Who Plays Complex Music? On the Correlations Between Structural and Behavioral Complexity Measures in Sign Boolean Networks.

Rémi Segretain, Laurent Trilling, Nicolas Glade, Sergiu Ivanov

Who Plays Complex Music?

On the Correlations Between Structural and Behavioral Complexity Measures in Sign Boolean Networks

Rémi Segretain*, Laurent Trilling*, Nicolas Glade*, Sergiu Ivanov†

* University Grenoble Alpes, CNRS UMR5525, CHU Grenoble Alpes, Grenoble INP
TIMC-IMAG, F-38000 Grenoble, France
Emails: {remi.segretain,laurent.trilling,nicolas.glade}@univ-grenoble-alpes.fr
† Université Paris-Saclay, Univ Évry, IBISC, 91020, Évry-Courcouronnes, France
Email: sergiu.ivanov@ibisc.univ-evry.fr

Abstract—Intuition tells us that highly complex structure should be strongly correlated with highly complex behavior. In this work, we show that, while complex behavior does require complex structure, the converse is not necessarily true. Indeed, structural complexity can be also used to implement robust behavior, or even a variety of different relatively simple behaviors. To obtain these results, we explored the spaces of sign Boolean networks (SBNs) containing 2, 3, and 4 nodes, and we used complexity measures introduced in our previous work to study the relationship between structural and behavioral complexities of these networks.

I. INTRODUCTION

When one observes a mechanical or an electronic device, or even a natural system that we consider complex, common sense suggests its function or behavior should be too. However this is not necessarily always the case. Complexity can also support robustness, or it can be allocated to obtaining multiple behaviors. Error correction, spatial probes and rovers, protein folding, gene regulatory networks, all need robustness, often based on redundancy but also on the re-usability of some functional parts [1]–[4]. Finally, a complex structure can carry out multiple different functions, neither of which is particularly complex nor robust. One wonders whether the unstructured protein p53 can be seen as an example of such a structure, since it may function as a relatively simple switch for multiple cellular processes [5], [6].

To properly consider the relation between the complexity of the structure of a system and its behavior, we rely on *Sign Boolean Networks* or SBNs. Boolean networks can be used to represent gene regulation, but can also be viewed as good abstraction of any system of interconnected entities whose states can actually be classified into "on" and "off" groups [7]–[11]. In this context, the behaviour of the modelled system can be reasonably associated with the entire asymptotic dynamics (attractors) of the Boolean network, or a part of it, *i.e.* limited to the dynamics of only few entities. Various aspects of dynamics can be precisely defined and measured for Boolean networks, such as the number of attractors—which may correspond in biological interaction systems to multiple behaviors, the average size of the basins of attraction—which can be viewed as a source of robustness, or the average complexity of the attractors themselves. SBNs are a particular class of Boolean networks, in which every node is a 0-threshold function, called signed Boolean function (SBF): the new state of an SBF is computed by comparing the weighted sum of its inputs to 0. This particularly homogeneous structure is very useful for formal analysis, as it allows for computing a natural measure of structural complexity by combining the individual complexities of the participating SBFs, capturing the local functionality, with the global complexity of their interactions, or the layout of the network. For a detailed discussion on the methodology for computing BN dynamics and the structural and dynamic complexity of SBNs we refer to [12]–[15].

In this paper, we consider the relationship between the structure and the behavior of SBNs. More precisely, we focus on the connection between the local (related to the SBFs themselves) and global (related to the connections between SBFs) structural complexities of SBNs and their number of attractors (*i.e.* multiplicity of behaviors), the average size of the basins of attraction (*i.e.* robustness), as well as the average complexity of the attractors (*i.e.* complexity of behaviors). Our goal is to investigate how these 3 observables of the dynamics are correlated with the structural complexity, *e.g.* to evaluate whether the most complex behaviors are observed in the most structurally complex SBNs. Considering the influence of structural complexity on various aspects of dynamic complexity helps understand the organisation of complex systems, of which the living organisms are a typical example. Intuitively, one would consider that structurally complex parts of an organism implement complex behaviors. In this work we suggest that this intuition may not always correspond to reality, as high structural complexity may also give rise to simpler but very robust behavior, or to multiple different behaviors, as illustrated in Figure 1. Furthermore, local complexity of individual SBFs and the global complexity of the layout of the SBN may have unequal contributions to the behavioral complexity.

(A) Complex attractor

Number of attractors:	3
Average attractor size:	$^8/_3$ (6, 1, and 1)
Average basin size:	$^8/_3$ (6, 1, and 1)
Rarity:	0.87
Centrality:	$\simeq 1$
Average behavior complexity:	90.9 (267.7, 2.5 and 2.5)

(B) Multiple simpler attractors

Number of attractors:	8
Average attractor size:	1
Average basin size:	1
Rarity:	0.94
Centrality:	0.8
Average behavior complexity:	2.5

(C) Large basins of attraction

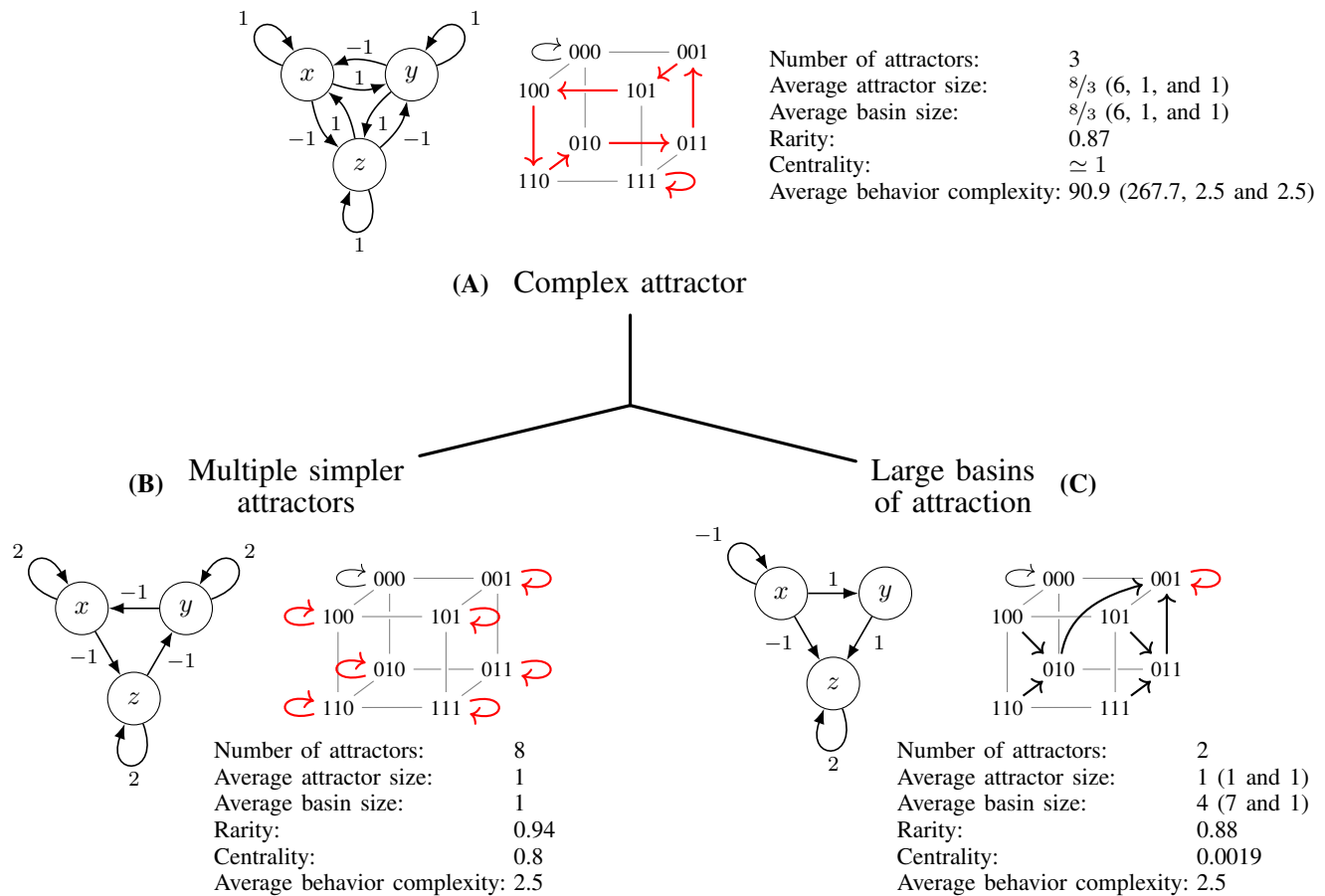| Number of attractors:	2
Average attractor size:	1 (1 and 1)
Average basin size:	4 (7 and 1)
Rarity:	0.88
Centrality:	0.0019
Average behavior complexity:	2.5

Fig. 1. **How does structural complexity distribute in different kinds of behaviors?** This figure illustrates 3 typical extreme situations in which structural complexity distributes differently into the 3 parameters of the dynamics evaluated in this work: the number of attractors (*i.e.* multiplicity of behaviors), their average complexity, and the average size of their basins of attraction (*i.e.* robustness). In the transition graphs, the red arrows indicate transitions belonging to an attractor, while the black ones indicate transitions not belonging to an attractor. A number of parameters are shown for each of the 3 SBNs, notably their structural data, *i.e.* the rarity of their SBFs and the centrality of the SBN. **(A)** illustrates an SBN and its transition graph containing a relatively complex 3-state attractor. In **(B)**, the SBN generates multiple simpler attractors with small basins of attraction. **(C)** shows a robust behavior with its rather simple attractor consisting only of the state 111, but whose basin of attraction covers the rest of the network, with the exception of the state 000.

II. METHODOLOGY

We computed the structural complexity of all sign Boolean networks (SBNs) of a given dimension d (60 SBNs for $d = 2$, 31065 SBNs for $d = 3$, and about 3.6 10^9 for $d = 4$) by combining the local complexities of the individual SBFs with the global complexity of the network structure. In [13], we thought the local complexity of an SBF as related to its *rarity*—the inverse of the probability of randomly picking the d particular SBFs that compose the SBN according to a uniform probability distribution, and the global complexity of the layout as a centralization index, aggregating the *centralities* of the nodes of the SBN—the probabilities that they influence each other. For each of these networks, we also computed the complexity of their behaviors, according to the following reasoning.

SBNs feature a deterministic evolution, necessarily ending up in a cyclic attractor or in a stable state. Therefore, the significant part of the dynamics is described by the sequence of states corresponding to a period of the attractor. In

addition, some nodes may be designated as observers, and the dynamics of the non-observer nodes may be discarded. In our work, we say that a binary sequence $s \in \{0, 1\}^*$ is a *behavior* of an SBN if this binary sequence is generated by at least one node during exactly one period of the cyclic attractor. We sometimes refer to a behavior of an SBN as a music it plays, since Boolean networks seem to play repeated binary musics. Observing the music played by only a subset of nodes makes a lot of sense biologically speaking, because only a subset of variables defining an observed living system is generally accessible.

Since the behaviors we consider in our work are at most of length 2^4, we use the Coding Theorem Method (CTM) in conjunction with the Block Decomposition Method (BDM), both described in [16], to compute their complexities. CTM essentially consists in enumerating all small Turing machines and counting how many of them produce the given binary sequence. BDM consists in splitting a longer sequence into blocks, computing the complexity of every block using CTM,

and then aggregating individual block complexities to obtain the complexity of the entire string. The combination of BDM with CTM was shown to approximate well the Kolmogorov complexity of the string.

For more details, we refer the reader to our previous works [13]–[15]).

III. RESULTS

We exhaustively explored the space of 2-, 3-, and 4-node networks ($d \in \{2, 3, 4\}$) and measured their structural complexity indices—average node centrality, average SBF rarity, as well as behavioral complexity indices—the number of attractors, average size of the basins of attraction, the complexity of the exhibited behaviors. We also complemented the structural complexity indices with the average number of inputs of the SBFs (arity). We explored the spaces of 2- and 3-node networks exhaustively, and the space of 4-node networks by sampling every 100th equivalence class. Figure 2 shows a number of 2D-histograms relating structural and behavioral complexities of 3-node networks ($d = 3$). The appendix also includes the histograms relating the same parameters for 4-node networks ($d = 4$).

Figure 2 (A) shows that networks covering the entire spectrum of centralities may generate simple behaviors. In fact, there are more than 10^4 3-node networks with centrality index equal to 1, but whose behavior is very simple. On the other hand, the higher the complexity of the behavior, the higher the centrality needs to be. In particular, the majority of complex behaviors (complexity > 250) are generated by networks with centrality index equal to 1.

One remarks in Figure 2 (A) that networks of high centrality tend to employ SBFs of lower rarity. Figure 2 (B) zooms in on the relation between arity, rarity, and centrality. It follows from this plot that rarer SBFs tend to have lower arity, and therefore belong to SBNs of lower centrality. On the other hand, high-centrality networks are moderately to highly frequent, and tend to have higher average arity.

Figures 2 (C) and (D) both illustrate the strong connection between the number of attractors of a network and the average size of their basins of attraction. Indeed, since the size of the state space is fixed ($2^3 = 8$ states for 3-node networks), larger basins of attraction imply fewer attractors. The distribution of SBF rarity and SBN centrality in these plots is in accordance with plot (B): more attractors with smaller basins of attraction tend to employ rarer functions in networks of lower centrality, while the largest basins of attraction employ moderately rare functions in high centrality networks.

Figures 2 (E) and (F) focus on the connection between the average number of attractors and behavioral complexity. It follows that higher complexity requires fewer attractors, and numerous attractors tend to produce simpler behavior on average. Remarkably, however, there exist networks with 3 attractors which produce behaviors as complex as those of some of the networks with 2 attractors. Average SBF rarity and SBN centrality is distributed in these graphs as expected from the previous histograms. In particular, attractors producing the most complex behavior employ SBFs of moderate rarity, but of high SBN centrality.

Figures 2 (G) and (H) complement the previous pair of plots by highlighting the connection between the average size of the basins of attraction and behavioral complexity. It turns out that highly complex behavior is generally produced by SBNs having attractors with larger basins, employing functions of moderate rarity, but having quite high SBN centrality. In fact, even moderately complex behaviors (of complexity approaching 50 from below) tend to require high centrality, while the arity of the SBFs involved in producing such behaviors covers almost the whole spectrum. We recall that, formally, the states of the attractor are part of the basin of attraction; attractors associated with complex behavior may occupy most of their respective basins of attraction. However, it is also noticeable that some SBNs including rare SBFs (*i.e.* with low arity), although having a low centrality, may display quite important complex behaviors, for example those that are shown in Figures 2 (G) and (H), and have a behavior complexity of 80 and 4 states in their basin of attraction.

Finally, we explicitly remark that most of the values appearing in these histograms are average: SBF rarities, arities, and centralities, as well as the sizes of the attractors and their basins of attraction. This is why all figures include violin plots showing the distribution of means and standard deviations of the measures defining the colors. For example, in Figure 2, the SBNs have average SBF rarity around 0.9, with low variation. On the other hand, SBNs having large variations of SBF rarity tend to employ relatively unrare SBFs, and are rather few. The violin plot in (B) shows that the variations of centralities in lower-centrality SBNs can be quite noticeable, while at the same time, in higher-centrality SBNs all nodes tend to have rather high centrality. In plots (C), (E), and (G), most SBNs tend to employ functions with rarity around 0.8, with negligible variations, except for SBNs containing lower-rarity SBFs. The spectrum of centralities is on the other hand richer in plots (D), (F), and (H). However, and in a remarkably similar way to rarity, when SBNs already contain some SBFs with high arity most of the SBN nodes tend to be of high arity too.

IV. CORRELATIONS BETWEEN STRUCTURAL AND BEHAVIORAL COMPLEXITY

The numerical data we present in the previous section give a rather detailed picture of the connections between different indices of structural and behavioral complexity. In particular, they show that while complex structure does sometimes correlate with complex behavior, it is quite more frequent for structural complexity to be invested into larger basins of attraction (higher robustness) or into more attractors (multiplicity of behaviors). In this work we do not bring forward definitive conclusions concerning the relationship between the structure and the function of the networks, but the numerical data and the observations we presented here lay down future research directions which will make this relationship clearer and more
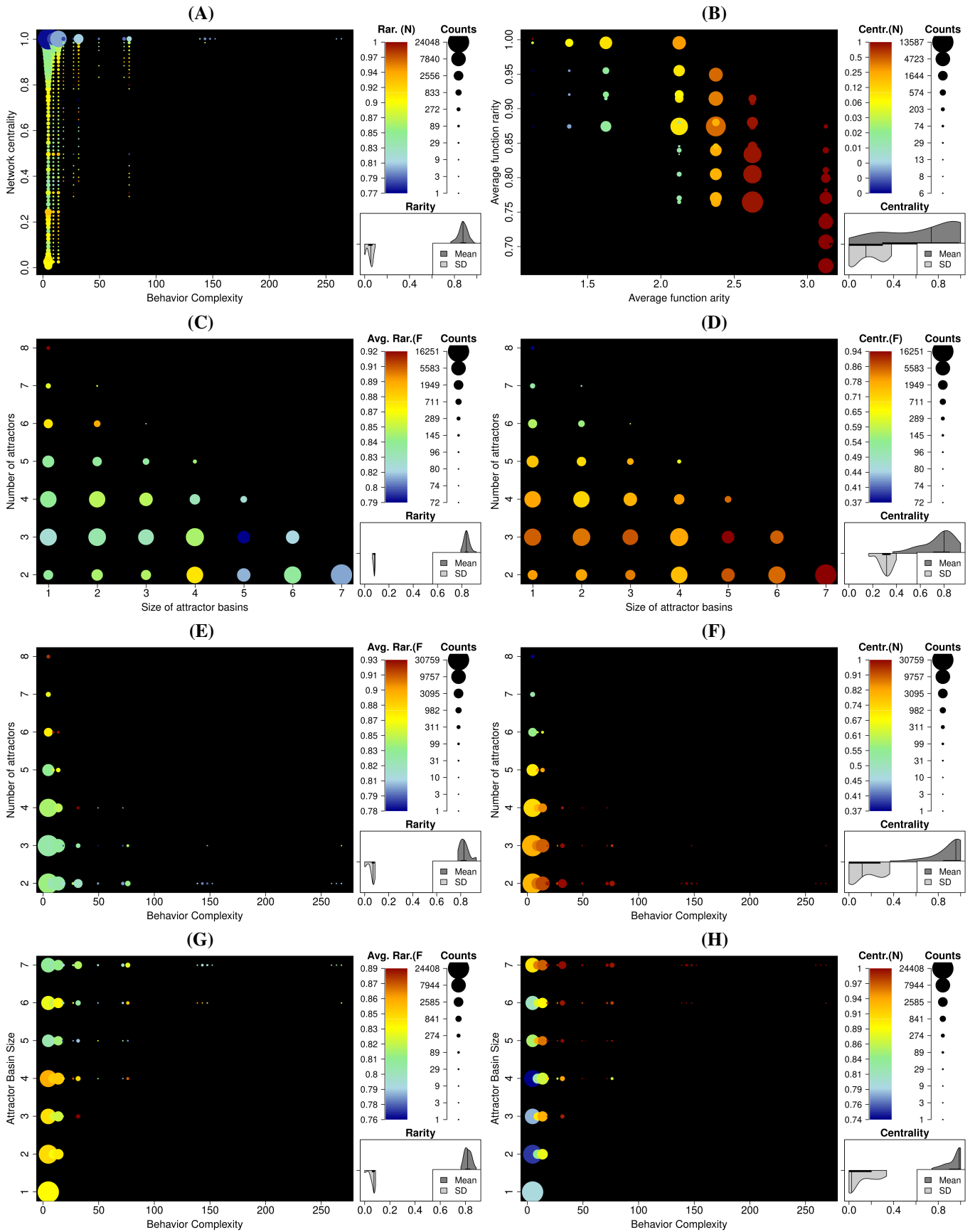
Fig. 2. **Parameters of structural and behavioral complexities for** $d = 3$. These histograms illustrate the connections between the structural complexities—centrality, rarity, and arity—and the number of attractors, the size of the basins of attraction, and the complexity of the behaviors, for all 3-node networks. Average values are shown. Each histogram is accompanied by a violin plot giving the distribution of the mean and the standard deviation of the corresponding measure. **(A)** shows that most networks produce simple behaviors, even the most structurally complex ones. **(B)** illustrates the relationship between rarity, arity, and centrality. **(C)** and **(D)** indicate that the number of attractors and the sizes of their basins of attraction are inversely related, and also show distributions of centrality and rarity. Finally, **(E)** and **(F)** show how rarity and centrality are connected to the number of attractors and the behavioral complexity, while **(G)** and **(H)** illustrate the relation with the size of the basins of attraction.

usable in the study of living systems. As an important example, it would be relevant to identify structural patterns leading specifically to high behavioral complexity, multiplicity of behaviors, and high robustness.

While the strategy of exhaustive exploration of the SBN space allows for making comprehensive observations, it is not necessarily representative of the networks which appear in concrete biological systems. For example, a gene in a regulatory network is typically influenced by a small number (< 10) of other genes (e.g. [17]), while neurons in neural networks tend to integrate many more inputs (often tens to hundreds, or even thousands in the case of Purkinje cells, *e.g.* [18]). Therefore, *subspaces* of the SBN space should instead be considered, based on the observed organisation of the systems of interest. In these subspaces, structural and behavioural complexity may be correlated differently, and new specific connections may emerge.

In this work, we focus on a particular set of structural and behavioral complexity measures, which are well suited for generic SBNs. Different new complexity measures may be introduced to capture the particularities of classes of SBNs sharing some properties with real-life networks. In fact, these measures should be stated explicitly as part of the modelling framework, and should be targeted towards a specific objective, as complexity is always a formal representation of a part of a question or problem.

Our observations have a number of methodological implications. One typically studies a complex living system part by part, trying to tackle its inherent complexity by identifying modules. Our work confirms the intuition that this approach is not always viable: Figures 2 (F) and (H) show for example that many behaviors, even the simplest ones, may be implemented in networks with high average node centrality. Finding modules in such structures may be hard or even impossible, as all nodes may contribute in some ways to all exhibited behaviours.

In fact, in many modelling scenarios, the actual structure is hard to identify correctly due to the complex interplay of different factors in the production of the behavior of interest. In these cases, one typically designs an ersatz structure, which is meant to represent the partially unknown mechanisms at a high level of abstraction. This is often the case with Boolean approaches, in which the structure of the network is derived from the observed time series (e.g., [17]). A finer-grained understanding of how structural complexity distributes into various aspects of behavioral complexity for the modelled networks may therefore help design better models.

References

[1] R.W. Hamming, "Error detecting and error correcting codes." *Bell System Technical Journal. USA: AT&T.* vol 29, no 2, pp 147–160, 1950

[2] S. Chien, A. Barrett, T. Estlin, G. Rabideau, "A comparison of co-ordinated planning methods for cooperating rovers." In AGENTS '00, Proceedings of the Fourth International Conference on Autonomous Agents, pp 100–101, 2000

[3] D.C. Ferreira, M.G. van der Linden, L.C. de Oliveira, J.N. Onuchic, A.F. de Araújo, "Information and redundancy in the burial folding code of globular proteins within a wide range of shapes and sizes." *Proteins*, vol 84, no 5, pp 515-531, 2016

[4] Y. Gao, N. Zhu, X. Zhu, M. Wu, C.Z. Jiang, D. Grierson, Y. Luo, W. Shen, S. Zhong, D.Q. Fu, and G. Qu, "Diversity and redundancy of the ripening regulatory networks revealed by the fruitENCODE and the new CRISPR/Cas9 CNR and NOR mutants." *Horticulture Research* vol 6, p39, 2019

[5] S. Surget, M.P. Khoury, and J.C. Bourdon. "Uncovering the role of p53 splice variants in human malignancy: a clinical perspective". *OncoTargets and therapy*, vol 7, pp 57–68, 2013

[6] E. Toufektchan and F. Toledo, "The Guardian of the Genome Revisited: p53 Downregulates Genes Required for Telomere Maintenance, DNA Repair, and Centromere Structure." *Cancers (Basel)*, vol 10, no 5, p 135, 2018

[7] W. S. McCulloch, W. Pitts, "A Logical Calculus of the Ideas Immanent in Nervous Activity." *Bulletin of Mathematical Biophysics*, vol 5, pp 115–133, 1943

[8] J. J. Hopfield, "Neural networks and physical systems with emergent collective computational abilities." *Proceedings of the National Academy of Sciences of the USA*, vol 79, pp 2554–2558, 1982

[9] L. Mendoza, E. R. Alvarez-Buylla, "Dynamics of the genetic regulatory network for *Arabidopsis thaliana* flower morphogenesis." *Journal of Theoretical Biology*, vol 193, pp 307–319, 1998

[10] S. Bornhold, "Boolean network models of cellular regulation: prospects and limitations." *Journal of the Royal Society Interface*, vol 5, pp S85–S94, 2008

[11] V. Tran, M. N. McCall, H. R. McMurray, A. Almudevar, "On the underlying assumptions of threshold Boolean networks as a model for genetic regulatory network behavior." *Frontiers in Genetics*, vol 4, no 263, 2013

[12] Q.-T. Vuong, R. Chauvin, S. Ivanov, N. Glade, L. Trilling, "A logical constraint-based approach to infer and explore diversity and composition in threshold Boolean automaton networks." In Complex Networks & Their Applications VI, Proceedings of the Complex Networks'17 conference, *Studies in Computational Intelligence Series*, vol 689. Springer, Cham., 2017, pp 567-579

[13] R. Segretain, S. Ivanov, L. Trilling, and N. Glade, "A Methodology for Evaluating the Extensibility of Boolean Networks' Structure and Function." In Complex Networks & Their Applications IX., Proceedings of the Complex Networks'20 conference. *Studies in Computational Intelligence*, vol 944. Springer, Cham., 2020, pp 372-385

[14] R. Segretain, S. Ivanov, L. Trilling, and N. Glade. "Implementation of a Computing Pipeline for Evaluating the Extensibility of Boolean Networks' Structure and Function." bioRxiv 2020.10.02.323949

[15] U. Christen, S. Ivanov, R. Segretain, L. Trilling, and N. Glade. "On computing structural and behavioral complexities of threshold Boolean networks." *Acta Biotheoretica*. vol 68, pp 119–138, 2020

[16] H. Zenil, S. Hernández-Orozco, N.A. Kiani, F. Soler-Toscano, A. Rueda-Toicen, and J. Tegnér, "A Decomposition Method for Global Evaluation of Shannon Entropy and Local Estimations of Algorithmic Complexity". *Entropy*, vol 20, no 8, 605, 2018

[17] J.G.T. Zañudo, S.N. Steinway, and R. Albert, "Discrete dynamic network modeling of oncogenic signaling: Mechanistic insights for personalized treatment of cancer." *Current Opinion in Systems Biology*, vol 9, pp 1–10, 2018

[18] K.E. Joyce, P.J. Laurienti, J.H. Burdette, S. Hayasaka, "A New Measure of Centrality for Brain Networks." *PLoS ONE*, vol 5, no 8, e12200, 2010
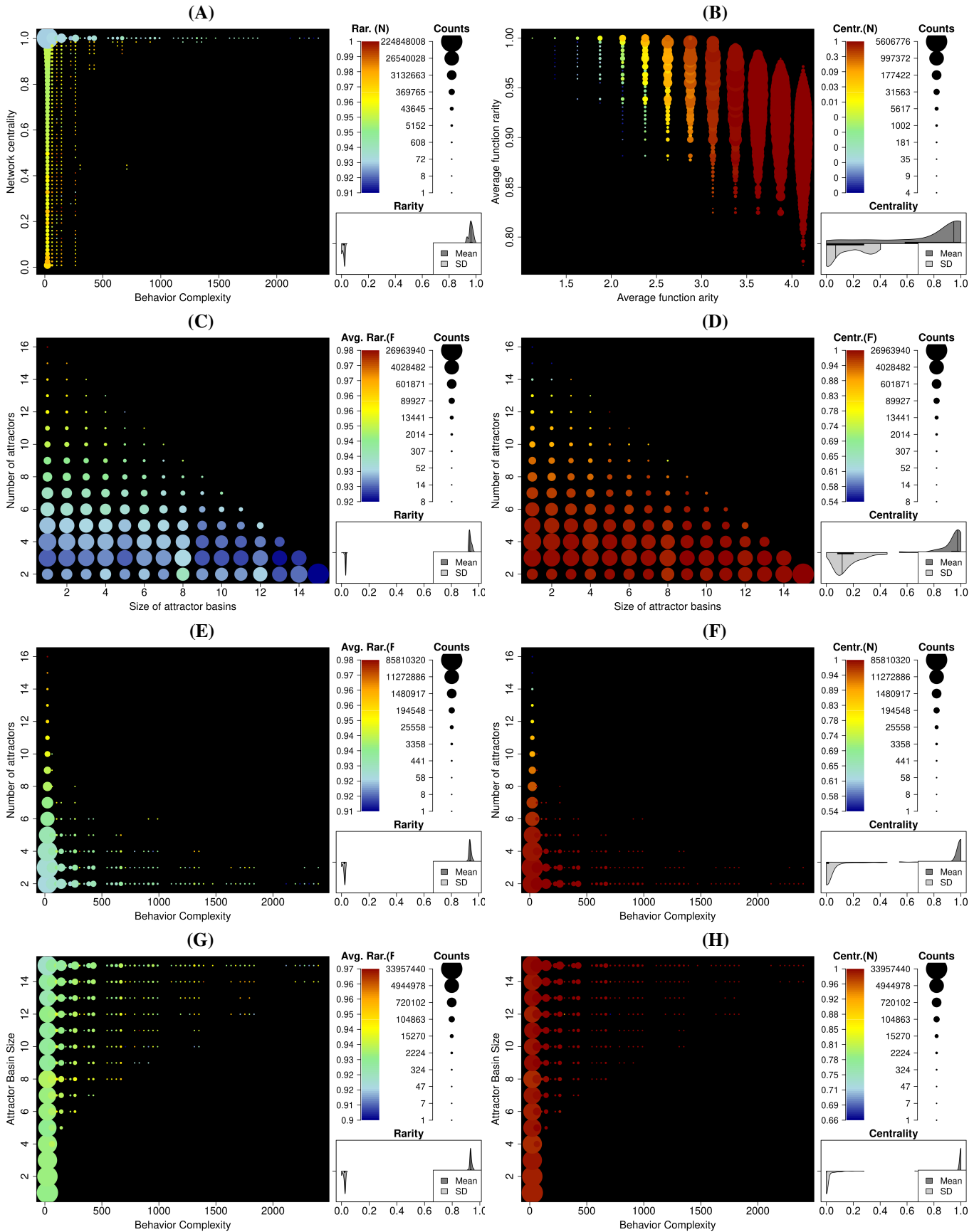
Fig. 3. **Parameters of structural and behavioral complexities for** $d = 4$. As in Figure 2 in the main text of the paper, these histograms illustrate the connections between the structural complexities—centrality, rarity, and arity—and the number of attractors, the size of the basins of attraction, and the complexity of the behaviors, for all 4-node networks. Average values are shown. Each histogram is accompanied by a violin plot giving the distribution of the mean and the standard deviation of the corresponding measure.