Accepted Manuscript

The basic principles of topology-dynamics relations in networks: An empirical approach

Havva Kohestani, Mahbubeh Totonkuban, Luisa Di Paola, Virginia Todde, Alessandro Giuliani

PII: S0378-4371(18)30589-2
DOI: https://doi.org/10.1016/j.physa.2018.05.045
Reference: PHYSA 19585

To appear in: *Physica A*

Received date : 14 July 2017
Revised date : 5 March 2018

Please cite this article as: H. Kohestani, M. Totonkuban, L. Di Paola, V. Todde, A. Giuliani, The basic principles of topology-dynamics relations in networks: An empirical approach, *Physica A* (2018), https://doi.org/10.1016/j.physa.2018.05.045

THE BASIC PRINCIPLES OF TOPOLOGY-DYNAMICS RELATIONS IN NETWORKS: AN EMPIRICAL APPROACH

Havva Kohestani[1], Mahbubeh Totonkuban[2], Luisa Di Paola[3,*], Virginia Todde[4], Alessandro Giuliani[4]

1. Department of Physics and Center for Molecular Study of Condensed Soft Matter, Illinois Institute of Technology, 3440 S Dearborn St., Chicago, IL 60616, USA
2. Department of Chemistry, Faculty of Basic Sciences, Payame Noor University of Isfahan, Iran
3. Unit of Chemical-Physics Fundamentals in Chemical Engineering, Department of Engineering, Università Campus Bio-Medico di Roma, via Álvaro del Portillo 21, 00128 Rome, Italy
4. Environment and Health Department, Istituto Superiore di Sanità, Viale Regina Elena 299, 00161, Rome, Italy

[*]Corresponding Author: email: l.dipaola@unicampus.it

Abstract

Inferring functional properties of a node from its topological role in the network is of utmost importance in many application fields. In this work, we try to give some hints on this issue focusing on the discovery of empirical topology/function rules located at different degrees of generalization going from a population of heterogeneous biological networks to a single wiring architecture.

Here we report a simulation approach to this problem based on the analysis of the relationship between dissipation kinetics across the network of a perturbative stimulus applied on a single node and its topological properties. The empirical character of the study prompted us to rely on real biological networks instead of 'ideal architectures' and to adopt a statistical attitude with a relatively low number of affected nodes for each network that is much more realistic in real-world biological applications. The analysis allowed us to recognize some general topology-dynamics correlation rules variously modulated by global wiring architecture.

Keywords: Networks dynamics, perturbative methods, signal transmission, biological networks.

Introduction

As aptly pointed out by Nicosia et al.[1] ''Networks are the fabric of complex systems''. Concepts including organized complexity[2], the middle way[3], and mesoscopic systems[4] all revolve around network paradigms.

Apparently disparate investigation fields – from protein science[5] to psychiatry[6]- share the principle that natural sciences must look for a unification frame not by investigating the deep structure of matter, but by exploiting the phenomenological consequences of shared organization rules governing the relations (interactions, correlations, ...) between the constitutive elements of a system[7]. As a matter of fact, in their seminal work[3], Laughlin and colleagues identify the frontiers of basic science in '*the search for the existence and universality of such rules, the proof or disproof of organizing principles appropriate to the mesoscopic domain'.*

This research program is particularly suited for biological sciences where the overwhelming role of boundary conditions with respect to fundamental laws and the presence of a plethora of equivalent solutions stemming from the complexity of the systems [8] make the classical 'differential equations style' of hard sciences too demanding.

In 1952, Bernard Tellegen[9] developed a conservation theorem that is valid for all network systems independently of their nature. This theorem motivates the quest for 'network laws' only stemming from wiring architectures.

The link existing between network topology and dynamics is still loosely understood[10], and inferring the dynamical behavior of a complex system from its structure still remains a cutting-edge task'[12]. Szalay and Csermely [13], by means of their 'Turbine' software, offered a precious operative tool for investigating the modulation exerted by network architecture on the consequences of a local (single node) perturbation. The basic idea behind this model is that intensive physical variables (e.g. temperature) tend to perform an equalization-like dynamics behaving like communicating vessels. Turbine allows the perturbation of any node of the network by an instantaneous pulse or by a fill-in mode. In the first mode (the one we adopted), a single node is displaced from its initial equilibrium value by an instantaneous perturbation that in turn is dissipated by both a built-in fixed decay rate and the spread of the perturbation across the network. Under the hypothesis that every affected node dissipates an equal amount of energy in each simulation step, the equations of the communicating vessels model suggest that when more nodes are affected by a given perturbation, the faster the perturbation becomes dissipated [13]. This implies that the time of decay (silencing time) is inversely proportional to the 'centrality' of the perturbed node in the network. Peripheral nodes have a long silencing time, while central nodes show a very fast decay. This simple paradigm creates an immediate topology-dynamics link (albeit confined to the special case of local perturbations).

We apply the Turbine paradigm to a collection of biological networks with different functionalities (metabolic networks, protein contact networks, neural networks; see Methods). We have seen in our previous work[14] that this dataset allows for a highly significant structure-function relationship on a pure topological (static) basis, justifying the effort to go more in depth into the mechanistic basis of this relationship.

The application of the Turbine approach to different network architectures allows us to shift from the 'entire network' to the 'single node' level of analysis and to identify the general principles linking topology to dynamics.

Turbine characterizes the dissipation of perturbation by two variables: the Extent (the number of nodes sensing the perturbation), Time (the number of simulation steps after which perturbed nodes come back to their initial position). Here we add another descriptor named Maxdis, which is the maximal distance in the number of links traveled by the perturbation. The first two variables were already analyzed by Szalay and Csermely[13] and account for the 'perturbation centrality' paradigm, whereas Maxdis is related to percolation properties of the system.

These dynamical descriptors were demonstrated to be modulated by the general different properties of the networks with the most evident modulation being the 'complexity' of the network graphs.

Here we are able to highlight the peculiar sub-diffusive[15] properties of amino acid-amino acid interaction networks (AAI) by means of the Maxdis descriptor. This parameter points to the relative importance of 'fast-and-specific communication lanes' inside networks.

Results and Discussion

General relations among dynamical descriptors

Maxdis describes the length of the path traveled by the perturbation. This descriptor is important in protein-drug interactions where targets are embedded into protein networks. Another possible application for this descriptor is the recognition of allosteric hotspots in a protein molecule.

We proved the independence of the Maxdis variable from the perturbation-centrality descriptors Extent and Time by means of Principal Component Analysis (PCA) on the whole data set, by using statistical units consisting of 1350 nodes relative to 45 networks and the variables Extent, Time, and Maxdis descriptors.

The first two components (PC1 and PC2) explained 50% and 34% of total variance, among the 1350 affected nodes of the entire data set. The major component (PC1) corresponded to the perturbation centrality paradigm set forth by Szalay and Csermely and was strongly correlated to both Extent and Time (Table1), while Maxdis generated a component of its own (PC2). Given the fact that principal components are orthogonal to each other by construction, we can assume that the maximal distance reached by a perturbation is independent of the general node impairment.

	PC1	PC2
Extent	**-0.857**	-0.161
Time	**0.842**	-0.244
Maxdis	0.069	**0.977**

Table 1 – Principal Components (PC) loading profile (Pearson correlation coefficients between original variables and components) of the complete dataset considered as a whole (no distinction among different network classes).

This implies the existence of privileged paths into the network that are able to transmit signals at elevated distances with low dissipation. If a node was located along one of these privileged tracks, it was able to 'communicate' with far-away nodes independently from its centrality in the network. The above result allows us to sketch the first network topology/function empirical rule:
The extent (number of nodes sensing the perturbation) and persistence (silencing time) of a perturbation is independent of the distance the effect of perturbation is sensed.
This rule is largely independent of the particular wiring architecture.

Network classes differential dynamics

In order to check whether different network classes (see Methods) are significantly different for the three dynamical descriptors, we applied Analysis of Variance with network class as a source of variation and the dependent variables of Extent, Time, and Maxdis. The results are reported in Table 2.

Variable	F-value	p-value
Extent	141,99	<0.0001
Time	90,64	<0.0001
Maxdis	139,31	<0.0001

Table 2 – ANOVA results relative to the effect of network class on the three dynamical variables.

The above result showed network functional classes widely differed for both topological (as already proven in[14]) and dynamical properties (as proved here). The extremely high F values showed that networks pertaining to the same class were very similar to each other in terms of their dynamical properties. Thus, different biological functions are closely related to different signal transmission

properties of the relative networks.

Having demonstrated the large importance of the 'class' effect over the 'within class' network variance, we have tackled the issue of node heterogeneity inside each network, (see Fig.1). In fact, although the three dynamical descriptors are associated with the correspondent network (by means of the average across the network nodes), they are originally measured in relation to a single affected node.

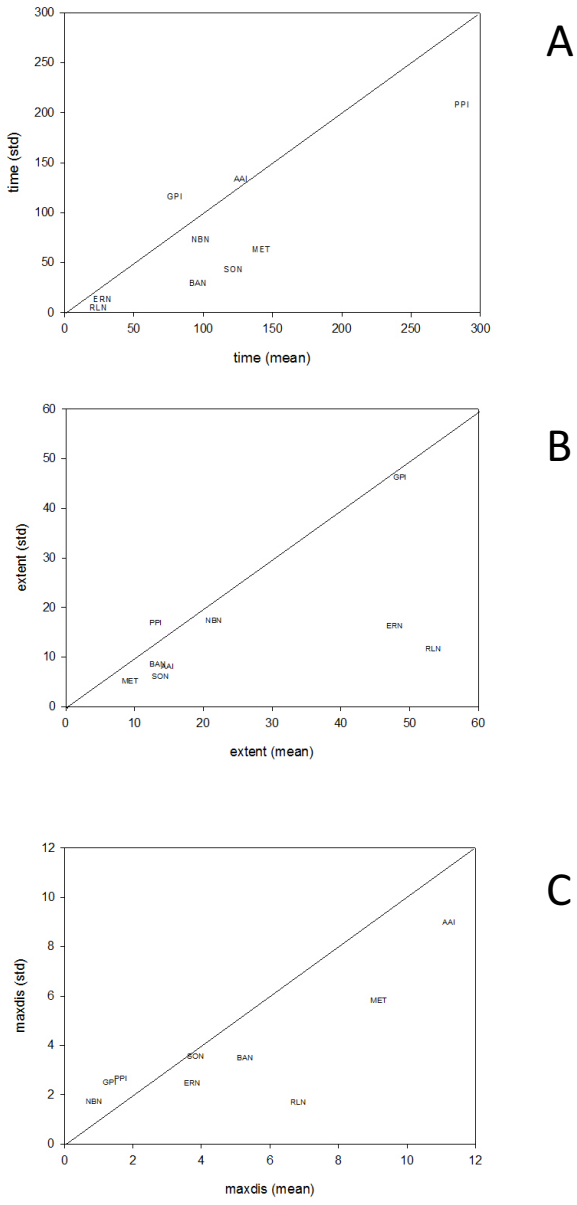

Figure 1 – Profiles of the mean (class average) and standard deviation (across nodes) for dynamical variables relative to different functional classes. A) Time; B) Extent; C) Maxdis.

In Figure 1, the mean and standard deviation of the dynamical parameters were computed over 150 nodes of the graphs corresponding to 30 affected nodes for 5 networks per class. Graphs suggest that

(with the only exception of two ideal and homogeneous by construction models like ERN and RLN) the existence of a link between the mean and among nodes standard deviation. This implies that an increase in both perturbation centrality (Extent, Time) and percolation (Maxdis) is related to an increase in node specialization within the network (within-network nodes heterogeneity).
Protein Contact Networks (AAI) showed an extremely high Maxdis consistent with their unique sub-diffusive[15] property with the contemporary presence of a 'slow lane' and a 'fast lane' of signal propagation[16]. The fast lane is known to be responsible for specific signal propagation in protein structures and for the allosteric effect[15-17]. This property is caught well by both the high Maxdis mean and standard deviation values. While the mean catches the exceptional signal transmission ability of protein molecules, the standard deviation highlights the possibility of isolating the single amino-acid residues responsible. Located on the fast lanes, responsible for the allosteric effect.
The above results are at the basis of the second topology/function rule:
The general function of a network corresponds to characteristic values of both 'average' and 'among-nodes variance' dynamical properties.
This rule stems from the existence of among networks architectural differences on both average and single node levels.

Comparison of the topological and dynamical views: Canonical correlation

Modulation of 'strength' of topology/dynamics relations across different architectures was estimated by means of canonical correlation analysis (CCA) having as X and Y variable sets the topological and the dynamical descriptors respectively (see Methods).
We performed nine independent CCAs for each network class, so to generate the 'mostly correlated' topology/dynamics model relative to each class (First canonical variate pair). The canonical correlation coefficient is thus an unbiased estimate of the 'amount of determinism' of the topology/dynamics relation for the class.
We adopted two different variables of choice for CCA:
1. X variables: all nine topological descriptors; Y variables: the three dynamical descriptors.
2. X variables: Average Shortest Path Length, Degree, Closeness; Y variables: the three dynamical descriptors.
The results of CCA according to choice 2 (the reduced set of topological X descriptors) are reported in Figure 2 in which vector points are the nodes.
The basic invariance between the results of the two CCAs (Figure 3, top panel) is a proof-of-concept of the robustness of the estimated topology/function correlation strength.
Figure 2 reports the topology/dynamics correlation for the different network classes in terms of U and V, the canonical variates pair for topological and dynamical spaces. The very strict correlation between

the topological and dynamical spaces is evident. Two (partial) exceptions to this strict link are the Global set and the AAI networks.

The Global set low topology/dynamics correlation (r=0.70) is a logical consequence of the heterogeneity of the network classes encompassing different architectures and thus different 'class specific' topology/dynamics relationships.

The low canonical correlation coefficient of the global set is a proof of the existence of 'class specific' modulation for the topology/dynamics link (Rule 2), which in turn influences the behavior of single nodes.

For the AAI network, the low canonical correlation (r=0.73) may be related to their peculiar properties. In fact, AAI are the only analyzed networks endowed with a 'physical' existence based on the geometry of the three-dimensional protein structure. In addition, the positioning of aminoacid residues along the sequence exerts an additional constraint on network architecture.

Apparently, AAI topology is influenced not only by the efficiency in signaling but also, by the fact that proteins need a structure stability in order to function properly, and this is independent to their communication needs. Thus, AAI complex topology may have evolved in order to accomplish their multitasking nature.

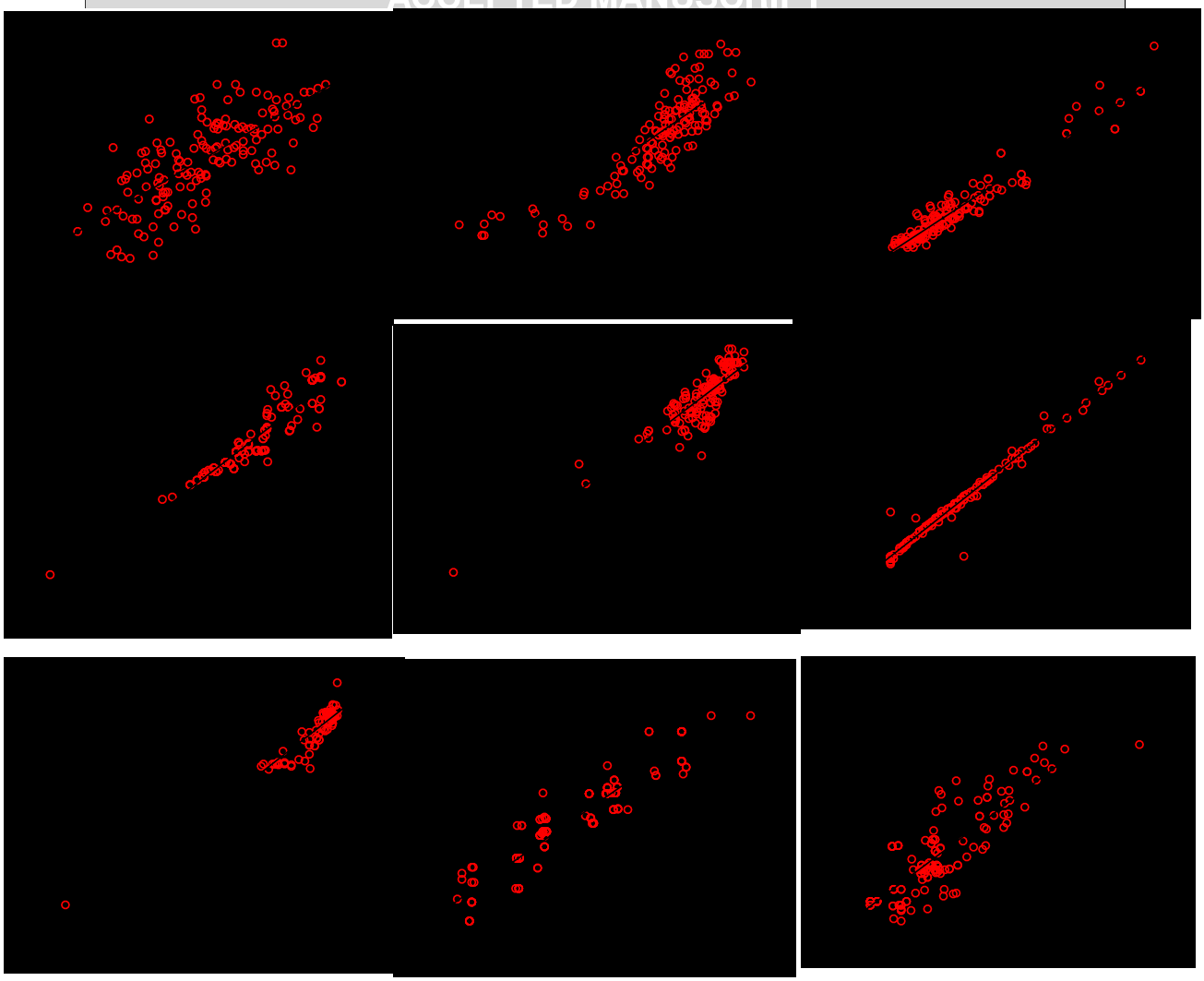

Figure 2 – Canonical Correlation plots: U and V are the first canonical variable pairs. The figure reports both the correlation coefficient and the linear plot. A) AAI; B) BAN; C) ERN; D) GPI; E) MET; F) NBN; G) PPI; H) RLN; I) SON.

CCA highlights the third topology/function rule: *The topological description of nodes is isomorphic with the dynamical consequences of their perturbation and has mainly to do with the relative centrality of the affected nodes.*

This rule (established at the single node level), while holding for all the different architectures, has different forms for different architectures.

Network complexity influences the strength of topology-dynamics correlation

The scoring of different amount of topology/dynamics correlation across different network architectures can be rationalized in terms of the amount of complexity of the network wiring with 'less complex' networks having a higher (quasi deterministic) correlation between single nodes 'static' topological profile and dynamical behavior.

Fig.3 summarizes the differences in the amount of 'determinism' of topology/dynamical rules across different network architectures.

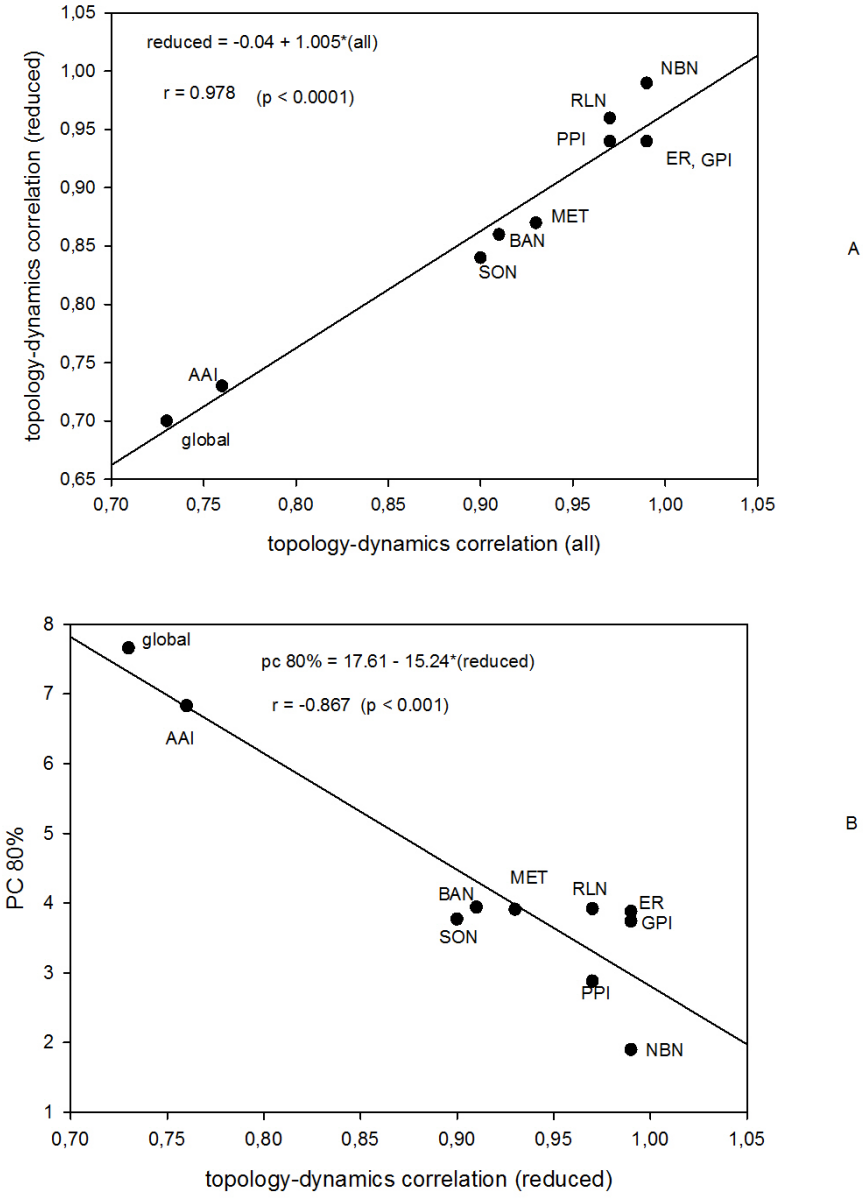

Figure 3 – Complexity of topology/dynamics rules across different network architectures: A) topology dynamics correlation, all vs reduced; B) PC 80% vs. topology-dynamics correlation (reduced).

The top-panel of Fig.3 has the topology/dynamics correlation value computed with paradigm 1 (nine topological vs. three dynamical descriptors) in the abscissa and the value corresponding to the reduced paradigm 2 (three topological vs. three dynamical descriptors) in ordinate. The high correlation between the two paradigms confirmed that the signal propagation through a network depends upon: Degree, Average Shortest Path Length, and the Centrality properties of the network.

This, in turn, proves the existence of a top-down modulation of the node topology/dynamics relations through the whole network wiring pattern.

In the bottom panel, the abscissa reports the amount of determinism of the topology/dynamics relation, while the ordinate is a network complexity score. This score (PC80%) stems from the eigenvalue distribution relative to each network class. For each class, a PCA was computed on the data set having 150 (30x5) nodes as statistical units and 12 (9 topological + 3 dynamical) descriptors as variables. In case of a Global Set, the nodes (i.e. the statistical units) are 1350. For each class specific (and Global) PCA, the number of components needed to explain 80% of the total variance was estimated by eigenvalue distribution. This means that the higher PCA80% is, the higher the complexity of the network class is in terms of among descriptor mutual relations. The strict relation between PCA solution and amount of complexity is explained in [17–19].

Network wiring complexity is maximal for the Global set and AAI (top left of the graph), while the most strict and deterministic topology-dynamics correlations are relative to less complex networks. Again, as for the canonical analysis, AAI demonstrated a higher complexity with respect to other networks.

On the other hand, the very specialized NBN showed a very low value of PCA80% and very high correlation coefficients between canonical variables to witness a high degree of specialization for NBN only purposed towards the signal transmission.

This allows to single out another topology/dynamics empirical rule (Rule 4): *The most complex a wiring architecture is, the less deterministic the relation between the topological characterization of a node and its dynamical properties.*

Conclusions

Topology and the dynamics of Networks are strictly connected. Connections may occur on multiple levels: general and local. Dynamical properties of node behavior when perturbed are influenced by the node topological identity and by the architectural properties of the network it is embedded in. We noted that the network wiring exerts a stronger effect on the node dynamics. Finally, it appears that the degree of network complexity severely impacts the topology and dynamics correlations (e.g. AAI networks), and this may be due to the fact that not all of the networks are only committed to signal

transmission tasks. For instance, at odds with signaling vocated networks like NBN, physical networks like AAI are known to be contemporarily constrained by signaling and structural tasks, hence they may evolve by multiple complex target functions.

Using a different perspective on signal transmission behavior of AAI networks, we were able to recognize two percolation levels[20] based on signal transmission: the slow lane level that accounts for local inter-residue communication (stability) and a fast lane level (functionality, allostery).

These two communication routes are based on the modular nature of protein structure, which is organized in domains (modules) whose mutual interactions are responsible for global dynamics. Once modules are identified, it is possible to address a topological role to each residue on the basis of purposed parameters, P (inter-module communication) and z (intra-module communication), corresponding respectively to the global (fast lane) and local (slow lane) dynamics[21].

We hypothesize, building upon previous literature (among others [22,23]) this bi-partite classification in terms of ‘long and specific’ and ‘short and aspecific' distance communication holds true even for different network systems. If confirmed, the discovery of ‘allosteric-like' behavior in communication networks can have a great impact on both the optimization of artificial network and the discovery of transition behavior in natural systems. On a more theoretical ground, it is worth noting the multi-scale character of the empirical rules we define. Rule 3 has to do directly with the topology/dynamics correlation and individuates in the concept of the topological centrality of the node at which the perturbation is applied the main determinant of the subsequent consequences of the perturbation. Rule 4 instead has to do with the ‘precision' and ‘strength' of such topology/dynamics relation for different network wiring architectures. Some network classes have more stringent ‘topology-dynamics' correlations than others (Fig.2), these differences can be rationalized in terms of network complexity: the more complex a network is, the weaker its topology-dynamics link (Fig.3). Rule 4 has to do with ‘top-down' causation: the node is influenced by the global structure of the network is embedded into, Rule 3 is a ‘bottom-up' one: the way a network is affected by a perturbation derives from the local properties of the node where the perturbation is applied. The presence of both ‘causal directions' is a hallmark of complex systems in which we have the simultaneous presence of both strong context dependencies (Rule 4) and singular properties of specific elements (Rule 3), this is a further proof of the relevance of complex network formalization for the investigation of organized complexity problems [2].

Methods

Networks description and perturbation

We selected nine classes of networks, which we listed and referenced in Table 3. A thorough description can be found in the reference[14].

Abbreviation	Description	Ref.
AAI	Amino acid-Amino acid Interaction networks, nodes are single amino acids	13
NBN	Neural Brain Networks, nodes are neurons or brain areas.	16
BAN	Barabasi-Albert Networks (models)	17
RLN	Regular Lattice Networks (models)	18
SON	Social Interaction Networks	19
PPI	Protein-Protein Interaction networks (nodes are single proteins)	20
GPI	Gene-Protein Interaction networks (signal transduction in cells)	21
MET	METabolic networks (nodes are single metabolites)	22
ERN	Erdos-Reny random Networks (models)	17

Table 3. Description of networks.

Network size has been set between 100 and up to 600 nodes across the five model networks within each class. Each class had the same average number of nodes to exclude relevant effects of size in our analyses (although it did not exert any recognizable role in our analysis in any case, data not shown). Primary network descriptors analysis was done by Cytoscape v.3.3.0[24] . We performed perturbation simulation on samples by Turbine_windows v.1.1[13] software. After normalizing every sample, the program performed the perturbation analysis by randomly choosing a single start point. The corresponding energy dissipation vectors, including the number of nodes affected (Extent), the distance traveled by the stimulus (Maxdis), and the silencing time of perturbation (Time), were collected for further analysis. For each network, 30 hit nodes were randomly selected: the perturbed node was excited by a pulse at 10000 units (low intensity, see[9]) and a $D_0 = 1$ constant dissipation was set. Silencing time (Time) was set as soon as all of the nodes had an energy value less than a pre-set threshold, set at D = 1. The Extent of the perturbation was equal to the number of nodes exceeding the threshold value in a given simulation. Maxdis corresponded to the maximum path length connecting

the starting perturbed node with the most remote one sensing the perturbation, it can be equated to the maximal distance (in number of links) traveled by the perturbation.

The 'low intensity' perturbation was chosen because it allowed for a sufficiently high variance across different hit nodes. In fact, both the medium and high-intensity modes in some cases had the effect of saturating the whole network (all nodes affected), thus shading the node-location and 'wiring pattern' effected on signal transmission.

The cycle was repeated 30 times (starting from different nodes) for every sample (independent network) and resulted in 1350 tests for whole data set; this corresponds to 45 networks (5 independent networks per class, for details, see Supplementary Information). Figure 4 reports two sample networks for AAI and Neural Brain Networks presented as appearing in Turbine upon perturbation (dynamical) descriptions.

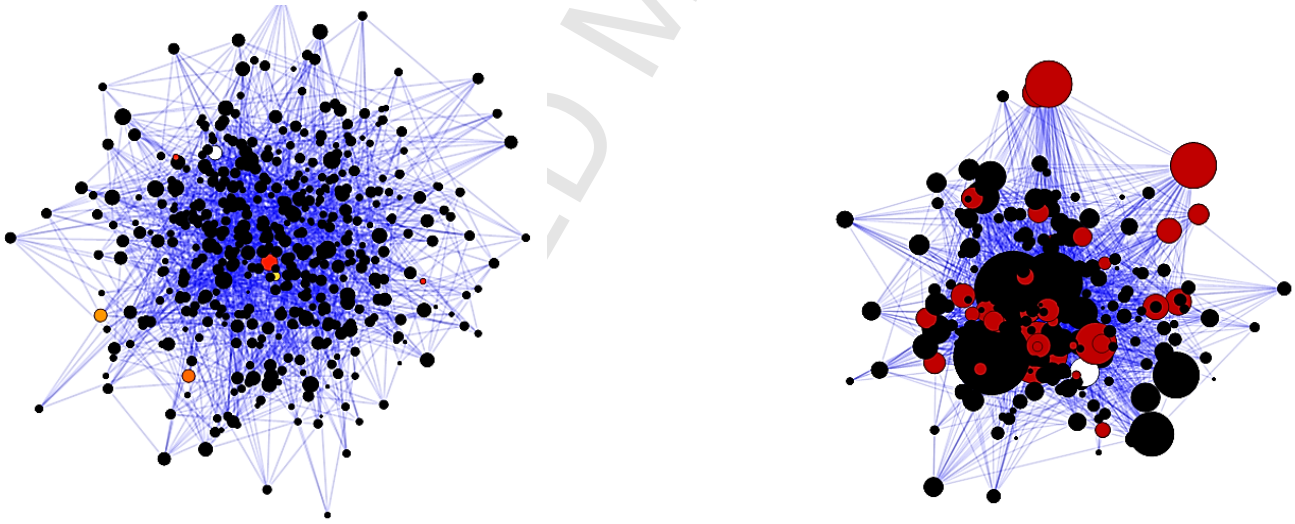

Figure 4: Sample network perturbation in Turbine. The NBN (right) and AAI (left) networks undergo random single point perturbation. The white node is the source of perturbation that released dissipating energy effects by other nodes (in red) based on rules governing the dynamics of networks. Black nodes are the ones not affected by perturbation.

Topological descriptors

We characterized each node by means of ~~nine~~ ten topological descriptors computed by NetworkAnalyzer, a Cytoscape plugin[25]:

- Node degree k_i is the simplest topological index; it reports the number of nodes adjacent to the i-th node, where adjacent means directly connected. The nodes directly connected to a node are also called first neighbors of the given node. Thus, the degree also corresponds to the number of adjacent incident edges.
- Clustering coefficient: in undirected networks for node *i*, it is defined as $C_i = {2e_i}/{k_i \cdot (k_i - 1)}$ where k_i is the node *i* degree and e_i is the number of connected pairs between all neighbors of *i*. Clearly, C_i is comprised between 0 and 1.
- The average shortest path length (or characteristic path length): the shortest sp_{ij} with the minimum number of links connecting the nodes *i* and *j*; the average shortest path length $\overline{sp_i}$ is the average value of the shortest path involving the node *i*;
- Closeness centrality for node *i* is defined as $CC_i = {1}/{\sum_j sp_{ij}}$ which is the sum extended to all nodes. It is null for isolated nodes and measures how fast information spreads from a given node to other reachable nodes in the network. We computed the closeness centrality using Network Analyzer.
- Betweenness centrality is defined as $BC_i = \sum_{i \neq j \neq k} {sp_{jk}(i)}/{sp_{jk}}$ where sp_{jk} denotes the shortest path between nodes *i* and *k* and $sp_{jk}(i)$ is the shortest path between the same nodes but passing by node *i*. The betweenness centrality can be computed only for networks that do not contain multiple edges. The betweenness value for node *i* is normalized by dividing by the number of node pairs excluding *i*: $\frac{(n-1)\cdot(n-2)}{2}$ where *n* is the total number of nodes in the connected component containing node *i*. The normalized value of the betweenness centrality lies between 1 and 0. The betweenness centrality of a node reflects the amount of control that this node exerts over the interactions of other nodes in the network. High values of betweenness centrality identify nodes that join communities (dense subnetworks), rather than nodes that lie inside a community.

  NetworkAnalyzer uses the fast algorithm by Brandes[26] to compute node betweenness centrality. This algorithm has a complexity of *O(nm)*, *n* being the number of nodes and *m* the number of edges in the network.
- Eccentricity centrality for node *i* is defined as the inverse of the longest shortest path: $ECC_i = {1}/{\max(sp_{ij})}$. Thus, if a node eccentricity is high, this means that all other nodes are in proximity. Conversely, if the eccentricity of a node is low, this means that there is at least one node (and all its neighbors) that is far from that node, although not excluding that several other nodes are in proximity. In other words, eccentricity is a more meaningful parameter if high.

- The stress of a node measures the number of shortest paths passing through that node. A node has a high stress if it is traversed by a high number of shortest paths. This parameter is defined only for networks without multiple edges.
- Neighborhood connectivity of node *i* is defined as the average degree of all nodes adjacent to node i. If the neighborhood connectivity distribution is a decreasing function of node degree *k*, there is a prevalence of edges between low connected and highly connected nodes.
- Radiality is a centrality index defined as $R_i = \sum_{j=1}^{n}(D - sp_{ij})$ having *n* as the nodes number and *D* as the network diameter, e.g. the longest shortest path. Nodes with high radiality show low shortest paths (close to all other nodes), while low radiality is a trait of nodes with high shortest paths (far from other nodes). The radiality is more significant when compared to the average value for the network (averaged on all nodes). Radiality must also be compared with other centrality metrics (closeness and eccentricity): a node with high closeness, eccentricity, and radiality is likely to play a central role in the graph.
- The topological coefficient for node *i* with degree k_i is defined as $T_i = {avg\ (J(i,j))}/{k_i}$ where $avg\ (J(i,j))$ is the average value of $J(i,j)$, which holds only for pairs of node *i* and *j* sharing at least a neighbor and stands for the number of neighbors shared by the two nodes *i* and *j*. Nodes with at most one neighbor are assigned a null value of T_i.

Sample size and aggregation level

The central aggregation level of the analysis corresponded to 45 independent networks subdivided into 9 classes (5 networks per class). Each network encompassed the data relative to 30 simulations (see 3.1), 30 is a recurrent choice in statistics, stemming from some considerations on central limit theorem[32] and allowing for a 96% confidence of catching the 'true' value of standard deviation (see Greenwood and Sandomire[27]). Moreover, this choice allows us to consider each topology-dynamics correlation as normally distributed inside each network. It is worth noting that the huge 'network class' effect (Table 2) allowed us to safely consider each of the 5 networks for each class as equivalent, thus we practically relied on a minimalistic 150 (30x5) nodes statistical base.

When the networks were grouped per class, we relied instead on 150 (30x5) different simulations to compute the mean and standard deviation of dynamical parameters. The initial definition of 'global relations' across the three dynamical descriptors by means of PCA was based on 1350 nodes (with no distinction among classes).

The topology-dynamics relationship in terms of both Canonical Correlation and Principal Component Analysis relied upon 150 and 1350 statistical units for single class and global analyses respectively.

Multivariate Statistical Analysis

Multivariate Data Analysis (MVDA) collects different tools[28] which clarify the relationship between variables exploiting correlation patterns. These tools apply to large data sets in order to identify single variables or their linear combination able to describe the behaviour of the whole system.
The very base of MVDA is the correlation analysis grounded on the computation of the Pearson correlation coefficient.

Principal Component Analysis (PCA)

The Principal Component Analysis (PCA)[29] transforms original variables into a new set of orthogonal variables, linear combinations of the original data set called Principal Components **PC**, that are ordered according to the descending values of explained variance. The number of components needed to explain a given amount of total variance is an index of the data set degree of complexity. Given the **PC** set, the generic element of the loadings matrix $\mathbf{L} = \{\boldsymbol{\lambda_{i,j}}\}$ represents the correlation coefficient between i-th original variable $\boldsymbol{x_i}$ and the j-th principal component $\boldsymbol{PC_j}$. PCA was computed by means of SAS v.8.1.

Canonical Correlation Analysis (CCA)

Canonical Correlation Analysis (CCA)[30] relied on a partition of the variables into two categories (topological and dynamical descriptors). Variables falling into each category are linearly combined to provide *k* variables (canonical variables), where *k* is the minimum rank (number of variables) among all categories. Linear combination scores were chosen in order to maximize the Pearson correlation coefficient between a canonical variable couple. CCA aims at computing a general correlation between different classes (categories) of variables describing the behavior of a system. Thus, a high correlation coefficient between the first canonical variables of two categories hints at a general interdependency between the two categories. Also in this case, the generic element of the loadings matrix $\mathbf{L} = \{\boldsymbol{\lambda_{i,j}}\}$ reports the correlation between the i-th canonical variable of a given category with the j-th original data set variable.

Author contribution statement

AG, LDP, and HK wrote the manuscript. HK and MT sorted and elaborated data. LDP, AG, and VT carried out the statistical analysis. All authors reviewed the manuscript.

Additional information

No Competing Financial Interests.

References

[1] V. Nicosia, M. De Domenico, V. Latora, Characteristic exponents of complex networks, EPL (Europhysics Lett. 106 (2014) 58005. doi:10.1209/0295-5075/106/58005.

[2] W. Weaver, Science and complexity, Am. Sci. 36 (1948) 536–544.

[3] R.B. Laughlin, D. Pines, J. Schmalian, B.P. Stojkovic, P. Wolynes, The middle way, PNAS. 1 (97AD).

[4] J. Imry, Introduction to mesoscopic physics, 2nd ed., Oxford University Press, 2002.

[5] P.G. Wolynes, Biomolecules: Where the Physics of Complexity and Simplicity Meet, Phys. Today. 47 (1994) 58–64.

[6] T.U. Hauser, V.G. Fiore, M. Moutoussis, R.J. Dolan, Computational Psychiatry of ADHD: Neural Gain Impairments across Marrian Levels of Analysis, Trends Neurosci. 39 (2016) 63–73.

[7] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.U. Hwang, Complex networks: Structure and dynamics, Phys. Rep. 424 (2006) 175–308. doi:10.1016/j.physrep.2005.10.009.

[8] M.K. Transtrum, B.B. Machta, K.S. Brown, B.C. Daniels, C.R. Myers, J.P. Sethna, Perspective: Sloppiness and emergent theories in physics, biology, and beyond, J. Chem. Phys. 143 (2015) 10901.

[9] B. Tellegen, A General Network Theorem with Application, Phillips Res. Reports. 7 (1952) 259–269.

[10] P. Oikonomou, P. Cluzel, Effects of topology on network evolution, Nat. Phys. 2 (2006) 532–536. doi:10.1038/nphys359.

[11] B. Barzel, A.-L. Barabási, Universality in network dynamics., Nat. Phys. 9 (2013) 673–681. doi:10.1038/nphys2741.

[12] J. Gao, B. Barzel, A.-L. Barabási, Universal resilience patterns in complex networks, Nature. 530 (2016) 307–312. doi:10.1038/nature16948.

[13] K.Z. Szalay, P. Csermely, Perturbation Centrality and Turbine: A Novel Centrality Measure Obtained Using a Versatile Network Dynamics Tool, PLoS One. 8 (2013) e78059. doi:10.1371/journal.pone.0078059.

[14] H. Kohestani, A. Giuliani, Organization principles of biological networks: An explorative study, BioSystems. 141 (2016) 31–39. doi:10.1016/j.biosystems.2016.01.004.

[15] L. Livi, E. Maiorino, A. Pinna, A. Sadeghian, A. Rizzi, A. Giuliani, Analysis of heat kernel highlights the strongly modular and heat-preserving structure of proteins, Phys. A Stat. Mech. Its Appl. 441 (2016) 199–214.

[16] D.M. Leitner, Energy flow in proteins., Annu. Rev. Phys. Chem. 59 (2008) 233–259.

doi:10.1146/annurev.physchem.59.032607.093606.

[17] E.S. Soofi, Capturing the Intangible Concept of Information, J. Am. Stat. Assoc. 89 (2016) 1243–1254.

[18] A. Giuliani, The Application of Principal Component Analysis to Drug Discovery and Biomedical Data, Drug Discov. Today. 22 (2017) 1069–1076. doi:10.1016/j.drudis.2017.01.005.

[19] A. Giuliani, M. Colafranceschi, C.L. Webber, J.P. Zbilut, A complexity score derived from principal components analysis of nonlinear order measures, Physica A. 301 (2001) 567–588.

[20] L.B. Caruso, A. Giuliani, A. Colosimo, Functional heterogeneity as reflected by topological parameters in a classical protein molecular model: T4 phage lysozyme, Curr. Protein Pept. Sci. 17 (2016) 52–61.

[21] S. Tasdighian, L. Di Paola, M. De Ruvo, P. Paci, D. Santoni, P. Palumbo, G. Mei, A. Di Venere, A. Giuliani, Modules Identification in Protein Structures: The Topological and Geometrical Solutions, J. Chem. Inf. Model. 54 (2013) 159–168.

[22] J. Goñi, M.P. van den Heuvel, A. Avena-Koenigsberger, N. Velez de Mendizabal, R.F. Betzel, A. Griffa, P. Hagmann, B. Corominas-Murtra, J.-P. Thiran, O. Sporns, Resting-brain functional connectivity predicted by analytic measures of network communication., Proc. Natl. Acad. Sci. U. S. A. 111 (2014) 833–8. doi:10.1073/pnas.1315529111.

[23] D.X. Horvath, J. Kertész, Spreading dynamics on networks: The role of burstiness, topology and non-stationarity, New J. Phys. 16 (2014). doi:10.1088/1367-2630/16/7/073037.

[24] P. Shannon, A. Markiel, O. Ozier, N.S. Baliga, J.T. Wang, D. Ramage, N. Amin, B. Schwikowski, T. Ideker, Cytoscape: A software Environment for integrated models of biomolecular interaction networks, Genome Res. 13 (2003) 2498–2504. doi:10.1101/gr.1239303.

[25] Y. Assenov, F. Ramírez, S.E.S.E. Schelhorn, T. Lengauer, M. Albrecht, Computing topological parameters of biological networks, Bioinformatics. 24 (2008) 282–284. doi:10.1093/bioinformatics/btm554.

[26] U. Brandes, A faster algorithm for betweenness centrality, J. Math. Sociol. 25 (2001) 163–177. doi:10.1080/0022250X.2001.9990249.

[27] J.A. Greenwood, M.M. Sandomire, Sample Size Required For Estimating The Standard Deviation as a Percent of Its True Value, J. Am. Stat. Assoc. 45 (1950) 257--260. doi:10.2307/2280683.

[28] P.J. Lewi, Multivariate Data Analysis in Industrial Practice, John Wiley & Sons, 1982.

[29] I. Joliffe, Principal Component Analysis and Factor Analysis, in: Princ. Compon. Anal. Factor Anal., Springer New York, 2002.

[30] B. Thompson, Canonical correlation analysis: uses and interpretation, Quant. Appl. Soc. Sci.

(1984) 1–71.

E-CAPTION- Supplementary Information
The Supplementary information reports the main features of data analysis steps.

Highlights

* The topological and dynamical features of experimental biological network nodes are each other related.

* The maximal distance at which a perturbation is sensed is independent on the extent (number of nodes affected) and silencing time of the perturbation.

* The empirical rules linking topology and dynamical behavior of nodes are modulated by network structure (and related function).

* The more complex a network wiring is, the less deterministic the topology/dynamics relations of its nodes are.