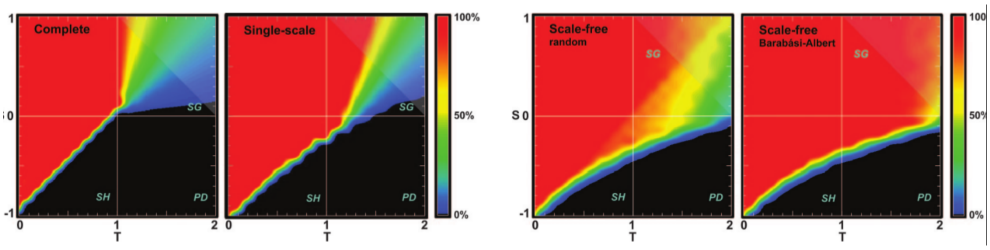


Figure 2.13: Effects of different underlying topology on the evolution of cooperation. Each graph corresponds to a distinct topology namely (from left to right): mean-field, Erdős-Rényi, Bárabasi-Albert, and configurational model. Results display the fraction of cooperators in the population plotted as a contour drawn as a function of two parameters: sucker S, and temptation T (from [131]).

another. The first kind of agent is always part of a bigger group of similar individuals forming a **core**, *i.e.* a connected component (subgraph) fully and permanently occupied by agents with the same strategy. Therefore, in our case, there are two types of core: *cooperator core* (CC), and *defector core* (DC). It is easy to see that a CC cannot be in direct contact with a DC but, instead, with a cloud of fluctuating elements that constitutes the frontier between these two cores.

It is worth to note that a CC is stable if none of its elements has a defector neighbor coupled to more than k_c/T cooperators where k_c is the number of cooperators linked to the element and T is the temptation. Thus, the stability of a CC is clearly enhanced by a high number of connections among pure cooperators, and explains why cooperative behavior is more successful in scale-free networks than in homogeneous graphs. Moreover, it has been verified that the cooperator core in SF networks contains the hubs. Due to the structure of SF networks, those hubs usually lay at the border of the core and thus act as a "screen" for the rest of the cooperators within the core that remain "unaware" of what is happening in the outer part of it. In particular, highly connected individuals survive as pure cooperators until the fraction of cooperators vanishes, thus keeping around them a highly robust cooperator core that loses more and more elements of its outer layer until cooperation is finally defeated by defection.

Groupal interactions Concerning groupal interactions, Santos *et al.* [141] have reformulated public goods games to be staged on complex networks. Every player i plays $k_i + 1$ public goods games, where k_i is the degree of player i, because she has to play in the group composed by her neighbors plus the k_i groups where she is part of her neighbors neighbourhood. Also, because the groups will also have different size, cooperators can contribute either a **fixed amount per game**, $c_i = z$, or a **fixed amount per member** of the group, $c_i = z/(k_i+1)$ (see Fig. 2.14). Identically to the traditional set-up, the contributions within different groups are multiplied by r and accumulated. However, the pay-off of an otherwise identical player is not the same for the two different options. In the extension of PGG to complex networks, the payoff of a generic

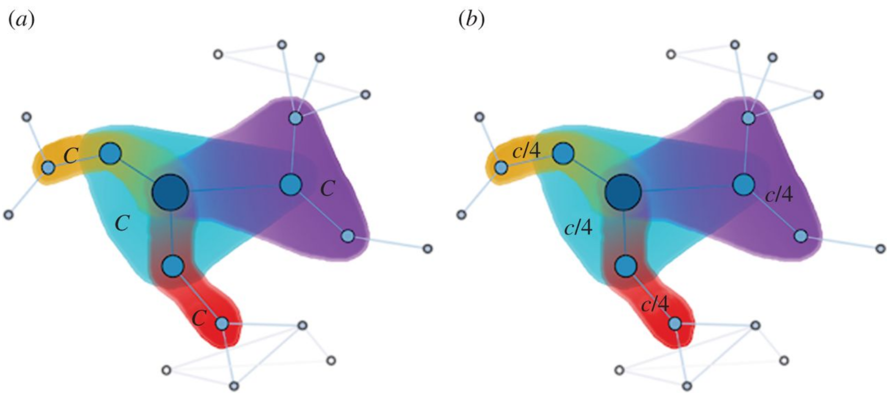


Figure 2.14: When the public goods game is staged on a complex network, cooperators can either bear a fixed cost per game (a), or a fixed amount per member (b) (from [141]).

agent i is equal to:

$$P_i = \sum_{j=1}^{N} a_{ij} \frac{r\left(\sum_{l=1}^{N} a_{jl}\, s_l\, c_l + s_j\, c_j\right)}{k_j+1} + \frac{r\left(\sum_{j=1}^{N} a_{ij}\, s_j\, c_j + s_i\, c_i\right)}{k_i+1} - (k_i+1)\; s_i\, c_i \,. \tag{2.42}$$

We obtain the following net benefit P_i for both versions of the game. Simply, once selected the version of the game, one has to change the values of contributions (costs) c_i according the choice made. Results presented in Santos *et al.* [141] show that heterogeneous networks promote the evolution of cooperation as shown in Fig. 2.15.

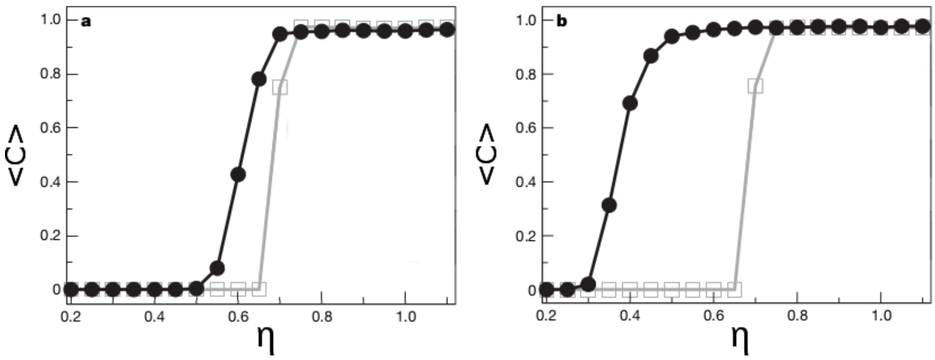


Figure 2.15: Evolution of cooperation in networked PGGs. Filled circles refers to scale-free graphs; open squares show results for regular graphs. (panel *a*) Fixed amount per game. (panel *b*) Fixed amount per member (from [141]).

Yet, this is particularly true when cooperators pay a fixed amount per member. In particular, in contrast with the mean field case, it has been shown that cooperation is viable already at $\eta \simeq r/(\langle k \rangle + 1) < 1$ (**normalized multiplication factor**) that accounts for the average group size. Finally, we can say that: phenomenologically, the promotion of cooperation is due to the diversity of investments, which is a direct consequence of the heterogeneity of the underlying network, which gives an evolutionary advantage to cooperative hubs, *i.e.* players with a high degree.

Chapter 3

Velocity-enhanced cooperation of moving agents playing public goods games

PHYSICAL REVIEW E **85**, 067101 (2012)

Velocity-enhanced cooperation of moving agents playing public goods games

Alessio Cardillo,[1,2] Sandro Meloni,[1] Jesús Gómez-Gardeñes,[1,2] and Yamir Moreno[1,3]
[1]*Instituto de Biocomputación y Física de Sistemas Complejos, Universidad de Zaragoza, E-50018 Zaragoza, Spain*
[2]*Departamento de Física de la Materia Condensada, Universidad de Zaragoza, E-50009 Zaragoza, Spain*
[3]*Departamento de Física Teórica, Universidad de Zaragoza, E-50009 Zaragoza, Spain*
(Received 8 March 2012; revised manuscript received 7 May 2012; published 18 June 2012)

In this paper we study the evolutionary dynamics of the public goods game in a population of mobile agents embedded in a two-dimensional space. In this framework, the backbone of interactions between agents changes in time, allowing us to study the impact that mobility has on the emergence of cooperation in structured populations. Our results point out that a low degree of mobility enhances cooperation in the system. In addition, we study the impact of the size of the groups in which games are played on cooperation. Again we find a rise and fall of cooperation related to the percolation point of the instant interaction networks created by the set of mobile agents.

DOI: 10.1103/PhysRevE.85.067101 PACS number(s): 89.75.Fb, 87.23.Ge

Despite its ubiquity in nature and human societies, the survival of cooperative behavior among unrelated agents (from bacteria to humans) when defection is the most advantageous strategy is not fully understood and constitutes one of the most fascinating theoretical challenges of evolutionary theory [1,2]. Recently, it has been pointed out that the integration of the microscopic patterns of interactions among the agents composing a large population into the evolutionary setting provides a way out for cooperation to survive in paradigmatic scenarios such as the prisoner's dilemma (PD) game [3–5]. The structural features studied span from simple regular lattices [6–8] to real patterns displayed by social networks [9], such as the small-world effect [10], scale-free (SF) patterns for the number of contacts per individual [11–15], the presence of clustering [16,17], or modularity [18].

Although the above studies mostly focus on the PD game, other paradigmatic settings have also been studied on top of network substrates, such as the public goods game (PGG). The public goods game is seen as the natural extension of a PD game when passing from pairwise to n-person games. After the work by Santos *et al.* [19] showing that SF architectures promote cooperation, many other works have continued this line of research by exploring the networked version of the PGG [20–26]. Moreover, as the PGG formulation introduces two structural scales, namely, individuals and the groups within which they interact, it has been shown that the structure of the mesoscale defined by the groups also plays an important role in the success of cooperation [27–30].

The assumption of a static graph that maps social ties, although still a coarse grained picture of the microscopic interactions, provides a useful approach for studying the dynamics of large social systems. However, when moving to smaller scales one has to consider additional microscopic ingredients that may influence the collective outcome of social dynamics. One of these ingredients is the mobility of individuals, a topic that has recently attracted a lot of attention, and that has been tackled from different perspectives. The range of studies in which mobile agents have been included spans from pure empirical studies [31–33] to theoretical ones that focus on the role that mobility patterns play in different dynamical processes such as disease spreading [34], synchronization [35,36] and evolutionary dynamics [37–42] in the context of the PD game. In addition, more complex representations in which an entanglement between agents mobility and evolutionary dynamics is introduced have been studied within the framework of the PD game [43–46] and, more recently, in the context of the PGG [47].

In this Brief Report we follow the setting introduced in Ref. [37] in which a population of N agents moves on a two-dimensional space. Simultaneously to the movement of the agents we consider that a PGG is played. To this end, the movement dynamics is frozen at equally spaced time steps, and each node engages its closest neighbors to participate in a group in which a PGG is played. Obviously, the mobility of individuals turns the usual static backbone of interactions into a time-evolving one, opening the door to novel effects on the evolution of cooperation. Our results point out a nontrivial dependence on the velocity of the agents and the group size in which PGG's are played, yielding optimum operation points at which cooperation is favored.

We start by introducing the dynamical setting in which the evolutionary dynamics of the PGG is implemented. Our population is composed of a set of N agents living in the area inside a square with side length L. Thus, the density of individuals is defined as $\rho = N/L^2$. Both the density and the number of agents remain constant along our simulations. Our agents are initially scattered at random on top of the surface by using two independent random variables uniformly distributed in $[0,L]$ for assigning the initial position $[x_i(0), y_i(0)]$ of each agent.

Once the initial configuration of the system is set, two dynamical processes co-evolve: movement and evolutionary dynamics. At each time step t, the movement of agents affects their current positions, $[x_i(t), y_i(t)]$ with $i = 1, \ldots, N$, by means of the following equations:

$$x_i(t+1) = x_i(t) + v \cos\theta_i(t), \quad (1)$$

$$y_i(t+1) = y_i(t) + v \sin\theta_i(t). \quad (2)$$

The value of each angular variable θ_i is randomly assigned for each agent at each time step from a uniform distribution in the interval $[-\pi,\pi]$. In addition to the above equations, we use periodic boundary conditions; if one agent reaches one side of the square, it reappears on the opposite one.

The second ingredient of the dynamical model is the evolutionary PGG played by the mobile agents. In addition to the random assignment of its initial position, each agent is assigned its initial strategy randomly, so that with equal probability an agent is set as cooperator [$s_i(0) = 1$] or defector [$s_i(0) = 0$]. After this initial stage, both movement and evolutionary dynamics evolve simultaneously. At each time step, just after each agent has updated its position in the plane as dictated by Eqs. (1) and (2), agents play a round of the PGG as follows: First a network of contacts is constructed as a random geometric graph (RGG) [9]. Each pair of agents (i,j) creates a link between them provided they are separated less than a certain threshold distance R: $\sqrt{[x_i(t) - x_j(t)]^2 + [y_i(t) - y_j(t)]^2} \leqslant R$. After all the nodes have established their connections with their nearest neighbors, a RGG for the network of contacts at time t emerges, whose topology is encoded in an adjacency matrix A^t_{ij} with entries $A^t_{ij} = 1$ when nodes i and j are connected at time t and $A^t_{ij} = 0$ otherwise.

Once the RGG is formed, each of the agents defines, together with her k_i nearest neighbors in the RGG, a group of size $k_i + 1$ in which one PGG is played. In each of the groups she participates in, a cooperator player contributes an amount c while a defector does not contribute. Besides, the total contribution of a group is multiplied by an enhancing factor r and distributed equally among all the participants. Thus the total payoff accumulated by an agent i at time t reads

$$P_i(t) = \sum_{j=1}^{N} \left(A^t_{ij} + \delta_{ij}\right) \frac{\sum_{l=1}^{N} \left(A^t_{jl} + \delta_{jl}\right) s_l(t) c r}{k_j(t) + 1} - [k_i(t) + 1] s_i(t) c. \tag{3}$$

After each round each of the agents can update her strategy. To this aim, an agent i chooses one of her instant neighbors j at random and with probability $\Pi[s_i(t+1) = s_j(t)]$, i will take the strategy of j during the next round of the PGG. The former probability reads:

$$\Pi[s_i(t+1) = s_j(t)] = \frac{\Theta[P_j(t) - P_i(t)]}{M[k_i(t), k_j(t)]}, \tag{4}$$

where $\Theta(x) = x$ when $x > 0$ while $\Theta(x) = 0$ otherwise, and $M(k_i, k_j)$ is the maximum possible payoff difference between two players with instant degrees $k_i(t)$ and $k_j(t)$. In our simulations, we let both movement and evolutionary dynamics co-evolve during 5×10^4 time steps. We take the first 25×10^3 steps as a transient period while the degree of cooperation of the system is measured during the second half of the simulations as $\langle c \rangle = \sum_{t=\tau}^{\tau+T} \sum_{i=1}^{N} s_i(t)/T$, with both $\tau = T = 25 \times 10^3$. The results reported below are averaged over different realizations (typically 50).

We start our analysis by considering the static case in which the velocity of the agents is set to $v = 0$. In this case, the RGG is fixed from the initial configuration while only the strategies of agents evolve. A RGG is described by a Poissonian distribution, $P(k) = \langle k \rangle^k e^{-\langle k \rangle}/k!$, for the probability of finding a node connected to k neighbors. This distribution corresponds to a homogeneous architecture in which the dispersion around the mean degree $\langle k \rangle$ is rather small. The same pattern for the degree distribution $P(k)$ is

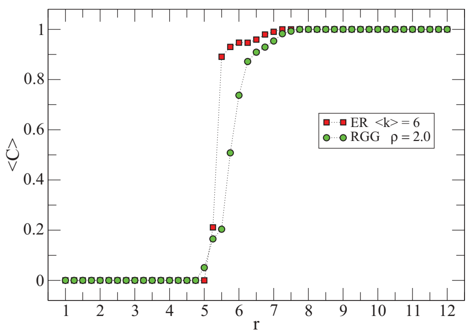


FIG. 1. (Color online) Average fraction of cooperators $\langle c \rangle$ with respect to the enhancement factor r for *static* RGG and ER networks. Both networks have the same number of nodes $N = 1000$ and average degree $\langle k \rangle = 6$.

obtained for the typical Erdős-Rényi (ER) random network model. However, the main differences between RGG and ER networks rely on the clustering coefficient, i.e., the probability that two nodes with a common neighbor share a connection, and their diameter. While in the case of ER graphs clustering vanishes as $N \to \infty$, the geometric nature of RGG boosts the density of triads leading to a finite and large clustering coefficient at the expense of being longer than ER graphs (since redundant links used to increase clustering do not contribute to create shortcuts). These differences are found to be relevant for the synchronization of RGG compared to ER graphs [48].

The results of the above analysis are shown in Fig. 1, where we represent the dependency of the average level of cooperation in the system $\langle c \rangle$ with respect to the enhancement factor r for both RGG's and ER graphs having the same number of elements N and the same average degrees $\langle k \rangle$. As expected, for low values of r defection dominates the system while for large r cooperation prevails. Between these two asymptotic regimes the transition from defection to cooperation occurs ($5 \leqslant r \leqslant 8$), pointing out slight differences between RGG's and ER graphs. In this region we observe that ER networks promote cooperation slightly more than RGG's, for which the transition curve toward full cooperation goes slower. This result seems to contradict previous observations in the context of the PD game [17] in SF networks. However, this latter result is related to the increase of heterogeneity when clustering is enlarged in SF networks. In our case, this effect is not present and, alternatively, clustering induces important differences between ER graphs and RGG's regarding the average path length. This quantity is shown to be much larger in RGG's than in ER graphs, thus making more difficult the percolation of cooperation in the whole system. On the other hand, the onset of both transitions are roughly the same.

We now focus on the impact that the motion of agents has on the level of cooperation with respect to the static case. Thus, from now on, we consider that agents move with constant velocity v following the rules given by Eqs. (1) and (2). Moreover, we set the value of the enhancement factor r to be in the region for which the transition from full defection to cooperation occurs in the static case, namely, $r = 5.75$. Then

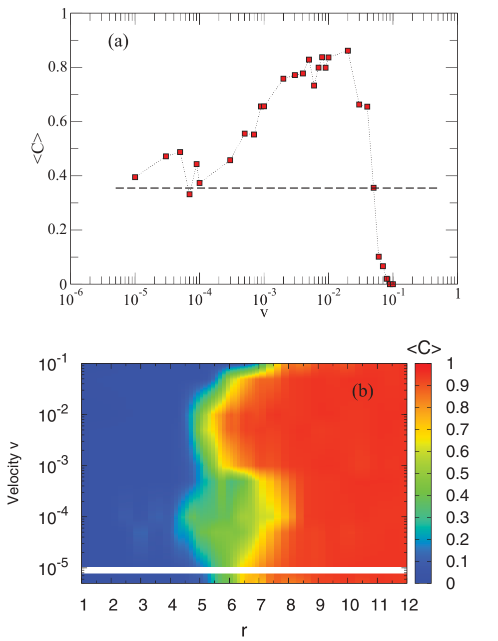


FIG. 2. (Color online) Effects of velocity on the promotion of cooperation. (a) Average level of cooperation $\langle c\rangle$ as a function of the velocity v of agents. The system has $\rho = 2.0$ and $R = 1.0$ and the enhancement factor r is set to $r = 5.75$. The velocity spans in the interval $[10^{-5}; 10^{-1}]$. The dashed line represents $\langle c\rangle$ in a static RGG with the same N and $\langle k\rangle$, and for the same value of r. (b) Cooperation level $\langle c\rangle$ as a function of v and r. The static case is displayed in the bottom strip of the panel (below the continuous thick band).

we monitor the degree of cooperation $\langle c\rangle$ in the system as a function of the velocity The results are shown in Fig. 2(a) together with the value (dashed line) for $\langle c\rangle$ in the static limit with $r = 5.75$. We observe a rise and fall of cooperation so that when the velocity v increases from very small values, the average level of cooperation increases significantly, reaching its maximum value for $v \simeq 2 \times 10^{-2}$. From this point on, the increase of v leads to the decay of cooperation so that $\langle c\rangle = 0$ beyond $v \simeq 10^{-1}$. The fall of cooperation for large values of the velocity of agents is a quite expected result: As the velocity increases, one approaches the well-mixed scenario for which cooperation is suppressed provided r is less than the typical size of groups in which the PGG is played (here $\langle k\rangle = 6$ so that groups are typically composed by 7 agents). The rise of cooperation for small values of v points out that there exists an optimal range for the velocity that allows a tradeoff between two important ingredients for cooperator clusters to form and resist the invasion of defectors, namely, the ability to explore the plane to find other cooperators and a large enough time to interact with them so as to allow for the growth of cohesive cooperator clusters.

A more extensive analysis on the effects of motion is found in Fig. 2(b), where a detailed exploration of the (v,r) parameter space is shown together with the cooperation level in the static case (bottom part of the panel) as obtained from the corresponding curve in Fig. 1. This panel confirms the results obtained in Fig. 2(a) and provides a more complete picture about the enhancement of cooperation produced by the mobility of agents. First, by comparing the bottom ($v = 0$) and top ($v = 10^{-1}$) parts of the panel, we observe that a large value of the velocity decreases the cooperation level of the static system. In particular, let us note that the transition region in the limit of large velocity is placed around $r \simeq 7$, thus recovering the well-mixed prediction. However, the relevant results are found between the static and large velocity limits. The effects of mobility in this region affect both the onset of the transition toward cooperation and its fixation. First, we observe that even for very low values of v the onset of cooperation is anticipated with respect to the static case at the expense of having a broader transition toward full cooperation as compared to the static RGG. However, when the velocity level is further increased, the transition becomes sharper and both the onset and the fixation of the full-cooperative state occur before with respect to the static case.

Finally, we focus on the influence of group size on the evolutionary success of cooperation. Considering a fixed velocity lying in the region for which the increase of cooperation is observed, namely, $v = 10^{-2}$, we compute the level of cooperation $\langle c\rangle$ as a function of the rescaled enhancement factor $\eta = r/(\langle k\rangle + 1)$, where the denominator is the average size of the groups. This rescaling is needed for the sake of comparing the cooperation levels for systems in which the group size is different. In this way, the well-mixed prediction is a sharp transition from full defection to full cooperation at $\eta = 1$. As anticipated above, the size of the groups can be written as $\rho\pi R^2 + 1$, thus we can vary either the radius of interaction R or the density of agents ρ.

In what follows we vary the radius and keep the density constant to $\rho = 2$. In Fig. 3 we observe that again, a rise and fall of cooperation is observed when going from low radii to large ones. Obviously, as the radius (and hence group size) increases, we approach the well-mixed case so that the transition point reaches the theoretical value $\eta = 1$. However,

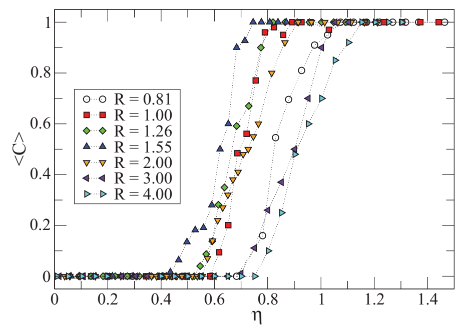


FIG. 3. (Color online) Cooperation level $\langle c\rangle$ as a function of $\eta = r/(\langle k\rangle + 1)$ varying the interaction radius of the agents (i.e., the size of the groups). In all the cases the density of players is $\rho = 2.0$ and their velocity is $v = 10^{-2}$.

this approach is not monotonous, and for intermediate values of the radius and group size the cooperation transition is anticipated with respect to lower values of R. The reason behind this behavior relies on the percolation of the effective network when increasing the radius of influence of each agent. For low values of R the effective network of contacts contains a number of disconnected clusters; however, reaching the percolation threshold (meaning that the $\langle k\rangle = \rho\pi R^2 > 2$) nearly all the agents are incorporated into a macroscopic giant component. In our case $\rho = 2$ so that the percolation point lies around $R \simeq \sqrt{\pi} \simeq 1.773$, which agrees with the numerical observation in Fig. 3. This connection with the percolation point and the cooperation level has been recently observed in Ref. [49] in the context of the PD game in regular lattices.

Summing up, the results presented in this Brief Report show that the mobility of the agents playing a PGG enhances cooperation provided their velocity is moderate. This enhancement is obtained by comparing the outcome of the evolutionary dynamics of the PGG with the results obtained in the static case. The addition of the random movement of agents produces the evolution in time of the original RGG, being the rate of creation and deletion of links controlled by the velocity of agents. When this rate is nonzero, allowing cooperators to explore the space, while moderate, so that cooperators clusters can be efficiently formed, we observe an optimal operation regime in which both the onset of cooperation and the fixation of cooperation in the system are enhanced. Finally, we have checked that group size shows a similar resonance phenomenon regarding the level of cooperation. However, in this case the point at which cooperation is enhanced is related to the percolation point of the effective network, i.e., with the point at which the size of the groups is large enough so as to have a macroscopic giant component for the network of contacts.

We acknowledge support from the Spanish DGICYT under projects FIS2008-01240, MTM2009-13848, FIS2009-13364-C02-01, and FIS2011-25167, and by the Comunidad de Aragón (FENOL). J. G. G. is supported through the Ramón y Cajal program.

[1] R. Axelrod and W. D. Hamilton, Science **211**, 1390 (1981).
[2] J. Maynard-Smith and E. Szathmáry, *The Major Transitions in Evolution* (Oxford University Press, Oxford, 1995).
[3] M. A. Nowak and R. M. May, Nature (London) **359**, 826 (1992).
[4] G. Szabó and G. Fáth, Phys. Rep. **447**, 97 (2007).
[5] C. P. Roca, J. Cuesta, and A. Sánchez, Phys. Life Rev. **6**, 208 (2009).
[6] M. A. Nowak, S. Bonhoeffer, and R. M. May, Int. J. Bif. Chaos **4**, 33 (1994).
[7] M. H. Vainstein and J. J. Arenzon, Phys. Rev. E **64**, 051905 (2001).
[8] S. Számadó, F. Szalai, and I. Scheuring, J. Theor. Biol. **253**, 221 (2008).
[9] S. Boccaletti *et al.*, Phys. Rep. **424**, 175 (2006).
[10] G. Abramson and M. Kuperman, Phys. Rev. E **63**, 030901(R) (2001).
[11] F. C. Santos and J. M. Pacheco, Phys. Rev. Lett. **95**, 098104 (2005).
[12] F. C. Santos, J. M. Pacheco, and T. Lenaerts, Proc. Natl. Acad. Sci. USA **103**, 3490 (2006).
[13] J. Gómez-Gardeñes, M. Campillo, L. M. Floría, and Y. Moreno, Phys. Rev. Lett. **98**, 108103 (2007).
[14] J. Poncela, J. Gómez-Gardeñes, L. M. Floría, and Y. Moreno, J. Theor. Biol. **253**, 296 (2008).
[15] C.-L. Tang, W.-X. Wang, X. Wu, and B.-H. Wang, Eur. Phys. J. B **53**, 411 (2006).
[16] A. Pusch, S. Weber, and M. Porto, Phys. Rev. E **77**, 036120 (2008).
[17] S. Assenza, J. Gómez-Gardeñes, and V. Latora, Phys. Rev. E **78**, 017101 (2008).
[18] S. Lozano, A. Arenas, and A. Sánchez, PLoS ONE **3**, e1892 (2008).
[19] F. C. Santos, M. D. Santos, and J. M. Pacheco, Nature **454**, 213 (2008).
[20] A. Szolnoki, M. Perc, and G. Szabó, Phys. Rev. E **80**, 056109 (2009).
[21] Z.-G. Huang, Z.-X. Wu, A.-C. Wu, L. Yang, and Y.-H. Wang, Europhys. Lett. **84**, 50008 (2008).
[22] Z. Rong and Z. X. Wu, Europhys. Lett. **87**, 30001 (2009).
[23] A. Szolnoki and M. Perc, Europhys. Lett. **92**, 38003 (2010).
[24] D. M. Shi and B. H. Wang, Europhys. Lett. **90**, 58003 (2010).
[25] Z. Rong, H.-X. Yang, and W.-X. Wang, Phys. Rev. E **82**, 047101 (2010).
[26] M. Perc, Phys. Rev. E **84**, 037102 (2011).
[27] J. Gómez-Gardeñes, M. Romance, R. Criado, D. Vilone, and A. Sánchez, Chaos **21**, 016113 (2011).
[28] J. Gómez-Gardeñes, D. Vilone, and A. Sánchez, Europhys. Lett. **95**, 68003 (2011).
[29] M. Perc, New J. Phys. **13**, 123027 (2011).
[30] A. Szolnoki and M. Perc, Phys. Rev. E **84**, 047102 (2011).
[31] C. Roth, S. M. Kang, M. Batty, and M. Barthélemy, PLoS One **6**, e15923 (2011).
[32] J. Stehle, N. Voirin, A. Barrat, C. Cattuto, and L. Isella, PLoS One **6**, e23176 (2011).
[33] A. Panisson *et al.*, Ad Hoc Networks (in press), doi: 10.1016/j.adhoc.2011.06.003.
[34] M. Frasca, A. Buscarino, A. Rizzo, L. Fortuna, and S. Boccaletti, Phys. Rev. E **74**, 036110 (2006).
[35] M. Frasca, A. Buscarino, A. Rizzo, L. Fortuna, and S. Boccaletti, Phys. Rev. Lett. **100**, 044102 (2008).
[36] N. Fujiwara, J. Kurths, and A. Díaz-Guilera, Phys. Rev. E **83**, 025101(R) (2011).
[37] S. Meloni, A. Buscarino, L. Fortuna, M. Frasca, J. Gómez-Gardeñes, V. Latora, and Y. Moreno, Phys. Rev. E **79**, 067101 (2009).
[38] J. Zhang, W.-Y. Wang, W.-B. Du, and X.-B. Cao, Physica A **390**, 2251 (2011).
[39] C. A. Aktipis, J. Theor. Biol. **231**, 249 (2004).
[40] M. H. Vainstein, A. T. C. Silva, and J. J. Arenzon, J. Theor. Biol. **244**, 722 (2007).

[41] C. P. Roca and D. Helbing, Proc. Nat. Acad. Sci. USA **108**, 11370 (2011).
[42] H. Cheng, H. Li, Q. Dai, Y. Zhu, and J. Yang, New J. Phys. **12**, 123014 (2010).
[43] D. Helbing and W. Yu, Adv. Compl. Syst. **8**, 87 (2008).
[44] D. Helbing and W. Yu, Proc. Nat. Acad. Sci. USA **106**, 3680 (2009).
[45] D. Helbing, Eur. Phys. J. B. **67**, 345 (2009).
[46] H. Y. Cheng, Q. Dai, H. Li, Y. Zhu, M. Zhang, and J. Yang, New J. Phys. **13**, 043032 (2011).
[47] C. A. Aktipis, Evol. Human Behav. **32**, 263 (2011).
[48] A. Díaz-Guilera, J. Gómez-Gardeñes, Y. Moreno, and M. Nekovee, Int. J. Bif. Chaos **19**, 687 (2009).
[49] Z. Wang, A. Szolnoki, and M. Perc, Sci. Rep. **2**, 369 (2012).

Chapter 4

Evolutionary dynamics of time-resolved social interactions

Evolutionary dynamics of time-resolved social interactions

Alessio Cardillo,[1,2] Giovanni Petri,[3] Vincenzo Nicosia,[4] Roberta Sinatra,[5] Jesús Gómez-Gardeñes,[1,2] and Vito Latora[4,6]

[1] *Departamento de Física de la Materia Condensada, Universidad de Zaragoza, E-50009 Zaragoza, Spain*
[2] *Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, E-50018 Zaragoza, Spain*
[3] *Institute for Scientific Interchange (ISI), via Alassio 11/c, 10126 Torino, Italy*
[4] *School of Mathematical Sciences, Queen Mary University of London, London, UK*
[5] *Center for Complex Network Research and Department of Physics, Northeastern University, Boston, MA 02115, USA*
[6] *Dipartimento di Fisica e Astronomia, Università di Catania, and INFN, Via S. Sofia 64, I-95123 Catania, Italy*

Cooperation among unrelated individuals is frequently observed in social groups when their members join efforts and resources to obtain a shared benefit which is unachievable by single ones. However, understanding why cooperation arises despite the natural tendency of individuals towards selfish behavior is still an open problem and represents one of the most fascinating challenges in evolutionary dynamics. Very recently, the structural characterization of the networks upon which social interactions take place has shed some light on the mechanisms by which cooperative behavior emerges and eventually overcome the individual temptation to defect. In particular, it has been found that the heterogeneity in the number of social ties and the presence of tightly-knit communities lead to a significant increase of cooperation as compared with the unstructured and homogeneous connection patterns considered in classical evolutionary dynamics. Here we investigate the role of social ties dynamics for the emergence of cooperation in a family of social dilemmas. Social interactions are in fact intrinsically dynamic, fluctuating and intermittent over time, and can be represented by time-varying networks, that is graphs where connections between nodes appear, disappear, or are rewired over time. By considering two experimental data sets of human interactions with detailed time information, we show that the temporal dynamics of social ties has a dramatic impact on the evolution of cooperation: the dynamics of pairwise interactions favor selfish behavior.

Popular Abstract.- **Why do animals (including humans) cooperate even when selfish actions provide larger benefits? This question has challenged the evolutionary theory during decades. The success of cooperation is essential to humankind and ubiquitous, no matter the cultural and religious traits of particular populations. Scientists have pointed out in the past a series of possible mechanisms that could favor cooperation between humans, as for example the peculiar way of establishing social relations which assumes the form of a complex network. Basically, the structural attributes of these networks, such as the presence of a few individuals that have a large number of social ties, promote cooperation and discourage the imitation of free-riders. However, social interactions are inherently dynamic and changing over time, an issue which has been usually disregarded in the study of cooperation on networks. We study evolutionary models in time-varying graphs and show that the volatility of social relations tends to decrease cooperation in respect to static graphs. Our results thus point out that the time-varying nature of social ties cannot be neglected and that the relative speed of graph evolution and strategy update is a crucial ingredient governing the evolutionary dynamics of social networks, having as much influence as the structural organization.**

I. INTRODUCTION

The organizational principles driving the evolution and development of natural and social large-scale systems, including populations of bacteria, ant colonies, herds of predators and human societies, rely on the cooperation of a large population of unrelated agents [1–3]. Even if cooperation seems to be a ubiquitous property of social systems, its spontaneous emergence is still a puzzle for scientists since cooperative behaviors are constantly threatened by the natural tendency of individuals towards self-preservation and the never-ceasing competition among agents for resources and success. The preference of selfishness over cooperation is also due to the higher short-term benefits that a single (defector) agent obtains by taking advantage of the efforts of cooperating agents. Obviously, the imitation of such a selfish (but rational) conduct drives the system towards a state in which the higher benefits associated to cooperation are no longer achievable, with dramatic consequences for the whole population. Consequently, the relevant question to address is why cooperative behavior is so common, and which are the circumstances and the mechanisms that allow it to emerge and persist.

In the last decades, the study of the elementary mechanisms fostering the emergence of cooperation in populations subjected to evolutionary dynamics has attracted a lot of interest in ecology, biology and social sciences [4, 5]. The problem has been tackled through the formulation of simple games that neglect the microscopic differences among distinct social and natural systems, thus providing a general framework for the analysis of evolutionary dynamics [6–8]. Most of the classical models studied within this framework made the simplifying assumption that social systems are characterized by homogeneous structures, in which the interaction probability is the same for any pair of agents and constant over time [9]. However, this assumption has been proven false for real systems, as the theory of complex networks has revealed that most natural and social networks exhibit large heterogeneity and non–trivial interconnection topologies [10–13]. It has been also shown that the structure of a network has dramatic effects on the dynamical processes taking place on it, so that complex networks analysis has become a fundamental tool in epidemiology, computer science, neuroscience and social sciences [14–16].

The study of evolutionary games on complex topologies has allowed a new way out for cooperation to survive in some paradigmatic cases such as the Prisoner's Dilemma [17–20] or the Public Goods games [21–23]. In particular, it has been pointed out that the complex patterns of interactions among the agents found in real social networks, such as scale-free distributions of the number of contacts per individual or the presence of tightly-knit social groups, tend to favor the emergence and persistence of cooperation. This line of research, which brings together the tools and methods from the statistical mechanics of complex networks and the classical models of evolutionary game dynamics, has effectively became a new discipline, known as Evolutionary Graph Theory [24–28].

Recently, the availability of longitudinal spatio-temporal information about human interactions and social relationships [29–32] has revealed that social systems are not static objects at all: contacts among individuals are usually volatile and fluctuate over time [33, 34], face-to-face interactions are bursty and intermittent [35, 36], agents motion exhibits long spatio-temporal correlations [37–39]. Consequently, static networks, constructed by aggregating in a single graph all the interactions observed among a group of individuals across a given period, can be only considered as simplified models of real networked systems. For this reason, time-varying graphs have been lately introduced as a more realistic framework to encode time-dependent relationships [40–44]. In particular, a time-varying graph is an ordered sequence of graphs defined over a fixed number of nodes, where each graph in the sequence aggregates all the edges observed among the nodes within a certain temporal interval. The introduction of time as a new dimension of the graph gives rise to a richer structure. Therefore, new metrics specifically designed to characterize the temporal properties of graph sequences have been proposed, and most of the classical metrics defined for static graphs have been extended to the time-varying case [44–49]. Lately, the study of dynamical processes taking place on time-evolving graphs has shown that temporal correlations and contact recurrence play a fundamental role in diverse settings such as random walks dynamics [50–52], the spreading of information and diseases [53–55] and synchronization [56].

Here we study how the level of cooperation is affected by taking into account the more realistic picture of social system provided by time-varying graphs instead of the classical (static) network representation of interactions. We consider a family of social dilemmas, including the Hawk-Dove, the Stag Hunt and the Prisoner's Dilemma games, played by agents connected through a time-evolving topology obtained from real traces of human interactions. We analyze the effect of temporal resolution and correlations on the emergence of cooperation in two paradigmatic data sets of human proximity, namely the MIT Reality Mining [29] and the INFOCOM'06 [30] co-location traces. We find that the level of cooperation achievable on time-varying graphs crucially depends on the interplay between the speed at which the network changes and the typical time-scale at which agents update their strategy. In particular, cooperation is facilitated when agents keep playing the same strategy for longer intervals, while too frequent strategy updates tend to favor defectors. Our results also suggest that the presence of temporal correlations in the creation and maintenance of interactions hinders cooperation, so that synthetic time-varying networks in which link persistence is broken usually exhibit a considerably higher level of cooperation. Finally, we show that both the average size of the giant component and the weighted temporal clustering calculated across different consecutive time-windows are indeed good predictors of the level of cooperation attainable on time-varying graphs.

II. RESULTS

A. Evolutionary Dynamics of Social Dilemmas

We focus on the emergence of cooperation in systems whose individuals face a social dilemma between two possible strategies: *Cooperation* (C) and *Defection* (D). A large class of social dilemmas can be formulated as in [18] via a two-parameter game described by the payoff matrix:

$$\begin{array}{c} \\ C \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} R & S \\ T & P \end{pmatrix} \end{array} = \begin{array}{c} \\ C \\ D \end{array}\!\!\begin{array}{c} \begin{array}{cc} C & D \end{array} \\ \begin{pmatrix} 1 & S \\ T & 0 \end{pmatrix} \end{array}, \qquad (1)$$

where R, S, T and P represent the payoffs corresponding to the various possible encounters between two players. Namely, when the two players choose to cooperate they both receive a payoff $R = 1$ (for *Reward*), while if they both decide to defect they get $P = 0$ (for *Punishment*).

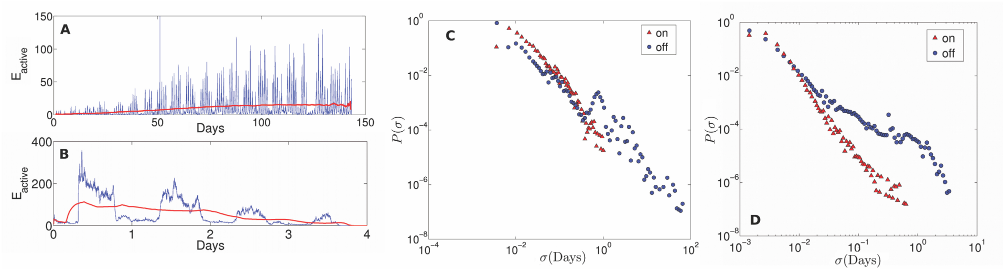


FIG. 1. **Activity patterns of human interactions.** The number E_{active} of links in the graph at time t is reported as a function of time (blue) for MIT Reality Mining **(A)** and INFOCOM'06 **(B)**. Weekly and daily periodicities are visible. Moving averages (red), respectively over a 1-month window and a 1-day window, reveal the non-stationarity of the sequences. The time distributions of edge active and inactive periods (respectively triangles and circles) for MIT Reality Mining **(C)** and INFOCOM'06 **(D)** The data were log-binned. The peak at $\sigma \sim 1$ for the inactive periods corresponds to 24 hours.

When a cooperator faces a defector it gets the payoff S (for *Sucker*) while the defector gets T (for *Temptation*). In this version of the game the payoffs S and T are the only two free parameters of the model, and their respective values induce an ordering of the four payoffs which determines the type of social dilemma. We have in fact three different scenarios. When $T > 1$ and $S > 0$, defecting against a cooperator provides the largest payoff, and this corresponds to the *Hawk-Dove* game. For $T < 1$ and $S < 0$, cooperating with a defector is the worst case, and we have the *Stag Hunt* game. Finally, for $T > 1$ and $S < 0$, when a defector plays with a cooperator we have at the same time the largest (for the defector) and the smallest (for the cooperator) payoffs, and the game corresponds to the *Prisoner's Dilemma*. In this work we consider the three types of games by exploring the parameter region $T \in [0, 2]$ and $S \in [-1, 1]$.

In real social systems, each individual has more than one social contact at the same time. This situation is usually represented [26] by associating each player i, $i = 1, 2, \ldots, N$ to a node of a *static* network, with adjacency matrix $A = \{a_{ij}\}$, whose edges indicate pairs of individuals playing the game. In this framework, a player i selects a strategy, plays a number of games equal to the number of her neighbors $k_i = \sum_j a_{ij}$ and accumulates the payoffs associated to each of these interactions. Obviously, the outcome of playing with a neighbor depends both on the strategiy selected by node i and that of the neighbor, according to the payoff matrix in Eq. 1. When all the individuals have played with all their neighbors in the network, they update their strategies as a result of an evolutionary process, i.e., according to the total collected payoff. Namely, each individual i compares her cumulated payoff p_i with that of one of her neighbors, say j, chosen at random. The probability $P_{i \to j}$ that agent i adopts the strategy of her neighbor j increases with the difference $(p_j - p_i)$ (see Methods for details).

The games defined by the payoff matrix in Eq. 1 and using a payoff-based strategy update rule have been thoroughly investigated in static networks with different topologies. The main result is that, when the network is fixed and agent strategies are allowed to evolve over time, the level of cooperation increases with the heterogeneity of the degree distribution of the network, being scale-free networks the most paradigmatic promoters of cooperation [17–19]. However, in most cases human contacts and social interactions are intrinsically dynamic and varying in time, a feature which has profound consequences on any process taking place over a social network. We explore here the role of time on the emergence of cooperation in time-varying networks.

B. Temporal patterns of social interactions

We consider two data sets describing the temporal patterns of human interactions at two different time scales. The first data set has been collected during the MIT Reality Mining experiment [29], and includes information about spatial proximity of a group of students, staff, and faculty members at the Massachusetts Institute of Technology, over a period of six months. The resulting time-dependent network has $N = 100$ nodes and consists of a time-ordered sequence $\{G_1, G_2, \ldots, G_M\}$ of $M = 41291$ graphs (snapshots), each graph representing proximity interactions during a time interval of $\tau = 5$ minutes. Remember that each graph $G_m, m = 1, \ldots, M$ accounts for all the instantaneous interactions taking place in the temporal interval $[(m-1)\tau, m\tau]$. The second data set describes co-location patterns, over a period of four days, between the participants of the INFOCOM'06 conference [30]. In this case, the resulting time-dependent network has $N = 78$ nodes, and contains a sequence of $M = 2880$ graphs obtained by detecting users co-location every $\tau = 2$ minutes.

The frequency of social contacts is illustrated in Fig. 1 (panels A and B), where we report the number of active links at time t, E_{active}, as a function of time. In the MIT

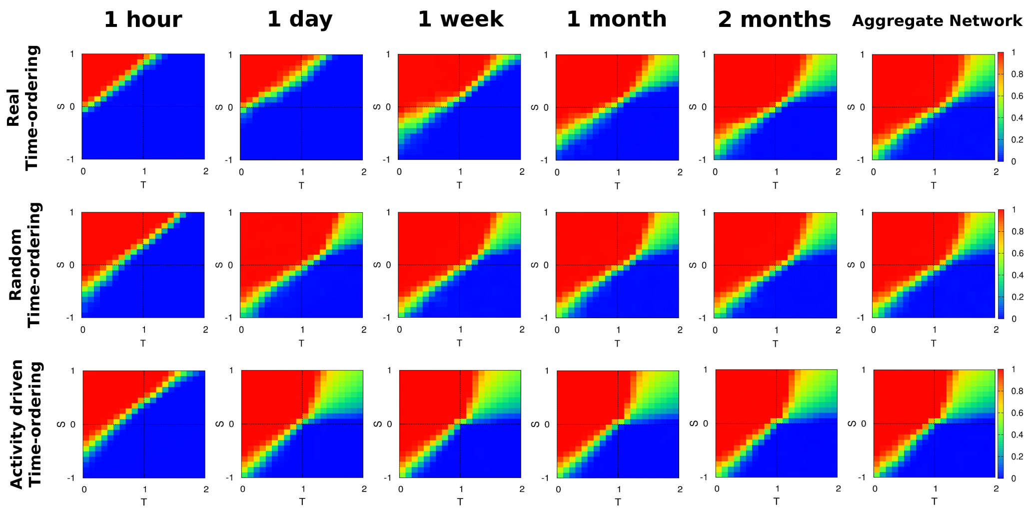


FIG. 2. **Cooperation diagrams for the MIT Reality Mining data set**. Fraction of cooperators at the equilibrium as a function of the temptation to defect (T) and of the sucker's score (S) for different values of the interval Δt between two successive strategy updates. From left to right, the diagrams correspond to Δt equal to 1 hour, 1 day, 1 week, 1 month, 2 months, and to the entire observation period $M\tau \simeq 5$ months. The diagrams in the top row correspond to time-varying graphs with original time ordering, those in the middle row are obtained for the same values of Δt but on randomized time-varying graphs, while the bottom row reports the results obtained on synthetic networks constructed through the activity-driven model. The results are averaged over 50 different realizations. Red corresponds to 100% of cooperators while blue indicates 100% defectors.

Reality Mining data set, social activity exhibits daily and weekly periodicities, respectively due to home-work and working days-weekends cycles. In addition to these rhythms, we notice a non-stationary behavior which is clearly visible when we plot the activity averaged over a 1-month moving window (red line in panel A). In the INFOCOM'06 data set we observe a daily periodicity and a non-stationary trend which is due, in this case, to a decreasing social activity in the last days of the conference as seen by aggregating activity over 24 hours (red line in panel B). We also report in Fig. 1 (panels C and D) the distributions $P(\sigma)$ of contact duration, $\sigma \equiv \sigma_{on}$, and of inter-contact time, $\sigma \equiv \sigma_{off}$ (i.e. the interval between two consecutive appearances of an edge). As it is often the case for human dynamics [35], the distributions of contact duration and inter-contact time are heterogeneous. For the MIT data set, an active edge can persist up to an entire day, while inactive intervals can last over multiple days and weeks; similar patterns are observed in the INFOCOM'06 data set, where some edges remain active up to one entire day and inter-contact times span almost the whole observation interval. Edge activity exhibits significant correlations over long periods of time. In particular, the autocorrelation function of the time series of edge activity shows a slow decay, up to lags of $10-12$ hours for the MIT data set, and of $6-8$ hours for INFOCOM'06, after which the daily periodicity becomes dominant (figure not reported).

C. Evolution of Cooperation in Time-varying Networks

To simulate the game on a time-varying topology $\{G_m\}_{m=1,\ldots,M}$, we start from a random distribution of strategies, so that each individual initially behaves either as a cooperator or as a defector, with equal probability. The simulation proceeds in *rounds*, where each round consists of a *playing* stage followed by a *strategy update*. In the first stage, each agent plays with all her neighbors on the first graph of the sequence, namely on G_1, and accumulates the payoff according to the matrix in Eq. 1. Then the graph changes, and the agents employ the same strategies to play with all their neighbors in the second graph of the sequence, G_2. The new payoffs are summed to those obtained in the previous iteration. The same procedure is then repeated n times with n such that $n \cdot \tau$ is equal to a chosen interval Δt, which is the *strategy update interval*. At this point, the playing stage terminates and agent strategies are updated. Namely, each agent compares the net payoff accumulated during the previous n time steps with that of one of her neighbors chosen at random. In particular, we adopt here the so-called *Fermi Rule* [57, 58] with a parameter $\beta = 1$ (see Methods). We have checked that the results are quite robust, and we obtain qualitatively similar outcomes for a wide range of β. After the agents have updated their strategy, their payoff is reset to 0 and they start another round,

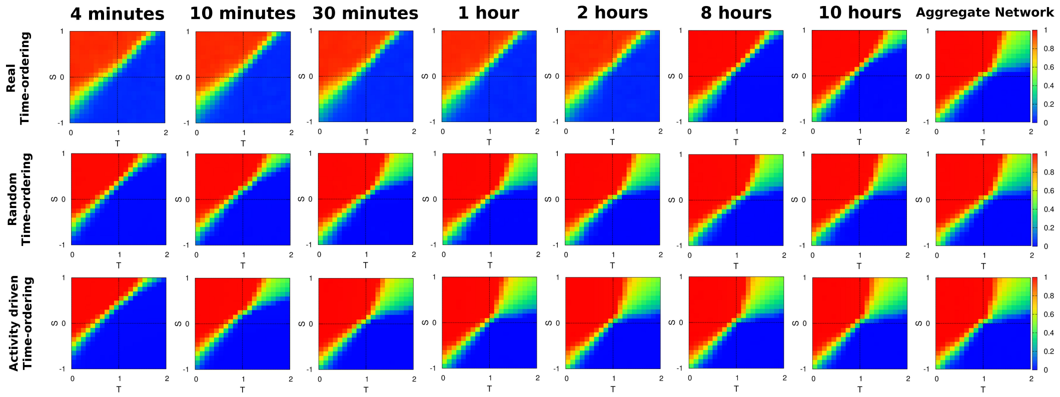


FIG. 3. **Cooperation diagrams for the INFOCOM data set.** Fraction of cooperators at the equilibrium as a function of the temptation to defect (T) and of the sucker's score (S) for different values of the interval Δt between two successive strategy updates. From left to right, the diagrams correspond to Δt equal to 4 minutes, 10 minutes, 30 minutes, 1 hour, 2 hours, 8 hours, 10 hours and $M\tau \simeq 4$ days. The top, middle and bottom row report, respectively, the results for the original data set, the reshuffled time-varying graph and synthetic graphs constructed through the activity-driven model. The results are averaged over 50 different realizations. Red corresponds to 100% of cooperators while blue indicates 100% defectors.

during the subsequent time interval of length $\Delta t = n \cdot \tau$, as described above.

To evaluate the degree of cooperation obtained for a given value of the strategy update interval Δt and a pair of values (T, S), we compute the average fraction of cooperators $\langle C(T,S)_{\Delta t} \rangle$:

$$\langle C(T,S)_{\Delta t} \rangle = \frac{1}{Q} \sum_{i=1}^{Q} \frac{N_c^i}{N} , \qquad (2)$$

where N_c^i is the number of cooperators found at time $i \cdot \Delta t$ and Q is the total number of rounds played. In general, we set Q large enough to guarantee that the system reaches a stationary state.

We have simulated the system using different values of Δt. Notice that for smaller value of Δt the time-scale of strategy update is comparable with that of the graph evolution, while when Δt is equal to the entire observation period $M\tau$ the game is effectively played on a static topology, namely the weighted aggregated graph corresponding to the whole observation interval. We focus here on the top panels of Fig. 2 and Fig. 3, where we show how the average fraction of cooperators depends on the parameters S and T and on the length Δt of the strategy update interval. We considered six values of Δt for the MIT Reality Mining data set, from $\Delta t = 1$ hour up to the whole observation interval, and eight values for INFOCOM'06, ranging from minutes up to the aggregate network.

At first glance, we notice that the rightmost diagrams in both figures, which correspond to $\Delta t = M\tau$, are in perfect agreement with the results about evolutionary games played on static topologies reported in the literature (see *e.g.* [18, 26]). If we look at the cooperation diagrams obtained by increasing the value of Δt in the original sequences of graphs (top panels of Fig. 2 and Fig. 3), we notice that, for any pair (T, S), a larger update interval corresponds to a higher fraction of cooperators. In particular, for MIT Reality Mining (Fig. 2) the fraction of cooperators increases up until $\Delta t = 2$ months, after which the cooperation diagram is practically indistinguishable from that obtained on the static aggregated graph. For INFOCOM'06, instead, a strategy update interval larger than 2 hours already produces a cooperation diagram similar to that obtained in the aggregated graph. These results indicate that defectors actually take advantage from the volatility of edges, and that cooperation can emerge only if the strategy update interval is large enough. We will check this issue later in section II D.

As we pointed out above, edge activation patterns show non-trivial correlations. To highlight the effects of temporal correlations and of periodicity in the appearance of links in the real data sets, we have simulated the games also on randomized time-varying graphs and on synthetic networks generated through the activity-driven model [60]. The results for randomized graphs and activity-driven graphs are reported, respectively, in the middle and in the bottom panels of Fig. 2 and Fig. 3.

Randomized time-varying graphs are obtained by uniformly reshuffling the original sequences of snapshots. In this case the frequency of each pairwise contact is preserved equal to that of the original data set. However, the temporal correlations of these contacts, namely the persistence of an edge during consecutive time snapshots, are completely washed out. As expected, for $\Delta t = M\tau$ the cooperation diagrams obtained on the reshuffled se-

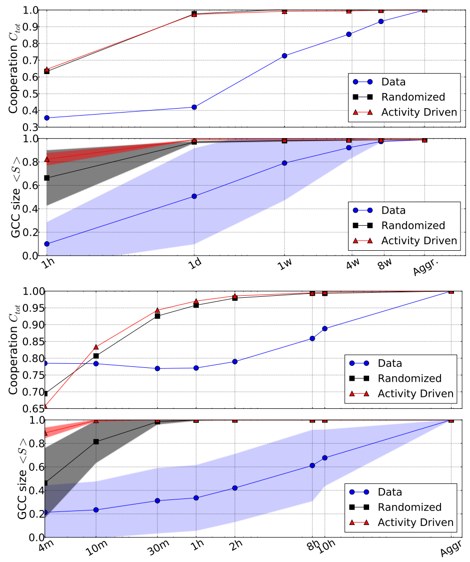


FIG. 4. **Cooperation level and size of the Giant Component**. Overall cooperation level $C_{tot}(\Delta t)$ and average size of the giant component $\langle \mathcal{S} \rangle$ as a function of the aggregation interval Δt for MIT Reality Mining (top panels) and INFOCOM'06 (bottom panels). Blue circles correspond to the original data, black squares to the reshuffled networks and red triangles to the activity-driven model. The shades indicate the standard deviation of $\langle \mathcal{S} \rangle$ across the sequence of graphs for each value of Δt. Notice that the typical size of the giant component at time-scale Δt correlates quite well with the observed cooperation level at the same time-scale.

quences (middle rightmost panels of Fig. 2 and Fig. 3) are identical to those obtained on the corresponding original data sets (top rightmost panels). In fact, when $\Delta t = M\tau$ each agent plays with all the contacts she has seen in the whole observation interval, with the corresponding weights, before updating her strategy. Since, in this particular case the frequencies of contacts are the only ingredients responsible for the emergence of cooperation, and such frequencies are the same both in the original data set and in the reshuffled version, the results of the original and randomized graphs have to be identical. Conversely, for smaller values of Δt, the importance of the temporal correlations of each pairwise contact becomes clear since the cooperation diagrams for randomized and original networks are very different in both MIT Reality Mining and INFOCOM'06. In fact for the randomized graphs, the cooperation levels at $\Delta t = 1$ week and $\Delta t = 1$ hour, respectively, become comparable to those obtained for $\Delta t = M\tau$. This points out that destroying the temporal correlations of each pairwise contact enhances cooperation.

Little differences are observed (compared to the case of randomized graphs) when using activity-driven synthetic networks (results shown in the bottom panels of Fig. 2 and Fig. 3). In this case not only temporal correlations are washed out, but also the microscopic structure of each snapshot is replaced by a graph having a similar density of links. This rewiring distributes links more heterogeneously than in the original and the randomized sequences (see Methods for details). The cooperation diagrams of activity-driven networks show a further increase of the cooperation levels for even smaller values of the strategy update interval, Δt, than in the case of Random graphs. Namely, for $\Delta t = 1$ day in Reality Mining (Fig. 2) and for $\Delta t = 30$ minutes in INFOCOM (Fig. 3) we already recover the cooperation levels of $\Delta t = M\tau$.

The results reported in Fig. 2 and Fig. 3 suggest that the ordering, persistence and distribution of edges over consecutive time-window are all fundamental ingredients for the success of cooperation. In general, a small value of Δt in the original data sets corresponds to playing the game on a sparse graph, possibly comprising a number of small components, in which agents are connected to a small neighborhood that persists rather unaltered over consecutive time windows. The small size of the isolated clusters and the persistence of the connections within them allow defectors to spread their strategy efficiently. In the following, we will analyze this hypothesis by characterizing the structural patterns of the original time varying graph, its randomized version and the activity-driven synthetic network.

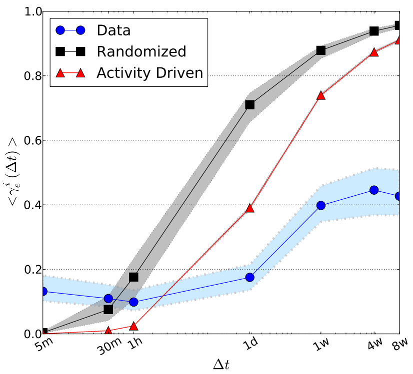
D. Structural analysis of time varying networks

In order to understand the dependence of cooperation on the strategy update interval Δt, we plot in Fig. 4 the average fraction, $\langle \mathcal{S} \rangle$, of the nodes found in the giant component of the graphs as a function of Δt, for the original data sets and for the reshuffled and synthetic sequences of snapshots. In general, for a given value of Δt, the giant component of graphs corresponding to randomized sequences or to the activity-driven model is larger than that of the graphs in the original ordering. The break down of temporal correlations between consecutive time snapshots in randomized and activity-driven networks produces an increase in the number of ties between different agents of the population even for small values of Δt. In addition, the more homogeneous distribution of links within the snapshots of the activity-driven network further increases the mixing of the agents and thus enlarges the size of the giant component compared to the randomized graphs.

In Fig. 4 we also plot the overall level of cooperation at a given aggregation scale Δt, $C_{tot}(\Delta t)$, defined as:

$$C_{tot}(\Delta t) = \frac{1}{C_{tot}(M\tau)} \int_0^2 \mathrm{d}T \int_{-1}^{1} C(T,S)\,\mathrm{d}S\,.$$

Notice that $C_{tot}(\Delta t)$ is divided by the value $C_{tot}(M\tau)$



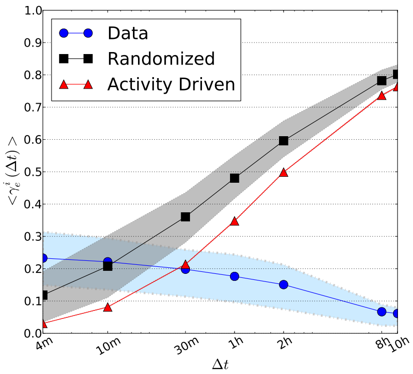


FIG. 5. Extremal temporal clustering γ_e^i as a function of the strategy update interval Δt on real data sets (blue dots). Top (bottom) panel refers to Reality Mining (INFOCOM). Black squares correspond to randomly reshuffled sequences and red triangles to activity-driven synthetic networks. In both datasets we notice that, for small values of Δt, real data display an higher value of clustering (persistence) than synthetic cases followed by a transition value of Δt above which we observe a rapid increase in the clustering of synthetic cases such that the previous situation is inverted.

corresponding to the whole observation interval, so that $C_{tot} \in [0, 1]$. The value of Δt at which $\langle \mathcal{S} \rangle$ is comparable with the number of nodes N, i.e. when $\langle \mathcal{S} \rangle \simeq 1$, coincides with the value of Δt at which the cooperation diagram becomes indistinguishable from that obtained for the aggregate network, $C_{tot}(\Delta t) \simeq 1$, for both the original and the reshuffled sequences of snapshots. This result confirms that the size of the giant connected component, of the graph corresponding to a given aggregation interval, plays a central role in determining the level of cooperation sustainable by the system, in agreement with the experiments discussed in [59] for the case of static complex networks.

It is also interesting to investigate the role of edge correlations on the observed cooperation level. To this aim, we analyze the temporal clustering (see Methods), which captures the average tendency of edges to persist over time. In Fig. 5 we plot the evolution of the temporal clustering as a function of the strategy update interval Δt. The results reveal clearly that, for small values Δt, the persistence of ties in the two original data sets is larger than in the randomized and the activity-driven graphs. In this regime the average giant component is small in both the real and randomized cases, thus pointing out that the temporally connected components are composed of small clusters. However, the larger temporal clustering observed for small Δt in the original data implies that the node composition of these small components changes very slowly compared to the faster mixing observed in the random data sets. These are then the two ingredients depressing the cooperation levels in the original data as compared to the random cases: the size of the giant component and how much such components change even at fixed size. As further confirmation, we notice that link persistence grows in a similar way as Δt increases in randomized and activity-driven networks. This growth points out that the randomization of snapshots in one null model and the redistribution of links in the other one make the ties more stable as Δt increases. Instead, the results found on the original data sets suggest (in particular for the case of INFOCOM) that ties are rather volatile, being active for a number of consecutive snapshots and then inactive for a large time interval. In randomized and activity-driven graphs the stabilization of ties together with the fast increase in the size of the giant component make the resulting time-varying graph much more similar to static networks, thus improving the survival of cooperation, as compared to the volatile and strongly fragmented scenario of the real time-varying graphs.

III. CONCLUSIONS

Although the impact of network topology on the onset and persistence of cooperation has been extensively studied in the last years, the recent availability of data sets with time-resolved information about social interactions allows a deeper investigation of the impact on evolutionary dynamics of time-evolving social structures. Here we have addressed two crucial questions: does the interplay between graph evolution and strategy update affect the classical results about the enhancement of cooperation driven by network reciprocity? And what is the role of the time-correlations of temporal networks in the evolution of cooperation? The results of the simulations confirm that, for all the four social dilemmas studied in this work, cooperation is seriously hindered when (i) agent strategy is updated too frequently with respect to the typical time-scale of agent interaction and (ii) real-world timecorrelations are present. This phenomenon is a consequence of the relatively small size of the giant component of the graphs obtained at small aggregation intervals. However, when the temporal sequence of social contacts is replaced by time-varying networks preserving the original activity attributes of links or nodes but breaking the original temporal correlations, the structural patterns of the network at a given time-scale of strategy update changes dramatically from those observed in real data. As a consequence, the effects of temporal resolution over cooperation are smoothed and, by breaking the real temporal correlations of social contacts, cooperation can emerge and persist also for moderately small strategy update frequencies. This result highlights that both the interplay of strategy update and graph evolution and the presence of temporal correlations, such as edge persistence and recurrence, seem to have fundamental effects on the emergence of cooperation.

Our findings suggest that the frequency at which the connectivity of a given system are sampled has to be carefully chosen, according with the typical time-scale of the social interaction dynamics. For instance, as stock brokers might decide to change strategy after just a couple of interactions, other processes like trust formation in business or collaboration networks are likely to be better described as the result of multiple subsequent interactions. These conclusions are also supported by the results of a recent paper of Ribeiro *et al.* [52] in which the effects of temporal aggregation interval Δt in the behavior of random walks are studied. Also, the fundamental role played by the real-data time correlations in dynamical processes on the graph calls for more models of temporal networks and for a better understanding of their nature.

In a nutshell, the arguments indicating network reciprocity as the social promoter of cooperator have to be revisited when considering time-varying graphs. In particular, one should always bear in mind that both the over- and the under-sampling of time-evolving social graph and the use of the finest/coarsest temporal resolution could substantially bias the results of a game-theoretic model played on the corresponding network. These results pave the way to a more detailed investigation of social dilemmas in systems where not only structural but also temporal correlations are incorporated in the interaction maps.

IV. METHODS

A. MIT Reality Mining data set

The data set describes proximity interactions collected through the use of Bluetooth-enabled phones [29]. The phones were distributed to a group of 100 users, composed by 75 MIT Media Laboratory students and 25 faculty members. Each device had a unique tag and was able to detect the presence and identity of other devices within a range of 5-10 meters. The interactions, intended as proximity of devices, were recorded over a period of about six months. In addition to the interaction data, the original dataset included also information regarding call logs, other Bluetooth devices within detection range, the cell tower to which the phones was connected and information about phone usage and status. Here, we consider only the contact network data, ignoring any other contextual metadata. The resulting time-varying network is an ordered sequence of 41291 graphs, each having N=100 nodes. Each graph corresponds to a proximity scan taken every 5 minutes. An edge between two nodes indicates that the two corresponding devices were within detection range of each other during that interval. We refer to such links as *active*. During the entire recorded period, 2114 different edges have been detected as active, at least once. This corresponds to the aggregate graph having a large average node degree $\langle k \rangle \simeq 42$. However, this is an artefact of the aggregation; the single snapshots tend to be very sparse, usually containing between 100 and 200 active edges.

B. INFOCOM'06 data set

The data set consists of proximity measurements collected during the IEEE INFOCOM'06 conference held in a hotel in Barcelona in 2006 [30]. A sample of 78 participants from a range of different companies and institutions were chosen and equipped with a portable Bluetooth device, Intel iMote, able to detect similar devices nearby. Area "inquiries" were performed by the devices every 2 minutes, with a random delay or anticipation of 20 seconds. The delay/anticipation mechanism was implemented in order to avoid synchronous measurements, because, while actively sweeping the area, devices could not be detected by other devices. A total number of 2730 distinct edges were recorded as active at least once in the observation interval, while the number of edges active at a given time is significantly lower, varying between 0 and 200, depending on the time of the day.

C. Strategy update rule

The Fermi Rule consists in the following updating strategy. A player i chooses one of her neighbors j at random and copies the strategy of j with a probability:

$$P_{i\to j} = \frac{1}{1+e^{-\beta(p_j-p_i)}} \ , \qquad (3)$$

where $(p_j - p_i)$ is the difference between the payoffs of the two players, and β is a parameter controlling the smoothness of the transition from $P_{i\to j} = 0$ for small values of $(p_j - p_i)$, to $P_{i\to j} = 1$ for large values of $(p_j - p_i)$. Notice that for $\beta \ll 1$ we obtain $P_{i\to j} \simeq 0.5$ regardless of the value of $(p_j - p_i)$, which effectively corresponds to a random strategy update. On the other hand, when $\beta \gg 1$ then $P_{i\to j} \simeq \Theta(p_j - p_i)$, being $\Theta(x)$ the Heaviside step function. In this limit, the strategy update is driven only by the ordering of the payoff values.

D. Activity-driven model

The activity-driven model, introduced in Ref. [60], is a simple model to generate time-varying graphs starting from the empirical observation of the activity of each node, in terms of number of contacts established per unit time. Given a characteristic time-window Δt, one measures the activity potential x_i of each agent i, defined as the total number of interactions (edges) established by i in a time-window of length Δt divided by the total number of interactions established on average by all agents in the same time interval. Then, each agent is assigned an activity $a_i = \eta x_i$, which is the probability per unit time to create a new connection or contact with any another agent j. The coefficient η is a rescaling factor, whose value is appropriately set in order to ensure that the total number of active nodes per unit time in the system is equal to $\eta\langle x\rangle N$, where N is the total number of agents. Notice that η effectively determines the average number of connections in a temporal snapshot whose length corresponds to the resolution of the original data set.

The model works as follows. At each time t the graph G_t starts with N disconnected nodes. Then, each node i becomes active with probability $a_i\Delta t$ and connects to m other randomly selected nodes. At the following time-step, all the connections in G_t are deleted, and a new snapshot is sampled.

Notice that time-varying graphs constructed through the activity-driven model preserve the average degree of nodes in each snapshot, but impose that connections have, on average, a duration equal to Δt, effectively washing out any temporal correlation among edges.

E. Temporal clustering

Several metrics have been lately proposed to measure the tendency of the edges of a time-varying graph to persist over time. One of the most widely used is the unweighted temporal clustering, introduced in Ref. [44], which for a node i of a time-varying graph is defined as:

$$\gamma^i = \frac{1}{T-1}\sum_{t=1}^{T-1}\frac{\sum_j a_{ij}^t a_{ij}^{t+1}}{\sqrt{k_i^t k_i^{t+1}}}\,, \tag{4}$$

where a_{ij}^t are the elements the adjacency matrix of the time-varying graph at snapshot t, k_i^t is the total number of edges incident on node i at snapshot t and T is the duration of the whole observation interval. Notice that γ^i takes values in $[0,1]$. In general, a higher value of γ^i is obtained when the interactions of node i persist longer in time, while γ^i tends to zero if the interactions of i are highly volatile.

If each snapshot of the time-varying graph is a weighted network, where the weight ω_{ij}^t represents the strength if the interaction between node i and node j at time t, we can define a weighted version of the temporal clustering coefficient as follows:

$$\gamma_w^i = \frac{1}{T-1}\sum_{t=1}^{T-1}\frac{\sum_j \omega_{ij}^t \omega_{ij}^{t+1}}{s_i^t s_i^{t+1}}\,. \tag{5}$$

Finally, if we focus more on the persistence of interaction strength across subsequent network snapshots, we can define the extremal temporal clustering as:

$$\gamma_e^i = \frac{1}{T-1}\sum_{t=1}^{T-1}\frac{\sum_j min(\omega_{ij}^t, \omega_{ij}^{t+1})}{\sqrt{s_i^t s_i^{t+1}}}\,, \tag{6}$$

where by considering the minimum between ω_{ij}^t and ω_{ij}^{t+1} one can distinguish between persistent interactions having constant strength over time and those interactions having more volatile strength. As in our case social interactions are seen to be highly volatile in real data sets, the extremal version of the temporal clustering seems to be the best choice to unveil the persistence of social ties at short time scales.

ACKNOWLEDGMENTS

This work was supported by the EU LASAGNE Project, Contract No.318132 (STREP), by the EU MULTIPLEX Project, Contract No.317532 (STREP), by the Spanish MINECO under projects MTM2009-13848 and FIS2011-25167 (co-financed by FEDER funds), by the Comunidad de Aragón (Grupo FENOL) and by the Italian TO61 INFN project. J.G.G. is supported by Spanish MINECO through the Ramón y Cajal program. G.P. is supported by the FET project "TOPDRIM" (IST-318121). R.S. is supported by the James S. McDonnell Foundation.

[1] E. Pennisi, "How did cooperative behavior evolve?" Science **309**, 93 (2005).
[2] E. Pennisi, "On the origin of cooperation." Science **325**, 1196–1199 (2009).
[3] J. Maynard-Smith, and E. Szathmary, "The Major Transitions in Evolution", Freeman, Oxford (UK), (1995).
[4] M.A. Nowak, "Evolutionary dynamics: exploring the equations of life." The Belknap Press of Harvard University Press, Cambridge, MA (2006).
[5] P. Kollock, "Social dilemmas: the anatomy of coopera-

tion." Annu. Rev. Sociol. **24**, 183–214 (1998).

[6] J. Maynard-Smith, and G.R. Price, "The logic of animal conflict." Nature **246**, 5427 (1973).

[7] J. Maynard-Smith, "Evolution and the Theory of Games", Cambridge Univ. Press, Cambridge (UK), (1982).

[8] H. Gintis, "Game theory evolving", Princeton University Press, Princeton, NJ (2009).

[9] L. Samuelson, "Evolution and game theory." J. Econ. Perspect. **16**, 47–66 (2002).

[10] D.J. Watts, and S.H. Strogatz, "Collective dynamics of small-world networks." Nature **393**, 440–442 (1998).

[11] S.H. Strogatz, "Exploring complex networks." Nature **410**, 268–276 (2001).

[12] R. Albert, and A-L Barabási, "Statistical mechanics of complex networks." Rev. Mod. Phys. **74**, 47–97 (2002).

[13] M.E.J. Newman, "The Structure and Function of Complex Networks." SIAM Rev. **45**, 167–256 (2003).

[14] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D-U Hwang, "Complex Networks: Structure and Dynamics." Phys. Rep. **424**, 175–308 (2006).

[15] S.N. Dorogovtsev, A.V. Goltsev, and J.F.F. Mendes, "Critical phenomena in complex networks." Rev Mod. Phys. **80**, 1275–1335 (2008).

[16] C. Castellano, S. Fortunato, and V. Loreto, "Statistical physics of social dynamics." Rev. Mod. Phys. **81** 591–646 (2009).

[17] F.C. Santos, and J.M. Pacheco, "Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation." Phys. Rev. Lett. **95**, 098104 (2005).

[18] F.C. Santos, J.M. Pacheco, and T. Lenaerts, "Evolutionary dynamics of social dilemmas in structured heterogeneous populations." Proc. Natl. Acad. Sci. U.S.A. **103**, 3490–3494 (2006).

[19] J. Gómez-Gardeñes, M. Campillo, L.M. Floría, and Y. Moreno, "Dynamical Organization of Cooperation in Complex Topologies." Phys. Rev. Lett. **98**, 108103 (2007).

[20] S. Assenza, J. Gómez-Gardeñes, and V. Latora, "Enhancement of cooperation in highly clustered scale-free networks." Phys. Rev. E **78**, 017101 (2008).

[21] F.C. Santos, M.D. Santos, and J.M. Pacheco, "Social diversity promotes the emergence of cooperation in public goods games." Nature **454**, 213–216 (2008).

[22] J. Gómez-Gardeñes, M. Romance, R. Criado, D. Vilone, and A. Sánchez, "Evolutionary games defined at the network mesoscale: The Public Goods game." Chaos **21**, 016113 (2011).

[23] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floria, and Y. Moreno, "Evolutionary dynamics of group interactions on structured populations: a review." J. Roy. Soc. Interface **10**, 20120997 (2013).

[24] G. Szabó, and G. Fáth, "Evolutionary games on graphs." Phys. Rep. **447**, 97 (2007).

[25] M.D. Jackson, "Social and economic networks" Princeton Univ. Press, Princeton, NJ (2008).

[26] C.P. Roca, J. Cuesta, and A. Sánchez, "Evolutionary game theory: temporal and spatial effects beyond replicator dynamics." Phys. Life Rev. **6**, 208 (2009).

[27] M. Perc, and A. Szolnoki, "Coevolutionary games - A mini review." BioSystems **99**, 109–125 (2010).

[28] T. Gross, and B. Blasius, "Adaptive coevolutionary networks: a review." J. R. Soc. Interface **5**, 259–271 (2008).

[29] N. Eagle, and A. Pentland, "Reality mining: sensing complex social systems." Personal and Ubiquitous Computing **10**, 255–268 (2006).

[30] J. Scott *et al.* , "CRAWDAD Trace", INFOCOM, Barcelona (2006).

[31] L. Isella, M. Romano, A. Barrat, C. Cattuto, V. Colizza V, *et al.* , "Close Encounters in a Pediatric Ward: Measuring Face-to-Face Proximity and Mixing Patterns with Wearable Sensors." PLoS ONE **6**, e17144 (2011).

[32] J. Stehle, N. Voirin, A. Barrat, C. Cattuto, L. Isella,*et al.* "High-Resolution Measurements of Face-to-Face Contact Patterns in a Primary School." PLoS ONE **6**, e23176 (2011).

[33] L. Isella, J. Stehlé, A. Barrat, C. Cattuto, J-F Pinton, and W. Van den Broeck, "What's in a crowd? Analysis of face-to-face behavioral networks." J. Theor. Biol. **271**, 166–180 (2011).

[34] M. Karsai, M Kivelä, R.K. Pan, K. Kaski, J. Kertész, A-L Barabási, and J. Saramäki, "Small But Slow World: How Network Topology and Burstiness Slow Down Spreading" Phys. Rev. E **83**, 025102(R) (2011).

[35] A-L Barabási, "The origin of bursts and heavy tails in human dynamics." Nature **435**, 207–211 (2005).

[36] J. Stehlé, A. Barrat, and G. Bianconi, "Dynamical and bursty interactions in social networks." Phys. Rev. E **81**, 035101(R) (2010).

[37] M.C. González, C.A. Hidalgo, and A-L Barabási, "Understanding individual human mobility patterns." Nature **453**, 779–782 (2008).

[38] M. Szell, R. Lambiotte, and S. Thurner, "Multi-relational Organization of Large-scale Social Networks in an Online World." Proc. Natl. Acad. Sci. U.S.A. **107**, 13636–13641 (2010).

[39] M. Szell, R. Sinatra, G. Petri, S. Thurner, and V. Latora, "Understanding mobility in a social petri dish" Scientific Reports **2**, 457 (2012).

[40] G. Kossinets, J. Kleinberg, and D. Watts, "The Structure of Information Pathways in a Social Communication Network," Proc. 14th ACM SIGKDD Intl. Conf. on Knowledge Discovery and Data Mining, (2008)

[41] V. Kostakos, "Temporal graphs." Physica A **388**, 1007–1023 (2009).

[42] J. Tang, M. Musolesi, C. Mascolo, and V. Latora, "Temporal Distance Metrics for Social Network Analysis." Proceedings of the 2nd ACM SIGCOMM Workshop on Online Social Networks (WOSN'09), (2009)

[43] P. Holme, and J. Saramäki, "Temporal networks" Phys. Rep. **519**, 97–125 (2012).

[44] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora, "Small-world behavior in time-varying graphs." Phys. Rev. E **81**, 055101(R) (2010).

[45] R.K. Pan, and J. Saramäki, "Path lengths, correlations, and centrality in temporal networks," Phys. Rev. E **84**, 016105 (2011).

[46] L. Kovanen, M. Karsai, K. Kaski, J. Kertesz, and J. Saramäki. "Temporal motifs in time-dependent networks." J. Stat. Mech., P11005 (2011).

[47] J. Tang, M. Musolesi, C. Mascolo, V. Latora, and V. Nicosia, "Analysing Information Flows and Key Mediators through Temporal Centrality Metrics" Proceedings of the 3rd ACM Workshop on Social Network Systems (SNS'10), (2010).

[48] V. Nicosia, J. Tang, M. Musolesi, G. Russo, C. Mascolo, and V. Latora, "Components in time-varying graphs" Chaos **22**, 023101 (2012).

[49] P.J. Mucha, T. Richardson, K. Macon, M.A. Porter, and J-P Onnela, "Community structure in time-dependent, multiscale, and multiplex networks." Science **328**, 876–878 (2010).
[50] M. Starnini, A. Baronchelli, A. Barrat, and R. Pastor-Satorras, "Random walks on temporal networks." Phys. Rev. E **85**, 056115 (2012).
[51] N. Perra, A. Baronchelli, D. Mocanu, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani, "Walking and searching in time-varying networks" Phys. Rev. Lett. **109**, 238701 (2012).
[52] B. Ribeiro, N. Perra, and A. Baronchelli, "Quantifying the effect of temporal resolution on time-varying networks." *Scientific Reports* **3**, 1 (2013).
[53] L.E.C. Rocha, F. Liljeros, and P. Holme, "Simulated Epidemics in an Empirical Spatiotemporal Network of 50,185 Sexual Contacts." PLoS Comput Biol **7**, e1001109 (2011).
[54] L.E.C. Rocha, A. Decuyper, and V.D. Blondel, "Epidemics on a stochastic model of temporal network." *arXiv:1204.5421* (2012).
[55] L.E.C. Rocha, and V.D. Blondel, "Bursts of vertex activation and epidemics in evolving networks." PLoS Comput Biol **9**, e1002974 (2013).
[56] N. Fujiwara, J. Kurths, and A. Díaz-Guilera, "Synchronization in networks of mobile oscillators." Phys. Rev. E **83**, 025101(R) (2011).
[57] L.E. Blume, "The Statistical Mechanics of Strategic Interaction." Games Econ. Behav. **5**, 387 (1993).
[58] G. Szabó, and C. Töke, "Evolutionary prisoner's dilemma game on a square lattice." Phys. Rev. E **58**, 69 (1998)
[59] Z. Wang, A. Szolnoki, and M. Perc, "If players are sparse social dilemmas are too: Importance of percolation for evolution of cooperation." Scientific Reports **2**, 369 (2012).
[60] N. Perra, B. Gonçalves, R. Pastor-Satorras, A. Vespignani, "Activity driven modeling of time varying networks", *Scientific Reports* **2**, 469 (2012).

Chapter 5

Evolutionary vaccination dilemma in complex networks

PHYSICAL REVIEW E **88**, 032803 (2013)

Evolutionary vaccination dilemma in complex networks

Alessio Cardillo,[1,2] Catalina Reyes-Suárez,[3] Fernando Naranjo,[3] and Jesús Gómez-Gardeñes[1,2]
[1]*Departamento de Física de la Materia Condensada, University of Zaragoza, Zaragoza 50009, Spain*
[2]*Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza 50018, Spain*
[3]*Departamento de Física, Universidad Pedagógica y Tecnológica de Colombia, Tunja, Colombia*
(Received 14 March 2013; revised manuscript received 15 August 2013; published 5 September 2013)

In this work we analyze the evolution of voluntary vaccination in networked populations by entangling the spreading dynamics of an influenza-like disease with an evolutionary framework taking place at the end of each influenza season so that individuals take or do not take the vaccine upon their previous experience. Our framework thus puts in competition two well-known dynamical properties of scale-free networks: the fast propagation of diseases and the promotion of cooperative behaviors. Our results show that when vaccine is perfect, scale-free networks enhance the vaccination behavior with respect to random graphs with homogeneous connectivity patterns. However, when imperfection appears we find a crossover effect so that the number of infected (vaccinated) individuals increases (decreases) with respect to homogeneous networks, thus showing the competition between the aforementioned properties of scale-free graphs.

DOI: 10.1103/PhysRevE.88.032803 PACS number(s): 89.75.Fb, 05.70.Fh

I. INTRODUCTION

The advent of network science [1,2] has provided an important set of computational and statistical physics tools for describing the problem of epidemic spreading by incorporating the realistic interaction patterns of the constituents of social and technological systems [3]. Classical approaches to epidemiology [4,5] rely on the use of the theory of phase transitions and critical phenomena, so as to unveil the onset and the macroscopic impact of epidemic outbreaks. Recently these techniques have been pervasively adapted to study a variety of critical phenomena on top of networks [6].

The main contribution of the former line of research to epidemiology has been the development of a generalized mean-field framework in which general patterns of interactions can be included. In particular, it was shown [7–12] that for scale-free networks [in which the probability distribution of having a node with k neighbors follows a power law, $P(k) \sim k^{-\alpha}$] the epidemic onset was anticipated as compared to substrates with more regular (or homogeneous) connectivity patterns. Moreover, when $\alpha < 3$ (as most of social and technological networks show [13,14]) and for large enough (thermodynamic limit) systems, the epidemic onset vanishes, meaning that even a very small fraction of infected elements with small infective power can spread a disease to a macroscopic part of the population by a sequence of contagions between neighbors of the network, as happens in human contacts [15–18].

Apart from the theoretical value of the above finding, its direct implications on public health campaigns and the security of technological networks such as the Internet demand a deeper understanding about the influence that diverse contact patterns have on disease dynamics, its co-evolution [19,20], and the design of new algorithms for immunization and vaccination policies. Typically, these studies aim at identifying the most efficient way for reducing the impact of an epidemic by the vaccination or immunization of the minimal number of nodes. To this aim, different methods to identify the most important nodes to be immunized have been proposed [21–24].

The former works concern the immunization of technological networks since in social contexts vaccination is typically voluntary. Thus, the study of the immunization of a population demands that we include the ways vaccination and risky behaviors compete and spread across individuals. To this aim, one may consider game theory to formulate a social dilemma in terms of the benefits associated to each of the behaviors: vaccination or not. Within this framework individuals act rationally, i.e., by choosing their strategy after an evaluation of their potential benefits. This evaluation is done by considering their perception of the risk to contract the disease. For well-mixed populations recent results show [25–30] that voluntary vaccination is not efficient to reach efficient immunization. However, this kind of approach was generalized to networks [31], unveiling an enhancement of voluntary vaccination.

The former game theoretical approach considers that agents aim at maximizing their own benefits. However, the decisions of individuals can evolve in time depending on the epidemic incidence observed in the population. In this framework agents are prone to adopt the strategies that are expected to perform better based on the information available. This evolutionary avenue has been recently adopted to the vaccination dilemma. A first evolutionary avenue is presented in Refs. [32–34] where both disease transmission and vaccinating behavior evolve in time simultaneously. The evolution for the fraction of vaccinated individuals is driven by the difference of payoffs between vaccinated and nonvaccinated agents (as in the case of the well-known replicator equation of evolutionary games [35,36]), with the latter determined by the epidemic incidence at that time. A second evolutionary approach is proposed in Ref. [37]. In this case, inspired by seasonal influenza, the number of vaccinated individuals remains constant during the duration of the influenza season. After each season, individuals evaluate the payoffs based on the incidence of the disease in the last season and decide whether to vaccinate or not for the next seasons.

Here we take a similar avenue to that of Ref. [37] regarding the dynamical setup and the motivation: the vaccination for seasonal influenza. However, the way in which payoffs are constructed and the way individuals choose their strategy follow the typical setup of evolutionary games [35,36]. This setup, originally presented in Ref. [38] for the vaccination dilemma, considers that individuals are assigned a payoff that is solely based on the personal experience during the last season. In addition, the strategic choice is based on the imitation of those individuals with better payoffs. Thus, we do not consider that individuals follow a rational derivation of the payoffs associated to vaccination and risky behavior based on the information available. This allow us to connect the vaccination dilemma with other studies on the evolutionary game dynamics of social dilemmas [35,36].

In recent years, the study of the evolutionary game dynamics of social dilemmas on structured populations [39–41] has shown that cooperation (here related to vaccination) is favored when the interactions among individuals take the form of scale-free networks [42,43]. Inspired by this result, in this work we explore the spread of vaccination behavior across networks with homogeneous and heterogeneous (scale-free) connectivity patterns. Our results show that when vaccine is perfect, scale-free networks enhance the vaccination behavior with respect to homogeneous graphs, thus reducing the impact of the disease on the population. However, when vaccine is imperfect, we find a crossover effect, and homogeneous networks outperform scale-free ones. This latter scenario reveals an interesting competition between the rapid spread of both diseases and cooperative behaviors in scale-free graphs.

II. THE MODEL

As introduced above, to incorporate the competition between disease spreading and evolutionary dynamics on top of a network we entangle these two dynamical frameworks by producing an iterative sequence of a two-stage process. In both stages the interaction pattern among individuals is described by a complex network (keeping the same network for both dynamical setups). This network is given by an $(N \times N)$ adjacency matrix A_{ij} so that when two individuals interact $A_{ij} = 1$, whereas $A_{ij} = 0$ otherwise. In this way, the number k_i of neighbors (contacts) of a given node, say, i, is given by $k_i = \sum_{j=1}^{N} A_{ij}$.

In this work we will consider two of the most paradigmatic network models: Erdős-Rényi (ER) graphs [44] and Barabási-Albert (BA) networks [45]. The former class of graphs are described by a Poisson degree distribution $P(k)$, so that most of the nodes have a connectivity close to the mean value $\langle k \rangle$. On the other hand, BA networks display a power-law degree distribution of the form, $P(k) \sim k^{-3}$, thus incorporating the scale-free (SF) property of real-world networks. The implementation of our dynamical setup aims at revealing the differences between the heterogeneous degree pattern displayed in SF and the rather homogeneous structure of ER graphs. To this aim, for both ER and SF networks, the average connectivity of the nodes is set to $\langle k \rangle = 6$. Below we introduce the rules governing the two-stage dynamics, also sketched in Fig. 1.

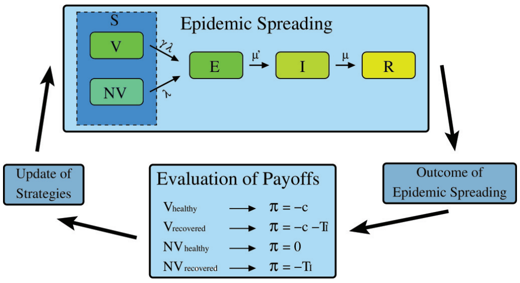


FIG. 1. (Color online) Resuming sequence of the evolutionary picture of our model. The top box describes the epidemic spreading process. The bottom one displays the payoffs accumulated by the agents according to their *strategy*. Arrows denote the causal order of the evolutionary process.

A. Disease spreading

The first of the stages of our dynamical setup is based on the evolution of a susceptible-exposed-infected-recovered (SEIR) model [4,5]. This model captures the dynamics of influenza-type infections. Susceptible nodes have not been infected and are healthy. They catch the disease via direct contact with exposed neighbors at a rate λ. Exposed nodes are supposed to carry the virus although they still do not display symptoms of the disease; thus these individuals are highly infectious during this incubation period. Exposed nodes become infected with some rate μ' which typically is the inverse time of the incubation period of the disease. Infected nodes, on the other hand, although still carrying the virus are here assumed not to be infectious. In particular, we consider that during this period they remain isolated from the rest of the population. Finally, infected nodes pass to the recovered state with rate μ that is the inverse duration time of the convalescence period.

With the above rules we consider that each node i interacts simultaneously with its k_i neighbors per unit time. Thus, for a network described by the adjacency matrix A_{ij} the effective probability that a susceptible node i gets the disease per unit time is given by

$$P^i_{S\to E} = 1 - (1-\lambda)^{\sum_{j=1}^{N} A_{ij} x_j}\ , \tag{1}$$

where $x_j = 1$ when node j is exposed and $x_j = 0$ otherwise. Here, in order to mimic the transmission of ordinary influenza, we have set $\mu' = 0.33$, since the time elapsed between exposure to the virus and development of symptoms is two to three days. In addition we take $\mu = 0.2$ since the symptoms of uncomplicated influenza illness resolve after a period of 3 to 7 days, so that the average permanence in the infected state is $\mu^{-1} = 5$ days.

The addition of vaccinated individuals to the formulation of our SEIR model implies that initially there is subset of susceptible individuals (representing a fraction N_V of the total population) that are less prone to catch the disease than nonvaccinated susceptible ones. In particular, we consider that a vaccinated individual is infected during a single contact with an exposed one at a rate $\lambda\gamma$, where $\gamma \in [0,1]$ is a parameter that modulates the quality of the vaccine, being perfect when $\gamma = 0$ and useless for $\gamma = 1$. In this way, the probability that a vaccinated individual i is infected per unit time reads

$$P^i_{S\to E} = 1 - (1-\gamma\,\lambda)^{\sum_{j=1}^N A_{ij}x_j} \,. \tag{2}$$

Once the values of the epidemic parameters μ and μ', the quality γ of the vaccine, and the fraction N_V of vaccinated individuals are set, we leave λ as the relevant control parameter of the SEIR model. In addition, the relevant order parameter of the dynamics is the fraction R of nodes that got infected once the epidemic process dies out, so that the macroscopic behavior is captured by the curve $R(\lambda)$. For a given value of λ one starts from an initial state in which a small fraction (here 5%) of the population is set as exposed. Then the SEIR dynamics is iterated until no individuals remain either as exposed or infected.

B. Evolutionary dynamics

Once the SEIR dynamics dies out we consider that the seasonal influenza period has passed. Before the next SEIR dynamics starts, individuals evaluate whether to vaccinate or not for the next season. At this point evolutionary dynamics takes place by assigning to each of the individuals a payoff π_i ($i = 1,\ldots,N$) that depends on their experience accumulated during the last SEIR propagation. As shown in Fig. 1, there are four possibilities:

(1) Vaccinated individuals that remain healthy during the last season have payoff $\pi = -c$ (where c is a cost associated to the vaccine).

(2) Vaccinated individuals that were infected during the last season have payoff $\pi = -c - T_I$ (where T_I is the time units that the individual remain in the infected state).

(3) Individuals that did not vaccinated and remain healthy during the season have payoff $\pi = 0$.

(4) Nonvaccinated individuals that were infected are assigned a payoff $\pi = -T_I$.

The cost c associated to the vaccination is related to different issues such as the time spent to get vaccinated (via Public Health Services) or the probability that the vaccine causes side effects. To illustrate the vaccination dilemma let us show a very simple situation of a susceptible agent i in contact with an exposed agent. In this situation the expected payoff of i when having taken the vaccine is $\pi_V^{\exp} = -(1-\gamma\lambda)c - \gamma\lambda(c + 1/\mu)$ (here we assume that $T_I \simeq 1/\mu$). On the other hand, if agent i has adopted a risky behavior, its expected payoff turns into $\pi_{NV}^{\exp} = -\lambda/\mu$. Thus, in this single pairwise encounter, the rational choice is not to take the vaccine for any costs $c > \lambda(1-\gamma)/\mu$. This simple situation clearly reveals the Vaccination Dilemma. However, in a networked population the situation is rather more complex, and, more importantly, here we assume that individuals are not fully rational and, instead of deciding their behavior on expectations, they evolve their strategies based on their previous experience.

Evolutionary dynamics provides the framework to implement the dynamical evolution of strategies. In particular, as is usually done in evolutionary social dilemmas on networks, each individual, say, i, chooses at random one of its first neighbors, say, j, and compares their payoffs π_i and π_j respectively. Then the probability that agent i takes the strategy of j, s_j, for the next season increases with their payoff difference, $(\pi_j - \pi_i)$. One of the most used frameworks to calculate this probability is that of the Fermi-like rule [46,47], in which the probability that the strategy of the neighbor j is adopted by i reads

$$P_{s_j\to s_i} = \frac{1}{1+\exp[-\beta(\pi_j - \pi_i)]}\,, \tag{3}$$

where β is a parameter that allows us to span between random ($\beta \ll 1$) and strong selection ($\beta \gg 1$). Here we adopted $\beta = 1$ and checked that our results are quite robust under changes of β. The update of strategies takes place simultaneously for all the agents. Once the new strategies are taken, the payoffs are set to zero, and the SEIR dynamics starts again with a new fraction N_V of vaccinated susceptible individuals.

Finally, let us note that we iterate the sequence of the two-stage process (SEIR dynamics and evolutionary dynamics) for a number of steps (generations) large enough to reach a steady state for the relevant observables: the average fraction of recovered, $\langle R\rangle$, and vaccinated individuals, $\langle N_V\rangle$. In addition, at the beginning of each generation we randomly assign the individuals that are vaccinated (so that they constitute 25% of the population) and those that are initially set as exposed (reaching 5% of the total population). It is worth mentioning that in real cases a small fraction of the population gain permanent immunity from exposure to the virus in the last generation. In our case we do not consider such inherited immunity to the new strain.

III. RESULTS

We start our discussion by briefly reporting the behavior of the SEIR model without vaccinated individuals. In the top panel of Fig. 2 we show the average fraction $\langle R\rangle$ of recovered individuals at the end of the SEIR dynamics as a function of the rate of infection per contact, λ, for ER and SF networks of $N = 1000$. From this figure it becomes clear that SF networks accelerates the onset λ_c of the epidemic regime as compared to ER graphs.

Let us now focus on the case of SF networks to evaluate the impact that voluntary vaccination (under an evolutionary framework) has on the immunization of the system. In the bottom panel of Fig. 2 we show the evolution of the

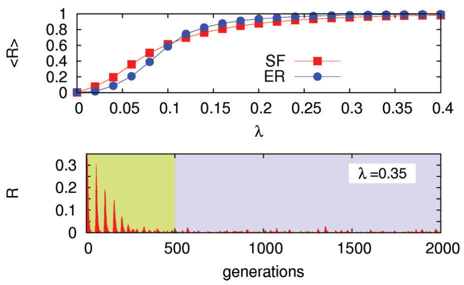


FIG. 2. (Color online) The top panel shows the epidemic diagram $\langle R\rangle(\lambda)$ for ER and SF networks when vaccination is not allowed. The bottom panel shows the evolution of the fraction of recovered individuals, R, with the generations of the evolutionary dynamics. The network is SF, and the rate of infection per contact is $\lambda = 0.35$, whereas vaccination is perfect $\gamma = 0$ and it has a cost $c = 0.1$.

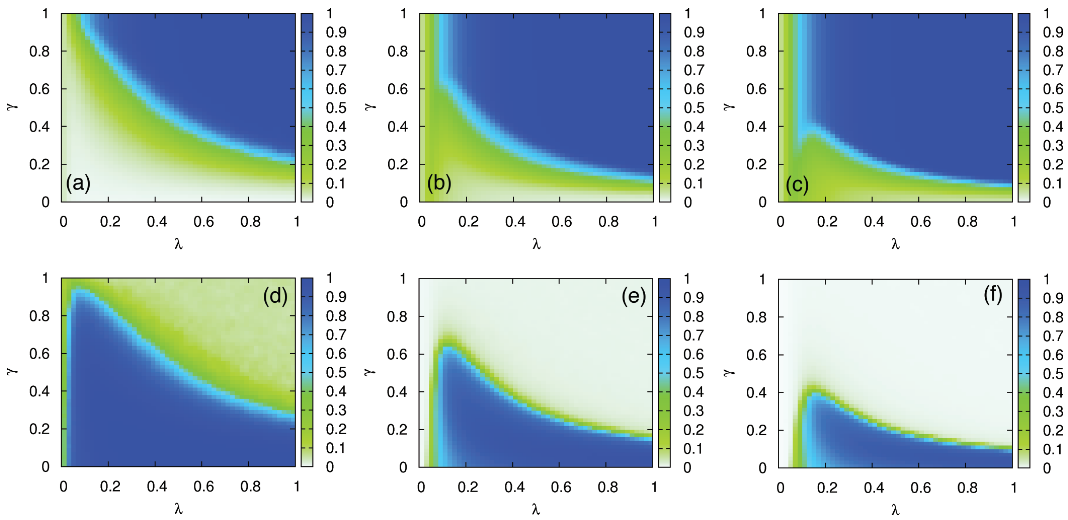


FIG. 3. (Color online) The contour plots show the average fraction of recovered $\langle R\rangle$ (top) and vaccinated $\langle N_V\rangle$ (bottom) individuals as a function of the infection rate λ and the vaccine quality γ for SF networks. From left to right the panels correspond to different vaccination costs: $c = 0.1$ [panels (a) and (d)], $c = 0.5$ [panels (b) and (e)], and $c = 1.0$ [panels (c) and (f)]. As the cost increases we note that the overall fraction of vaccinated individuals decreases while that of recovered nodes increases. Interestingly when $c = 0.1$ there is a range of low γ values ($\gamma < 0.1$) for which the epidemic threshold disappears and the disease cannot spread for any value of λ.

fraction of recovered individuals R for a sequence of 2000 generations. The rate of infection used in this simulation is set to $\lambda = 0.35$, which, as the top panel shows, corresponds to a situation in which almost all the population has been infected $\langle R\rangle \simeq 1$ when no vaccination is allowed. Instead, when individuals can decide whether to take the vaccination (under the aforementioned evolutionary rules) we show that the epidemic phase does not appear ($R \simeq 0$) since the population has evolutionarily adopted the vaccination strategy.

Remarkably, the transient regime (lasting around 500 generations) shows an interesting pattern of rise and falls for the number of recovered individuals R. This behavior points out that, before vaccination prevails, the population displays an oscillating behavior between vaccination and risky behavior. Obviously, when many people vaccinate (falls in R) the epidemic falls, but vaccinated individuals are tempted not to take the vaccine due to the higher benefits of risky individuals. This leads to a progressive increase of the infections (denoted by the increase of R) that reverse the balance of benefits between risky and vaccinated individuals. This rise-and-fall behavior together with the significant duration of this transient regime reveal the importance of risk perception in voluntary vaccination.

A. Macroscopic behavior of vaccine taking in SF networks

Now we analyze the behavior after the transient regime. To this aim we compute the average fraction of vaccinated $\langle N_V\rangle$ and recovered $\langle R\rangle$ individual in the steady state as a function of λ and the quality γ of the vaccine. For each couple of values (λ, γ) we have run 100 simulations (each of them comprising 2000 generations). In Fig. 3 we report these functions for several vaccine costs c in SF networks. In particular, the panels in the top show the diagrams $\langle R\rangle(\lambda,\gamma)$ and those in the bottom show $\langle N_V\rangle(\lambda,\gamma)$. From left to right the panels correspond to the following vaccine costs: $c = 0.1$, 0.5, and 1.0.

Let us focus on those diagrams corresponding to $c = 0.1$ [panels (a) and (d)]. The function $R(\lambda,\gamma)$ shows that for values of $\gamma \in [0,0.1]$ (roughly perfect vaccination) the epidemic threshold disappears since $\langle R\rangle \simeq 0$ for all the values of λ. In its turn, we note from panel (d) that for this latter region the fraction of vaccinated individuals is roughly $\langle N_V\rangle \simeq 1$ except for very low values of λ for which the disease cannot spread even when no immunization is present. If we increase further the value of γ we recover the epidemic onset λ_c whose values decrease as the vaccine get worse, i.e., as γ increases. In addition, the vaccination behavior decreases so that for a given value of γ the advantage provided by vaccines is not useful anymore for $\lambda > \lambda_c$. Obviously, for $\gamma = 1$ we recover the usual diagram $R(\lambda)$, shown in the top panel of Fig. 1, for SF networks since the vaccine provides no advantage, and, as shown in panel (d), almost no individual in the network holds the vaccination strategy giving $\langle N_V\rangle \simeq 0$ for all λ values.

As we increase the cost of the vaccine to $c = 0.5$ [panels (b) and (e)] and $c = 1.0$ [panels (c) and (f)] we observe that the overall fraction of recovered (vaccinated) individuals increases (decreases). Remarkably, the maximum value of γ for which there is no epidemic threshold decreases with c, and for $c = 1.0$ we cannot appreciate this effect. It is interesting to note that the usual epidemic diagram of SF networks without immunization is recovered for lower values of γ. For instance, in panel (b) we note that for $\gamma > 0.6$ the curve $R(\lambda)$ does not change, whereas from panel (e) we note that, within this region, individuals do not vaccinate anymore ($\langle N_V \rangle = 0$).

B. SF versus ER networks: The importance of vaccine quality

Having reported the macroscopic behavior in SF networks as concerns the influence of the vaccine quality and its cost, we now focus on the dependence on the networked substrate in which both the disease and the vaccination strategies spread. To this aim, we compare the behavior in SF and ER networks in order to measure the role of degree heterogeneity on the vaccination behavior. Importantly, we have considered SF networks as obtained from the Barabási-Albert model [45] after a complete randomization preserving the degree sequence of the nodes. In this way, we obtain SF networks with $P(k) \sim k^{-3}$ without any kind of degree-degree correlations that could influence the dynamical behavior. In addition, we have increased the size of the networks considered (in order to fully exploit the heterogeneous property of SF networks) to $N = 10^4$ nodes.

We first explore the case of perfect vaccination, $\gamma = 0$. In Fig. 4 we show the diagrams $\langle R \rangle(\lambda)$ (top) and $\langle N_V \rangle(\lambda)$ (bottom) for two different vaccination costs: $c = 0.5$ [panels (a) and (c)] and $c = 1.0$ [panels (b) and (d)]. In these panels we also show the standard deviations around the average values reported. From the panels we observe that SF networks outperform ER graphs since the overall average number of recovered (vaccinated) individuals is smaller (higher) in SF networks. In particular, the epidemic diagrams $\langle R \rangle(\lambda)$ display a clear peak around the respective epidemic thresholds, λ_c, of the original (without vaccination) graphs. Up to this point $\lambda < \lambda_c$, the epidemic does not spread, and thus vaccination behavior is not observed either as shown in the diagrams $\langle N_V \rangle(\lambda)$. The peak thus points out that the risk is so small that vaccination behavior does not show up, leading to a burst of infections, which reaches higher values in ER graphs. This result seems counterintuitive, since from the literature on epidemics on networks, SF graphs are always more prone to the spread of diseases than ER ones. Furthermore, from the diagrams $\langle N_V \rangle(\lambda)$ we note that the vaccination onset starts earlier for SF graphs, as their natural epidemic threshold is smaller than that of ER ones.

For values of λ above the natural epidemic threshold, the number of recovered nodes decreases dramatically in both networks. Here the risk of infection becomes larger, and individuals start to adopt the vaccination strategy as diagrams $\langle N_V \rangle(\lambda)$ in panels (c) and (d) show. However, vaccination behavior spreads more easily in SF networks than in ER graphs, and it is quite remarkable that, for this regime, the number of recovered nodes in ER graphs is always (for any value of λ) higher than in SF networks. Thus cooperative behavior, by taking the vaccine, spreads better in SF networks, in agreement with those studies about cooperation and social dilemmas in complex networks [42,43].

In Fig. 5 we explore the scenario of imperfect vaccination considering $\gamma = 0.12$. This regime shows the competition between two well-known effects: the aforementioned prevalence of cooperative behaviors in SF networks (with respect to ER

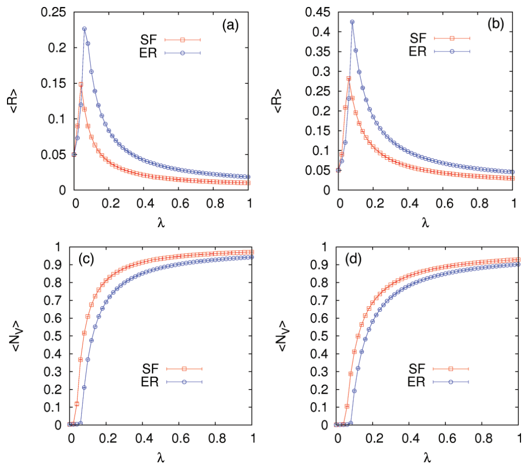


FIG. 4. (Color online) Epidemic $\langle R \rangle(\lambda)$ (top panels) and vaccination $\langle N_V \rangle(\lambda)$ (bottom panels) diagrams for ER and SF networks ($N = 10^4$, $\langle k \rangle = 6$) when the vaccine is perfect ($\gamma = 0$). The cost associated to the vaccine are $c = 0.5$ (left panels) and $c = 1.0$ (right panels).

graphs) and their weakness to the spread of diseases (again with respect to ER graphs). This competition appears as a crossover between the behavior of both $\langle R \rangle(\lambda)$ and $\langle N_V \rangle(\lambda)$ in SF and ER networks. In panels (a) and (b) we show that the curves $\langle R \rangle(\lambda)$ (after the peak close to the natural epidemic thresholds of both networks) cross at some λ^* values, which decreases with the cost of the vaccine c. Panels (c) and (d) show also a crossover behavior for $\langle N_V \rangle(\lambda)$, which appears with some delay with respect to that occurring at λ^* for $\langle R \rangle(\lambda)$. Note that this crossover is well defined since the standard deviations around the average values $\langle R \rangle$ and $\langle N_V \rangle$ are extremely low.

The behavior for $\lambda < \lambda^*$ shows the same trend as for the perfect vaccination case. SF networks outperform ER graphs showing a larger number of vaccinated individuals and a smaller number of infections. However, for the imperfect vaccine ($\gamma > 0$) the growth of λ affects both nonvaccinated and vaccinated individuals. Under such conditions, the virus finds in the SF networks a better backbone to propagate. In this way, panels (a) and (b) show that the failure of vaccination starts to become evident in SF networks at λ^*. The smaller benefits provided by the imperfection of the vaccine cause the number of vaccinated individuals to start to decrease after λ^*. Being larger the number of infections due to the imperfect vaccine in SF networks, as shown for $\lambda > \lambda^*$, the fall of vaccinated individuals occurs in SF networks at smaller values of λ than in ER graphs, giving rise to the crossover for $\langle N_V \rangle$ shown in panels (c) and (d).

It is quite remarkable that for large λ values and for $c = 1.0$ [panels (b) and (d)] the number of vaccinated individuals vanishes and the values of $\langle R \rangle$ goes close to one in a similar way as in the original network (without vaccination). Obviously, as the vaccine cost c increases, the solution $\langle R \rangle \simeq 1$ spans across a larger interval of λ values so that for large

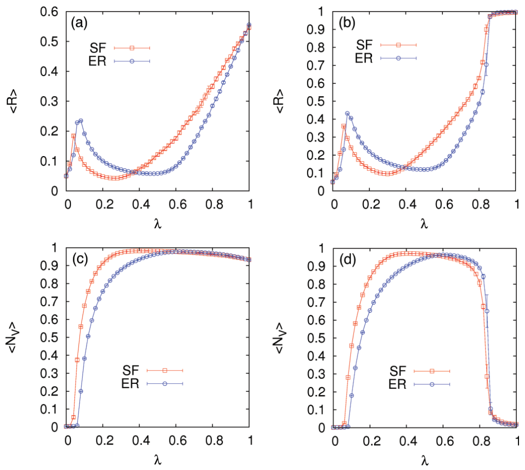


FIG. 5. (Color online) Epidemic $\langle R\rangle(\lambda)$ (top panels) and vaccination $\langle N_V\rangle(\lambda)$ (bottom panels) diagrams for ER and SF networks ($N = 10^4$, $\langle k\rangle = 6$) when the vaccine is not perfect ($\gamma = 0.12$). The cost associated to the vaccine are $c = 0.5$ (left panels) and $c = 1.0$ (right panels). The imperfection of the vaccine causes two crossovers, one for $\langle R\rangle$ and the other one for $\langle N_V\rangle$, between the performance of SF networks and ER graphs.

enough c there is no vaccinated individual in the population and one finally recovers the typical $\langle R\rangle(\lambda)$ diagram of Fig. 2(a).

IV. CONCLUSIONS

In this work we have analyzed the evolution of voluntary vaccination in networked populations. At variance with classical approaches we have considered an evolutionary framework so that individuals facing the vaccination dilemma do not take the most rational strategy by considering the benefits associated to each choice. On the contrary, they are considered as replicating agents that imitate the strategies based on their previous experience. To this aim we have entangled the spreading dynamics of an influenza-like disease with an evolutionary framework taking place at the end of each season. Our results show that when vaccine is perfect (so that vaccinated individuals do not get infected) scale-free networks enhance both the vaccination behavior and the effective immunization of the population as compared with random graphs with homogeneous connectivity patterns.

By considering vaccine imperfection we obtain two remarkable results. First, we have shown that, for scale-free networks and low vaccine costs, there is a threshold value for the vaccine imperfection so that, for values lower than this threshold, vaccination behavior spans across the population, and it is possible to suppress the disease for all the infection probabilities. Instead, when vaccine imperfection becomes large, agents are less prone to take it, and the disease takes advantage of this risky behavior to spread more efficiently across the population.

The other interesting result concerns the comparison between scale-free and homogeneous networks. We have shown that when imperfection appears the better performance of scale-free network is broken and there is a crossover effect so that the number of infected (vaccinated) individuals increases (decreases) with respect to homogeneous networks when λ is large enough. This crossover results from the competition of two well-known dynamical properties of scale-free networks: the fast propagation of diseases and the promotion of cooperative behaviors. Thus, the ability of scale-free networks in promoting cooperative behaviors (here represented as paying the cost of taking vaccine) is threatened when payoffs are dependent on a related dynamical process (here the spreading of a disease) whose evolution is also affected (here enhanced) by the heterogeneity of the network.

ACKNOWLEDGMENTS

J.G.G. acknowledges the hospitality of UPTC and useful discussions with S.L. Dorado and L.M. Floría. This work has been partially supported by the Spanish MINECO under projects FIS2011-25167 and FIS2012-38266-C02-01 and by the Comunidad de Aragón (Grupo FENOL) and the UPTC (Proyecto Capital Semilla). J.G.G. is supported by MINECO through the Ramón y Cajal program.

[1] R. Albert and A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002).

[2] M. E. J. Newman, SIAM Rev. **45**, 167 (2003).

[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez and, D.-U. Hwang, Phys. Rep. **424**, 175 (2006).

[4] D. J. Daley and J. Gani, *Epidemic Modeling* (Cambridge University Press, Cambridge, 1999).

[5] R. M. Anderson, R. M. May, and B. Anderson, *Infectious Diseases of Humans: Dynamics and Control* (Oxford University Press, Oxford, 1992).

[6] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes, Rev Mod. Phys. **80**, 1275 (2008).

[7] R. Pastor-Satorras and A. Vespignani, Phys. Rev. Lett. **86**, 3200 (2001).

[8] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **63**, 066117 (2001).

[9] A. L. Lloyd and R. M. May, Science **292**, 1316 (2001).

[10] Y. Moreno, R. Pastor-Satorras, and A. Vespignani, Eur. Phys. J. B **26**, 521 (2002).

[11] M. Barthélémy, A. Barrat, R. Pastor-Satorras, and A. Vespignani, Phys. Rev. Lett. **92**, 178701 (2004).

[12] S. Gómez, J. Gómez-Gardeñes, Y. Moreno, and A. Arenas, Phys. Rev. E **84**, 036105 (2011).

[13] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. E. Stanley, and Y. Aberg, Nature (London) **411**, 907 (2001).

[14] L. Amaral, A. Scala, M. Barthélémy, and H. E. Stanley, Proc. Natl. Acad. Sci. USA **97**, 11149 (2000).

[15] J. M. Read and M. J. Keeling, Proc. R. Soc. London B **270**, 699 (2002).

[16] J. M. Read and M. J. Keeling, Theo. Pop. Biol. **70**, 201 (2006).

[17] K. T. D. Eames and M. J. Keeling, Proc. Natl. Acad. Sci. USA **99**, 13330 (2002).
[18] J. Gómez-Gardeñes, V. Latora, Y. Moreno, and E. Profumo, Proc. Natl. Acad. Sci. USA **105**, 1399 (2008).
[19] T. Gross, C. J. Dommar D'Lima, and B. Blasius, Phys. Rev. Lett. **96**, 208701 (2006).
[20] B. Guerra and J. Gómez-Gardeñes, Phys. Rev. E **82**, 035101(R) (2010).
[21] R. Pastor-Satorras and A. Vespignani, Phys. Rev. E **65**, 036104 (2002).
[22] R. Cohen, S. Havlin, and D. ben-Avraham, Phys. Rev. Lett. **91**, 247901 (2003).
[23] P. Echenique, J. Gómez-Gardeñes, Y. Moreno, and A. Vázquez, Phys. Rev. E **71**, 035102 (2005).
[24] J. Gómez-Gardeñes, P. Echenique, and Y. Moreno, Eur. Phys. J. B **49**, 259 (2006).
[25] C. T. Bauch, A. P. Galvani, and D. J. D. Earn, Proc. Natl. Acad. Sci. USA **100**, 10564 (2003).
[26] C. T. Bauch and D. J. D. Earn, Proc. Natl. Acad. Sci. USA **101**, 13391 (2004).
[27] R. Breban, R. Vardavas, and S. Blower, Phys. Rev. E **76**, 031127 (2007).
[28] R. Breban, R. Vardavas, and S. Blower, PLoS Comput. Biol. **3**, e85 (2007).
[29] A. Perisic and C. T. Bauch, PLoS Comput. Biol. **5**, e1000280 (2008).
[30] D. M. Cornforth, T. C. Reluga, E. Shim, C. T. Bauch, A. P. Galvani, and L. A. Meyers, PLoS Comput. Biol. **7**, e1001062 (2011).
[31] H. Zhang, J. Zhang, C. Zhou, M. Small, and B. Wang, New J. Phys. **12**, 023015 (2010).
[32] S. Bhattacharyya and C. T. Bauch, J. Theor. Biol **257**, 276 (2010).
[33] C. T. Bauch, and S. Bhattacharyya, PLoS Comput. Biol. **8**, e1002452 (2012).
[34] A. D'Onofrio, P. Manfredi, and P. Poletti, PLoS One **7**, e45653 (2012).
[35] H. Gintis, *Game Theory Evolving* (Princeton University Press, Princeton, NJ, 2000).
[36] M. A. Nowak and R. M. May, Nature (London) **359**, 826 (1992).
[37] C. R. Wells, E. Y. Klein, and C. T. Bauch, PLoS Comput. Biol. **9**, e1002945 (2013).
[38] F. Fu, D. I. Rosenbloom, L. Wang, and M. A. Nowak, Proc. Roy. Soc. B **278**, 42 (2011).
[39] G. Szabó and G. Fath, Phys. Rep. **446**, 97 (2007).
[40] C. P. Roca, J. A. Cuesta, and A. Sánchez, Phys. Life Rev. **6**, 208 (2009).
[41] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, and Y. Moreno, J. Roy. Soc. Interface **10**, 20120997 (2013).
[42] F. C. Santos and J. M. Pacheco, Phys. Rev. Lett. **95**, 098104 (2005).
[43] J. Gómez-Gardeñes, M. Campillo, L. M. Floría, and Y. Moreno, Phys. Rev. Lett. **98**, 108103 (2007).
[44] P. Erdős and A. Rényi, Publ. Math. Inst. Hung. Acad. Sci. **5**, 17 (1960).
[45] A. L. Barabási and R. Albert, Science **286**, 509 (1999).
[46] L. E. Blume, Games Econ. Behav. **5**, 387 (1993).
[47] G. Szabó and C. Töke, Phys. Rev. E **58**, 69 (1998).

Chapter 6

Emergence of network features from multiplexity

SCIENTIFIC REPORTS

OPEN

SUBJECT AREAS:
STATISTICAL PHYSICS, THERMODYNAMICS AND NONLINEAR DYNAMICS
AEROSPACE ENGINEERING
APPLIED MATHEMATICS
APPLIED PHYSICS

Received
10 December 2012

Accepted
14 February 2013

Published
27 February 2013

Correspondence and requests for materials should be addressed to J.G.-G. (gardenes@gmail.com)

Emergence of network features from multiplexity

Alessio Cardillo[1,2], Jesús Gómez-Gardeñes[1,2], Massimiliano Zanin[3,4,5], Miguel Romance[3,6], David Papo[3], Francisco del Pozo[3] & Stefano Boccaletti[3]

[1]Institute for Biocomputation and Physics of Complex Systems (BIFI), Universidad de Zaragoza, E-50018 Zaragoza, Spain, [2]Departamento de Física de Materia Condensada, Universidad de Zaragoza, E-50009 Zaragoza, Spain, [3]Center for Biomedical Technology, Universidad Politécnica de Madrid, E-28223 Pozuelo de Alarcón, Madrid, Spain, [4]Innaxis Foundation & Research Institute, José Ortega y Gasset 20, 28006, Madrid, Spain, [5]Faculdade de Ciências e Tecnologia, Departamento de Engenharia Electrotécnica, Universidade Nova de Lisboa, Portugal, [6]Departmento de Matemática Aplicada, Universidad Rey Juan Carlos, E-28933 Móstoles, Madrid, Spain.

Many biological and man-made networked systems are characterized by the simultaneous presence of different sub-networks organized in separate layers, with links and nodes of qualitatively different types. While during the past few years theoretical studies have examined a variety of structural features of complex networks, the outstanding question is whether such features are characterizing all single layers, or rather emerge as a result of coarse-graining, i.e. when going from the multilayered to the aggregate network representation. Here we address this issue with the help of real data. We analyze the structural properties of an intrinsically multilayered real network, the European Air Transportation Multiplex Network in which each commercial airline defines a network layer. We examine how several structural measures evolve as layers are progressively merged together. In particular, we discuss how the topology of each layer affects the emergence of structural properties in the aggregate network.

In the past fifteen years, network theory[1–3] has successfully characterized the interaction among the constituents of a variety of complex systems[4,5], ranging from biological[6] to technological[7], and social[8] systems. However, up until recently, attention was almost exclusively given to networks in which all components were treated on equivalent footing, while neglecting all the extra information about the temporal- or context-related properties of the interactions under study. Only in the last three years, taking advantage of the enhanced resolution in real data sets, network scientists have directed their attention to the multiplex character of real-world systems, and explicitly considered the time-varying[9–14] and multi-layered[15–26] nature of networks.

A paradigmatic example of intrinsically multiplex system is represented by the Air Transportation Network (ATN). The ATNs have undergone a very significant growth during the last decades, giving rise to the dense and redundant system we know nowadays[27]. In the ATN, nodes represent airports, while links stand for direct flights between two airports. On the other hand, each commercial airline corresponds to a different layer, containing all the connections operated by the same company. While a considerable effort has recently been devoted to the characterization of the structural properties[28–30] of ATNs and their role in the dynamical processes taking place on them[31–34], their multiplex nature has remained almost unexplored.

When studying systems that can be represented as a graph made of diverse relationships (layers) between its constituents, an important question, typical of complex systems analysis, arises: can the topological properties of the whole system be traced to those of its layers or do they *emerge* from the simultaneous presence of multiple layers? Emergence is said to happen when the focus is switched from one scale to a coarser level of description. This question can be addressed by comparing the most usual structural properties of the multiple layers composing a network[35] and their analogue in the aggregate representation of the network, in which the layer structure is disregarded.

To address the above question we resort to the European ATN data set. Taking advantage of the high-resolution of these data, comprising a number of airlines (layers) operating in Europe during the year 2011, we succeed to extract the multiplex character of the system, and we investigate how the structural properties usually observed in the ATN are here emerging as a result of progressive layer merging. To this end, we quantify various topological measures, such as the degree distribution, the clustering coefficient or the presence of rich-club effect, in networks obtained by merging together a growing number of layers, from the lowest level of resolution of a single layer, up to the fully *aggregate* network. In addition, we compare two different types of

www.nature.com/scientificreports

layers, those corresponding to major (national) airlines and those labeled as low-cost companies. We analyze their structural differences, and their different contribution to the properties of the global ATN.

Results

The European ATN can be represented as a graph composed of $M = 37$ different layers each representing a different European airline (see Methods for details). Each layer m has the same number of nodes, N, as all European airports are represented in each layer. Furthermore, the data set allows extracting two main subsets, comprising all major, and low-cost airlines, with 18 and 10 layers respectively (See Fig. 1). In particular, panels **(a)** and **(b)** display the structure of the aggregate network focusing first on its *redundancy* by sketching those links belonging to more than one layer and on its *unicity* by reporting those links that only exist in a specific layer. Panels **(c)** and **(d)** show, instead, the single-layer ATN corresponding to a given major and low-cost airlines, respectively. In each of the panels we highlighted the nodes with the highest number of connections.

Topological measures. To characterize the structural properties of both the aggregate ATN and its layers, we consider several features widely used in network literature[35], i.e. cumulative degree distribution $P_>(k)$, clustering coefficient C, size of the giant component S, average path length L and Rich-club coefficient R. We briefly describe below the specific meaning of each of these measures in our context. The interested reader will find a complete description of all those quantities in the Methods section.

- The cumulative degree distribution $P_>(k)$, gives the probability of finding a node with a number of connections (or degree) equal or greater than k. The degree distribution is a powerful tool which allows understanding both structural and dynamical characteristics of a system as, for instance, its tolerance to attacks or failures[36,37] so it represents a cornerstone in the characterization of critical infrastructures, such as the ATN.
- The average path length $\langle L\rangle$, measures the average number of hops one has to make to go from a node to another. In the context of ATNs, it indicates the *average number of flights* a passenger has to take to go from his/her origin to his/her destination. However, if the system is not connected, this quantity diverges and it is preferable to restrict attention to the giant (largest) component of the system (see below).
- The clustering coefficient C, measures the probability, $C \in [0, 1]$, that two nodes with a common neighbor are connected together. C is a typical measure in systems made of social acquaintances[8], but in our case it is useful to estimate the density of triangular motifs (denoting the possibility of performing round trips of length 3).
- The size of the giant component[38] S, denotes the largest fraction of overall nodes such that any pair of them is connected through a path of finite length. In our case, it estimates the largest *coverage* that a given airline (or a combination of them) provides in terms of the available destinations that a passenger can reach from an origin inside the giant component.
- The Rich-club coefficient[39] R, measures the tendency of highly connected nodes, *i.e.* the *hubs*, to be connected among themselves. To measure it, one has to compute the abundance of links, $\phi(k)$, among nodes with a number of connections equal or greater than a certain value k, and the maximum possible number of links among those nodes, $\phi(k)_{max}$. Then, the ratio between these two quantities gives the relative abundance of links among nodes with at least k connections. Finally, $R(k)$ is given by the ratio between the abundance of links in the real case $\phi(k)/\phi(k)_{max}$ and the same quantity calculated in a proper randomized version of the original network. Colizza *et al.*[28] measured R for the ATN, and found that world air transportation network displays indeed a Rich-club effect, *i.e.* for large values of k the value of $R(k)$ is larger than 1.

Emergence of topological properties of the European ATN. We now analyze the evolution of the former measures as more and more layers are merged (independently of whether they do correspond to major or low-cost companies), until the complete aggregate ATN, comprising all the available layers, is reached (see the Methods section for the details on the layer merging procedure). The results are shown in Fig. 2.

In panel **(a)** we show the evolution for the cumulative degree distribution of the aggregate ATN and those networks obtained by merging 1, 5 and 20 randomly chosen layers. Since right-skewed distributions often display high noise levels at the end of their tails due to the lack of statistics, it is convenient to consider the cumulative distribution instead of the distribution $P(k)$ itself[35]. A power-law behavior $P_>(k) \propto k^{-\alpha}$ is observed in all the situations considered, with a decrease in the exponent α, ranging from $\alpha = 1.84$ in the single layer case ($m = 1$) to $\alpha = 1.39$ for the aggregate ATN. The increase in heterogeneity with the number of layers considered points to a richer-gets-richer phenomenon different from the one seen in classical models for growing scale-free networks: while in the latter case, it results from the addition of new nodes, in the present case it emerges from the addition new layers.

In panel **(b)** we report the clustering coefficient. In this case, we show the behavior of $\langle C\rangle$ as a function of the number of layers used to construct the aggregate ATN, averaged over the number of different combinations of m elements ($m = 1, \ldots, M$). Interestingly, we see how the clustering suddenly increases as we merge just a few layers:

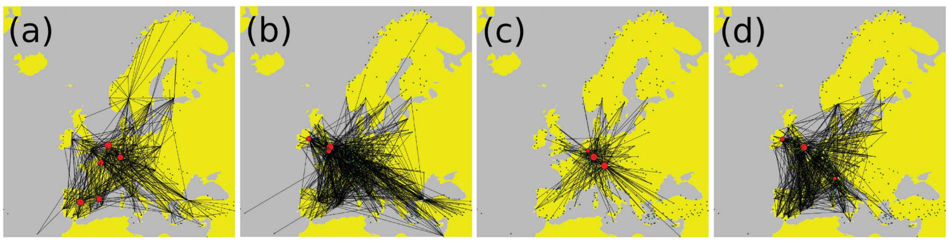


Figure 1 | Visual representation of the ATN. From left to right: the aggregate network of all the layers in which only links belonging to more than one layer are displayed. The same network but in which we display those links which belongs to only one layer and connecting at least one node with degree greater than or equal to 75. An example of ATN network of a major airline and, finally, the network of a low-fare (low-cost) airline. In each network, the airports with the highest degree are highlighted.

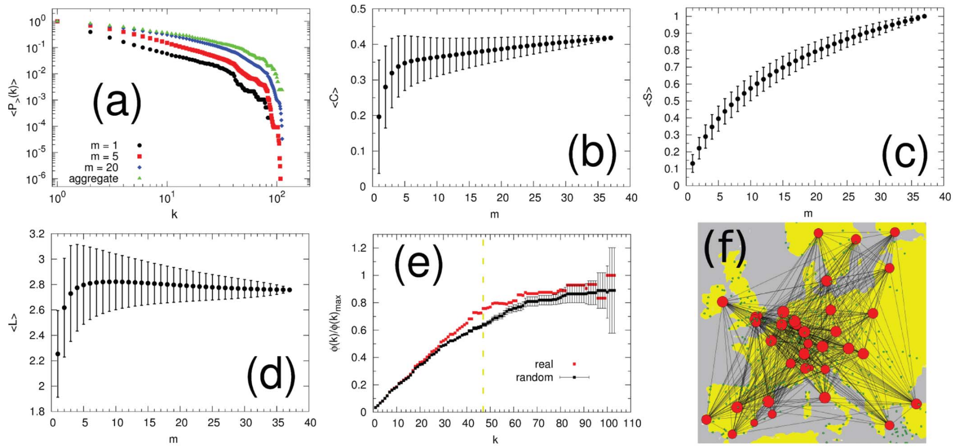


Figure 2 | Evolution of topological properties of the complete ATN network. (a) Average cumulative degree distribution $P_>(k)$ for groups of layers merged together: single layers (●), five layers (■), twenty layers (◆) and the aggregate (▲). (b–c–d) Average clustering $\langle C \rangle$, size of giant component $\langle S \rangle$, path length $\langle L \rangle$ as a function of m. (e) Link abundance for nodes of degree k or greater, $\phi(k)$ divided by its maximum $\phi(k)_{max}$ for the aggregate network in both real case (■) and its randomized version (●). The vertical dashed line represent the value of k at which the difference among the two curves is maximal. (f) A subset of the aggregate network showing the connections among those nodes whose degree is greater than (or equal to) 47. The size of the nodes is proportional to the degree.

to achieve more than 80% of the final clustering value, we only need to randomly merge together five layers. This result indicates that the large density of triangles present in the ATN is a consequence of the merging of different layers rather than a single-layer property. Thus, in order to make round trips of length 3 one should make use, most of the times, of more than one airline.

The former result contrasts with the picture obtained for the evolution of the size of the giant component $\langle S \rangle$. Panel **(c)** describes a monotonous and progressive increase of the *coverage* as more layers are aggregated. In fact, around 40% of the European cities are covered when merging together five randomly chosen layers. It is worth noticing that $\langle S \rangle$ also tells us that we are considering a system which is already above the percolation threshold, so that every step towards the aggregate network produces an increment in the collection of reachable destinations (see the value of $\langle S \rangle$ for $m = 1$). However, the behavior of the transition for the average path length $\langle L \rangle$ (restricted to those nodes in the giant component) in panel **(d)** shows a rise-and-fall behavior indicating that combining few layers results in the merging of unconnected components at the aggregate level, causing a fast increase in its length. On the other hand, after the maximum for L is reached, the addition of new layers has a twofold effect on the giant components: it incorporates new nodes, but also creates alternative links between already present nodes. Thus, the average path length of the giant component balances the addition of new destinations with the creation of new links, and suffers a slow decrease when increasing m.

Finally, panel **(e)** shows, only for the aggregate network, the existence of a Rich-club effect quantifying the abundance of links between nodes with degree larger or equal to k, $\phi(k)$, normalized with respect to its maximum. This quantity is computed both for the real ATN and for a set of randomized versions of the network in which all the links are rewired keeping the same degree sequence of the original network. This randomization aims at destroying any kind of correlation between the local properties of connected nodes. From the figure it is clear that initially the two curves coincide indicating that the existence of flights between airports with few connections (less than $k = 30$) is equally probable in the ATN and in its randomized version. Instead, for $k \in [30,60]$ the points corresponding to the real ATN stand above those corresponding to the randomized network. This result points out that the aggregate ATN displays Rich-club effect (the largest effect being found for $k = 47$), thus confirming for the European case the findings of Colizza *et al.*[28] for the ATN. The existence of such effect is quite logical, as usually highly connected nodes correspond to the principal airports of the main European cities which, in most of the cases, are connected among themselves via direct flights. Finally, for $k > 60$ the fluctuations of the randomized case are too large for any statement to be made on the existence of a Rich-club effect.

Major *versus* low-cost layers. The European ATN is composed of layers corresponding to airlines of different types. In particular, we find among them *major* (national, such as Lufthansa), *low-cost* fares (such as Easyjet), *regional* (such as Norwegian Air Shuttle) or *cargo* (such as Fed-Ex) airlines. These kinds of airlines have developed according to different structural/commercial constraints. For instance, it is known that major airlines are designed following the so-called *hub and spoke* structure, to provide an almost complete coverage of the airports belonging to a given country[40,41] and maximize efficiency in terms of national transportation interests. Low-cost companies, instead, tend to avoid overly centralized structures and, to be more competitive, typically cover more than one country simultaneously. To unveil the role that each type of airline plays in the emergence of the topological features of the aggregate ATN, we considered two subsets of layers respectively comprising only it majors and *low-cost* airlines. The results of this study are shown in Fig. 3.

We first address the cumulative degree distribution $P_>(k)$. In the two panels **(a)** and **(b)** we show the distributions $P_>(k)$ for major **(a)** and low-cost **(b)** layers when considering different levels for the merging of the layers of the same kind. For major airlines, the typical

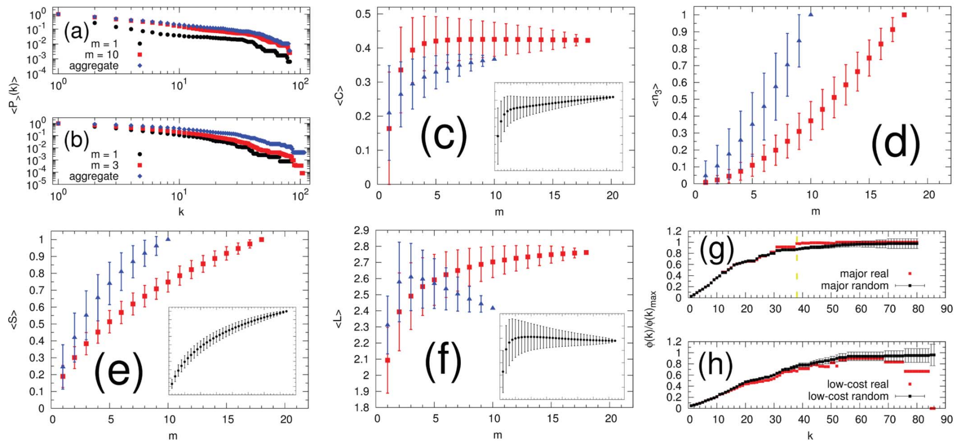


Figure 3 | Evolution of topological properties of major (■) and low-cost (▲) subsets. (a–b) Average cumulative degree distribution $P_{>}(k)$ for different number of layers merged together. (c–d–e–f) Average clustering $\langle C \rangle$, number of triangles $\langle n_3 \rangle$, size of giant component $\langle S \rangle$, path length $\langle L \rangle$ as a function of the number of layers merged. The insets display the same quantities in the case of the complete set. (g–h) link abundance for the aggregate network. The vertical dashed line represents the value of k at which the difference among the two curves is maximal.

trend of a single layer ($m = 1$) displays a plateau for moderate values of k, indicating a centralized character of this kind of layers, with few hubs having remarkably higher than average connectivity. In addition, when merging more layers ($m = 10$ or all the major airlines) the trend shows a rather continuous decay due to the combination of hubs of different size (depending on the nation of the airline). Notice that a hub of a single layer (a single national airline) is highly connected within the same country, but also has some flights to capitals of other European countries which, in turn, are hubs of their corresponding major layers. On the other hand, the cumulative distribution of typical low-cost airlines shows a rather different pattern, as its decay is rather progressive, and airports of different size coexist within the same layer.

The differences in organization of low-cost and major airlines is further highlighted by the behavior of the clustering coefficient $\langle C \rangle$. Panel **(c)** shows how major airlines display sharp increases in $\langle C \rangle$ as more major layers are merged, followed by a plateau for $m > 5$. This saturation of C is due to the fact that, when merging major layers randomly, national hubs tend to connect together (we have already discussed this fact when introducing the Rich-club effect) in the aggregate network. The saturation of clustering is, however, not observed for the aggregate ATN [see Fig. 2.**(b)** or the inset in panel **(c)**] for which $C(m)$ always increases. This is due to the fact that the merging of low-cost layers leads to a continuous formation of new triangles, thus increasing the clustering with m. In addition, in panel **(d)** we show the evolution with m of the average number of triangles, $\langle n_3 \rangle$, normalized with respect to the total number of triads in the aggregate network for both major and low-cost layers. Interestingly, the monotonic growth of $\langle n_3 \rangle$ reveals that the saturation of the clustering coefficient when $m = 5$ for major layers is not due to the fact that new triangles are not added when $m > 5$ but to a balance between the new triads and the new connections added when merging additional layers.

The behavior of the giant component $\langle S \rangle$, normalized with respect to the total number of destinations covered by each kind of airline (see panel **(e)**), does not give any particular insight in terms of differences between low-cost and major airlines, except for the fact that in the low-cost case we observe larger fluctuations, mainly due to the large variability in size of the giant component of single layers. On the other hand, the picture described by the average path length $\langle L \rangle$ in panel **(f)** is very interesting. Major and low-cost subsets behave rather differently not only between them, but also with respect to the evolution of the complete set (see inset). For layers corresponding to major companies, $\langle L \rangle$ increases with the number of merged layers. The interpretation of this continuous growth is straightforward: each time a layer corresponding to a major airline is added, even if it shares some common destinations (say some European capitals having their corresponding major airlines within the original set of merged layers), the number of new available nodes (small destinations only available through the new added major layer) is large enough to generate an increase in L. On the contrary, the case of low-cost displays a rise-and-fall in the behavior of $\langle L \rangle$, due to the large coverage of European countries/cities that already each single low-cost layer displays. Thus, as we merge some of them together, they already cover nearly all the low-cost destinations, and merging of additional layers just adds new connections between them. When combined into the original ATN, these two different trends lead to the saturated evolution of $\langle L \rangle(m)$ shown in the inset.

Finally, we examine once again the onset of the Rich-club effect. From panels **(g)** and **(h)** we notice how the graph corresponding to the aggregate network constructed by merging layers corresponding to major airlines **(g)** displays the presence of a rich club for $k = 38$ (almost the same value as in the case of the total aggregate ATN). Interestingly, the Rich-club effect is absent when merging low-cost layers so that, while in the case of major airlines the merge of layers containing large hubs ends up in a system composed of a connected core of highly connected nodes, the more distributed nature of the low-cost layers prevents the formation of a Rich-club. Thus, a relevant conclusion is that the well-known[28] Rich-club effect observed in ATNs is *exclusively* related to the presence of major airlines.

Discussion

The characterization of the interaction patterns in large systems has recently been spurred by the incorporation of the paradigm of multiplexity. Taking advantage of the European ATN data set, with details of the airlines operating each flight, we showed that the topological properties of the ATN are generally not present in single layers, rather they are the consequence of an emerging phenomenon intimately related to the multilayer character of the system. We also pointed out that the merging of low-cost and major (national) layers leads to the emergence of qualitatively different aggregate networks. Finally, we demonstrated that the combination of these two different behaviors accounts for the many important structural features of the global ATN, such as the Rich-club effect (mainly due to the layers of major airlines), path redundancy (resulting from a cooperative combination of the clustering of low-cost and major layers), or small-worldness (remarkably enhanced by the presence low-cost layers).

Our study highlights the importance of considering the multiplex character of most real networked systems, and shows that considering layers as relevant entities of a network (such as nodes and links at the micro-scale or communities at the meso-scale) will contribute to a better understanding and modeling of dynamical processes taking place at the level of aggregate network.

Methods

Dataset. The data analyzed in this paper are taken from the complete list of airlines operating Instrumental Flight Rules (IFR) flights between European airports on a certain day obtained from EUROCON-TROL and the Complex World Network in the context of the SESAR Work Package E[42]. We selected only those airlines whose number of *destinations* is above the average (which is 32), obtaining $\mathcal{L}=37$ different airlines (layers), that include both *major* companies (like Lufthansa or Air France), and *low-fares* (low-cost) companies (as Ryanair or Easyjet). Each layer ℓ in this multiplex representation is a graph $G^\ell=\left(\mathcal{N}^\ell,\mathcal{E}^\ell\right)=(\mathcal{N},\mathcal{E}^\ell)$ with $\mathcal{N}^\ell=\mathcal{N}=450$ nodes and K^ℓ links that models a single airline. An example of such networks is shown in Fig. 1. The ensemble of all these layers constitutes our multilayer system, that we will call the *complete set*. We will also consider the subset of *major* airlines, that will be a multiplex network made of $\mathcal{L}'=18$ layers, and the subset of *low-cost* companies, with $\mathcal{L}''=10$ layers. Note that the remaining airlines, such as cargo airlines, constitutes a marginal small subset and therefore its analysis is residual.

Topological indexes. In this section, we present a summary of the topological measures used throughout the paper. Note that the considered topological measures are essentially defined for classic monoplex networks, and their extensions to the multiplex setting is an exercise, whose details are here shown.

One of the most basic topological parameter of a complex network $G=(\mathcal{N},\mathcal{E})$ is the *degree distribution* $P(k)$ which is defined as the probability that a node chosen uniformly at random has degree k, or equivalently the fraction of nodes in the network having degree k[3,35]. Since broad distributions often display high noise levels at the end of their tails, here related to the low abundance of highly connected nodes, it is convenient to consider the *cumulative distribution* $P_>(k)$. Cumulative distribution $P_>(k)$ is the probability that a randomly chosen node has a degree equal or greater than k, i.e.

$$P_>(k)=\frac{1}{N}\sum_{k'=k}^{\infty}N(k), \tag{1}$$

where $N(k)$ is the number of nodes with degree k and $N=|\mathcal{N}|$ is the total number of nodes in the network.

The *average path length*[4] $L(G)$ is the average length of the shortest paths among all the couples of nodes in the network, i.e.

$$L(G)=\frac{1}{N(N-1)}\sum_{i,j\in\mathcal{N}}d_{ij}, \tag{2}$$

where d_{ij} is the minimum number of hops one has to make to go from node i to node j in G (the *distance* from i to j). Note that this definition diverges if G is not connected, since d_{ij} may be infinite. One way to avoid this divergence is considering the average only on the largest connected component, and an alternative approach that has been shown very useful in many cases is considering the harmonic mean of the distances.

The (local) *clustering coefficient*[4] c_i of a node $i\in\mathcal{N}$ is defined as

$$c_i=\frac{2\,e_i}{k_i(k_i-1)}, \tag{3}$$

where e_i is the number of neighbors of i which are mutual neighbors, and k_i is the degree of node i. Therefore the (local) clustering coefficient of a node i is the ratio between the number of neighbors of i which are mutual neighbors and the maximal possible number of edges between neighbors of i. The (average) clustering coefficient C of a graph is the arithmetic mean of c_i over all its nodes.

The *giant component* $S(G)$ is the largest connected component of G and the *size of the giant component* is the proportion of nodes in the network that belong to the giant component, i.e.,

$$S(G)=\max_{i\in\mathcal{N}}\frac{N_i}{N} \tag{4}$$

where N_i is the number of nodes of the maximal connected subnetwork of G containing node i.

If we take a node with *degree* $0\le k\le|\mathcal{N}|$, the *Rich-club coefficient* $R(k)$[39] is given by

$$R(k)=\frac{\phi(k)}{\phi(k)_{max}}\left(\frac{\phi'(k)}{\phi'(k)_{max}}\right)^{-1}=\frac{2\phi(k)}{N_{>k}(N_{>k}-1)}\frac{N_{>k}(N_{>k}-1)}{2\phi'(k)}=\frac{\phi(k)}{\phi'(k)}, \tag{5}$$

where

(i) $\phi(k)$ is the number of edges connecting nodes of degree greater or equal to k (called the *link abundance*),

(ii) $\phi(k)_{max}$ is the maximum number of links that can exist between nodes of degree k,

(iii) $\phi'(k)$ is the *link abundance* on a network with the same degree sequence of the original but with connections randomly shuffled.

(iv) $\phi'(k)_{max}$ is the maximum number of links that can exist between nodes of degree k on a network with the same degree sequence of the original but with connections randomly shuffled.

(v) $N_{>k}$ is the number of nodes with degree greater or equal to k.

If, for a certain value of k, $R(k)>1$ for some $0\le k\le N=|\mathcal{N}|$, then we say that G has a *Rich-club*. Note that in the plots presented in this paper, we decided to present the ratios $\phi(k)/\phi(k)_{max}$ and $\phi'(k)/\phi'(k)_{max}$ instead of $R(k)$. The randomization, in our case, is repeated 1,000 times, while the shuffling is repeated 10,000 times to ensure a robust statistical sampling. Note that for the ATN network, having a size of $N=450$ nodes, the number of random shuffling steps is large enough to guarantee that the resulting network is fully randomized. This randomization method is known as *Markov Chain Monte Carlo Algorithm*[43]. However, for bigger graphs other methods are recommended so to minimize the computation cost for producing reliable randomized networks, see the work by Del Genio *et al.*[44].

Next, we describe the *layer merging* procedure used to study the evolution of the topological measures and the behavior of the layers in the major airline and low-cost multiplex sub-networks.

If we fix a subset of layers $\left\{G^\ell;\ \ell=\ell_1,\cdots,\ell_m\right\}$ to merge together, we construct a monoplex network $G'=(\mathcal{N},\mathcal{E}')$ (i.e. a classic network with only one layer) given by

$$G'=\bigcup_{j=1}^{m}G^{\ell_j}.$$

This network G' is obtained by projecting all the m layers onto one and by converting multiple links into single ones.

Now if we fix m, we look for all the possible mergings of m layers, The number of different configurations to arrange n layers into groups of size m without repetitions is given by $C_m^n=\binom{n}{m}$, therefore if we want to compute a topological measure on the ensemble of m layers, we should first compute it on each of the C_m^n mergings, and then average over all C_m^n possible configurations. However, when the number of possible configurations exceeded a certain threshold, we operated a random sampling over 500,000 mergings in order to avoid the growth of the computation time. Throughout the paper the operator $\langle\cdot\rangle$ denotes the average over the elements of the ensemble. As an example, if we want to compute the clustering coefficient over an ensemble $\mathcal{C}$, we compute:

$$\langle C\rangle=\frac{1}{N_{comb}}\sum_{i\in\mathcal{C}}C_i, \tag{6}$$

where N_{comb} is the number of elements of $\mathcal{C}$ and C_i is the average clustering of the network obtained merging together the layers corresponding to $i\in\mathcal{C}$.

1. Albert, R. & Barabási, A.-L. Statistical mechanics of complex networks. *Rev. Mod. Phys.* **74**, 47–97 (2002).
2. Newman, M. E. J. The Structure and Function of Complex Networks. *SIAM Rev.* **45**, 167–256 (2003).
3. Boccaletti, S., Latora, V., Moreno, Y., Chavez, M. & Hwang, D. Complex networks: Structure and dynamics. *Physics Reports* **424**, 175–308 (2006).
4. Watts, D. J. & Strogatz, S. H. Collective dynamics of small-world networks. *Nature* **393**, 440–442 (1998).
5. Strogatz, S. H. Exploring complex networks. *Nature* **410**, 268–276 (2001).
6. Milo, R., Shen-Orr, S., Itzkovitz, S., Kashtan, N., Chklovskii, D. & Alon, U. Network motifs: simple building blocks of complex networks. *Science* **298**, 824–827 (2002).
7. Albert, R., Jeong, H. & Barabási, A.-L. Diameter of the World-Wide Web. *Nature* **401**, 398–399 (1999).
8. Wasserman, S. & Faust, K. Social Network analysis. *Cambridge University Press* Cambridge (1994).
9. Onnela, J.-P., Saramaki, J., Hyvonen, J., Szabó, G., Lazer, D., Kaski, K., Kertesz, J. & Barabási, A.-L. Structure and tie strengths in mobile communication networks. *Proc. Nat. Acad. Sci. U.S.A.* **104**, 7332–7336 (2007).
10. Tang, J., Scellato, S., Musolesi, M., Mascolo, C. & Latora, V. Small-world behavior in time-varying graphs. *Phys. Rev. E* **81**, 055101R) (2010).

11. Zhao, K., Stehlé, J., Bianconi, G. & Barrat, A. Social network dynamics of face-to-face interactions. *Phys. Rev. E* **83**, 056109 (2011).
12. Isella, L., Stehlé, J., Barrat, A., Cattuto, C., Pinton, J.-F. & Van den Broeck, W. What's in a crowd? Analysis of face-to-face behavioral networks. *J. Theor. Biol.* **271**, 166–180 (2011).
13. Mucha, P. J., Richardson, T., Macon, K., Porter, M. A. & Onnela, J.-P. Community structure in time-dependent, multiscale, and multiplex networks. *Science* **328**, 876–878 (2010).
14. Szell, M., Lambiotte, R. & Thurner, S. Multi-relational Organization of Large-scale Social Networks in an Online World. *Proc. Natl. Acad. Sci. (U.S.A.)* **107**, 13636–13641 (2010).
15. Kurant, M. & Thiran, P. Layered complex networks. *Phys. Rev. Lett.* **96**, 138701 (2006).
16. Buldyrev, S. V., Parshani, R., Paul, G., Stanley, H. E. & Havlin, S. Catastrophic cascade of failures in interdependent networks. *Nature* **464**, 1025–1028 (2010).
17. Parshani, R., Buldyrev, S. & Havlin, S. Critical effect of dependency groups on the function of networks. *Proc. Natl. Acad. Sci. (U.S.A.)* **108**, 1007–1010 (2011).
18. Gao, J., Buldyrev, S. V., Stanley, H. E. & Havlin, S. Networks formed from interdependent networks. *Nature Phys.* **8**, 40–48 (2012).
19. Lee, K.-M., Kim, J. Y., Cho, W.-K., Goh, K.-I. & Kim, I.-M. Correlated multiplexity and connectivity of multiplex random networks. *New J. Phys.* **14**, 033027 (2012).
20. Brummitt, C. D., D'Souza, R. M. & Leicht, E. A. Suppressing casades of load in interdependent networks. *Proc. Natl. Acad. Sci. (U.S.A.)* **109**, E680–E689 (2012).
21. Brummitt, C. D., Lee, K.-M. & Goh, K.-I. Multiplexity-facilitated cascades in networks. *Phys. Rev. E* **85**, 045102(R) (2012).
22. Gómez-Gardeñes, J., Reinares, I., Arenas, A. & Floría, L. M. Evolution of cooperation in Multiplex Networks. *Scientific Rep.* **2**, 620 (2012).
23. Gómez, S., Diaz-Guilera, A., Gómez-Gardeñes, J., Perez-Vicente, C., Moreno, Y. & Arenas, A. Diffusion Dynamics on Multiplex Networks. *Phys. Rev. Lett.* **110**, 028701 (2013).
24. Criado, R., Flores, J., García del Amo, A., Gómez-Gardeñes, J. & Romance, J. A mathematical model for networks with structures in the mesoscale. *Int. J. Comp. Math.* **89** (3), 291–309 (2012).
25. Ronhovde, P. & Nussinov, Z. Multiresolution community detection for megascale networks by information-based replica correlations. *Phys. Rev. E* **80**, 016109 (2009).
26. Ronhovde, P., Chakrabarty, S., Hu, D., Sahu, M., Sahu, K. K., Kelton, K. F., Mauro, N. A. & Nussinov, Z. Detecting hidden spatial and spatio-temporal structures in glasses and complex physical systems by multiresolution network clustering. *Eur. Phys. J. E* **34**, 105 (2011).
27. EUROCONTROL, Long-Term Forecast of Annual Number of IFR Flights (2010–2030) (2010). Available at: http://www.eurocontrol.int/articles/forecasts. Accessed on June 2012.
28. Colizza, V., Flammini, A., Serrano, M. A. & Vespignani, A. Detecting rich-club ordering in complex networks. *Nature Physics* **2**, 110–115 (2006).
29. Guimera, R., Mossa, S., Turtschi, A. & Amaral, L. A. N. The worldwide air transportation network: Anomalous centrality, community structure, and cities' global roles. *Proc. Nat. Acad. Sci. (U.S.A.)* **102**, 7794–7799 (2005).
30. Wuellner, D., Roy, S. & D'Souza, R. M. Resilience and rewiring of the passenger airline networks in the United States. *Phys. Rev. E* **82**, 056101 (2010).
31. Colizza, V., Pastor-Satorras, R. & Vespignani, A. Reaction-diffusion processes and metapopulation models in heterogeneous networks. *Nature Phys.* **3**, 276–282 (2007).
32. Colizza, V. & Vespignani, A. Epidemic modeling in metapopulation systems with heterogeneous coupling pattern: Theory and simulations. *J. Theor. Biol.* **251**, 450–467 (2008).
33. Lacasa, L., Cea, M. & Zanin, M. Jamming transition in air transportation networks. *Physica A* **388**, 3948–3954 (2009).
34. Meloni, S., Perra, N., Arenas, A., Gómez, S., Moreno, Y. & Vespignani, A. Modeling human mobility responses to the large-scale spreading of infectious diseases. *Scientific Rep.* **1**, 62 (2011).
35. Newman, M. Networks: An Introduction. *Oxford University Press* (2010).
36. Cohen, R., Erez, K., ben Avraham, D. & Havlin, S. Resilience of the Internet to Random Breakdowns. *Phys. Rev. Lett.* **85**, 4626 (2000).
37. Cohen, R., Erez, K., ben Avraham, D. & Havlin, S. Breakdown of the Internet under Intentional Attack. *Phys. Rev. Lett.* **86**, 3682 (2001).
38. Stauffer, D. & Aharony, A. Introduction to percolation theory. *Taylor & Francis*, London, UK (1994).
39. Zhou, S. & Mondragon, R. J. The rich-club phenomenon in the Internet topology. *IEEE Commun. Lett.* **8**, 180–182 (2004).
40. Barthélemy, M. Spatial networks. *Physics Reports* **499**, 1–101 (2011).
41. Bryan, D. L. & O'kelly, M. E. Hub-and-spoke networks in air transportation: an analytical review. *J. Reg. Sci.* **39**, 275–295 (1999).
42. SESAR and WP-E. Available at: http://www.complexworld.eu/sesar-and-wp-e/. Accessed on February 2012.
43. Blitzstein, J. & Diaconis, P. A sequential importance sampling algorithm for generating random graphs with prescribed degrees. *Internet Math.* **6**, 489 (2011).
44. Del Genio, C. I., Kim, H., Toroczkai, Z. & Bassler, K. E. Efficient and exact sampling of simple graphs with given arbitrary degree sequence. *PloS one* **5**, e10012 (2010).

Acknowledgments

This work has been partially supported by the Spanish MINECO under projects FIS2011-25167 and FIS2012-38266-C02-01; European FET project MULTIPLEX (317532); and by the Comunidad de Aragon (FENOL group). J.G.G is supported by MINECO through the Ramon y Cajal program.

Author contributions

A.C., J.G.G., M.Z., M.R. and S.B. devised the model and designed the study. A.C. carried out the numerical simulations. A.C., J.G.G., F.d.P. and M.Z. analyzed the data and prepared the figures. A.C., J.G.G., D.P. and S.B. wrote the main text of the manuscript.

Additional information

Competing financial interests: The authors declare no competing financial interests.

How to cite this article: Cardillo, A. *et al.* Emergence of network features from multiplexity. *Sci. Rep.* **3**, 1344; DOI:10.1038/srep01344 (2013).

Chapter 7

Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures

Eur. Phys. J. Special Topics **215**, 23–33 (2013)
© EDP Sciences, Springer-Verlag 2013
DOI: 10.1140/epjst/e2013-01712-8

THE EUROPEAN PHYSICAL JOURNAL SPECIAL TOPICS

Regular Article

Modeling the multi-layer nature of the European Air Transport Network: Resilience and passengers re-scheduling under random failures

Alessio Cardillo[1,2,a], Massimiliano Zanin[3,4,5,b], Jesús Gómez-Gardeñes[1,2,c], Miguel Romance[3,6,d], Alejandro J. García del Amo[3,6,e], and Stefano Boccaletti[3,f]

[1] Department of Condensed Matter Physics, University of Zaragoza, Spain
[2] Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Spain
[3] Center for Biomedical Technology (CTB), Technical University of Madrid, Spain
[4] INNAXIS Foundation and Research Institute, Madrid, Spain
[5] Departamento de Engenharia Electrotécnica, Universidade Nova de Lisboa, Portugal
[6] Department of Applied Mathematics, Rey Juan Carlos University, Madrid, Spain

Received 30 July 2012 / Received in final form 29 November 2012
Published online 29 January 2013

Abstract. We study the dynamics of the European Air Transport Network by using a multiplex network formalism. We will consider the set of flights of each airline as an interdependent network and we analyze the resilience of the system against random flight failures in the passenger's rescheduling problem. A comparison between the single-plex approach and the corresponding multiplex one is presented illustrating that the multiplexity strongly affects the robustness of the European Air Network.

1 Introduction

In the last century, the application of aeronautics to the transportation of people and goods has witnessed an uninterrupted growth [1]. In less than a hundred years we have moved from a sparsely connected system, to a redundant one capable of moving 2.7 billion passengers in 2011. During the last decade, scientists have studied the properties of airline transportation systems by means of network theory, unveiling their structural characteristics as done with other natural and technological complex networks. Along this period, complex networks [2,3] have extensively been

[a] e-mail: `alessio.cardillo@ct.infn.it`
[b] e-mail: `massimiliano.zanin@ctb.upm.es`
[c] e-mail: `gardenes@gmail.com`
[d] e-mail: `miguel.romance@urjc.es`
[e] e-mail: `alejandro.garciadelamo@urjc.es`
[f] e-mail: `stefano.boccaletti@fi.isc.cnr.it`

used to model and understand the structures of relations beyond many real-world systems [4,5], but only recently some limitation of this approach have been highlighted. One of the most important limitations refers to the multi-layer nature of real-world systems: nodes usually belong to different *layers* at the same time, and may have different neighbourhoods depending on the layer being considered. It is clear that nodes, in some complex systems, often have interactions of different kinds, which take place upon several interacting networks, i.e. constituting a so-called multiplex network. An example of this concept is represented by social networks [6]. Traditionally, social networks have been modelled as simple graphs; yet, it should be noticed that each node, representing an actor in the social network, may have different types of connections with other nodes, such as friendship, professional relationships etc. For such kind of systems, a multiplex model represents better the real situation, as it can better catch the different dynamics developing in each layer: for instance, usually the information transmitted to friends will not be the same as the one shared with colleagues. Therefore, in order to understand how the structure is affecting the global dynamics of a system, it is of utmost importance to take into account the presence of interactions at multiple levels [7,8]. There are other concepts strongly related to the multiplex networks that have recently been introduced in the literature, such as interacting [9,10], interdependent [11] and multilevel networks [12].

Recently, several works have focused on the vulnerability of networks to cascading failures, and especially how a multi-layer structure effectively reduces the resilience of the system. For instance, ref. [13] analyzes different communication and transportation networks, composed of two layers: a physical and a logical network, the latter representing the flows of information and people. In ref. [11], the Italian power grid and the Internet network are modeled as a single dual-layer system; the interconnections between both layers drammatically increase the vulnerability of the system, as a failure of a node may propagate to the other layer and generate a cascade dynamics. In ref. [14], a generalization of the threshold cascade model is studied, in which nodes are deactivated too if at least a given fraction of the neighbors have been deactivated. The generalization consists in introducing a multi-layer structure, that was not considered as part of the original model [15]: thanks to that, some topologies that were initially stable generate cascade dynamics when connected in a multi-layer paradigm.

In this contribution, we tackle the problem of the resilience of the Air Transport Network (ATN) from a multi-layer point of view, against the deletion of a connection, that is, the cancellation of a flight. The ATN is clearly one of the tenets of our societies. In 2010, the global air transport dealt with 2.4 billion passengers and 43 million tonnes of cargo, has been responsible for 32 million jobs, 2% of global carbon emissions and $545 billion in revenue [16]. It embraces the whole world and tightly links together the different regions, with all their individual differences.

The importance of the ATN is especially relevant when its dynamics is disturbed by external events; even when these events have only a local impact, like, for instance, a thunderstorm that forces the cancellation of a few flights, the indirect consequences (in terms of delays, passengers loosing connections, and so forth) may affect the overall performance of the system. This situation is expected to worsen in the future, as forecasted growth rates (about 5% per year [17], with crises, like the WTC attack, SARS or the financial crisis [18], only having a temporary impact) will imply a tightening of the room for manoeuvre available to cope with such disturbances. The relevance of the resilience of ATN has recently been recognized in the policy-making context, as, for instance, in the European Commission's new roadmap (White Paper) to a Single European Transport Area for 2050 [19,20].

The dynamics and resilience of the ATN has already been studied in the past by considering the usual single layer network formalism in which all the connections between airports are considered to be equivalent [21,22]. Yet, a study of ATN under

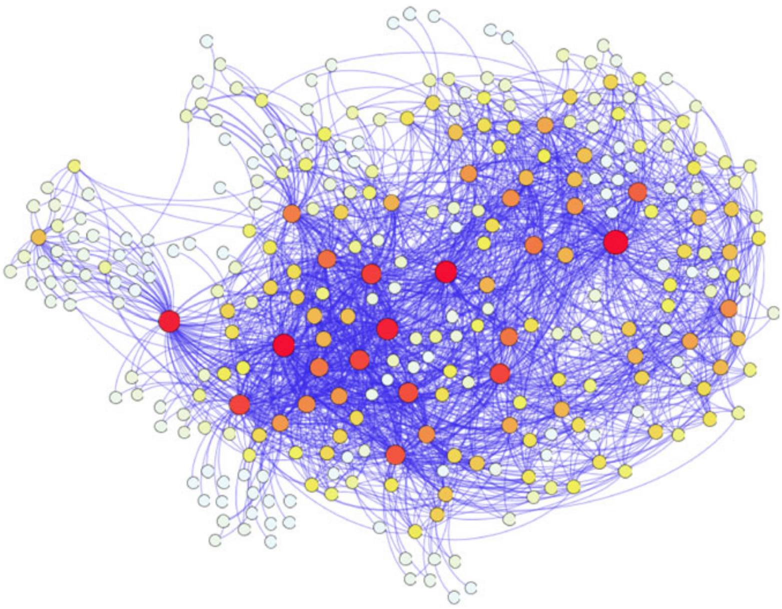


Fig. 1. The European Air Transport Network (ATN). The network has been constructed by considering only commercial (both regular and charter) flights operated between two European airports the 1st of June 2011. Size and color of nodes accounts for their degree.

the multi-layer approach is still missing. The intrinsic multi-layer nature of ATN is validated by the fact that passengers cannot use all the possible sequences of links between the airports bypassing the cost associated to the use of different airlines. This may negatively affect the resilience of the system, as well as the tools available to the system to reduce the impact of failures on the flow passengers. This demands for a study in which network and transportation sciences tackle the influence that the multi-layer architecture of the ATN has in its robustness under random failures.

2 The European ATN as a multilayer network

We start by describing the structural multiplex backbone of the European Air Transport Network (ATN) considered in our model. We consider a set of 15 layers, each of which representing one of the 15 biggest airline companies operating in Europe. In each layer ℓ (that represents an airline A), the set of nodes corresponds to the set of airports operated by the airline A and the links (denoted by $(i, j; \ell)$) are the flights between the airports i and j that are operated by airline A. Data corresponds to commercial IFR (Instrumental Flight Rules) operations for the 1st of June 2011. The resulting multi-layer network (see Fig. 1) is an undirected system, composed of 15 layers and 308 nodes, corresponding to the 20% of the operations in the European airspace.

Looking at the structural properties of the different layers, we realize that they are organized in two main families: (1) networks corresponding to *major* airlines (such as Lufthansa, Air France, or Iberia), being scale-free with hubs representing the airline headquarters; and (2) networks corresponding to the so-called *low-cost* (or *low-fares*)

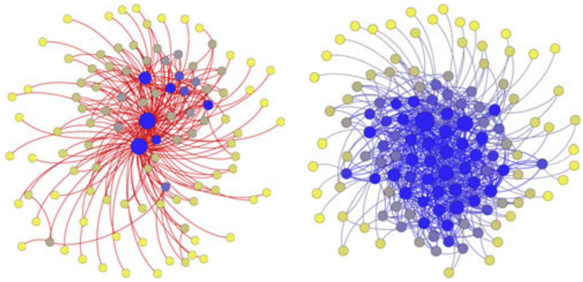


Fig. 2. The Air Transport Networks of a traditional major company (on the left) and of a low-cost airline (on the right). The hubs of each network are indicated by blue big circles.

airlines, showing a more uniform structure due to a *point-to-point* organization of their business [23]. Figure 2 illustrates two of such layers: a traditional major company on the left, and of a low-cost company on the right. In this figure the hubs of each network are indicated by blue big circles; notice that the heterogeneity is much stronger in traditional companies.

The introduction of a multiplex-type network for the European ATN produces structural properties that differ from the corresponding single-mode network, i.e., the single layer projection of the transport network. For instance, we focus here on the global degree distribution $P(k^A)$ of the multi-layer network; the global degree of each node i (denoted by k_i^A) is calculated as the sum of the number of connections of that node over all the layers. Therefore, k_i^A is defined as:

$$k_i^A = \sum_{\ell=1}^{L} k_i^{\ell}, \tag{1}$$

where k_i^{ℓ} is the degree of node i in the layer ℓ. Figure 3 illustrates the cumulative probability distribution of degrees of the European ATN in log-log scale. Clearly, there are strong differences between the distribution for the multi-layer network model (top left panel) and the average of the cumulative degree distribution over all the layers considered (top right panel). Note that the degree of nodes in the case of the multi-layer model is greater than the corresponding degree in the classic one-layer approach. This phenomenon is even more explicit in the case of the hubs, since a link that could happen in different layers is counted as many times as it is present in the multi-layer network, while it is only counted only once in the classic model. Despite this fact, one could expect that this enhancement of the degree is uniform along the network, but the real situation is quite far from this. The heterogeneity of the structure and distribution of each layer makes that the effects of the enhancement of the degree of each node in the multi-layer network is very disperse and therefore the degree distribution in the multi-layer model is very different from the corresponding classic model. Furthermore, if we compute the average degree distribution along all the layers in the network (see the top right panel in Fig. 3), the result is quite different from the corresponding figure for the multi-layer model. Note that the average degree distribution illustrate the average degree distribution if we pick up a layer at random and we look at its degree distribution. Hence, the significant differences between the top panels in Fig. 3 illustrate the different behavior of the multiplex model and the corresponding for each single layer, that comes from the heterogeneity of the structure and distribution of the network. A similar situation occurs if we consider

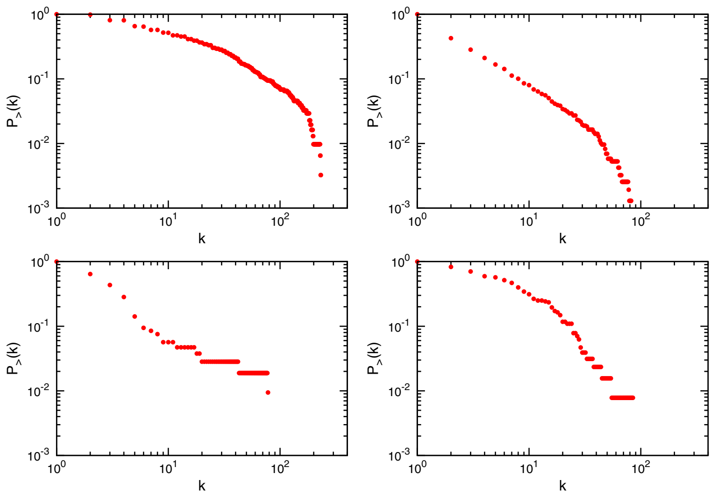


Fig. 3. Example of different cumulative degree probability distributions $P_>(k)$ for the European ATN in log-log scale. Top panels show the degree distribution for the multi-layer network model (on the left), and the average of the degree distributions of the 15 airlines under study (on the right). Bottom panels illustrate the cumulative distributions for a single traditional major company of 106 nodes (on the left) and for a low-cost company of 128 nodes (on the right).

the cumulative probability distribution of each single layer individually (see Fig. 3 bottom).

3 The model

As anticipated above, we will consider the set of direct flights of the same airline as the links of one independent network, *i.e.*, a single layer. On the other hand, each of the N nodes of the ATN will be present in each of the layers. Thus the collection of the L layers composes a multiplex representation of the ATN. Each of the layers will be denoted by a super index $\ell = 1, \dots, L$ so that the shortest distance between each couple of nodes within the same layer is denoted as d_{ij}^{ℓ} and the degree of a node i within layer ℓ is k_i^{ℓ}.

Once the topology of each layer of the multiplex ATN is characterized, we implement a model for the flow of a set of N_p passengers. First we assign the routes followed by each of the N_p passengers that move across the ATN. To this aim, and for each of the N_p passengers, we randomly choose two nodes of the ATN (one accounting for the origin and one for the destination of the passenger). Both nodes are selected proportionally to their global degrees k_i^A, as defined in Eq. 1. In this way, a node i will be selected as origin or destination of a given passenger with a probability:

$$P(i) = \frac{k_i^A}{\sum_{j=1}^{N} k_j^A}. \tag{2}$$

Obviously, paths starting from and ending at the same node are not allowed. Once the origin and destination of a passenger have been chosen, we search among all the layers the one for which the distance between the origin i and the destination j is minimal, *i.e.*, $d_{ij} = \min\{d^{\ell}_{ij}, \ell = 1, \ldots, L\}$. The distance between two airports, d_{ij}, is here defined as the hopping distance, *i.e.*, the sum of the number of *jumps* needed to reach the destinations; other factors usually taken into account by the passenger (like duration of the flight, cost, and so forth) are here disregarded. If there is more than one layer with the same minimum distance, one of them is randomly selected with equal probability; notice that this is equivalent to a passenger selecting one of the multiple airlines available to reach its destination. Finally, after the layer is selected, we compute the shortest path between origin and destination; a shortest path is randomly chosen when more than one was available.

The above process ends when all the N_p passengers have selected a couple of nodes (that is, their origin and destination), an airline (the layer) and a route (the shortest path between origin and destination nodes in the selected layer). Then, we can compute the load of each link $(i, j; \ell)$, $L(i, j; \ell)$, in each of the layers of the ATN, defined as the number of passengers whose path pass through it. In addition, for each link in the system, we assign a *maximum load* $L^M(i, j; \ell)$ as:

$$L^M(i, j; \ell) = L(i, j; \ell)(1 + f_{tol}), \tag{3}$$

where f_{tol} accounts for the fraction of additional load that each link can handle. For instance, a value of $f_{tol} = 0.2$ implies that airlines leave a number of vacant seats equal to the 20% of the real load. In what follows, we analyze situations in which $0 \leq f_{tol} \leq 0.3$, in line with the load factors observed in real operations (70% for short flights, and 80% for long-range connections [24]).

Once the model has been initialized, we simulate a random failure of the system by randomly removing a fraction of the links. With this aim, we visit each link connecting two nodes i and j in a given layer ℓ, and with some probability p we remove that link. As a consequence, all the passengers whose original paths passed through one (or more) of the removed links have to be re-scheduled, *i.e.*, they are forced to look for an alternative route between their departure and destination airports. As a previous step to the re-scheduling of a passenger, we decrease by one the load of the remaining active links in the passenger's original path.

3.1 Re-scheduling algorithm

After simulating the perturbation of the original system, we proceed with the re-scheduling phase. For each affected passenger, we try to find a new path between the origin i and the destination j of distance $d_{ij}(n) = d_{ij} + n$ (being d_{ij} the original distance in the unperturbed ATN), with $n = \{0, 1, 2, \ldots\}$. Obviously, we start by trying to allocate passengers in new routes with $n = 0$, so that the number of connections required for completing the trip (a proxy for the cost incurred by passengers) is not increased. Thus, for a given value of n, we proceed as follows:

(i) We recalculate all the *active paths* between each pair of nodes (i, j), within the same layer ℓ. We impose the distance between the pairs of nodes to be $d_{ij}(n) = d_{ij} + n$. Two situations may lead to the absence of *active paths* of length d_{ij} in a layer ℓ:
  (a) there are no paths of this length in the original multiplex graph.
  (b) there are some paths of length $d_{ij}(n)$, but all of them contains removed or full (see below) links.

  After this stage, each passenger is classified as either *fly* (he/she already has a route assigned, not affected by the removal of links), *re-scheduling* (he/she has the possibility of being assigned to an active route) or *no-fly* (there is no active path of distance less or equal to $d_{ij}(n)$ in any of the layers).

(ii) We take all the passengers one at the time in the *re-scheduling* group. For each of them:
   (a) We take their original layer and try to construct an (active) alternative path enabling the passenger to reach its destination whenever possible. If the chosen active path does not contain any full (see below) link, then the passenger is classified as *fly* and the load of each link in the chosen path is increased by one. Those links that reach their total capacity $L^M(i,j;\ell)$ with the addition of this last passenger are then classified as full.
   (b) If after step (*ii*.a) the passenger remains as re-scheduling, we repeat this last step for all the layers that contain at least one active path of length $d_{ij}(n)$ between its origin i and destination j. Again, if the passenger is successfully re-scheduled, it goes to the *fly* club and we add 1 to the load of each links used. If any of these links reaches its capacity $L^M(i,j;\ell)$, then it is classified as full.

(iii) Once all the passengers in the *re-scheduling* compartment have been processed, we check the remaining number of *re-scheduling* passengers. If it is non-zero, we go again to step *(i)*.

At the end, of the above iterative process, we partitioned the set of passengers into the subsets of *fly* and *no-fly*. We then perform the above process for different values of n increasing from $n = 0$. After each round n, passengers classified as *no-fly* are introduced again in the model as *re-scheduling* at the beginning of round $(n+1)$. In principle, n can be increased as many times as desired; nevertheless, to be realistic, we stop the algorithm at $n = 2$, meaning that passengers could, at most, look for alternative paths that are up to two steps longer than their original ones. Thus, at the end of round $n = 2$, *no-fly* passengers are those for which no active path of the former length exists between their origins and destinations. The rest of the passengers have been efficiently re-scheduled and take part of the final *fly* club[1]. It is worth noticing that this re-scheduling algorithm does not include any information about alliances between airlines; this means that (i) passengers cannot plan their trip by connecting flights of different airlines, and (ii) that the re-scheduling is unbiased, while in the real world airlines try first to accomodate passengers in flights of the same alliance.

4 Results

In this section we will explore the effects that link deletion causes on the flow of passengers across the multiplex ATN. To shed light on the effects on multiplexity, we compare the results obtained in the multiplex network with those of the aggregate ATN. The model introduced in the previous section has two parameters, namely the probability p that a link is deleted, and the fraction of tolerance f_{tol} that airlines assign to their connections. In what follows, we explore the robustness of the ATN as a function of the former two parameters.

4.1 Robustness of the ATN as a multiplex network

In order to characterize the effects that link deletion has on the re-organization of the flow in the multiplex ATN, we considered the partition into different groups of the

[1] Let us remark that we are assuming that passengers try to move to other airlines (layers) in order to avoid longer trips than those originally planned, *i.e.*, the case $n = 0$. Only when this latter attempt fails, they consider to perform longer trips ($n > 0$).

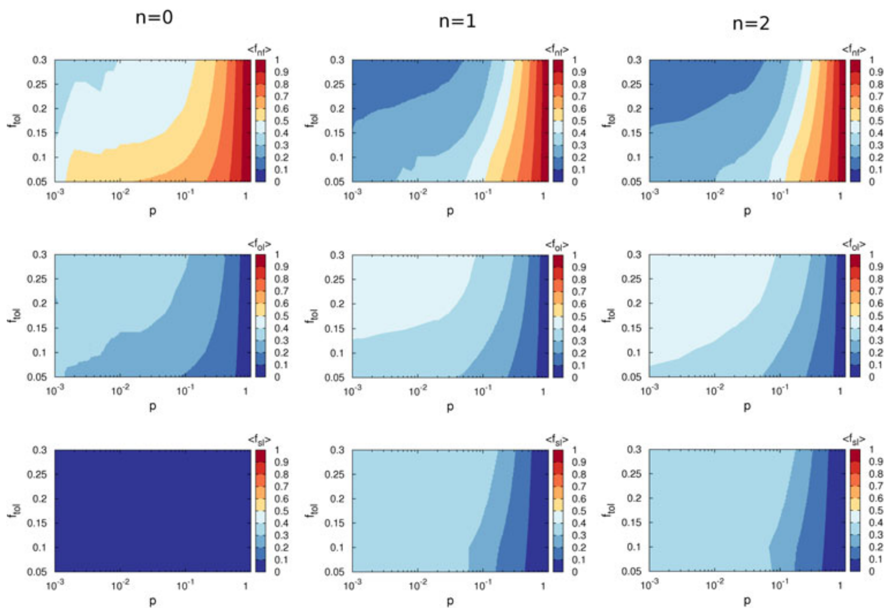


Fig. 4. Outcome of the re-scheduling process as a function of the probability of link failure, p, and load tolerance, f_{tol}. Each column displays (from top to bottom): the average fraction of passengers that cannot fly (f_{nf}), that of those that are re-scheduled in other layers (f_{ol}), and that for those re-scheduled within the original layer (f_{sl}). Each column accounts for the possibility of scheduling passengers on paths with length up to $d_{ij}(n)$ with $n = 0$ (left), $n = 1$ (center), and $n = 2$ (right). Results shown here refer to a population of $N_p = 50000$ passengers and are averaged over 50 different realizations. p spans logarithmically in the range $[10^{-3}, 1]$, while f_{tol} spans in the range $[0, 0.3]$.

total population of $N_A (\leq N_p)$ passengers affected by link deletion. This population can be divided into two groups: one, of size N_f, composed of those passengers that can reach their destination thanks to the re-scheduling process; and another group, of size N_{nf}, composed of those passengers that cannot be accommodated after the random failure of the system. Following this classification, we have that $N_A = N_{nf} + N_f$. In order to clearly monitor the effect of having different layers (airlines) in the multiplex ATN, we further split the group of N_f passengers that have been successfully re-scheduled into two other groups: those N_{sl} passengers that are re-scheduled within the same layer as originally planned, and those N_{ol} passengers that were forced to change layer in order to reach their corresponding destinations. With this new division we obtain the following equality: $N_A = N_{nf} + N_{sl} + N_{ol}$.

The three groups (*no-fly*, *same-layer* and *other-layer*) completely describe the final state of the population of affected passengers. In Fig. 4 we plot the fraction of passengers belonging to each compartment: $f_{nf} = N_{nf}/N_A$, $f_{ol} = N_{ol}/N_A$ and $f_{sl} = N_{sl}/N_A$, as a function of the two parameters p and f_{tol}. We also show how these quantities behave by iterating the re-scheduling algorithm for several values of n. Namely, in the left column of the figure we show (from top to bottom) the panels corresponding to $f_{nf}(p, f_{tol})$, $f_{ol}(p, f_{tol})$ and $f_{sl}(p, f_{tol})$ for $n = 0$, *i.e.*, when passengers are allowed to perform alternative trips only if their lengths are equal to the original one. In this plot we observe that there are almost no re-scheduled passengers flying across their original layers. This points out the low degree of redundant shortest-paths between two nodes in a given layer. As a consequence, almost all the successfully re-scheduled passengers are forced to change airline. The number of efficiently re-scheduled passengers ($N_{sl} + N_{ol}$) decreases with p and increase with f_{tol}. However, as can be observed in the top panel, the number of *no-fly*-passengers is extremely large even for a low rate of link deletion and a high degree of tolerance. Namely, for a value of $p \sim 10^{-2}$ and a degree of tolerance of about 10%, the fraction of *no-fly* passengers lies over the 50% of the population initially affected by the removal of links.

The constraint imposed in the case $n = 0$ seems too restrictive to achieve an efficient re-allocation of passengers, as it does not allow passengers to perform alternative paths in their respective original layers, at the cost of increasing the total length. Therefore, we relax this constraint and explore the cases $n = 1$ and $n = 2$ in the middle and right columns respectively. From these panels we observe that the average fraction of *no-fly* passengers (upper panels) is much lower than in the previous case $n = 0$. The decrease becomes more apparent for those regions corresponding to high values of load tolerance and low values of p. Remarkably, contrary to the case for $n = 0$, both for $n = 1$ and $n = 2$ some of the re-scheduled passengers succeded in traveling through alternative routes within their original layer. Besides, we observe that the plots corresponding to $n = 1$ and $n = 2$ are quite similar, pointing out that allowing the search for routes with $n > 2$ would not improve the results. Therefore, the plot of $f_{nf}(p, f_{tol})$ for $n = 2$ indicates a relative weakness of multiplex ATN with respect to perturbations, given that, even for very large values of tolerance and very low values of p, there is always some non-zero fraction of *no-fly* passengers.

4.2 Aggregate network results

In order to gain more insight on the effects of the multiplex structure of our system, we now show the results obtained with the same re-scheduling algorithm on the aggregate version of the ATN. Such aggregate network is obtained by merging all the layers of the multiplex representation into a single one, *i.e.*, by projecting the multiplex graph into a simplex one. This projection produces a complex network with the presence of multiple links between those couples of nodes that were connected in more than two layers; in other words, the number of connections between two airports is given by the number of airlines operating between them. In order to test whether the robustness of the aggregate network is larger than that of the multiplex network, we have performed the same link removal process followed by the re-scheduling program described in Sect. 3, this time considering the single layer comprising all the links in the aggregate ATN.

Figure 5 shows the final state of the system for the same three scenarios explored for the multiplex ATN, namely $n = 0$, 1 and 2. Since the aggregate ATN is composed of a single layer, in this case we only focus on the fraction of passengers affected by link deletion that are not able to be efficiently re-scheduled, $f_{nf}(p, f_{tol})$. As observed from the three panels in Fig. 5, compared with the corresponding panels $f_{nf}(p, f_{tol})$ for the multiplex ATN, the fraction of no-fly passengers decreases considerably in the three studied cases. In particular, while for those regions of the plot corresponding to $p > 10^{-1}$ remains roughly the same as in the case of the multiplex ATN, the main differences show up for low values of p; specifically, for $n > 0$ we can observe regions for which almost all the affected passengers can be re-scheduled, with an almost empty *no-fly* set. Again the panels corresponding to $n = 1$ and $n = 2$ are identical pointing out that the system is unable to achieve a better balance of affected passenger by increasing the length of the alternative trips. As a conclusion, the aggregate network shows an improved robustness with respect to the multiplex one, and a null impact of link deletion for some range of parameters.

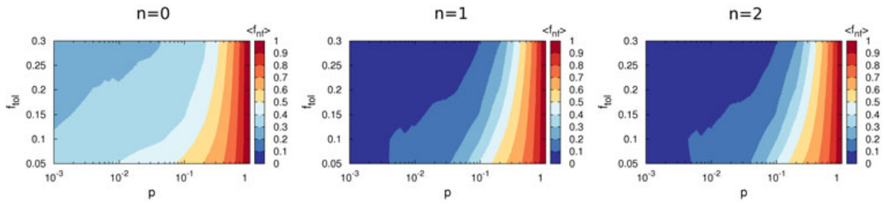


Fig. 5. Effect of link removal on the final state for the case of the aggregate network. We plot the average fraction of passengers that are not able to fly (f_{nf}) as a function of the probability of link removal, p, and the load tolerance, f_{tol}. Each column accounts of the possibility of re-scheduling passengers by means of paths of length up to $d_{ij}(n)$ with $n = 0, 1, 2$. Simulations are run with the same parameters and under the same conditions of those shown in Fig. 4.

This comparison confirms that multiplexity affects the robustness of the ATN. The root of the differences between the performance of both topologies is the constraint imposed by the multiplex architecture, which forces passengers to move within single layers. Therefore, in order to find an efficient alternative path, the affected passenger cannot mix connections of different airlines (layers) into the same path, thus reducing his capacity of optimizing the movement. This constraint disappears in the aggregate network, allowing affected passengers to make use of *hybrid* alternative paths. This provides the aggregate system way out to re-schedule the affected population of passengers in an efficient manner.

5 Conclusions

We presented a model for studying the re-scheduling problem in the European Air Transport Network using the paradigm of multiple layers structure, where each layer is made by the flights of a given airline. it is worth noting that, when comparing this multi-layer network with an equivalent single layer representation, topological characteristics differ, both qualitatively and quantitatively, as exposed in Sect. 2. Furthermore, on top of this multiplex network, we built a dynamical model, accounting for the re-scheduling problem of a group of passengers affected by the random failures of a set of connections. The affected passengers are then re-scheduled on new itineraries according to the availability of new routes (and free seats) in their former airline first, or eventually in a different one. The availability of routes is modulated by the probability of link failure p, and the tolerance on the load of a link f_{tol}. We presented our results in terms of the number of those passengers that are successfully re-scheduled and those for which the re-scheduling procedure fails. To achieve a deeper insight on the effects of dealing with a layered structure, we further subdivided passengers who are successfully re-scheduled into two subcategories: those which continue their trip using the same airline and those who, instead, are forced to switch to a different one. In addition, in order to increase the realism of our model, we allowed passengers to be re-allocated also on paths which are longer than the former ones. When compared to those corresponding to the single-layer representation, our results indicate that the multi-layer structure strongly reduces the resilience of the system against perturbations. In other words, the use of a *projection* of the ATN system is an over-simplification that results in an over-estimation of the resilience of the ATN. While it is known that a multi-layer structure can drammatically change the resilience of the system [11], to the best of our knowledge this is the first application of such representation to the air transport system; all previous studies (see, for instance, [22]) only considers projections of the network. We anticipate that this framework may be an important tool for policy makers in the near future, especially when other elements (e.g., more real estimation of the distance between airports, airline alliances, estimation of the costs of re-routing, etc., here excluded for the sake of simplicity) would be included. We also believe that these results could also be valid in other real-world complex systems, which have been widely studied in the last decade under the single layer network paradigm, when their multiplex nature is taken into account.

This work has been partially supported by the Spanish DGICYT under projects FIS2008-01240, MTM2009-13848 and FIS2011-25167; by the Comunidad de Aragón through Project No. FMI22/10. FET project PLEXMATH (317614) J.G.G. is supported by MICINN through the Ramón y Cajal program. Authors gratefully acknowledge EUROCONTROL and the ComplexWorld Network (`www.complexworld.eu`) in the context of the SESAR Work Package E for sharing the operational data set. The paper reflects only the authors' views. EUROCONTROL and the ComplexWorld Network are not liable for any use that may be made of the information contained therein. Authors also thank David Papo for his suggestions and corrections on the Manuscript.

References

1. T.A. Heppenheimer, *Turbulent Skies: The History of Commercial Aviation* (Wiley, 1998)
2. S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Physics Reports **424**, 175 (2006)
3. R. Albert, A.-L. Barabási, Rev. Mod. Phys. **74**, 47 (2002)
4. M. Barthélemy, Physics Reports **499**, 1 (2011)
5. L. da F. Costa, O.N. Oliveira, G. Travieso, F.A. Rodrigues, P.R.V. Boas, L. Antiqueira, M.P. Viana, L.E.C. Rocha, Adv. Phys. **60**, 329 (2011)
6. S. Wasserman, K. Faust, *Social Network analysis* (Cambridge University Press, Cambridge, 1994)
7. M. Kurant, P. Thiran, Phys. Rev. Lett. **96**, 138701 (2006)
8. J. Gómez-Gardeñes, I. Reinares, A. Arenas, L.M. Floría, Scientific Rep. **2**, 620 (2012)
9. E.A. Leicht, R. D'Souza [`arXiv: 0907.0894`] (2009)
10. J. Gómez-Gardeñes, C. Gracia-Lázaro, L.M. Floría, Y. Moreno, Phys. Rev. E **86**, 056113 (2012)
11. S.V. Buldyrev, R. Parshani, G. Paul, H.E. Stanley, S. Havlin, Nature **464**, 1025 (2010)
12. R. Criado, J. Flores, A. Garcia del Amo, J. Gomez-Gardeñes, M. Romance, Int. J. Comput. Math. **89**, 291 (2012)
13. M. Kurant, P. Thiran, P. Hagmann, Phys. Rev. E **76**, 026103 (2007)
14. C.D. Brummitt, K.-M. Lee, K.-I. Goh, Phys. Rev. E **85**, 045102(R) (2012)
15. D.J. Watts, Proc. Nat. Acad. Sci. USA **99**, 5766 (2002)
16. `http://www.iata.org/pressroom/speeches/Pages/2010-06-07-01.aspx`
17. EUROCONTROL, Long-Term Forecast of Annual Number of IFR Flights (2010-2030) (2010)
18. Airbus, flying smart, think big, Global Market Forecast 2009 – 2028
19. European Commission, White Paper: Roadmap to a Single European Transport Area Towards a competitive and resource efficient transport system, Brussels (2011)
20. European Commission, Flightpath 2050 – Europes Vision for Aviation (Report of the High Level Group on Aviation Research) (2011)
21. B. Sridhar, K. Sheth, Network Characteristics of Air Traffic in the Continental United States, Proceedings of the 17th World Congress of The International Federation of Automatic Control Seoul, Korea, July 6–11 (2008)
22. L. Lacasa, M. Cea, M. Zanin, Physica A **388**, 3948 (2009)
23. Y.H. Chou, Transp. Plann. Technol. **14**, 243 (1990)
24. Association of European Airlines, *Yearbook 2007* (2007)

Chapter 8

Analysis of remote synchronization in complex networks

CHAOS 23, 043103 (2013)

Analysis of remote synchronization in complex networks

Lucia Valentina Gambuzza,[1] Alessio Cardillo,[2,3] Alessandro Fiasconaro,[2,4] Luigi Fortuna,[1] Jesus Gómez-Gardeñes,[2,3] and Mattia Frasca[1]
[1]*Dipartimento di Ingegneria Elettrica Elettronica e Informatica, Università degli Studi di Catania, viale A. Doria 6, 95125 Catania, Italy*
[2]*Departamento de Física de la Materia Condensada, University of Zaragoza, Zaragoza 50009, Spain*
[3]*Institute for Biocomputation and Physics of Complex Systems (BIFI), University of Zaragoza, Zaragoza 50018, Spain*
[4]*Instituto de Ciencia de Materiales de Aragón, CSIC–University of Zaragoza, Zaragoza 50009, Spain*

(Received 17 May 2013; accepted 23 September 2013; published online 2 October 2013)

A novel regime of synchronization, called remote synchronization, where the peripheral nodes form a phase synchronized cluster not including the hub, was recently observed in star motifs [Bergner *et al.*, Phys. Rev. E **85**, 026208 (2012)]. We show the existence of a more general dynamical state of remote synchronization in arbitrary networks of coupled oscillators. This state is characterized by the synchronization of pairs of nodes that are not directly connected via a physical link or any sequence of synchronized nodes. This phenomenon is almost negligible in networks of phase oscillators as its underlying mechanism is the modulation of the amplitude of those intermediary nodes between the remotely synchronized units. Our findings thus show the ubiquity and robustness of these states and bridge the gap from their recent observation in simple toy graphs to complex networks. © *2013 AIP Publishing LLC*. [http://dx.doi.org/10.1063/1.4824312]

In this work we show a novel synchronization state in networks of coupled oscillators. This state, called Remote Synchronization, is characterized by the synchronization of pairs of nodes that are not directly connected via a physical link or any sequence of synchronized nodes. Moreover, remote synchronization is manifested when considering oscillators having amplitude and phase as dynamical variables, in contrast to the usual setting in which phase oscillators are considered, as its underlying mechanism is the modulation of the amplitude of those intermediary nodes allowing the exchange of information between remotely synchronized units. Although some previous observations of such phenomenon were made in simple star-like graphs, here we show its ubiquity in the general framework of complex networks. To this end we analyze its existence as a robust dynamical state that appears before global synchronization shows up. Our findings thus open the door for experimental observations of this novel state in which the existence of a synchronized pair cannot be associated to a given physical interaction through a single link of the network. In addition, our results highlight the important difference between the real (i.e., associated to physical links) and the functional (i.e., emerging from synchronization) connectivity of a network.

I. INTRODUCTION

Synchronization constitutes one of the most paradigmatic examples of emergence of collective behavior in natural, social, and man-made systems.[2–4] Its ubiquity relies on the general framework in which it occurs: the interaction between two or more nonidentical dynamical units that, as a consequence, adjust a given property of their motion. As coupling between units increases, synchronization shows up as a collective state in which the units behave in a coordinated way. Synchronization phenomena span across many life scales, ranging from the development of cognitive tasks in neural systems[5] to the onset of social consensus in human societies.[6]

The ubiquity of synchronization in real systems together with the recent discovery[7–11] of their real architecture of interactions has motivated its study when units are embedded in a complex network.[12] In this way, each unit is represented as a node of a network while it only interacts with those adjacent units, i.e., those directly coupled via an edge. In the last decade many studies have unveiled the impact that diverse interaction topologies have on the onset of synchrony[13–17] and its stability.[18–21] In addition, related issues such as that of adaptive networks, in which the interaction pattern changes according to the degree of synchronization of the system, have also attracted the attention of the community.[22–25]

The former studies mainly rely on the study of coupled phase oscillators, such as the Kuramoto model,[26,27] which produces globally synchronized systems as a result of the direct interaction of pairs of adjacent units. However, it has been recently found[1] that, for more general oscillator models (in which both amplitude and phase are dynamical variables) such as the Stuart-Landau (SL) model,[3] two oscillators, which are not directly linked but are both connected to a third unit, can become synchronized even if the third oscillator does not synchronize with them. This novel phenomenon, termed *remote synchronization*, relies on the modulation of the amplitude parameter of an intermediary node allowing the passage of information between two of its neighbors for their synchronization, even when the former is not synchronized with them. Thus, this tunnel-like mechanism is out of reach in ensembles of phase oscillators. Although the term remote synchronization has been used in quite different contexts, as, for example, in computer science where it refers to synchronization of two or more files located in two, remotely connected, computers or in some synchronization schemes for dynamical systems,[28] to emphasize the remote location of the receiver with respect to the transmitter, we will use it to refer to the novel form of synchronization as reported in Ref. 1.

Remote synchronization has been found to occur in very specific and simple topologies, such as star-like networks in which the central node has a natural frequency different from that of the leaves. Within this particular setting, it was numerically and experimentally shown[1] that leaves become mutually synchronized without the need of the synchronization of the central node. In this paper, we aim at showing that remote synchronization is not limited to the particular configuration of a star-like motif or a tight specification of the node frequencies. To this end, we introduce a general procedure for detecting remote synchronization in arbitrary networks and then discuss the results of our analysis on arbitrary complex networks.

II. MEASURES OF REMOTE SYNCHRONIZATION

In Ref. 1, where star motifs were dealt with, remote synchronization is detected by observing that for intermediate values of the coupling coefficient the synchronization level among the leaves (measured with the so called *indirect* Kuramoto parameter) is higher than that between the hub and the leaves (measured with the so called *direct* Kuramoto parameter). We note that such measures are not applicable to the general case of arbitrary topologies, since they are based on an *a priori* analysis of the network structure which allows one to establish which nodes can remotely synchronize. Therefore, in this paper we first introduce a general procedure for detecting remote synchronization in arbitrary networks and then show ubiquity and robustness of remote synchronization in the general case of complex networks.

To this end, we consider a network of N coupled Stuart-Landau oscillators.[3] Each node i is characterized by two variables, $(x_i, y_i)^T$, whose dynamical evolution is as follows:

$$\begin{pmatrix} \dot{x}_i \\ \dot{y}_i \end{pmatrix} = \begin{pmatrix} \alpha - x_i^2 - y_i^2 & -\omega_i \\ \omega_i & \alpha - x_i^2 - y_i^2 \end{pmatrix} \begin{pmatrix} x_i \\ y_i \end{pmatrix} + \frac{\lambda}{k_i} \sum_{j=1}^{N} a_{ij} \left[\begin{pmatrix} x_j \\ y_j \end{pmatrix} - \begin{pmatrix} x_i \\ y_i \end{pmatrix} \right], \qquad (1)$$

where $\sqrt{\alpha}$ and ω_i are, respectively, the amplitude and the (natural) frequency of oscillator i when uncoupled. The second term on the right accounts for the coupling of the dynamics of node i with its k_i neighbors. The strength of the coupling is controlled by λ ($\lambda = 0$ in the uncoupled limit) while $\mathcal{A} = \{a_{ij}\}$ represents the adjacency matrix of the network defined as: (i) for $i \neq j$, $a_{ij} = 1$ when nodes i and j are connected while $a_{ij} = 0$ otherwise and (ii) $a_{ii} = 0$.

To study the synchronization properties of system (1), we work with the phase variable of each oscillator, defined as $\theta_i = \tan^{-1}(y_i/x_i)$. Then we can measure the degree of synchronization of any (connected or not) pair of oscillators by means of the time averaged order parameter

$$r_{ij} = |\langle e^{\imath[\theta_i(t) - \theta_j(t)]} \rangle_t|, \qquad (2)$$

where $\langle \cdot \rangle_t$ means an average over a large enough time interval and $\imath = \sqrt{-1}$. We will consider two nodes as synchronized when $r_{ij} > \delta$, where δ is a constant threshold that we fix to $\delta = 0.8$. Nonetheless, we checked that the results presented are robust as other values of δ yield qualitatively the same outcomes.

Once two nodes i and j are classified as mutually synchronized we label their relationship according to the following three situations: (i) i and j are directly connected ($a_{ij} = a_{ji} = 1$), (ii) there is a path of mutually synchronized nodes between them, and (iii) neither of the former two situations hold. While the first two cases are similar, as both are examples of synchronization through *physical links*, the third case is analogous to the observed remote synchronization in a star-like network, but in the more general context of a complex network. Thus, we define that two nodes i and j are remotely synchronized (RS) when they are synchronized ($r_{ij} > \delta$) and they are not connected by either a direct link or a path of synchronized nodes.

To quantify systematically the extent of remote synchronization we count the number of RS nodes, defined as the number N_{RS} of nodes that appear RS with at least another node in the network. This allows us to introduce the following order parameter: $n_{RS} = N_{RS}/N$, representing the normalized number of RS nodes with respect to the total number of nodes N. Finally, to quantify the importance that remote synchronization has on the dynamics of the system we also measure the global level of synchronization through the usual Kuramoto-like order parameter

$$r = \frac{1}{N^2} \sum_{i,j=1}^{N} r_{ij}. \qquad (3)$$

Note that r takes into account the contribution of both synchronized ($r_{ij} > \delta$) and not synchronized ($r_{ij} \leq \delta$) nodes.

III. RESULTS

As two well-known paradigmatic network topologies we have analyzed both Erdős-Rényi (ER) and Scale-free (SF) graphs. The former type of networks is characterized by a Poisson distribution $P(k)$ for the probability of finding a node with k contacts while SF graphs show a power-law distribution, $P(k) \sim k^{-\gamma}$. Thus, while in ER graphs most of the nodes are close to the mean connectivity $\langle k \rangle$, SF networks display a large heterogeneity in the number of contacts per node as revealed from the existence of hubs having $k_i \gg \langle k \rangle$. For their construction we have made use of the model introduced in Ref. 29 that allows one to control the mean connectivity of both networks in order to be exactly the same. In the networks reported in this paper the size and mean connectivity are fixed to $N = 100$ and $\langle k \rangle = 2$, respectively. The SF networks generated with this model have $\gamma = 3$. Finally, in order to stay close to the framework used in Ref. 1, we have considered a bimodal distribution for the natural frequencies of the oscillators so that nodes with high degree (those analogous to the central nodes in a star graph) present a larger frequency, ω_h, than that, ω_l, of less connected (the ones playing the role of leaves in the star topology). In particular, we labeled as hubs those nodes having $k_i > k^*$ (Ref. 30) and assigned them $\omega_i = \omega_h + \xi_i\omega_h$, while for the rest of nodes $\omega_i = \omega_l + \xi_i\omega_l$. In the former expressions ξ_i is a random variable uniformly distributed between -0.025 and 0.025.

In Fig. 1, we show the emergence of remote synchronization as a function of the two relevant parameters: the coupling strength λ and the frequency mismatch of the network hubs $\Delta\omega = \omega_h - \omega_l$. In particular, we report the behavior of the global synchronization, r [panels 1(a) and 1(c)] and the fraction of RS nodes, n_{RS} [panels 1(b) and 1(d)] for SF (top) and ER (bottom) networks. The results are averaged over 50 different network instances and, for each network, we average the results over 10 different realizations of the distribution of natural frequencies.

We find that remote synchronization occurs in both types of networks in a region of parameters characterized by a strong frequency mismatch $\Delta\omega$ and moderate coupling λ. In fact, for low values of the coupling parameter, nodes cannot synchronize (either in a direct or remote way) as observed from the low values of r and n_{RS}. On the contrary, for large values of λ the network is fully synchronized ($r \simeq 1$) and, accordingly, n_{RS} assumes values close to zero since all the nodes are mutually synchronized with their neighbors. As panels 1(a) and 1(c) reveal, the onset of full synchronization requires greater values of the coupling as the frequency mismatch increases. In fact, a large frequency mismatch together with values of coupling under the threshold for complete synchronization favors the onset of remote synchronization, as observed from the behavior of n_{RS} in panels 1(b) and 1(d).

Compared to SF networks, the values of n_{RS} in ER networks are greater, thus indicating that remote synchronization in ER networks involves a larger number of nodes. Moreover, in ER networks the onset of remote synchronization occurs for lower values of λ. ER and SF networks also show qualitative differences in the appearance of remote synchronization: by keeping fixed $\Delta\omega$ and increasing the value of λ, we find that n_{RS} in SF networks show two peaks, while for ER networks it shows a rise-and-fall behavior.

In both (ER and SF) cases remote synchronization appears as an intermediary state before full synchronization is achieved. However, from the analysis of panels 1(a) and 1(c) one observes that the behavior of r *vs.* λ for a fixed value of $\Delta\omega$ is qualitatively different in SF and ER networks. In particular, in ER networks (panel 1(c)) a large plateau around $r \simeq 0.5$ is set in the region where remote synchronization shows up. In this region, the increase of λ does not contribute to the overall synchronization level, but to a redistribution of the average oscillation frequencies of the network nodes.

This is evident in Fig. 2, where the average values (over the simulation time T) of the instantaneous frequency of each oscillator are reported along with the parameter n_{RS}. The results are obtained by increasing λ adiabatically from $\lambda = 0$ so that the system starts from a bimodal distribution as dictated by the configuration for the natural oscillations. As λ increases, the gap between the two main frequency values of the bimodal distribution decreases until the network reaches full synchronization and the nodes oscillate at a common frequency. The readjustment of frequencies reveals

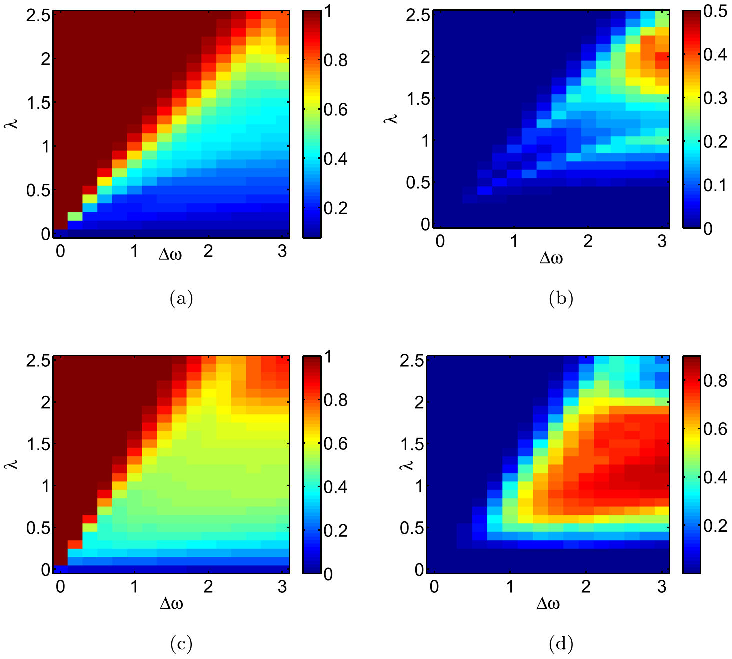


FIG. 1. Evolution of the degree of global synchronization r [panels (a) and (c)] and the number of remotely synchronized nodes n_{RS} [panels (b) and (d)] for SF (upper panels) and ER (bottom panels) networks as a function of the coupling strength λ and the frequency mismatch $\Delta\omega$. In both cases the networks have $N = 100$ and $\langle k \rangle = 2$. The other relevant parameters are fixed to $\alpha = 1$, $\omega_l = 1$. Remote synchronization (high values of n_{RS}) is found for strong frequency mismatch $\Delta\omega$ and moderate coupling λ, while, for low values of the coupling parameter, nodes cannot synchronize (r and n_{RS} have low values), and, for large values of λ, the network is fully synchronized ($r \simeq 1$).

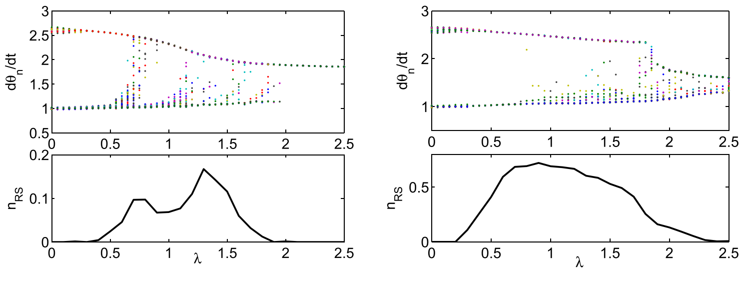


FIG. 2. Evolution of the average oscillation frequency of each oscillator and n_{RS} as a function of λ for SF (a) and ER (b) networks. The mismatch of natural frequencies is $\Delta\omega = 1.5$ while the rest of parameters are the same as in Fig. 1. The average oscillation frequencies, which for $\lambda = 0$ start from a bimodal distribution as dictated by the configuration for the natural frequencies, as λ is increased tend towards a common value, characterizing full synchronization. The strong reorganization of the frequencies (characterized by a spread of the oscillation frequencies between the two extreme values) corresponds to the values of coupling for which n_{RS} is peaked.

that, for some values of the coupling, the system undergoes a strong reorganization, as shown by the spread of the oscillation frequencies between the two extreme values. This readjustment coincides with the peaks displayed by n_{RS} in both SF and ER networks. However, the readjustment seems to occur faster in SF networks for which the plateau of r is not observed.

Now we illustrate the role of parameter α. To this end, we consider a general graph and show that for $\alpha \gg 1$ the SL model transforms into a network of Kuramoto oscillators, so that the amplitude of the oscillators become decoupled and stationary. We consider Eq. (1) in polar coordinates

$$\dot{\rho}_i = \alpha\rho_i - \rho_i^3 + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}(\rho_j \cos(\theta_j - \theta_i) - \rho_i),$$
$$\dot{\theta}_i = \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} \frac{\rho_j}{\rho_i} a_{ij} \sin(\theta_j - \theta_i), \tag{4}$$

where $\rho_i e^{\imath\theta_i} = x_i + \imath y_i$. Defining $R_i = \frac{\rho_i}{\sqrt{\alpha}}$, where $\sqrt{\alpha}$ is the value of the amplitude at the equilibrium, Eq. (4) can be rewritten as follows:

$$\dot{R}_i = \alpha R_i - \alpha R_i^3 + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij}(R_j \cos(\theta_j - \theta_i) - R_i),$$
$$\dot{\theta}_i = \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} \frac{R_j}{R_i} a_{ij} \sin(\theta_j - \theta_i). \tag{5}$$

In the first equation we can rescale time according to $dT = \alpha dt$ (while the second equation remains unchanged)

$$\frac{dR_i}{dT} = R_i - R_i^3 + \frac{\lambda}{\alpha k_i}\sum_{j=1}^{N} a_{ij}(R_j \cos(\theta_j - \theta_i) - R_i),$$
$$\dot{\theta}_i = \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} \frac{R_j}{R_i} a_{ij} \sin(\theta_j - \theta_i). \tag{6}$$

Now as $\alpha \to \infty$ the coupling term in the amplitude equation vanishes, and from the analysis of the first equation we derive that $R_i \to 1$ for all i (in fact $R_i = 1$ is the only equilibrium and the dynamics evolve very fast as $dT = \alpha dt$ and α is large). In the second equation, $R_i \to 1$ leads to $\frac{R_i}{R_j} = 1$ and thus the second equation becomes

$$\dot{\theta}_i = \omega_i + \frac{\lambda}{k_i}\sum_{j=1}^{N} a_{ij} \sin(\theta_j - \theta_i). \tag{7}$$

Therefore, as $\alpha \to \infty$, we recover the model of Kuramoto purely phase oscillators coupled into a network. In this limit, we observe that the amplitude equation plays no role. In this case, the level of RS is very low, as for instance reported in Fig. 3, where a network of Stuart-Landau oscillators with $\alpha = 1$ is compared with a network of Stuart-Landau oscillators with $\alpha = 1000$ and with a network of Kuramoto purely phase oscillators. We note that for $\alpha = 1000$ the network of Stuart-Landau oscillators is already a good approximation of the network of Kuramoto purely phase oscillators. In both the two examples of networks (SF and ER), for Kuramoto oscillators n_{RS} (Figs. 3(a) and 3(b)) is lower than in Stuart-Landau oscillators (with $\alpha = 1$). The lower level of RS in Kuramoto oscillators is more evident when the number of RS links, labeled as L_{RS}, is examined as in Figs. 3(c) and 3(d), which shows how the number of RS links is decreased by an order of magnitude with respect to the case of Stuart-Landau oscillators (with $\alpha = 1$). This suggests that amplitude modulation is the main mechanism underlying RS (this was also shown with other arguments in Ref. 1 for star-like networks).

To gain more insight into the relation between the regime of remote synchronization and the onset of global synchrony, we now consider the analysis of all synchronized pairs and its partition into those corresponding to remote synchronization and those for which a synchronized physical connection (either a direct link or a path of synchronized nodes) exists. To this aim, we define $\eta_{ij} = 1$ if nodes i and j are connected either by a physical link or by a path of synchronized nodes and $\eta_{ij} = 0$, otherwise. We then introduce the following quantities:

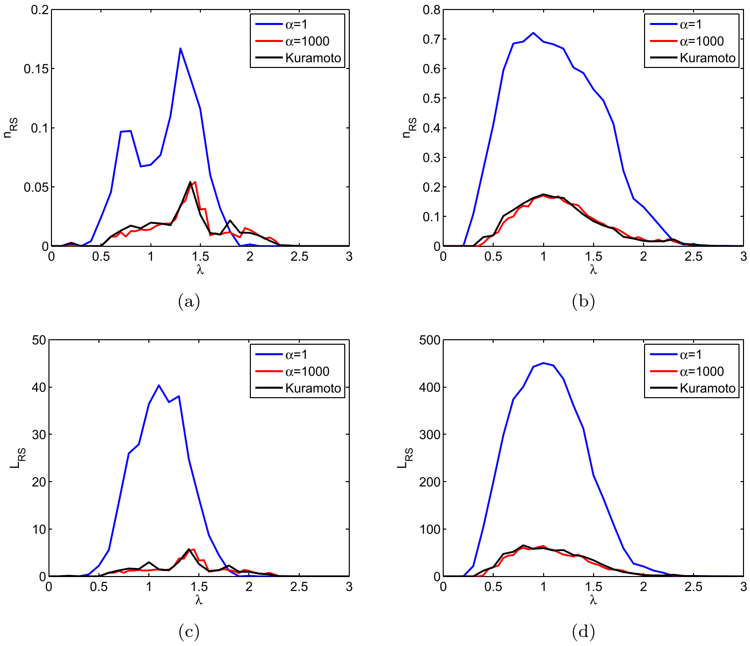


FIG. 3. Comparison of n_{RS} (a) and (b) and L_{RS} (c) and (d) in a SF (a) and (c) and ER (b) and (d) network of Stuart-Landau oscillators with $\alpha = 1$ or $\alpha = 1000$ and a network of Kuramoto purely phase oscillators for $\Delta\omega = 2.6$.

$$f_P = \frac{\sum_{i,j=1}^{N} \eta_{ij} H(r_{ij} - \delta)}{\sum_{i,j=1}^{N} H(r_{ij} - \delta)} \tag{8}$$

and

$$f_{RS} = \frac{\sum_{i,j=1}^{N} (1 - \eta_{ij}) H(r_{ij} - \delta)}{\sum_{i,j=1}^{N} H(r_{ij} - \delta)}, \tag{9}$$

where $H(x)$ is the Heaviside function. Thus, f_P and f_{RS} represent the fraction of synchronized links due to a physical or remote connection, respectively. Obviously, as $f_P + f_{RS} = 1$, it is enough to report the behavior of f_{RS}.

In Fig. 4 we show the evolution of f_{RS} vs. λ for several values of $\Delta\omega$. The presence of two peaks in the evolution of f_{RS} in SF networks reveals a similar behavior to that found for n_{RS}. As $\Delta\omega$ increases, the percentage of RS links increases and the two peaks shift towards increasing values of λ. On the other hand, for ER networks the percentage of RS links is higher than in SF networks and f_{RS} shows, as in the case of n_{RS}, a rise-and-fall trend. The fall in the number of RS links points out that the network is able to recruit physical links to get synchronized and thus those regions that appeared as RS become merged into a single component made of physically synchronized links.

To visualize the progressive substitution of RS links by physical ones in the path towards full synchronization, we show in Fig. 5 for an ER network (with $\Delta\omega = 2.8$) snapshots of both remotely and physically synchronized links for two values of the coupling λ. In Figs. 5(a) and 5(c) we plot two networks corresponding to physically and remotely synchronized links, respectively, when $\lambda = 1.65$. In this case the network is divided into several clusters of physically synchronized nodes (the color of the nodes corresponds to the cluster of physically synchronized links they belong to)

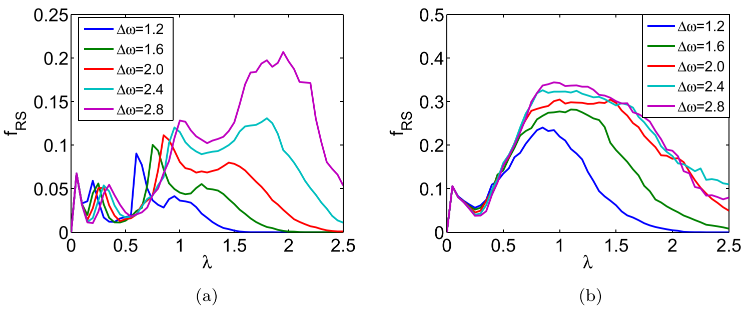


FIG. 4. Evolution of the fraction of RS links, f_{RS}, in SF (a) and ER (b) networks as a function of the coupling strength λ, and for different values of $\Delta\omega$. The remaining parameters are set as in Fig. 1. The fraction of RS links first increased as λ is increased, with one (in ER networks) or two peaks (in SF networks) as observed for the evolution of n_{RS}, and then falls as networks recruit physical (instead of RS) links to get synchronized.

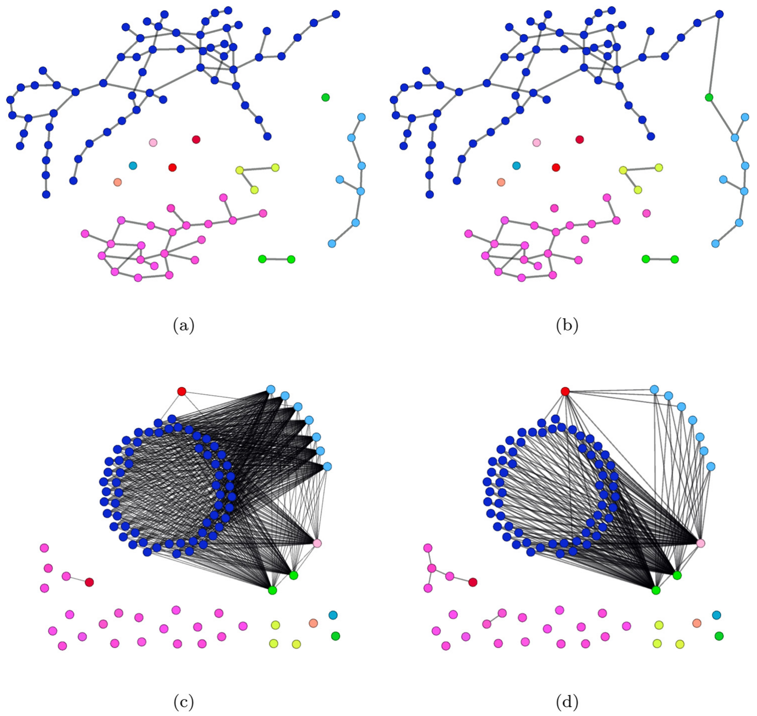


FIG. 5. Evolution of components of physically (a) and (b) and remotely (c) and (d) synchronized nodes in an ER network with $\Delta\omega = 2.8$, when λ is increased by adiabatic continuation from $\lambda = 1.65$ (a)–(c) to $\lambda = 1.70$ (b)–(d). Nodes are colored according to the physically synchronized component they belong to when $\lambda = 1.65$, i.e., in (a). The remaining parameters are the same as in Fig. 1. At $\lambda = 1.70$ two communities (the one with blue nodes and the one with cyan nodes), that were remotely synchronized at $\lambda = 1.65$, fuse into a single one and, as a consequence, the RS links between the two communities existing for $\lambda = 1.65$ disappear at $\lambda = 1.70$.

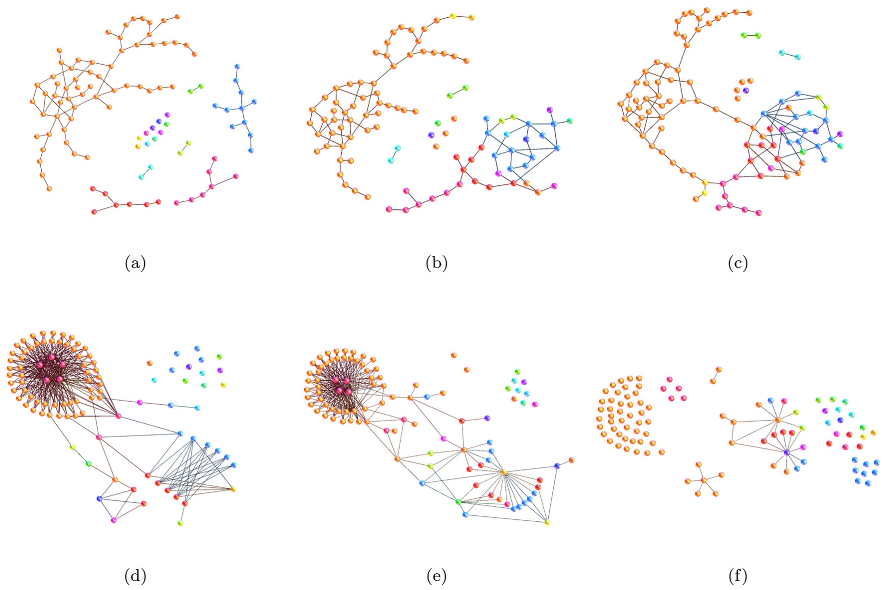


FIG. 6. Evolution of components of physically (a), (b), (c) and remotely (d), (e), (f) synchronized nodes in an ER network with $\Delta\omega = 1.5$ and $\langle k \rangle = 2$, when λ is adiabatically increased. Nodes are colored according to the component to which they belong in (a) and, then, represented with the same colors in (b)–(f). The networks correspond to the following value of λ: (a) and (b) $\lambda = 1.75$; (c) and (d) $\lambda = 1.80$; (e) and (f) $\lambda = 1.85$. For progressive increase of the coupling coefficient, first three of the four communities, existing at $\lambda = 1.75$ and synchronized thanks to RS links, merge and, correspondingly, the RS links between these communities disappear, and, then, the fourth community (synchronized with the other three, already at $\lambda = 1.80$ thanks to RS links) aggregates to the previous ones.

and some nodes of these clusters appear remotely synchronized with nodes belonging to different clusters [as shown in Fig. 5(c)]. When λ is increased to $\lambda = 1.70$, two of these clusters merge together [Fig. 5(b)] through two physically synchronized links that connect each cluster to a new node synchronized to each of them. Thus, at $\lambda = 1.70$ two communities that were remotely synchronized at $\lambda = 1.65$, fuse into a single one and, as a consequence, those RS links between the nodes of the two communities reported for $\lambda = 1.65$ in Fig. 5(c) disappear at $\lambda = 1.70$ [Fig. 5(d)]. We note that the choice of the threshold δ may impact in which nodes are assigned to which groups, although we have observed qualitatively similar results when the threshold is changed. In fact, the evolution of communities remains the same, although the value of λ at which they merge may be slightly different.

A further example of the merging of RS clusters is shown in Fig. 6. We consider an ER network with $\Delta\omega = 1.5$ and $\langle k \rangle = 2$, when λ is increased with continuation from $\lambda = 1.75$ to $\lambda = 1.85$. For $\lambda = 1.75$, the network is divided into four main components of physically synchronized nodes plus some small communities and isolated nodes (Fig. 6(a)). The analysis of the components of RS nodes (Fig. 6(d)) reveals that there are RS links between the four communities. In fact, increasing the coupling to $\lambda = 1.80$ three of these communities merge (Fig. 6(b)) and, correspondingly, the RS links between these communities disappear (Fig. 6(e)). Finally, a further increase of λ ($\lambda = 1.85$ in Fig. 6(c)) leads to the aggregation of the fourth community (the bigger one) to the previous ones. Also in this case, almost all the RS links disappear (Fig. 6(f)) and very few RS links are observed for $\lambda = 1.85$.

IV. CONCLUSIONS

In this paper we have provided measures to study remote synchronization in general complex networks. This phenomenon relies on the mutual synchronization of pairs of uncoupled nodes. Each remotely synchronized pair of nodes is thus physically connected through an intermediary node (not synchronized with them) or a sequence (path) of intermediary nodes. This is an important difference with another form of remote synchronization reported in Ref. 31, where the analysis focused on the distribution of phase lags in a network of homogeneous oscillators (all oscillating at the same frequency) and a relationship between modules appearing in the network structure and the pattern of phase lags was revealed. The analysis we have presented reveals a stronger condition in that, according to our results, two RS nodes do not show any form of synchronization with intermediate nodes.

Although the original discovery of remote synchronization was restricted to a rather particular setup, a star graph, the analysis carried out in this paper, through the introduction of appropriate indicators, reveals that remote synchronization is common in general complex networks, such as Erdős-Rényi and Scale-free graphs of coupled oscillators having amplitude and phase as dynamical variables. The addition of amplitude as a dynamical variable, in contrast with the typical framework of networks of coupled phase-oscillators, provides the observation of remote synchronization and elucidates an important role played by it. In fact, we have found that remote synchronization constitutes a mechanism anticipating synchronization by physical links in networks with heterogeneous distribution of natural frequencies. Our results indicate that, in these networks, communities of nodes synchronized through RS links appear for values of coupling just lower than those allowing the merging of these communities through physical links. As synchronization is ubiquitous in natural and man-made systems, we suggest that this can be an important mechanism to explain the emergence of communities of synchronized nodes, not connected by physical links.

Our work suggests that remote synchronization is not significant for ensembles of phase oscillators, since its main underlying mechanism seems to be the modulation of the amplitude of intermediary nodes allowing information transfer between uncoupled pairs of nodes. In fact, when similar settings are applied to phase oscillators, a different phenomenon is observed, namely, that of *explosive synchronization*[17] in which the typical second-order synchronization transition transforms into a first-order one. In its turn, remote synchronization appears as a rather robust state prior the onset of global synchronization since for a wide range of coupling strengths almost all the nodes are remotely synchronized with, at least, another one while the level of global synchronization remains small. Thus, our results open the door for experimental observations of this novel state in which the existence of a synchronized pair cannot be associated to a given physical interaction through a single link of the network and highlight the important difference between the real (i.e., associated with physical links) and the functional (i.e., emerging from synchronization) connectivity of a network.

ACKNOWLEDGMENTS

This work has been partially supported by the Spanish MINECO under projects FIS2011-25167 and FIS2012-38266-C02-01 and by the Comunidad de Aragon (Grupo FENOL). J.G.G. is supported by MINECo through the Ramon y Cajal program.

[1]A. Bergner, M. Frasca, G. Sciuto, A. Buscarino, E. J. Ngamga, L. Fortuna, and J. Kurths, "Remote synchronization in star networks," Phys. Rev. E **85**, 026208 (2012).
[2]S. H. Strogatz, "From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators," Physica D **143**, 1–20 (2000).
[3]A. Pikovsky, M. Rosenblum, and J. Kurths, *Synchronization: A Universal Concept in Nonlinear Sciences* (Cambridge University Press, 2003).
[4]S. Boccaletti, *The Synchronized Dynamics of Complex Systems* (Elsevier, 2008).
[5]E. Bullmore and O. Sporns, "Complex brain networks: Graph theoretical analysis of structural and functional systems," Nat. Rev. Neurosci. **10**, 186–198 (2009).
[6]A. Pluchino, V. Latora, and A. Rapisarda, "Changing opinions in a changing world: A new perspective in sociophysics," Int. J. Mod. Phys. C **16**, 515 (2005).
[7]S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. U. Hwang, "Complex networks: Structure and dynamics," Phys. Rep. **424**, 175–308 (2006).

[8]S. N. Dorogovtsev, A. V. Goltsev, and F. F. Mendes, "Critical phenomena in complex networks," Rev. Mod. Phys. **80**, 1275–1335 (2008).
[9]S. H. Strogatz, "Exploring complex networks," Nature **410**, 268–276 (2001).
[10]Y. Wu, J. Xiao, G. Hu, and M. Zhan, "Synchronizing large number of non-identical oscillators with small coupling," Europhys. Lett. **97**, 40005 (2012).
[11]S. Acharyya and R. E. Amritkar, "Synchronization of coupled nonidentical dynamical systems," Europhys. Lett. **99**, 40005 (2012).
[12]A. Arenas, A. Diaz-Guilera, J. Kurths, Y. Moreno, and C. S. Zhou, "Synchronization in complex networks," Phys. Rep. **469**, 93–153 (2008).
[13]Y. Moreno and A. F. Pacheco, "Synchronization of Kuramoto oscillators in scale-free networks," Europhys. Lett. **68**, 603–609 (2004).
[14]A. Arenas, A. Diaz-Guilera, and C. J. Pérez-Vicente, "Synchronization reveals topological scales in complex networks," Phys. Rev. Lett. **96**, 114102 (2006).
[15]J. Gómez-Gardeñes, Y. Moreno, and A. Arenas, "Paths to Synchronization on Complex Networks," Phys. Rev. Lett. **98**, 034101 (2007).
[16]I. Lodato, S. Boccaletti, and V. Latora, "Synchronization properties of network motifs," Europhys. Lett. **78**, 28001 (2007).
[17]J. Gómez-Gardeñes, S. Gómez, A. Arenas, and Y. Moreno, "Explosive synchronization transitions in scale-free networks," Phys. Rev. Lett. **106**, 128701 (2011).
[18]T. Nishikawa, A. E. Motter, Y. C. Lai, and F. C. Hoppensteadt, "Heterogeneity in oscillator networks: Are smaller worlds easier to synchronize?," Phys. Rev. Lett. **91**, 014101 (2003).
[19]M. Chavez, D. U. Hwang, A. Amann, H. G. E. Hentschel, and S. Boccaletti, "Synchronization is enhanced in weighted complex networks," Phys. Rev. Lett. **94**, 218701 (2005).
[20]C. Zhou, A. E. Motter, and J. Kurths, "Universality in the synchronization of weighted random networks," Phys. Rev. Lett. **96**, 034101 (2006).
[21]D. Gfeller and De Los Rios, "Spectral coarse graining and synchronization in oscillator networks," Phys. Rev. Lett. **100**, 174104 (2008).
[22]C. Zhou and J. Kurths, "Dynamical weights and enhanced synchronization in adaptive complex networks," Phys. Rev. Lett. **96**, 164102 (2006).
[23]F. Sorrentino and E. Ott, "Adaptive synchronization of dynamics on evolving complex networks," Phys. Rev. Lett. **100**, 114101 (2008).
[24]T. Aoki and T. Aoyagi, "Co-evolution of phases and connection strengths in a network of phase oscillators," Phys. Rev. Lett. **102**, 034101 (2009).
[25]R. Gutierrez, A. Amann, S. Assenza, J. Gómez-Gardeñes, V. Latora, and S. Boccaletti, "Emerging meso- and macroscales from synchronization of adaptive networks," Phys. Rev. Lett. **107**, 234103 (2011).
[26]Y. Kuramoto, "Self-entrainment of a population of coupled nonlinear oscillators," Lect. Notes Phys. **39**, 420–422 (1975).
[27]J. A. Acebrón, L. L. Bonilla, C. J. Pérez-Vicente, F. Ritort, and R. Spigler, "The kuramoto model: A simple paradigm for synchronization phenomena," Rev. Mod. Phys. **77**, 137 (2005).
[28]Y. Zhang and Y. Qu, "Remote synchronization of coupled dynamic networks with neutral type neural network nodes," in *Proceedings of the 31st Chinese Control Conference, 2012*, pp. 3407–3412.
[29]J. Gómez-Gardeñes and Y. Moreno, "From scale-free to Erdos-Renyi networks," Phys. Rev. E **73**, 056124 (2006).
[30]The value of $k*$ is set so that the percentage of hubs is around the 20% of the total number of nodes in the network.
[31]V. Nicosia, M. Valencia, M. Chavez, A. Diaz-Guilera, and V. Latora, "Remote synchronization reveals network symmetries anf functional modules," Phys. Rev. Lett. **110**, 174102 (2013).

Chapter 9

Conclusions

We have arrived at the end of this journey, and it is now time to summarize the main achievement obtained and profile the next steps to do. We will start providing an overview of the results obtained throughout the doctorate and of the consequences that these have inside the global scenario of network science. Then, we will focus more in detail on the results obtained by each paper. In this way, at the end of the reading, the reader will achieve both a global and a "local" knowledge about the results obtained. Also, before commenting on the future perspectives, we will spend few words on other works that are not included in this manuscript, but that also form part of the scientific production of the candidate.

9.1. Main results

9.1.1. General achievements

Considering what has been shown so far in the previous chapters, we can try to summarize it and extract some general conclusions about the work exposed in the present thesis. To do so, we will lean on the three main objectives listed in the introduction, using them as a guidance. Concerning objective number 1 (*Extension of known models used to study complex systems to achieve a more realistic description of them*), we have seen how the extension on known models (like that of Stuart Landau), has allowed to observe new phenomena and, in some cases, to update the set of conditions necessary to observe them (like, for example, including a range of velocity inside which cooperation is promoted more), thus enriching the description of previously studied systems. Such kind of work can be included into the more general context of improvement of existing models with the aim of provide a description of complex systems as a whole; since, up to now they have been studied taking into account just one component (kind of interaction) per time, or by considering the dynamics in a separate way albeit, *de facto*, these are usually coupled.

With respect to objective 2 (*Retrieval, analysis, and management of big amount of data originated by real complex systems*), we can say that studying complex systems without the opportunity to make real comparisons between models and real data, reduces such study to a mere mathematical exercise getting rid of every physical meaning. For this reason, whenever possible, we have tried to use in our works as much real data as possible. At the present time, it would also be not very meaningful thinking of simulating systems with sizes of the same order of the real ones using only toy models. The management of big amount of data is becoming, day after day, an unavoidable element to take into account for the comprehension of real system (just think, as an example, to the number of time varying interactions among people occurring every day). Due to the increase in the size of datasets, new ways to store, manage, and analyze efficiently those data have to be found. The scientific community is thus putting a lot of efforts into the re-design of a many algorithms in order to study both topological and dynamical properties of complex systems of increasing size and bearing a greater amount of information encoded within them. In a nutshell, big-data related problems are not just a whim of the scientific community but, rather, a necessary step to make in order to accrue a deeper comprehension of complex systems.

Finally, the achievement of objective 3 (*Study emerging properties and phenomena in systems described using the new frameworks of time-varying and multilevel interactions.*), places us inside the new age of network science (and complex systems one as well). We have seen, in fact, how the inclusion of the time dimension in the description of the interactions dramatically alters the behaviour of systems previously studied under "*static*" conditions. As an example of that, we have seen how changing the time resolution used to study the system changes the emergence of cooperation in the system due to the presence of time correlations in the interactions. If we really want to achieve a better and more realistic description of complex systems, we have to encode the highest possible amount of information in their representation. The appearance of new formalisms of time-varying and multilevel networks, has surely allowed us to make a big leap forward in this sense. Nevertheless, a previous step is necessary: a systematic study of the topological properties (and their emergence or disappearance when one falls back to the "usual" description) of such systems is due. In this sense, a significant contribution to this aspect has been given in one of the works included in this thesis. In fact, in Chap. 6, we have proposed a study of how the topological properties of multilevel networks emerge when these are converted into single layer ones. As concluding remark, we can say that, in general, in the time-varying case (as well as in the multilevel one) the behaviour of well understood dynamical processes (like diffusion/routing, or evolutionary game theory) taking place on top of such kind of topologies is completely different from that observed so far as has been shown in Chaps. 4 and 7. These results, together with others present in the literature, represent an authentic discontinuity with the past, and will demand a lot more of efforts to acquire a complete masterhood about the behaviour of real systems mappable with such formalisms.

Arrived at this point, and keeping in mind this global vision on the results obtained, we are now ready to make a further step forward by analyzing more in detail the results found in each manuscript presented in this thesis.

9.1.2. Achievements of each paper

We are now ready to resume the main achievement obtained in each paper. The aim of this part is double: provide a detailed view of the main results obtained, and pave the road to the future development of the work discussed so far. So, in a glance, we can conclude that:

In Chap. 3, we have seen how the presence of time-varying interactions generated by the movement of the agents, alters the scenario of the onset of cooperation displaying an intermediate region of values of velocity for which the onset of cooperation occurs under less favorable conditions than in the static case. Also, the transition from the fully defective state toward the fully cooperative one depends on the value of the velocity. The reasons behind such phenomenon are rooted in the fact that the rate of creation and deletion of links is controlled by the velocity of agents. When such rate is moderate, cooperators are able to explore the space and clusters of cooperators can form and survive. Also, the group size shows a similar resonance phenomenon although, in this case, the point at which cooperation is enhanced is related to the percolation point of the effective network, *i.e.*, with the point at which the size of the groups is large enough so as to have a macroscopic giant component for the network of contacts.

In Chap. 4, we focused on a different aspect of time-varying interactions. We tried to uncover, in fact, not simply the effects on cooperation originated by time varying interactions *per se*, but rather, what is the role played by time-correlations in the evolution of cooperation. In order to do so, we considered an evolutionary game dynamic on top of a time-varying topology corresponding to real social interactions. The numerical results have shown that cooperation is seriously hindered when agent strategy is updated too frequently with respect to the typical time-scale of agent interaction, and also when real time correlations are present. The reasons behind such behaviour may be found in the correlation between the size of the giant component, and the presence of temporal correlations, such as edge persistence and recurrence, as a function of the aggregation interval.

In Chap. 5, we present a co-evolutionary approach involving two different dynamical processes: evolutionary game theory, and spreading of diseases. Considering a population of agents, and thinking of the act of getting vaccinated voluntarily as cooperating, it is possible to set up a social dilemma involving the vaccination. Starting from the fact that scale-free topologies foster more both the spread of infections and the emergence of cooperation, we can ask ourselves what happens if we put these to processes into competition. We observe that, depending on the circumstances (basically: the quality of the vaccine, its cost, and the probability of being infected) the spread of infection may be overcome by vaccination. This implies the suppression of the infection and the appearance of a macroscopic fraction of immunized individuals providing herd immunization to the whole system. Moreover, we can observe a crossover effect in both the fraction of immunized and infected subjects as a function of infection probability among systems with interactions patterns described by Erdős-Rényi or Barabási-Albert topologies.

In Chap. 6, we try to understand if the topological properties of a multiplex system remain the same, or not, when the layers are progressively projected onto a single one. We are able to find that, in general, the properties displayed by the single layers and the same quantities calculated in the multiplex system are not the same. They are the consequence of an emerging phenomenon intimately related to the multilayer character of the system. In the case of European Airlines multiplex, this is confirmed also by the fact that if we consider layers of different kind, the merging of low-cost and major (national) layers leads to the emergence of qualitatively different aggregate networks. This demonstrates that multiplex nature of many systems cannot be ignored and has deep consequences also on the dynamical processes acting on multilayer networks.

In Chap. 7, following the idea of Chap. 6, we study the robustness of the European airlines under the problem of re-scheduling passengers of a flight that has been cancelled in an airlines multiplex network composed of many layers, each of which corresponds to a different airline. The inclusion of a structure composed of separate layers alters the resilience of the system compared to the case of a single layer network. To account for the multilayer structure, we measure the resilience of the system as a function of the number of passengers that have been successfully re-schedule within the same airline (layer) or using a different one. In particular, the ability to re-schedule passengers strongly depends on the variation of the shortest path length as a function of the number of deleted flights, but also on the capacity tolerance of each flight.

Finally, in Chap. 8, we studied the problem of finding remote synchronization in networks of coupled Stuart-Landau oscillators. Apart from confirming that remote synchronization exists also for topologies different from a star, we observe that remote synchronization constitutes a mechanism anticipating synchronization by physical links in networks with heterogeneous distribution of natural frequencies. Also, we confirm that remote synchronization is not significant for ensembles of phase oscillators, since its main underlying mechanism seems to be the modulation of the amplitude of intermediary nodes allowing information transfer between uncoupled pairs of nodes. In fact, when similar settings are applied to phase oscillators, a different phenomenon is observed, namely, that of *explosive synchronization*.

9.2. Other publications

During the doctorate, the candidate has also worked, and published, on other subjects apart from those appearing in this manuscript. Such works, do not fall within the route followed by this thesis, albeit being related with complex networks as well. Below, we provide a brief comment about the main results found in each of these papers.

The first paper, entitled: "*Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks*" studies the emergence of cooperation in the prisoner's dilemma game in the case of players that can coevolve both their strategy and the rule used to update (change) it. Moreover, even the role played by different underlying topologies in such coevolution is studied, considering network topologies spanning from single-scale random (Erdős-Rényi) networks, to scale-free (Barabási-Albert) ones passing through intermediate ones. The interested reader could look at [145].

The second paper, entitled: "*Urban street networks, a comparative analysis of ten European cities*" belongs to a line of research undertaken prior to the start of the Ph.D, concerning the study of the structural properties of a special kind of spatial networks: those generated from urban street patterns. In this work, in particular, the structural properties, like average path length L, of ten European cities are compared in order to find analogies, regularities, and differences among them. For further information, please consult [146].

The third paper, entitled: "*Information sharing in quantum complex networks*" may be classified within a highly innovative line of research that has come to light very recently, and that is capturing the attention of two separate communities: the one of “network scientists” and that of quantum physicists. In particular, in this work, we study how the presence of an interacting pattern described by a complex network could affect the behaviour of the entanglement entropy. Different network substrates have been used (regular, random, scale-free) in order to study such behaviour and the results are resumed in [147].

9.3. Future perspectives

If the results contained in this thesis represent the “point of arrival” of this *going beyond* process, it is worth spending some time speculating on what could be the *next* steps. Of course, the most intuitive one is to continue following the path taking advantage of coevolutive modelling in order to design more realistic models able to capture more complex scenarios. As an example, one could think of combining synchronization and evolutionary game theory in order to model systems where the fully synchronized state can be achieved only if all the agents cooperate but where two agents that try to get synchronized have to pay a cost to do that.

Another direction worth of being explored is that of considering well known dynamical processes (like, again, synchronization) on top of new kind of topologies like time-varying and multiplex ones. The former case, for example, will allow the time to play a role not only in the evolution of the dynamical state of the nodes but also in the way they interacts among them. The latter, instead, will permit to deal with different kind of interactions at the same time that may exert different influences on the state of the system. With respect to the previous point, the use of new kind of topologies represents more a leap in the dark because we still lack a lot of information due to the novelty of these approaches.

Last, but not least, an interesting direction that could be followed is that offered by the so called “*big data*”. In the last few years, in fact, an extremely higher amount of data have become available to the scientists. The reasons behind the appearance of such big amount of data deal, from one side, with the availability of new techniques to retrieve them (experiments, inquiries, web crawling, and so on so forth) and, from another one, with the explosion of information available in context that did not exist before; think, for example, to the case of the so-called *social networks* or the World Wide Web. Until few years ago platforms like Facebook or applications like FourSquare did not exist and the information available on the friendship relations among people (or on how people move) were significantly smaller that those available nowadays.

Thanks to big data, we will have the opportunity to: make better comparisons between theory and experiment, extend the size of the systems studied toward the thermodynamic limit, and set up ad hoc experiments with the aim of proofing hypothesis that up to now we were able to check only through theory or computation. Of course, we must not forget, that there are still a lot of complex systems that have not been studied by means of complex networks. Those systems can be thought as goldmines because may lead to the formulation of new measures, or the discovery of phenomena that have not seen before.

In conclusion, complex networks are a very powerful tool to unveil the secrets of many different complex systems. Since its birth as a discipline, complex networks science has been able to explain many phenomena but there are still a lot of questions that await for an answer. This is the challenge that network science is called to face.

Acknowledgements
Agradecimientos
Ringraziamenti

Writing the acknowledgements of a thesis is always an authentic ordeal because "*is the only part that people, except for the referees of course, will read*". So, one has to choose very carefully the words he is going to use.

Almost four years have passed since my decision to leave my land and my country to begin an adventure (or I would rather say that it was more a challenge with myself) that could let me grow both under human and professional point of view. Now that I am at the end of this adventure; I can say that, in the end, things have gone better than expected.

During these years many things have happened: some good, and some others not so much. I had to renounce to some things but, in exchange for that, I had obtained others. In short, a lot of water has flowed beneath the bridge since then. I have been under the effect of both homeostasis and transistasis and, because of that, I am a bit the same as always and also different from before. Now, new challenges await me but, before preparing to face them, I believe that I have to give a look back to pay an acknowledgement to all those people that, in a way or another, left a trace in my life and helped me to arrive up to this point. I also want to thank all those people that, instead, tried to push me down and made my journey less pleasant. Sorry guys, try again (maybe you will be luckier next time)!

Now, I would like to spend some words of gratitude towards some special people.

- A mis directores de tesis: Jesús y Sandro. El primero por haber apostado (y confiado) en mí aceptándome como su estudiante y el segundo por la paciencia que tuvo a lo largo de estos años con mi torpeza. La verdad es que definir mi relación con ellos como la de jefe-estudiante sería reductivo. Cuando Jesús estuvo de post-doc en Catania pensé que me hubiera gustado poder hacer el doctorado bajo su dirección. Hoy puedo decir que he cumplido ese deseo y le doy las gracias por todo lo que hizo por mi incluso cuando le tenía gana de matarme por mi pereza o por mis errores. A Sandro les agradezco por haberme echado las broncas un día sí y otro, también con el único objetivo de convertirme en algo mejor de lo que era. En mi mente quedarán clavadas las miles de comidas juntos, los días de esquiar en el Pirineo, las películas en el cine (excepto "Headhunters", por la cual me disculpo públicamente con él), y muchísimas cosas más.

- La mia famiglia. Per loro, avermi visto andar via ed essere stato lontano tutti questi anni (ed ancora di tornare a casa non se ne parla manco per scherzo) non dev'essere stato facile. Li ringrazio per il loro appoggio incondizionato e per avermi permesso di poter inseguire i miei sogni. Una menzione speciale va ai miei cugini Lucio e Gabriella.
  Me gustaría mencionar también a mi "familia española". Gracias a Doña Lola y Don Jesús por haberme acogido como un hijo.

- A los miembros del grupo FENOL de la Universidad de Zaragoza. Tener buenos compañeros de trabajo es un requisito fundamental para poder disfrutar laboralmente. Por eso, gracias a: Fernando, Juanjo, Mari Carmen, Mario, Pedro, Yamir, Alessandro, Carlos, David, Javi, Uta, Ana Elisa, Emanuele, Fernando, Joaquín, Pablo, Rafa y especialmente a Raquel, por haber aguantado todas mis bromas y charlas.

- Un ringraziamento speciale va ad alcuni collaboratori che hanno avuto la sfortuna di dover lavorare con me durante il mio dottorato. Un grazie quindi a: Vito, Roberta (GGIOIA), Enzo, Mattia, Valentina, Manlio, Serafina, Oscar (con el que he compartido tanto comidas como el mantenimiento del cluster) e Giulia.

- Siempre he pensado que el rugby es más que un deporte, es una manera de enfrentarse a la vida. Aquí en Zaragoza he tenido la suerte de formar parte de un club formado por personas extraordinarias, una verdadera FAMILIA (no consigo encontrar ninguna otra palabra). La cosa que me ha dejado más sorprendido ha sido la manera en que mis compañeros me han incorporado dentro del club. He estado aquí jugando solo dos años, y me parecía como si formase parte del club desde toda mi vida. A lo largo de mi carrera como deportista puedo asegurar que no había sentido muchas veces este sentimiento de "pertenencia". Gracias entonces a tod@s los miembros del Club Deportivo Universitario Rugby Zaragoza. En particular, quiero darle las gracias a algunas personas especiales: Juki (si no te hubiese conocido no estaría escribiendo estos agradecimientos), Caco (por el ejemplo que das con tu actitud), Andrés (por haber confiado en mí) y (last but not least) Waldo (un auténtico "hermano de armas").

- La vita mi ha portato lontano dalla mia terra ma non ho dimenticato quelle persone che sono rimaste "a casa" e che non mancano mai di farmi sentire il loro appoggio e la loro amicizia anche a distanza e nonostante non possa incontrarmi spesso con loro. Vorrei fare un ringraziamento speciale a: Paola, Cristina e Luca, Peppe Angilella, Stefano, Antonio e Lory, Alex (meglio noto come Maverick), Rauni, Daniela ed Emiliano, Francesca e Giacomo.

- Una menzione speciale va ai membri del **quintetto etilo-lescano**. Come sempre, sono l'ultimo del gruppo ad arrivare al traguardo ma so che non abbiamo mai dato troppo peso a queste formalità ma piuttosto al bisogno di incrociare periodicamente le nostre strade. Nonostante tutti i cazzi della vita so che potrò sempre contare su di voi e sul fatto che almeno una volta l'anno in qualche parte del mondo si celebrerá una reunion. Un grazie speciale quindi a Thänô Tchaikovsky (futuro truffatore delegato dell'Eni), Salvuccio (Tombeur de femmes in quel di Paris), Gabriele (simpatico omino di pastafrolla) e Danilo (ti vogliamo bene uguale anche se sei un hipster).

- Si en el curro soy un torpe y un vago, en la vida común soy incluso peor. Aguantarme, con lo pesado, soso, borde y friqui que soy; y con todas las manías (in primis aquella de la limpieza) que tengo debe de ser verdaderamente una pesadilla. Por lo tanto, quiero agradecer aquellos "santos" que en estos años han compartido hogar conmigo sin echarme o tirarme una bala en la cabeza. En modo especial: Terruccio (my son!), Lety (nunca olvidaré tus risas), Annuzza (mi alemana favorida), Mar (casi me matabas incendiandos el piso) y Mer (por sus "cumplidos cariñosos").

- Il penultimo ringraziamento va a lei, Angela. Per lei sono stati anni molto difficili avendo dovuto accettare la mia decisione unilaterale di emigrare all'estero. La distanza mi ha costretto a non poter stare al suo fianco quando magari avrebbe avuto bisogno di me, e del mio sostegno, e per questo le chiedo scusa. Spesso mi sono chiesto cosa le ha fornito la forza con la quale ha atteso con pazienza ogni nostro incontro, oppure sopportato ogni mia partenza. Ho scelto da tempo che è lei la persona con cui voglio spendere il resto dei miei giorni, ed avevo persino immaginato ad una proposta "speciale" da farle il giorno in cui fossi diventato dottore. Purtroppo, nella vita non sempre tutto va come vorremmo e mi scuso quindi anche per averle "rovinato la sorpresa" avendo dovuto accelerare i tempi. Non basterebbe una tesi per elencare tutti i motivi per cui dovrei ringraziarla ma se volesse farmi il piacere di diventare mia moglie magari un pò di tempo per elencarglieli poco a poco penso che riesco a trovarlo. Vorrei concludere dicendo che senza di lei questa tesi sarebbe comunque stata scritta ma sicuramente ci sarebbero state molte meno risate nella mia vita. Ti amo.

- My last acknowledgement goes to Spain, its people, colors, and flavours. Here I have learned what I believe is the true meaning of the word *fiesta*. After three and a half years, I still consider myself (and I always be) as an Italian but I am now also a bit Spanish (or, more precisely: *¡Aragonés!*).

Before letting you go, I want to reward you for your patience leaving some citations that have meant a lot during these years.

> On the theoretical side, small-world networks are turning out to be a Rorschach test – different scientists see different problems here, depending on their disciplines.
>
> Steven H. Strogatz - Nature, 2001.

> [...] En este momento yo, de verdad, no tendría ningún inconveniente en morir. No tengo ningun interés en seguir. Lo que pasa es que quiero mi mujer, sé que le hago falta, sé que le conviene y estaré todo el tiempo que le haga falta. Pero hoy ya hay la serie de molestias que tiene la vida de un viejo sordo, medio ciego y con otros inconvenientes que no se le voy a describir. Todos estos inconvenientes ya todas las mañanas: me tengo que levantar, poner la camisa y tal, me tengo que poner las molas, me tengo que pulir los ojos, las orejas. Todo eso no me interesa ya porqué además ya he visto el espectáculo y me importa tres pepinos, ¿verdad? Pero, yo tengo que vivir porqué: se habla mucho del derecho a la vida, pero es que hay más. Hay el deber de vivirla. Hemos recibido de la vida una vida; pues, ¡vamos a vivirla!
>
> José Luis Sampedro – entrevistado por Jordi Evolé 29/01/2012

> Life does not end with your death. What will survive of you is the message you send to other people. This is the immortality of a human being.
>
> Rita Levi Montalcini

It takes a heap o' livin' in a house t' make it home,
A heap o' sun an' shadder, an' ye sometimes have t' roam
Afore ye really 'preciate the things ye lef' behind,
An' hunger fer 'em somehow, with 'em allus on yer mind.
It don't make any differunce how rich ye get t' be,
How much yer chairs an' tables cost, how great yer luxury;
It ain't home t' ye, though it be the palace of a king,
Until somehow yer soul is sort o' wrapped round everything.

It takes a lot of living to make a house a home,
It doesn't make any difference how rich you get to be
How much your chairs and tables cost, how great your luxury;
It isn't home to you, though it be the palace of a king,
Until somehow your soul is wrapped round everything.

Edgar Albert Guest – Home

Bibliography

[1] M. Gell-Mann. *The Quark and the Jaguar: Adventures in the Simple and the Complex.* Freeman, San Francisco (1994).

[2] R. Badii and A. Politi. *Complexity: Hierarchical Structures and Scaling in Physics.* Cambridge Nonlinear Science Series. Cambridge University Press, Cambridge, UK (1999).

[3] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, and D. Hwang. *Complex networks: Structure and dynamics.* Physics Reports **424**(4-5), 175 (2006).

[4] M. E. J. Newman. *Networks: An Introduction.* Oxford University Press, London (2010).

[5] M. S. John. *Evolution and the Theory of Games.* Cambridge University Press, Cambridge, UK (1982).

[6] R. M. Anderson and R. M. May. *Infectious diseases of humans: Dynamics and control.*, volume 15. Oxford: Oxford University Press, Oxford (1991).

[7] R. Pastor-Satorras and A. Vespignani. *Epidemic Spreading in Scale-Free Networks.* Physical Review Letters **86**(14), 3200 (2001).

[8] D. Balcan, V. Colizza, B. Gonçalves, H. Hu, J. J. Ramasco, and A. Vespignani. *Multiscale mobility networks and the spatial spreading of infectious diseases.* Proceedings of the National Academy of Sciences of the United States of America **106**(51), 21484 (2009).

[9] D. Stauffer and A. Aharony. *Introduction To Percolation Theory.* CRC Press, USA (1992).

[10] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin. *Resilience of the Internet to Random Breakdowns.* Physical Review Letters **85**(21), 4626 (2000).

[11] R. Cohen, K. Erez, D. Ben-Avraham, and S. Havlin. *Breakdown of the Internet under Intentional Attack.* Physical Review Letters **86**(16), 3682 (2001).

[12] S. V. Buldyrev, R. Parshani, G. Paul, H. E. Stanley, and S. Havlin. *Catastrophic cascade of failures in interdependent networks.* Nature **464**(7291), 1025 (2010).

[13] J. Gómez-Gardeñes, I. Reinares, A. Arenas, and L. M. Floría. *Evolution of cooperation in multiplex networks.* Scientific reports **2**, 620 (2012).

[14] S. Gómez, A. Díaz-Guilera, J. Gómez-Gardeñes, C. J. Pérez-Vicente, Y. Moreno, and A. Arenas. *Diffusion Dynamics on Multiplex Networks.* Physical Review Letters **110**(2), 028701 (2013).

[15] P. Holme and J. Saramäki. *Temporal networks.* Physics Reports **519**(3), 97 (2012).

[16] M. Barthélemy. *Spatial networks.* Physics Reports **499**(1-3), 1 (2011).

[17] S. Fortunato. *Community detection in graphs.* Physics Reports **486**(3-5), 75 (2010).

[18] C. P. Roca, J. a. Cuesta, and A. Sánchez. *Evolutionary game theory: Temporal and spatial effects beyond replicator dynamics.* Physics of life reviews **6**(4), 208 (2009).

[19] A. Arenas, A. Díaz-Guilera, J. Kurths, Y. Moreno, and C. Zhou. *Synchronization in complex networks.* Physics Reports **469**(3), 93 (2008).

[20] M. C. González, C. Hidalgo, and A.-L. Barabási. *Understanding individual human mobility patterns.* Nature **453**(7196), 779 (2008).

[21] A. Bazzani, B. Giorgini, S. Rambaldi, R. Gallotti, and L. Giovannini. *Statistical laws in urban mobility from microscopic gps data in area of Florence.* Journal of Statistical Mechanics **2010**, P05001 (2010).

[22] L. D. Benedictis and L. Tajoli. *The World Trade Network.* The World Economy **34**, 1417 (2011).

[23] M. De Domenico, A. Solé-Ribalta, E. Cozzo, M. Kivelä, Y. Moreno, M. A. Porter, S. Gómez, and A. Arenas. *Mathematical Formulation of Multilayer Networks.* Physical Review X **3**(4), 041022 (2013).

[24] S. Wasserman and K. Faust. *Social network analysis : methods and applications*, volume 24. Cambridge University Press, New York (1994).

[25] M. Kurant and P. Thiran. *Layered Complex Networks.* Physical Review Letters **96**(13), 138701 (2006).

[26] P. J. Mucha, T. Richardson, K. Macon, M. a. Porter, and J.-P. Onnela. *Community structure in time-dependent, multiscale, and multiplex networks.* Science (New York, N.Y.) **328**(5980), 876 (2010).

[27] B. Bollobas. *Random Graphs.* Cambridge University Press, London, 2nd edition (2001).

[28] P. Diaconis and B. Bollobas. *Modern Graph Theory*, volume 95 of *Graduate Texts in Mathematics.* Springer Verlag, New York (1998).

[29] D. West. *Introduction to Graph Theory.* Pearson, London, 2nd edition (2000).

[30] J. B. U. Murty. *Graph Theory*, volume 244 of *Graduated Texts in Mathematics.* Springer, London (UK) (2008).

[31] A. Barrat, M. Barthélemy, R. Pastor-Satorras, and A. Vespignani. *The architecture of complex weighted networks.* Proceedings of the National Academy of Sciences of the United States of America **101**(11), 3747 (2004).

[32] E. Katifori, G. J. Szöllősi, and M. O. Magnasco. *Damage and Fluctuations Induce Loops in Optimal Transport Networks.* Physical Review Letters **104**, 048704 (2010).

[33] M. Newman and M. Girvan. *Finding and evaluating community structure in networks.* Physical Review E **69**(2), 026113 (2004).

[34] M. Marchiori and V. Latora. *Harmony in the Small World.* Physica A **285**, 539 (2000).

[35] V. Latora and M. Marchiori. *Efficient Behavior of Small-World Networks.* Physical Review Letters **87**(19), 198701 (2001).

[36] M. Granovetter. *The strength of weak ties.* American J. Sociology **78**, 1360 (1973).

[37] M. E. J. Newman, S. H. Strogatz, and D. J. Watts. *Random graphs with arbitrary degree distributions and their applications.* Physical Review E **64**(2), 026118 (2001).

[38] L. A. Amaral, A. Scala, M. Barthelemy, and H. E. Stanley. *Classes of small-world networks.* Proceedings of the National Academy of Sciences of the United States of America **97**(21), 11149 (2000).

[39] M. Boguñá and R. Pastor-Satorras. *Class of correlated random networks with hidden variables.* Physical Review E **68**(3), 036112 (2003).

[40] M. Newman. *Assortative Mixing in Networks.* Physical Review Letters **89**(20), 208701 (2002).

[41] S. Zhou and R. Mondragon. *The Rich-Club Phenomenon in the Internet Topology.* IEEE Communications Letters **8**(3), 180 (2004).

[42] V. Colizza, A. Flammini, M. A. Serrano, and A. Vespignani. *Detecting rich-club ordering in complex networks.* Nature Physics **2**(2), 110 (2006).

[43] V. Latora and M. Marchiori. *Economic small-world behavior in weighted networks.* The European Physical Journal B - Condensed Matter **32**(2), 249 (2003).

[44] R. Milo, S. Shen-Orr, S. Itzkovitz, N. Kashtan, D. Chklovskii, and U. Alon. *Network motifs: simple building blocks of complex networks.* Science (New York, N.Y.) **298**(5594), 824 (2002).

[45] I. Lodato, S. Boccaletti, and V. Latora. *Synchronization properties of network motifs.* Europhysics Letters **78**(2), 28001 (2007).

[46] R. Sinatra, D. Condorelli, and V. Latora. *Networks of Motifs from Sequences of Symbols.* Physical Review Letters **105**(17), 178702 (2010).

[47] L. Kovanen, K. Kaski, J. Kertész, and J. Saramäki. *Temporal motifs reveal homophily, gender-specific patterns, and group talk in call sequences.* Proceedings of the National Academy of Sciences of the United States of America **110**(45), 18070 (2013).

[48] S. S. Shen-Orr, R. Milo, S. Mangan, and U. Alon. *Network motifs in the transcriptional regulation network of Escherichia coli.* Nature Genetics **31**, 64 (2002).

[49] R. Milo, S. Itzkovitz, N. Kashtan, R. Levitt, S. Shen-Orr, I. Ayzenshtat, M. Sheffer, and U. Alon. *Superfamilies of evolved and designed networks.* Science (New York, N.Y.) **303**(5663), 1538 (2004).

[50] S. Mangan and U. Alon. *Structure and function of the feed-forward loop network motif.* Proceedings of the National Academy of Sciences of the United States of America **100**, 11980 (2003).

[51] N. Kashtan, S. Itzkovitz, R. Milo, and U. Alon. *Efficient sampling algorithm for estimating subgraph concentrations and detecting network motifs.* Bioinformatics (Oxford, England) **20**(11), 1746 (2004).

[52] E. Lieberman, C. Hauert, and M. A. Nowak. *Evolutionary dynamics on graphs.* Nature **433**(7023), 312 (2005).

[53] F. Débarre, C. Hauert, and M. Doebeli. *Social evolution in structured populations.* Nature communications **5**, 3409 (2014).

[54] D. J. Watts and S. H. Strogatz. *Collective dynamics of 'small-world' networks.* Nature **393**(6684), 440 (1998).

[55] P. Erdős and A. Rényi. *On the evolution of random graphs.* Evolution **5**(1), 17 (1960).

[56] D. D. S. Price. *A general theory of bibliometric and other cumulative advantage processes.* Journal of the American Society for Information Science **27**(5), 292 (1976).

[57] M. Molloy and B. Reed. *A critical point for random graphs with a given degree sequence.* Random Structures & Algorithms **6**, 161 (1995).

[58] M. Molloy and B. Reed. *The size of the largest component of a random graph on a fixed degree sequence.* Combinatorics Probability and Computing **7**, 295 (1998).

[59] J. Travers and S. Milgram. *An Experimental Study of the Small World Problem.* Sociometry **32**(4), 425 (1969).

[60] A.-L. Barabási and R. Albert. *Emergence of Scaling in Random Networks.* Science **286**(5439), 509 (1999).

[61] F. Liljeros, C. R. Edling, L. A. N. Amaral, H. Stanley, and Y. Aberg. *The web of human sexual contacts.* Nature **411**(6840), 907 (2001).

[62] P. Erdős and A. Rényi. *On random graphs.* Publicationes Mathematicae Debrecen **6**, 290 (1959).

[63] Z. Burda, J. Jurkiewicz, and A. Krzywicki. *Statistical mechanics of random graphs.* Physica A: Statistical Mechanics and its Applications **344**(1-2), 56 (2004).

[64] J. Park and M. E. J. Newman. *Statistical mechanics of networks.* Physical Review E **70**(6), 066117 (2004).

[65] J. Dall and M. Christensen. *Random geometric graphs.* Physical Review E **66**(1), 016121 (2002).

[66] M. Penrose. *Random Geometric Graphs.* Oxford scholarship online. Oxford University Press, Oxford, UK (2003).

[67] A. Antonioni and M. Tomassini. *Degree correlations in random geometric graphs.* Physical Review E **86**(3), 1 (2012).

[68] S. N. Dorogovtsev, J. F. F. Mendes, and A. N. Samukhin. *Structure of Growing Networks with Preferential Linking.* Physical Review Letters **85**, 4633 (2000).

[69] G. Caldarelli, A. Capocci, P. De Los Rios, and M. Muñoz. *Scale-Free Networks from Varying Vertex Intrinsic Fitness.* Physical Review Letters **89**(25), 258702 (2002).

[70] G. D'Agostino and A. Scala. *Networks of Networks: The Last Frontier of Complexity.* Understanding Complex Systems. Springer International Publishing, Cham (2014).

[71] J. Gao, S. V. Buldyrev, H. Stanley, and S. Havlin. *Networks formed from interdependent networks.* Nature Physics **8**(1), 40 (2011).

[72] A. Cardillo, J. Gómez-Gardeñes, M. Zanin, M. Romance, D. Papo, F. del Pozo, and S. Boccaletti. *Emergence of network features from multiplexity.* Scientific reports **3**, 1344 (2013).

[73] J. Y. Kim and K.-I. Goh. *Coevolution and Correlated Multiplexity in Multiplex Networks.* Physical Review Letters **111**(5), 058702 (2013).

[74] G. Bianconi. *Statistical mechanics of multiplex networks: Entropy and overlap.* Physical Review E **87**(6), 062806 (2013).

[75] V. Nicosia, G. Bianconi, V. Latora, and M. Barthélemy. *Growing Multiplex Networks.* Physical Review Letters **111**(5), 058701 (2013).

[76] A. Solé-Ribalta, M. De Domenico, N. E. Kouvaris, A. Díaz-Guilera, S. Gómez, and A. Arenas. *Spectral properties of the Laplacian of multiplex networks.* Physical Review E **88**(3), 032807 (2013).

[77] F. Battiston, V. Nicosia, and V. Latora. *Structural measures for multiplex networks.* Physical Review E **89**(3), 032804 (2014).

[78] M. Kivelä, A. Arenas, M. Barthelemy, J. P. Gleeson, Y. Moreno, and M. A. Porter. *Multilayer Networks.* ArXiv e-prints (2013), `1309.7233`.

[79] J. Gao, S. V. Buldyrev, S. Havlin, and H. E. Stanley. *Robustness of a Network of Networks.* Physical Review Letters **107**(19), 195701 (2011).

[80] A. Saumell-Mendiola, M. A. Serrano, and M. Boguñá. *Epidemic spreading on interconnected networks.* Physical Review E **86**(2), 026106 (2012).

[81] E. Cozzo, M. Kivelä, M. De Domenico, A. Solé, A. Arenas, S. Gómez, M. A. Porter, and Y. Moreno. *Clustering Coefficients in Multiplex Networks*. ArXiv e-prints (2013), `1307.6780`.

[82] M. De Domenico, A. Sole, S. Gomez, and A. Arenas. *Random Walks on Multiplex Networks*. ArXiv e-prints (2013), `1306.0519`.

[83] A. Gautreau, A. Barrat, and M. Barthélemy. *Microdynamics in stationary complex networks*. Proceedings of the National Academy of Sciences of the United States of America **106**(22), 8847 (2009).

[84] P. Holme. *Network reachability of real-world contact sequences*. Physical Review E **71**, 046119 (2005).

[85] N. Eagle and A. (Sandy) Pentland. *Reality Mining: Sensing Complex Social Systems*. Personal Ubiquitous Comput. **10**(4), 255 (2006).

[86] A. Clauset and N. Eagle. *Persistence and periodicity in a dynamic proximity network*. DIMACS Workshop on Computational Methods for Dynamic Interaction Networks (2007).

[87] J. Tang, S. Scellato, M. Musolesi, C. Mascolo, and V. Latora. *Small-world behavior in time-varying graphs*. Physical Review E **81**(5), 055101 (2010).

[88] J. Saramäki, M. Kivelä, J.-P. Onnela, K. Kaski, and J. Kertész. *Generalizations of the clustering coefficient to weighted complex networks*. Physical Review E **75**(2), 027105 (2007).

[89] D. Kempe, J. Kleinberg, and A. Kumar. *Connectivity and Inference Problems for Temporal Networks*. Journal of Computer and System Sciences **64**(4), 820 (2002).

[90] P. Holme, C. R. Edling, and F. Liljeros. *Structure and time evolution of an Internet dating community*. Social Networks **26**(2), 155 (2004).

[91] R. K. Pan and J. Saramäki. *Path lengths, correlations, and centrality in temporal networks*. Physical Review E **84**(1), 016105 (2011).

[92] E. Cheng, J. W. Grossman, and M. J. Lipman. *Time-stamped graphs and their associated influence digraphs*. Discrete Applied Mathematics **128**(2-3), 317 (2003).

[93] J. Moody. *The Importance of Relationship Timing for Diffusion*. Social Forces **81**(1), 25 (2002).

[94] A.-L. Barabási. *The origin of bursts and heavy tails in human dynamics.* Nature **435**(7039), 207 (2005).

[95] J. Iribarren and E. Moro. *Impact of Human Activity Patterns on the Dynamics of Information Diffusion.* Physical Review Letters **103**(3), 038702 (2009).

[96] Y. Wu, C. Zhou, J. Xiao, J. Kurths, and H. J. Schellnhuber. *Evidence for a bimodal distribution in human communication.* Proceedings of the National Academy of Sciences **107**(44), 18803 (2010).

[97] R. D. Malmgren, D. B. Stouffer, A. S. L. O. Campanharo, and L. A. N. Amaral. *On Universality in Human Correspondence Activity.* Science **325**(5948), 1696 (2009).

[98] K.-I. Goh and A.-L. Barabási. *Burstiness and memory in complex systems.* EPL (Europhysics Letters) **81**(4), 48002 (2008).

[99] N. Perra, B. Gonçalves, R. Pastor-Satorras, and A. Vespignani. *Activity driven modeling of time varying networks.* Scientific reports **2**, 469 (2012).

[100] B. Ribeiro, N. Perra, and A. Baronchelli. *Quantifying the effect of temporal resolution on time-varying networks.* Scientific Reports **3**, 1 (2013).

[101] P. Grassberger. *On the critical behavior of the general epidemic process and dynamical percolation.* Mathematical Biosciences **63**(2), 157 (1983).

[102] D. Daley, J. Gani, and J. Gani. *Epidemic Modelling: An Introduction.* Cambridge Studies in Mathematical Biology. Cambridge University Press, Cambridge, UK (2001).

[103] J. Murray. *Mathematical Biology.* Springer-Verlag, Berlin, 3rd edition (2007).

[104] R. Pastor-Satorras and A. Vespignani. *Epidemic dynamics and endemic states in complex networks.* Physical Review E **63**(6), 066117 (2001).

[105] V. Colizza and A. Vespignani. *Invasion Threshold in Heterogeneous Metapopulation Networks.* Physical Review Letters **99**, 148701 (2007).

[106] C. Poletto, M. Tizzoni, and V. Colizza. *Human mobility and time spent at destination: impact on spatial epidemic spreading.* Journal of theoretical biology **338**, 41 (2013).

[107] A. Vespignani and *et al.* *The Global Epidemic and Mobility Model project homepage.* Available at: http://www.gleamviz.org/.

[108] H. C. Huygens. *Horologium Oscillatorium.* Apud F., Parisiis, France (1673).

[109] A. T. Winfree. *Biological rhythms and the behavior of populations of coupled oscillators.* Journal of Theoretical Biology **16**(1), 15 (1967).

[110] S. Strogatz and R. Kronauer. *Circadian wake-maintenance zones and insomnia in man.* Sleep Research **14**, 219 (1985).

[111] Y. Kuramoto. *Chemical Oscillations, Waves, and Turbulence.*, volume 39. Springer-Verlag, New York, New York, USA (1984).

[112] J. Acebrón, L. Bonilla, C. Pérez Vicente, F. Ritort, and R. Spigler. *The Kuramoto model: A simple paradigm for synchronization phenomena.* Reviews of Modern Physics **77**(1), 137 (2005).

[113] S. H. Strogatz. *From Kuramoto to Crawford: exploring the onset of synchronization in populations of coupled oscillators.* Physica D: Nonlinear Phenomena **143**(1-4), 1 (2000).

[114] I. Hermoso de Mendoza, L. A. Pachón, J. Gómez-Gardeñes, and D. Zueco. *The Quantum Kuramoto Model.* ArXiv e-prints (2013), `1309.3972`.

[115] H. Hong, M. Y. Choi, and B. J. Kim. *Synchronization on small-world networks.* Physical Review E **65**, 026139 (2002).

[116] Y. Moreno and A. F. Pacheco. *Synchronization of Kuramoto oscillators in scale-free networks.* Europhys. Lett. **68**(4), 603 (2004).

[117] T. Ichinomiya. *Frequency synchronization in a random oscillator network.* Physical Review E **70**, 026116 (2004).

[118] J. G. Restrepo, E. Ott, and B. R. Hunt. *Onset of synchronization in large networks of coupled oscillators.* Physical Review E **71**, 036151 (2005).

[119] S. N. Dorogovtsev, A. V. Goltsev, and J. F. F. Mendes. *Critical phenomena in complex networks.* Rev. Mod. Phys. **80**, 1275 (2008).

[120] J. Gómez-Gardeñes, Y. Moreno, and A. Arenas. *Paths to Synchronization on Complex Networks.* Physical Review Letters **98**(3), 034101 (2007).

[121] A. Pikovsky, M. Rosenblum, and J. Kurths. *Synchronization: A Universal Concept in Nonlinear Sciences.* Cambridge University Press, Cambridge (2003).

[122] A. E. Motter, C. Zhou, and J. Kurths. *Network synchronization, diffusion, and the paradox of heterogeneity.* Physical Review E **71**, 016116 (2005).

[123] C. Darwin. *On the Origin of Species by Means of Natural Selection.* Murray, London. Or the Preservation of Favored Races in the Struggle for Life (1859).

[124] E. Pennisi. *Origins. On the origin of cooperation.* Science (New York, N.Y.) **325**(5945), 1196 (2009).

[125] M. A. Nowak. *Five rules for the evolution of cooperation.* Science (New York, N.Y.) **314**(5805), 1560 (2006).

[126] J. Hofbauer and K. Sigmund. *Evolutionary Games and Population Dynamics*, volume 273. Cambridge University Press, Cambridge, UK (1998).

[127] M. A. Nowak and K. Sigmund. *Evolutionary Dynamics of Biological Games.* Science **303**(5659), 793 (2004).

[128] C. Hauert and M. Doebeli. *Spatial structure often inhibits the evolution of cooperation in the snowdrift game.* Nature **428**, 643 (2004).

[129] M. A. Nowak and R. M. May. *Evolutionary games and spatial chaos.* Nature **359**(6398), 826 (1992).

[130] M. W. Macy and A. Flache. *Learning dynamics in social dilemmas.* Proceedings of the National Academy of Sciences of the United States of America **99**, 7229 (2002).

[131] F. C. Santos, J. M. Pacheco, and T. Lenaerts. *Evolutionary dynamics of social dilemmas in structured heterogeneous populations.* Proceedings of the National Academy of Sciences of the United States of America **103**(9), 3490 (2006).

[132] P. W. Anderson. *More Is Different.* Science **177**(4047), 393 (1972).

[133] A. Szolnoki, M. Perc, and G. Szabó. *Topology-independent impact of noise on cooperation in spatial public goods games.* Physical Review E **80**, 056109 (2009).

[134] M. Archetti and I. Scheuring. *Review: Game theory of public goods in one-shot social dilemmas without assortment.* Journal of Theoretical Biology **299**(0), 9. Evolution of Cooperation (2012).

[135] A. Szolnoki and M. Perc. *Group-size effects on the evolution of cooperation in the spatial public goods game.* Physical Review E **84**, 047102 (2011).

[136] G. Szabó and A. Szolnoki. *Cooperation in spatial prisoner's dilemma with two types of players for increasing number of neighbors.* Physical Review E **79**, 016106 (2009).

[137] J. H. Kagel and A. E. Roth. *"The Handbook of Experimental Economics".* Princeton University Press, Princeton (1995).

[138] G. Hardin. *The tragedy of the commons.* Science (New York, N.Y.) **162**(3859), 1243 (1968).

[139] G. Abramson and M. Kuperman. *Social games in a social network.* Physical Review E **63**(3), 030901 (2001).

[140] F. Santos and J. Pacheco. *Scale-Free Networks Provide a Unifying Framework for the Emergence of Cooperation.* Physical Review Letters **95**(9), 098104 (2005).

[141] F. C. Santos, M. D. Santos, and J. M. Pacheco. *Social diversity promotes the emergence of cooperation in public goods games.* Nature **454**(7201), 213 (2008).

[142] M. Perc and A. Szolnoki. *Coevolutionary games–a mini review.* Bio Systems **99**(2), 109 (2010).

[143] M. Perc, J. Gómez-Gardeñes, A. Szolnoki, L. M. Floría, and Y. Moreno. *Evolutionary dynamics of group interactions on structured populations: a review.* Journal of The Royal Society Interface **10**(80), 20120997 (2013).

[144] J. Gómez-Gardeñes, M. Campillo, L. Floría, and Y. Moreno. *Dynamical Organization of Cooperation in Complex Topologies.* Physical Review Letters **98**(10), 108103 (2007).

[145] A. Cardillo, J. Gómez-Gardeñes, D. Vilone, and A. Sánchez. *Co-evolution of strategies and update rules in the prisoner's dilemma game on complex networks.* New Journal of Physics **12**(10), 103034 (2010).

[146] E. Strano, M. Viana, L. da Fontoura Costa, A. Cardillo, S. Porta, and V. Latora. *Urban street networks, a comparative analysis of ten European cities.* Environment and Planning B: Planning and Design **40**(6), 1071 (2013).

[147] A. Cardillo, F. Galve, D. Zueco, and J. Gómez-Gardeñes. *Information sharing in quantum complex networks*. Physical Review A **87**(5), 052312 (2013).