Metadata of the chapter that will be visualized in SpringerLink

Book Title	Artificial Life and Evolutionary Computation	
Series Title		
Chapter Title	A Relevance Index Method to Infer Global Properties of Biological Networks	
Copyright Year	2018	
Copyright HolderName	Springer International Publishing AG, part of Springer Nature	
Corresponding Author	Family Name	**Villani**
	Particle	
	Given Name	**Marco**
	Prefix	
	Suffix	
	Role	
	Division	Dip. Scienze Fisiche, Informatiche e Matematiche
	Organization	Università di Modena e Reggio Emilia
	Address	Modena, Italy
	Email	marco.villani@unimore.it
Author	Family Name	**Sani**
	Particle	
	Given Name	**Laura**
	Prefix	
	Suffix	
	Role	
	Division	Dip. di Ingegneria e Architettura
	Organization	Università di Parma
	Address	Parma, Italy
	Email	
Author	Family Name	**Amoretti**
	Particle	
	Given Name	**Michele**
	Prefix	
	Suffix	
	Role	
	Division	Dip. di Ingegneria e Architettura
	Organization	Università di Parma
	Address	Parma, Italy
	Email	
Author	Family Name	**Vicari**
	Particle	
	Given Name	**Emilio**
	Prefix	
	Suffix	

	Role	
	Division	
	Organization	CAMLIN
	Address	Parma, Italy
	Email	
Author	Family Name	**Pecori**
	Particle	
	Given Name	**Riccardo**
	Prefix	
	Suffix	
	Role	
	Division	Dip. di Ingegneria e Architettura
	Organization	Università di Parma
	Address	Parma, Italy
	Division	SMARTEST Research Centre
	Organization	Università eCAMPUS
	Address	Novedrate, CO, Italy
	Email	
Author	Family Name	**Mordonini**
	Particle	
	Given Name	**Monica**
	Prefix	
	Suffix	
	Role	
	Division	Dip. di Ingegneria e Architettura
	Organization	Università di Parma
	Address	Parma, Italy
	Email	
Author	Family Name	**Cagnoni**
	Particle	
	Given Name	**Stefano**
	Prefix	
	Suffix	
	Role	
	Division	Dip. di Ingegneria e Architettura
	Organization	Università di Parma
	Address	Parma, Italy
	Email	
Author	Family Name	**Serra**
	Particle	
	Given Name	**Roberto**
	Prefix	
	Suffix	
	Role	
	Division	Dip. Scienze Fisiche, Informatiche e Matematiche

Organization	Università di Modena e Reggio Emilia
Address	Modena, Italy
Email	

Abstract

Many complex systems, both natural and artificial, may be represented by networks of interacting nodes. Nevertheless, it is often difficult to find meaningful correspondences between the dynamics expressed by these systems and the topological description of their networks. In contrast, many of these systems may be well described in terms of coordinated behavior of their dynamically relevant parts. In this paper we use the recently proposed Relevance Index approach, based on information-theoretic measures. Starting from the observation of the dynamical states of any system, the Relevance Index is able to provide information about its organization. Moreover, we show how the application of the proposed approach leads to novel and effective interpretations in the T helper network case study.

Keywords (separated by '-')

Complex systems - Biological networks - Dynamical behavior - Relevance index - T helper cells

A Relevance Index Method to Infer Global Properties of Biological Networks

Marco Villani[2(✉)], Laura Sani[1], Michele Amoretti[1], Emilio Vicari[4], Riccardo Pecori[1,3], Monica Mordonini[1], Stefano Cagnoni[1], and Roberto Serra[2]

[1] Dip. di Ingegneria e Architettura, Università di Parma, Parma, Italy
[2] Dip. Scienze Fisiche, Informatiche e Matematiche, Università di Modena e Reggio Emilia, Modena, Italy
marco.villani@unimore.it
[3] SMARTEST Research Centre, Università eCAMPUS, Novedrate, CO, Italy
[4] CAMLIN, Parma, Italy

Abstract. Many complex systems, both natural and artificial, may be represented by networks of interacting nodes. Nevertheless, it is often difficult to find meaningful correspondences between the dynamics expressed by these systems and the topological description of their networks. In contrast, many of these systems may be well described in terms of coordinated behavior of their dynamically relevant parts. In this paper we use the recently proposed Relevance Index approach, based on information-theoretic measures. Starting from the observation of the dynamical states of any system, the Relevance Index is able to provide information about its organization. Moreover, we show how the application of the proposed approach leads to novel and effective interpretations in the T helper network case study.

Keywords: Complex systems · Biological networks
Dynamical behavior · Relevance index · T helper cells

1 Introduction

Nowadays a plethora of molecular data results in a vast amount of pathways, networks of interactions and molecular scenarios. A large quantity of information is available on many biological systems, and researchers use it to infer global properties of biological networks [15,21]. In spite of the strong representational power and flexibility of networks, there are, however, two major limitations which affect most studies in the field [16,23]:

- the information about the underlying true interactions is often incomplete, so the inferred networks do not provide a complete picture of the interactions in the system under study;
- network studies are often concerned with "static" topological information, like connectivity and betweenness, whereas, in order to understand the functionality of a system, it is important to study its *dynamical properties*.

M. Pelillo et al. (Eds.): WIVACE 2017, CCIS 830, pp. 1–13, 2018.
https://doi.org/10.1007/978-3-319-78658-2_10

Modeling the dynamic behavior of such systems is difficult, due to the lack of kinetic data and to computational limitations. Among the methods for facing this problem, those based on steady-state approximations are widely used [13,25]. Nevertheless, these kinds of analysis do not provide enough constraints to find a unique solution to the problem: thus researchers support these techniques by means of suitable hypotheses as, for example, minimization or maximization issues [25]. This drawback, in terms of modeling, has turned out to be particularly relevant when controlling the steady-state behavior of complex networked dynamical systems. In this respect, some efficient model-free methods based on multi-agent reinforcement learning [5] and on mean-field game theory [3] are rapidly emerging in several domains, such as telecommunications.

In order to overcome the aforementioned limitations of steady-state methods, it is worthwhile to resort to methods able to directly deal with the dynamical repertoire of the system. In this paper, we use a recently proposed approach, the *Relevance Index* (RI for short) method [11,32,33], which has the following features:

1. It is based on the observation of the dynamical states of the system (whether simulated or real), without requiring any *a priori* knowledge of the interactions among variables (whenever such knowledge is available, it can be used to complement the proposed method);
2. It can be applied to states coming from different steady state conditions, or even to states obtained from perturbation of these conditions (it does not require fixed asymptotic states);
3. It provides information about the organization of the system itself; indeed, complex systems often display complex organizational features that cannot be captured by a simple tree-like structure;
4. It is robust against noisy or incomplete data, being based on information-theoretic measures.

The overall contribution of this paper is twofold.

On one hand, we show that (i) the dynamically relevant groups of variables identified using the RI index in a biological network are extremely useful in describing the overall dynamics of the system and that (ii) this description could significantly enlarge the explicative power of the graph description of a biological system, by highlighting the links that are really effective.

On the other hand, we present a novel method for creating the homogeneous system used as a reference to evaluate the significance of the RI results. This method considers non-zero pairwise correlations among the variables of the system and is based on the NORTA technique.

The rest of the paper is structured as follows. Section 2 presents the context about complex systems and related works. Section 3 provides a brief review of the Relevance Index method and of the improvement in computing the homogeneous system. Section 4 shows how the application of the RI method leads to novel and effective interpretations in a biological network (T helper case study). Finally, Sect. 5 seals up the work.

2 Context and Related Work

In most natural or artificial dynamical systems, there are groups of variables showing highly coordinated internal dynamics able to significantly influence other groups or even the whole system (Relevant Sets, or shortly RS in the following). The capacity of detecting their presence can often lead to a high-level description of the dynamical organization of the system, and thus to its understanding [32].

However, the identification and monitoring of the significant or relevant portions of dynamical systems is very difficult, especially if these systems exhibit emergent or self-organizing phenomena, the latter being the most interesting and prominent situation for complex dynamical systems [7].

Indeed, most theories and models take into account only two-level systems and describe the formation of relatively simple dynamical patterns as, for example, the creation of the well-known Bénard-Marangoni hexagonal convection pattern [12]. In this case the two levels involved are those of the water particles and of the hexagonal convection cells. Indeed, the apparatus where the phenomenon takes place (which is, of course, necessary, since it determines some major features of the phenomenon itself) is not affected by what happens at the lower levels: in other words, it just provides the fixed boundary conditions that allow the phenomenon to occur.

However, the most interesting recurrent patterns of interaction [18] take place very often at levels that can be regarded as intermediate between pre-existing layers, which are, in turn, affected by the dynamics of these patterns. There are several examples of these "sandwiched" phenomena in physics, biology and social sciences [18]. Perhaps the most evident cases are the presence of vortexes on fluids surfaces, the presence of organs and tissues in multi-cellular organisms, or the action of various groups of humans (such as companies, cooperatives, associations, factions, communities) within societies[1]. Note that the formation of structures or patterns not explicitly designed is frequent even in artificial systems, as for example power grids [34], e-mail networks [6], Internet [1,8], and so on. Thus, the detection of intermediate-level structures and patterns is a very central issue in complex dynamical systems.

Many interesting systems can be represented, at least partially, by means of graphs. In this case, a widespread property is the presence of the so-called communities, portions of system elements within which the connections are dense, but between which they are sparser [20]. Their identification sometimes could detect groups that can be good relevant set candidates.

A method that mixes static and dynamic issues was proposed by Thomas et al. [27,28] for regulatory networks, the focus of this paper, to capture the main qualitative features of the dynamics of such systems.

[1] The lower and upper level being constituted by the fluid particles and their global stream, by cells and the organism to which they belong, and by human beings and societies, respectively.

Works that use dynamical features in order to detect functional groups are not so frequent; many of them rely on similarity measures and clustering algorithms. This is what is done by Feldt et al. [9], for example.

An interesting approach uses methods introduced in information theory and applied in neurosciences by Edelman and Tononi in 1994 and 1998 [29,30] to detect functional groups of brain regions. In our previous works, we extended the approach to non-stationary dynamical regimes, in order to apply the method to a broad range of systems, including abstract models of gene regulatory networks and simulated social [10], chemical [32], and biological [33] systems. The resulting approach could also be used to identify the critical states of complex dynamical systems [24].

Finally, an interesting literature review about the reconstruction of gene regulatory networks and the development of mathematical models of how the patterns of activation and inhibition determine the state of activation of the network can be found in [4]. The T helper regulatory network considered in this paper is based on the one described in [19].

3 Method

The technique employed in this paper to identify subsets of nodes that are good candidates as RSs is mainly based on the Relevance Index (RI) method. For a complete overview of the methodology adopted in this work please refer to Villani et al. [33]. In the following we will only summarize it briefly.

Main assumptions:

- the values of the system nodes, or variables, express the observed states of the system;
- there exist one or more subsets where these variables are acting in a coordinated way;
- the variables of each subset interact with the other system variables more weakly than among one another internally;
- The computation of the RI is usually based on observational data, and probabilities are estimated as the relative frequencies of the values observed for each variable.

Consider a system U composed of n random variables $(X_1, X_2, ..., X_n)$, and a subset S_k composed of k of them, with $k < n$. The $RI(S_k)$ value is defined as the ratio between the *integration* I of S_k and the *mutual information* MI between S_k and the rest of the system:

$$RI(S_k) = \frac{I(S_k)}{MI(S_k; U \backslash S_k)} \tag{1}$$

where $I(S_k)$, the integration, measures the statistical independence of the k elements in S_k and $M(S_k; U \backslash S_k)$, the mutual information, expresses the mutual

dependence between the subset S_k and the rest of the system $U \backslash S_k$. The integration is defined by the following formula:

$$I(S_k) = \sum_{s \in S_k} H(s) - H(S_k) \tag{2}$$

Values of MI equal to zero indicate that the Candidate Relevant Set (CRS in the following) does not communicate with the rest of the system, i.e., it is a separate system and its variables can be neglected. The RI scales with the size of the CRS, thus it needs to be normalized by dividing each member of the quotient in Eq. 1 by its average value within a system where no dynamical structures are present, i.e., a *homogeneous system* where no specific interaction within groups of variables can be highlighted. Moreover, the statistical significance of RI differences should be assessed by means of an appropriate test. For these reasons, a statistical significance index T_c was introduced, which measures how much larger (or smaller) the RI of a subset of variables S_k is with respect to the average RI of groups of the same size within the homogeneous system:

$$T_c(S_k) = \frac{RI(S_k) - \langle RI_h \rangle}{\sigma(RI_h)} = \frac{\nu RI - \nu \langle RI_h \rangle}{\nu \sigma(RI_h)} \tag{3}$$

where $\langle RI_h \rangle$ and $\sigma(RI_h)$ are, respectively, the average and the standard deviation of the RI of a sample of subsets of size k extracted from a reference homogeneous system U_h, and $\nu = \langle MI_h \rangle / \langle I_h \rangle$ is its normalization constant. A more detailed description can be found in previous work [26,31].

The generation of the homogeneous system is critical, and often, in past papers, a simple but general and easy to compute solution was chosen. This solution encompassed the computation of the frequency of each variable, given the available observations, and the generation of a new random series of samples, where each variable had a prior probability equal to the frequency of the original observations. The homogeneity required by Tononi was achieved by considering the components of the random vector U_h, representing the homogeneous system, to be independent. This produced:

1. A unity correlation matrix of the homogeneous system, i.e., with pairwise correlations set to zero;
2. An integration $I(S_k) = 0$ for all subsets of the homogeneous system.

In this paper, we introduce, for the first time, a novelty in the generation of the homogeneous system compared to previous works: homogeneity is maintained by forcing all off-diagonal elements of the correlation matrix to have the same constant value ρ different from zero:

$$CORR(U_h) = \begin{bmatrix} 1 & \rho & \dots & \rho \\ \rho & \ddots & \ddots & \vdots \\ \vdots & \ddots & \ddots & \rho \\ \rho & \dots & \rho & 1 \end{bmatrix}$$

Such a value ρ is computed as the average of all pairwise correlations of the observed variables. In this way we preserve both homogeneity and dependence among the different variables.

In order to generate a homogeneous system with the aforementioned features, we use the NORTA method [2], a mathematical procedure that solves the issue of creating random vectors of correlated samples, given the set of their marginal distributions (marginals) and a measure of the dependence among them. The dependence measure we used in NORTA is the usual *product-moment* correlation matrix, based on the linear Pearson correlation coefficient.

As a final step of our methodology, a further *sieving algorithm* [11] can be used to isolate the most representative CRSs, i.e., those having the highest T_c. This procedure is based on the following criterion: if CRS C_1 is a proper subset of C_2 and ranks higher than CRS C_2, then C_1 is considered to be more relevant than C_2. Thus it is possible to keep only those CRSs not included in or not including any other CRS with higher T_c. The sieving activity stops when no more eliminations are possible and the remaining sets of variables are the true relevant sets.

4 Experimental Results

4.1 The T Helper Cell Differentiation System

The vertebrate immune system is composed of several cell populations, including antigen presenting cells, natural killer cells, and B and T lymphocytes. There are two main kinds of T lymphocytes: the T cytotoxic cells that actively destroy virus-infected cells and tumor cells and the T helper cells (Th) that take part in cell- and antibody-mediated immune responses by secreting various cytokines, differently distributed in the two main T helper cell sub-types Th1 and Th2. Both sub-types derive from a common precursor Th0 through a rather complex differentiation path, modeled in [19,22]. In this work, we use the discretization of an updated version of these paths described in [19] (Fig. 1).

The nodes TCR, IL_{18}, IFN_b and IL_{12} receive their input from outside the Th differentiation system and constitute the way the system is aware of its context (in other words, they constitute the system "sensors"). Several signalling pathways are stimulated by their activation [14].

4.2 RI Results

We simulated the gene regulatory network described in Fig. 1 by means of a synchronous Boolean system. There are 2^{19} different initial conditions for each of the 2^4 different scenarios identified by the "sensor" nodes. However, we found only 33 different asymptotic behaviors (all fixed points). Three of these attractors coincide with the gene expression of Th0, Th1 and Th2 cells. These attractors are presented in [19] as the only really stable states, according to the information derived from the application of the so-called generalized logical analysis [28] to the Th differentiation system.

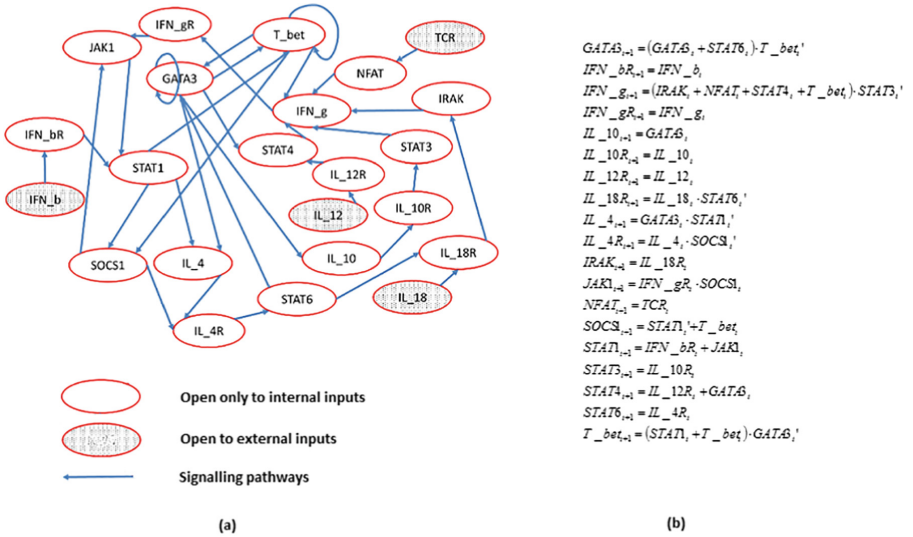

Fig. 1. (a) A graph representation of the Th differentiation system. Note that all the gray-filled nodes (TCR, IL_18, IFN_b, and IL_12) do not receive their input from the network regulating the differentiation system. Thus, in this representation, they do not have incoming links. (b) The dynamical rules of the Th differentiation system as described in [19].

However, the gene regulatory network can express 33 different asymptotic behaviors. Indeed, this fact should give us some information about the dynamical organization of the system[2]. Therefore, to extract this information, we tried to apply the RI methodology (i) to the mere juxtaposition of these attractors or (ii) by weighting their presence proportionally to the size of their basins of attraction, i.e., the width of the neighborhood from which the system converges into the state represented by the attractor under consideration.

In both cases the relevant subsets that were found are composed by TCR and NFAT nodes (Group1 in Fig. 2) and all the other nodes (Group2 in Fig. 2)[3].

[2] In this work we do not make hypotheses about the biological plausibility (or stability or biological function, if any) of these attractors, suggesting the interested readers to refer to Mendoza and Xenarios [19] and to the references quoted therein. Rather we highlight that, once a mathematical model has been established, its structure implies the presence of a well-defined set of attractors: so, an analysis that takes into account their presence (and therefore which highlights their interrelated dynamical relationships) should provide better results than a method that does not act in this way.

[3] The node JAK1 is constantly inactive in all attractors. Thus, its presence is useless for the purposes of a dynamical analysis and no CRS include it. Indeed, it is active in transient states, but this kind of analysis is out of the scope of this work (see [24] for a first comparison of the results of RI application to transients and asymptotic states).

Process	Group	TCR	IL_18	IFN_b	IL_12	GATA3	IFN_bR	IFN_g	IFN_gR	IL_10	IL_10R	IL_12R	IL_18R	IL_4	IL_4R	IRAK	JAK1	NFAT	SOCS1	STAT1	STAT3	STAT4	STAT6	T_bet	Tci
Sieve1	Group1	■																	■						61379.40
	Group2		■	■	■	■	■	■	■	■	■	■	■	■	■	■	■			■	■	■	■	■	3519.58
Sieve2	Group3			■			■												■						1416.54
	Group4	■											■		■										1186.76
	Group5				■			■	■	■	■											■		■	784.15
	Group6					■						■											■		780.80
Pre-Sieve2	Group7			■			■							■	■			■						■	642.18
	Group8	■											■	■											632.36

Fig. 2. The table shows the groups detected by the application of the RI methodology followed by the sieving algorithm (groups 1–6): each group is represented as a row where black boxes denote the variables belonging to it. Group7 and Group8 have been discarded by the sieving algorithm, because they include the stronger relevant subsets indicated as Group3 and Group4: however their observation is important, because it traces a significant coupling among Group3 and Group4 and the other system variables. Indeed, a second application of the iterated RI method fixes this strong association (data not shown).

This fact indicates that the Th differentiation machinery is indeed highly integrated. We can register the presence of these two first CRSs and successively filter them out, in order to apply the sieving algorithm to all the remaining groups. In this case, the two approaches produce different results.

The simple attractor juxtaposition separates Group2 into two big subsets (see Fig. 3, left), whereas the application of RI to an extended set of observations obtained by repeating input data related with the 33 attractors a number of times proportional to the width of their basins of attraction is able to identify (i) the four chains that transmit the external signals toward the inner core of the Th differentiation system (the TCR-NFAT chain, i.e., the Group1, already identified during the first RI application) and (ii) a "circle" of nodes that appears to be the "dynamical engine" of the Th differentiation system, denoted as Group5 (Fig. 3, right).

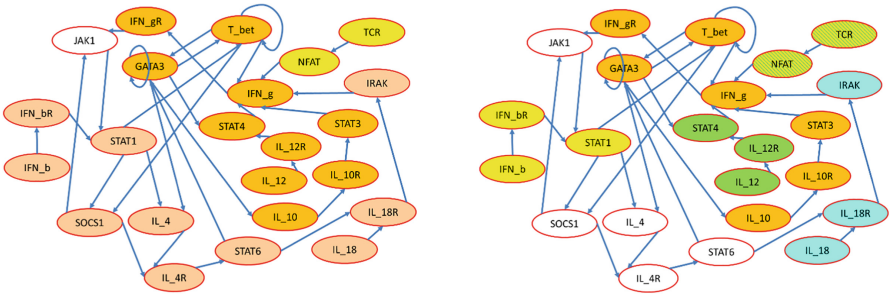

Fig. 3. The main relevant subsets identified using the simple juxtaposition of the attractors of the Th differentiation system (left) or by weighting their presence proportionally to their basins of attraction - the right part of the figure. In this part we highlight the presence of Group1, Group3, Group4, Group5, and Group6, respectively in striped, yellow, blue, orange, and green background. (Color figure online)

It appears that nodes SOCS1, IL_4, IL_4R, and STAT6 do not belong to any relevant subsets (Fig. 3, right), if we strictly adhere to the relevant subset definition. However, before the application of the sieving algorithm, the RI analysis reports two highly-ranked groups in the top positions, namely Group7 (composed by the aforementioned nodes and by Group3) and Group8 (composed of IL_4,IL_4R, STAT6, and by Group4). Indeed, these two groups are discarded by the sieving algorithm because they include two already identified and slightly stronger relevant subsets. Vice-versa, we can use this information in order to identify the nodes influenced by (or influencing) Group3 and Group4. Thus, given the directions of the links of the Th system, it appears that the information acquired by Group3 (in particular by node IFN_b) is transmitted to the nodes belonging to the "white group", which, in turn, passes it to Group4. Therefore, the white group is composed by elements that seem to act as a sort of "transmission engine" for the Th differentiation system. Figure 4 highlights such an information flow from the "yellow" region (group 7) to the "blue" region (group 8).

The RI analysis therefore induces an interesting interpretation of the dynamical data which, when mapped on the already available topological knowledge, provides an expressive explanation of the system functioning. The same knowledge (the identification of groups of variables and of their relationships) is not derivable from the static analysis alone. The usual algorithms for the search of communities [17,20] identify only the pair GATA3-T_bet. Moreover, only one of the identifiable 27 circuits is highlighted (Group5, which involves nodes T_bet, GATA3, IL_10, IL_10R, STAT3, IFN_g, IFN_gR, JAK1 and STAT1).[4]

On the other hand, the usual dynamical analyses are mainly focused on the detailed reproduction or prediction of the system's behaviors [19] and therefore are not suitable for a highly abstracted and "global" vision of the system functioning. The same generalized logical analysis [28] that mixes topological and dynamical issues identifies chains of positive and negative feedbacks, eventually providing clues for the identification of stable attractors, but does not give the overall vision of the RI method, which identifies the genes involved in injecting information into the system (the groups 1, 3, 4 and 6) and the main circuit responsible of the information processing (group5).

Obviously, this method cannot be used to reconstruct the detailed topology of the investigated system (though it could suggest useful groupings). It is worth mentioning, however, that the RI method can be applied directly to the experimental data, if these are available. In this respect, we can note that while the collection of time series is an experimentally difficult and costly task, the RI methodology can be applied merely by comparing different steady states (whose data could derive even from different beings), in such a way taking advantage from more common data sources. In case experimental data are available, the RI method can provide an effective idea of the dynamical organization of the observed system without requiring any knowledge of topology, dynamical rules, or parameters [26,31,32].

[4] Note that the node STAT1 participates in Group 3, one of the "sensors groups" of the Th differentiation system.

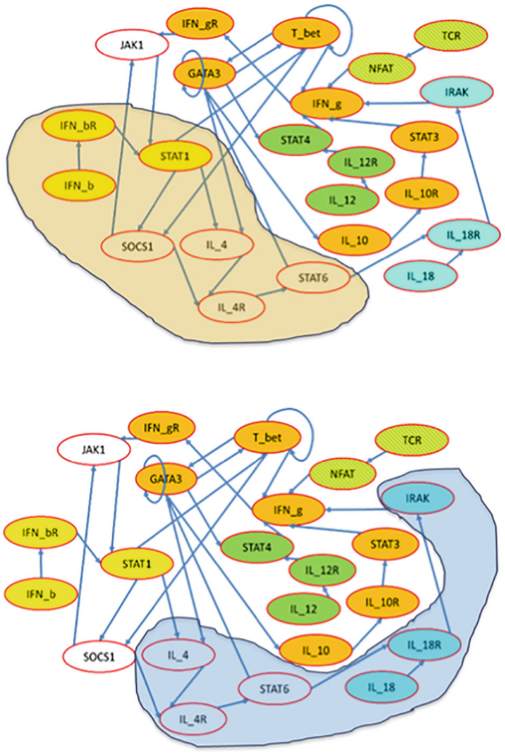

Fig. 4. Same as Fig. 3, but highlighting the correlation of Group3 and Group4 with other Th differentiation variables (SOCS1, IL_4, IL_4R and STAT6 – for brevity indicated in this caption as "WhiteGroup"). With reference to the table reported in Fig. 2, one can see that, indeed, the first (and unique) significant appearance of these variables as a block occurs along with Group3 and Group4, with which they compose Group7 and Group8, as shown by the third block of results in the table. Given the directions of the links, in this example assumed to be known, it appears that the graph structure of the system could allow the signal transmission from Group3 and Group5 to the WhiteGroup (first row). However, the RI index indicates as evident the influence of just Group3. In turn, the information acquired by the WhiteGroup from Group3 is transmitted to Group4, in such a way modulating the external signals coming from node IL_18 (second row). (Color figure online)

5 Conclusion

In this paper, we proposed to use the RI method, improved through a novel technique for computing the correlation matrix of the homogeneous system, as a means to infer global properties of biological networks. With respect to steady-state approximation approaches, the RI method, which is based on the observation of the dynamical states of the system, provides information about the organization of the system itself and is robust against noisy or incomplete

data, being based on information-theoretic measures. The RI method can be applied directly to the experimental data, if available. In this case, it can sketch an effective picture of the dynamical organization of the observed system. As a use case, we illustrated the analysis of the T helper network.

Regarding future work, we plan to apply the RI method to several biological networks. This can be done quite easily because it can be applied to system characterized by both continuous and discrete (Boolean or multi-valued) variables. The ultimate objective is twofold and encompasses both finding new insights about those systems and refining the method itself. In particular, we are interested in studying systems with a large number of nodes, which cannot be explored exhaustively, even with parallel computing approaches. For such systems, the adoption of meta-heuristics is necessary in order to find the relevant groups of nodes in a reasonable amount of time.

Acknowledgments. The work of Michele Amoretti was supported by the University of Parma Research Fund - FIL 2016 - Project "NEXTALGO: Efficient Algorithms for Next-Generation Distributed Systems".

This work greatly benefited from discussions with Andrea Roli, to whom the authors are warmly thankful.

References

1. Albert, R., Jeong, H., Barabási, A.L.: Internet: diameter of the world-wide web. Nature **401**(6749), 130–131 (1999)
2. Cario, M.C., Nelson, B.L.: Modeling and generating random vectors with arbitrary marginal distributions and correlation matrix. Technical report (1997)
3. Cimorelli, F., Priscoli, F.D., Pietrabissa, A., Celsi, L.R., Suraci, V., Zuccaro, L.: A distributed load balancing algorithm for the control plane in software defined networking. In: 2016 24th Mediterranean Conference on Control and Automation (MED), pp. 1033–1040, June 2016
4. De Jong, H.: Modeling and simulation of genetic regulatory systems: a literature review. J. Comput. Biol. **9**(1), 67–103 (2002)
5. Delli Priscoli, F., Di Giorgio, A., Lisi, F., Monaco, S., Pietrabissa, A., Celsi, L.R., Suraci, V.: Multi-agent quality of experience control. Int. J. Control Autom. Syst. **15**, 892–904 (2017)
6. Ebel, H., Mielsch, L.I., Bornholdt, S.: Scale-free topology of e-mail networks. Phys. Rev. E **66**, 035103 (2002). http://www.citebase.org/cgi-bin/citations?id=oai:arXiv.org:cond-mat/0201476
7. Emmeche, C., Køppe, S., Stjernfelt, F.: Explaining emergence: towards an ontology of levels. J. Gen. Philos. Sci. **28**(1), 83–117 (1997)
8. Faloutsos, M., Faloutsos, P., Faloutsos, C.: On power-law relationships of the Internet topology. SIGCOMM Comput. Commun. Rev. **29**(4), 251–262 (1999)
9. Feldt, S., Waddell, J., Hetrick, V., Berke, J., Żochowski, M.: Functional clustering algorithm for the analysis of dynamic network data. Phys. Rev. E **79**(5), 056104 (2009)
10. Filisetti, A., Villani, M., Roli, A., Fiorucci, M., Poli, I., Serra, R.: On some properties of information theoretical measures for the study of complex systems. In: Pizzuti, C., Spezzano, G. (eds.) WIVACE 2014. CCIS, vol. 445, pp. 140–150. Springer, Cham (2014). https://doi.org/10.1007/978-3-319-12745-3_12

11. Filisetti, A., Villani, M., Roli, A., Fiorucci, M., Serra, R.: Exploring the organisation of complex systems through the dynamical interactions among their relevant subsets. In: Proceedings of the European Conference on Artificial Life, pp. 286–293 (2015)
12. Haken, H.: An introduction. Synergetics, pp. 1–387. Springer, Heidelberg (2004). https://doi.org/10.1007/978-3-662-10184-1_1
13. Herrgård, M.J., Covert, M.W., Palsson, B.Ø.: Reconstruction of microbial transcriptional regulatory networks. Curr. Opin. Biotechnol. **15**(1), 70–77 (2004)
14. Huang, Y., Wange, R.L.: T cell receptor signaling: beyond complex complexes. J. Biol. Chem. **279**(28), 28827–28830 (2004)
15. Jeong, H., Tombor, B., Albert, R., Oltvai, Z.N., Barabási, A.L.: The large-scale organization of metabolic networks. Nature **407**(6804), 651–654 (2000)
16. Johnson, J.: Hypernetworks in the Science of Complex Systems, vol. 3. World Scientific, Singapore (2013)
17. Johnston, H.: Cliques of a graph-variations on the Bron-Kerbosch algorithm. Int. J. Parallel Prog. **5**(3), 209–238 (1976)
18. Lane, D., Pumain, D., van der Leeuw, S.E., West, G.: Complexity Perspectives in Innovation and Social Change, vol. 7. Springer Science and Business Media, Berlin (2009). https://doi.org/10.1007/978-1-4020-9663-1
19. Mendoza, L., Xenarios, I.: A method for the generation of standardized qualitative dynamical systems of regulatory networks. Theor. Biol. Med. Model. **3**(1), 13 (2006)
20. Newman, M.E., Girvan, M.: Finding and evaluating community structure in networks. Phys. Rev. E **69**(2), 026113 (2004)
21. Ravasz, E., Somera, A.L., Mongru, D.A., Oltvai, Z.N., Barabási, A.L.: Hierarchical organization of modularity in metabolic networks. Science **297**(5586), 1551–1555 (2002)
22. Remy, E., Ruet, P., Mendoza, L., Thieffry, D., Chaouiya, C.: From logical regulatory graphs to standard petri nets: dynamical roles and functionality of feedback circuits. In: Priami, C., Ingólfsdóttir, A., Mishra, B., Riis Nielson, H. (eds.) Transactions on Computational Systems Biology VII. LNCS, vol. 4230, pp. 56–72. Springer, Heidelberg (2006). https://doi.org/10.1007/11905455_3
23. Serra, R., Villani, M.: Modelling Protocells. Springer Science and Business Media, Dordrecht (2017). https://doi.org/10.1007/978-94-024-1160-7
24. Roli, A., Villani, M., Caprari, R., Serra, R.: Identifying critical states through the relevance index. Entropy **19**(2), 73 (2017)
25. Ruppin, E., Papin, J.A., De Figueiredo, L.F., Schuster, S.: Metabolic reconstruction, constraint-based analysis and game theory to probe genome-scale metabolic networks. Curr. Opin. Biotechnol. **21**(4), 502–510 (2010)
26. Sani, L., Amoretti, M., Vicari, E., Mordonini, M., Pecori, R., Roli, A., Villani, M., Cagnoni, S., Serra, R.: Efficient search of relevant structures in complex systems. In: Adorni, G., Cagnoni, S., Gori, M., Maratea, M. (eds.) AI*IA 2016. LNCS (LNAI), vol. 10037, pp. 35–48. Springer, Cham (2016). https://doi.org/10.1007/978-3-319-49130-1_4
27. Thomas, R., Kaufman, M.: Multistationarity, the basis of cell differentiation and memory. l. structural conditions of multistationarity and other nontrivial behavior. Chaos Interdisc. J. Nonlinear Sci. **11**(1), 170–179 (2001)
28. Thomas, R., Thieffry, D., Kaufman, M.: Dynamical behaviour of biological regulatory networks-l. Biological role of feedback loops and practical use of the concept of the loop-characteristic state. Bull. Math. Biol. **57**(2), 247–276 (1995)

29. Tononi, G., McIntosh, A.R., Russell, D.P., Edelman, G.M.: Functional clustering: identifying strongly interactive brain regions in neuroimaging data. Neuroimage **7**(2), 133–149 (1998)
30. Tononi, G., Sporns, O., Edelman, G.M.: A measure for brain complexity: relating functional segregation and integration in the nervous system. Proc. Nat. Acad. Sci. **91**(11), 5033–5037 (1994)
31. Vicari, E., Amoretti, M., Sani, L., Mordonini, M., Pecori, R., Roli, A., Villani, M., Cagnoni, S., Serra, R.: GPU-based parallel search of relevant variable sets in complex systems. In: Rossi, F., Piotto, S., Concilio, S. (eds.) WIVACE 2016. CCIS, vol. 708, pp. 14–25. Springer, Cham (2017). https://doi.org/10.1007/978-3-319-57711-1_2
32. Villani, M., Filisetti, A., Benedettini, S., Roli, A., Lane, D., Serra, R.: The detection of intermediate-level emergent structures and patterns. In: Proceedings of the European Conference on Artificial Life, pp. 372–378 (2013)
33. Villani, M., Roli, A., Filisetti, A., Fiorucci, M., Poli, I., Serra, R.: The search for candidate relevant subsets of variables in complex systems. Artif. Life **21**, 412–431 (2015)
34. Watts, D.J., Strogatz, S.H.: Collective dynamics of "small-world" networks. Nature **393**(6684), 440–442 (1998)

Author Queries

Chapter 10

Query Refs.	Details Required	Author's response
AQ1	This is to inform you that corresponding author has been identified as per the information available in the Copyright form.	
AQ2	As per Springer style, both city and country names must be present in the affiliations. Accordingly, we have inserted the city and country names in all affiliation. Please check and confirm if the inserted city and country names are correct. If not, please provide us with the correct city and country names.	