Boolean Threshold Networks as Models of Genotype-Phenotype Maps

Chico Q. Camargo[1(✉)] and Ard A. Louis[2]

[1] Oxford Internet Institute, University of Oxford, Oxford, UK
chico.camargo@oii.ox.ac.uk
[2] Rudolf Peierls Centre for Theoretical Physics, University of Oxford, Oxford, UK

Abstract. Boolean threshold networks (BTNs) are a class of mathematical models used to describe complex dynamics on networks. They have been used to study gene regulation, but also to model the brain, and are similar to artificial neural networks used in machine learning applications. In this paper we study BTNs from the perspective of genotype-phenotype maps, by treating the network's set of nodes and connections as its genotype, and dynamic behaviour of the model as its phenotype. We show that these systems exhibit (1) Redundancy, that is many genotypes map to the same phenotypes; (2) Bias, the number of genotypes per phenotypes varies over many orders of magnitude; (3) Simplicity bias, simpler phenotypes are exponentially more likely to occur than complex ones; (4) Large robustness, many phenotypes are surprisingly robust to random perturbations in the parameters, and (5) this robustness correlates positively with the evolvability, the ability of the system to find other phenotypes by point mutations of the parameters. These properties should be relevant for the wide range of systems that can be modelled by BTNs.

Keywords: Boolean networks · Gene regulatory networks · Genotype-phenotype maps · Input-output maps

1 Introduction

Boolean networks (BN) were first introduced by Stuart Kauffman as models of gene regulatory networks (GRNs) [28]. Each gene is abstracted as a binary (Boolean) variable that can be either on or off. A temporal dynamics is imposed, where each gene interacts with others through a Boolean function. For example, if $S_i(t)$ is the state of a node i at time t, then

$$S_1(t+1) \;=\; (\, S_1(t) \;\texttt{AND}\; S_2(t) \,) \;\texttt{OR NOT}\; (\, S_3(t) \;\texttt{AND}\; S_4(t) \,) \tag{1}$$

is the state of the node S_1 at time $t+1$, and this is influenced by the states of nodes S_2, S_3, and S_4. The state of every node in the network is then updated synchronously with all other nodes in the network with rules similar to the example above. The requirement of synchronicity is sometimes dropped, in what

H. Barbosa et al. (Eds.): *Complex Networks XI*, SPCOM, pp. 143–155, 2020.
https://doi.org/10.1007/978-3-030-40943-2_13

is called asynchronous updating. In Kauffman's original formulation [28], he defined an NK network of N nodes with K connections. Figure 1 shows an example of this kind of network, for $N = 3$ and $K = 3$, showing how different initial configurations of the network can lead to different attractors in its state space. Here attractors refer to steady state behaviour that the system falls into.

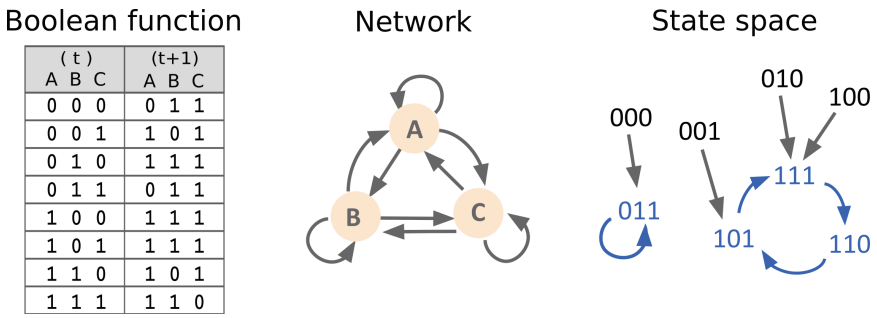

Fig. 1. Example of Boolean network with $N = 3$, $K = 3$. In this network, all nodes are regulated by the whole network ($N = K$). The table on the left defines the Boolean functions that describe how the state of nodes A, B and C will change according to their states in the previous time step. The $2^3 = 8$ states in state space are organised in two attractors, one fixed point and one 3-cycle, both shown in blue.

Even though they are highly simplified models, BNs are known to be extremely versatile. They have been used to represent complex systems at multiple scales, ranging from gene regulatory networks to ecosystems [29].

In the context of gene regulation, one popular application has been to model pluripotent cells as a gene network with many attractors [29]. Each cell type then corresponds to a different attractor. Another famous application of the BN model was to segment polarity in *Drosophila* [2], where it was shown that simplifying a highly complex differential equation model of the GRN to a much simpler BN model nevertheless led to good agreement with experiments. Similar successful applications have since proliferated, including BN models for the GRNs regulating flower development [18], signal transduction in human fibroblasts [25], plant cell signalling [33], and mammalian cortical development [20]. BNs have also been used to study more general properties of GRNs, such as robustness [5,11,12,41] and evolvability [6,11,12,41].

In the context of neural networks, different attractors have been interpreted as different memories stored in the network [26,27,35]. Attractors can also represent different heartbeat rhythms for cardiac systems [21], or alternative species distributions when modelling ecosystems [34].

The versatility of BN models comes at the price of having a large parameter space. As each node in a NK network is assigned a Boolean function on the 2^K possible states of those K nodes, there are 2^{2^K} possible functions per node, meaning that for $N = 3$ and $K = 3$, as in Fig. 1, there are over 3×10^9 possible

networks, whereas for a network with $N = 5$ and $K = 3$ this number grows to approximately 10^{17}. One way to overcome the impractical size of the parameter spaces of BNs is by focussing on a subset of all BNs. Boolean threshold networks (BTNs) correspond to one of these subsets, and are characterised by simple interaction rule:

$$S_i(t+1) = \begin{cases} 1, & \text{if } \sum_{j=0}^{N} w_{ij}\, S_j(t) > 0 \\ 0, & \text{if } \sum_{j=0}^{N} w_{ij}\, S_j(t) < 0 \end{cases} \tag{2}$$

where the weights w_{ij} indicate the strength and sign of the regulation of node i by node j. The value of $S_i(t+1)$ when $\sum_{j=0}^{N} w_{ij}\, S_j(t) = 0$ varies across studies, having values such as $S_i(t+1) = 1$, 0, or $S_i(t)$ itself. This choice has been shown not to have a big impact on results [14,31,36]. In this work, we will model GRNs as defined in Eq. (2), with $S_i(t+1)$ when $\sum_{j=0}^{N} w_{ij}\, S_j(t) = 0$, and $w_{ij} \in \{1, 0, 1\}$, respectively indicating upregulation, downregulation and lack of interaction between genes. This choice reflects the strong assumption that the effect of activatory and inhibitory interaction between genes is essentially additive, and that all gene-gene interactions are equally strong [48].

The threshold function in Eq. (2) inevitably reduces the range of behaviours of a GRN [48], but it also reduces its parameter space by orders of magnitude. Instead of requiring N Boolean functions to be defined, as would be the case for a Boolean network, a BTN can be completely specified by its $N \times N$ adjacency matrix w_{ij}. Each GRN or genotype can also be treated as a node in a genotype network where two genotypes are neighbours if they differ in only one term in their w_{ij} matrix [12]. Since every connection can be either -1, 0 or 1 the whole genotype space is composed of $3^{N \times N}$ gene networks, each one connected to $2N^2$ neighbours.

BTNs have also been used to successfully model gene expression, in GRNs such as the ones regulating lymphocyte differentiation [38], signal transduction in human fibroblasts [25], and the mammalian cell cycle [19]. One of the most successful applications of BTNs is in modelling the yeast cell cycle [14,15,31], which led to studies predicting knockout mutant phenotypes [7,15], as well as multiple studies providing explanations for the designability and robustness of the wild-type phenotype [3,4,8,10].

2 Genotype-Phenotype Maps

When used to study the evolutionary properties of GRNs, the mapping from a Boolean network's wiring to its dynamic properties can be studied as a genotype-phenotype (GP) map. GP maps have proven to be a useful lens in understanding the origins of phenotypic variation, and how this variation can steer evolution even before natural selection comes into play [39,46]. Much of this understanding comes from a growing body of research structural properties shared by multiple GP maps [1,24,46]. These properties include: (1) Redundancy, meaning that many genotypes often map to the same phenotypes, (2) Bias, referring to how the

majority of genotypes map only to a handful of phenotypes [1,23], (3) Simplicity bias, i.e. how simple phenotypes correspond to larger fractions of genotype space [16], (4) a degree of mutational robustness which is much larger than what would be expected for a random uncorrelated GP map [1,22,24,24], and (5) a positive correlation between robustness and evolvability [23,45,46].

Provided that the genotype of a GRN can be described by its topology and represented by its adjacency matrix w_{ij}, one still needs to choose how to represent its phenotype. Different definitions of phenotype might be more or less appropriate depending on what is being measured. For example, when modelling the regulation of circadian rhythm by a GRN that produces oscillating behaviour for multiple initial conditions, it might be convenient to define the phenotype of that GRN as its cyclic attractor in state space. If one is interested in the possible cell fates of a pluripotent cell, it might be better to look at the whole list of attractors. Alternatively, one might define a phenotype concerning a specific set of initial conditions, or the transitions between certain states. With that in mind, we explore two different phenotype definitions, that capture aspects of different kinds of biologically relevant phenotypes:

Phenotype 1: we define the phenotype of a GRN as the attractor occupying the largest fraction of its state space.

Phenotype 2: we define the phenotype as a list of the attractors corresponding to all possible initial states of the network.

The two phenotype definitions described above are very general, and might not be suited for all GRNs. For instance, it might be the case that not all initial states of the network are biologically relevant, or that BTNs that produce very similar behaviour from a dynamical systems point of view might represent very different biological behaviours. In spite of these limitations, we find that all phenotype definitions studied here result in GP maps which show many of the structural properties of GP maps listed above.

3 Results for Phenotype 1: Dominant Attractor

To identify the attractor with the largest basin for each genotype, we use the same method as Nochomovitz and Li [36]: first, we use the update function from Eq. (2) to produce a list of 2^N "next states", which we then connect as a directed graph from state to state, as illustrated in Fig. 1. We then use Tarjan's algorithm to calculate the strongly connected components of this directed graph, that is, the sections of state space where every state can reach every other state [42]. As the threshold update function is deterministic, every state will only lead to one other state, and the strongly connected components of the state graph will be the attractors of the BTN state space. Finally, we also take into account how multiple attractors might be equivalent under symmetry operations. These operations include permuting gene order, shifting the cyclic attractors by any number of steps, or swapping 1s for 0s for a given gene. This last form of symmetry has been described as a "gauge symmetry" [11,12]. We calculate the

dominant attractor for every GRN in the genotype space of all BTNs with N genes. In the sections below, we present the properties observed for this GP map, for the full enumeration of all $3^{16} \approx 43$ million $N = 4$ networks.

3.1 Phenotype Frequencies Vary over Orders of Magnitude

As there are many more genotypes than phenotypes in this GP map, it is natural that the map will show some redundancy – genotypes which map to the same phenotypes. In the GP map literature, the concept of redundancy is measured in terms of *neutral networks*, i.e. regions of a discrete genotype space that map to the same phenotype. The number of genotypes in a neutral network has been given different names in the literature, such as the phenotype's degeneracy level [9], abundance [13], designability [32,36], genotype set size [37], neutral set size [39] or neutral network size (NNS) [40,45]. In this work, we will also use the word frequency, meaning the NSS of a phenotype divided by the size of the whole genotype space. The frequency of a phenotype can also be understood as the probability that a randomly chosen genotype will map to that phenotype.

We performed a full enumeration of all $3^{16} \approx 4.3 \times 10^7$ GRNs with $N = 4$ genes. As can be seen in Fig. 2a, the frequencies of 2759 phenotypes observed for $N = 4$ networks range over many orders of magnitude. The distribution of phenotype frequencies is very skewed, with the most common attractor type, which is a single fixed point, covering 67% of genotype space. This uneven distribution of phenotype frequency is also observed within cyclic outputs of the

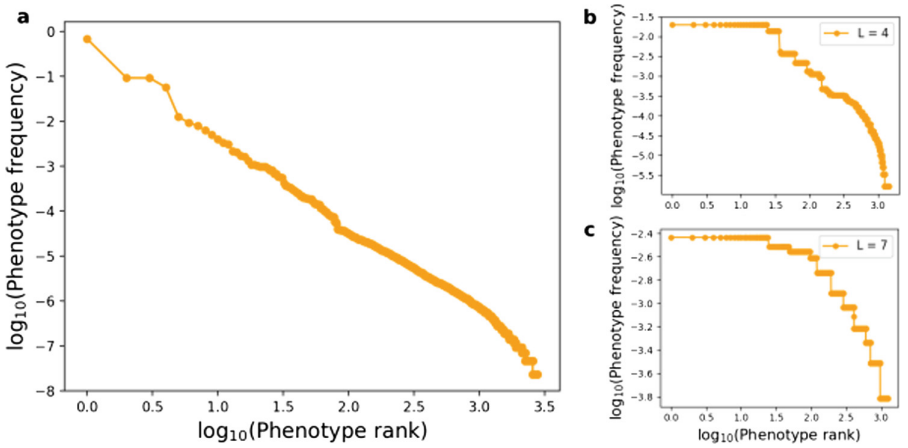

Fig. 2. This GP map shows redundancy and bias. **(a)** Rank plot for the number of gene networks that produce each one of all the 2759 dominant attractors found from a full enumeration of of $N = 4$ gene networks. It is a very uneven distribution, with 99% of all genotypes mapping to 0.38% of all phenotypes, and the most designable phenotype, a single fixed point, corresponding to 67% of the genotype space. The skewed distribution of neutral network sizes is also observed within cyclic outputs of the same length, shown in **(b)** and **(c)** for lengths $L = 4$ to 7 respectively.

same length, as already noted for $N = 4$ BTNs by Nochomovitz and Li [36]. Figures 2b and c illustrate this pattern for cyclic attractors of lengths $L = 4$ and $L = 7$ respectively.

3.2 Low-Complexity Attractors Correspond to More Genotypes

Not only do Figs. 2b and c show that there is a wide distribution of phenotype frequencies even among cyclic attractors of the same length, but they also show that the frequency of $L = 7$ cyclic attractors is orders of magnitude lower than that of $L = 4$ attractors. This is shown in Fig. 3c for other attractor lengths. Overall, the larger the cycle, i.e. the more states in the cycle, the smaller the number of GRNs which will produce that phenotype.

This negative relation between phenotype frequency and cycle length is part of a wider pattern, where low-complexity phenotypes take over larger fractions of genotype space. This behaviour, which Dingle et al. call *simplicity bias* [16], is found in numerous input-output maps, including but not limited to RNA sequence-to-structure maps [17], stochastic models in financial mathematics [16] and parameter-to-behaviour maps in deep neural networks [44]. This pattern has also been reported for genotype-phenotype maps for protein quaternary structures and macromolecular self-assembly and generative models of tree geometries [16]. And even though the specific definition of a phenotype depends on

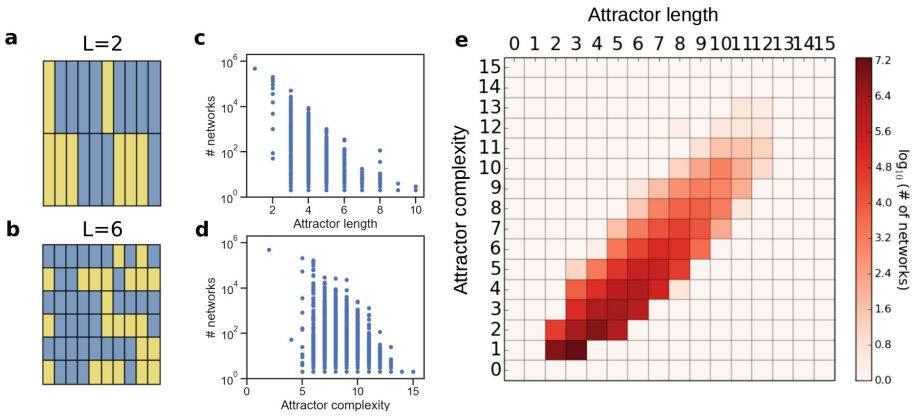

Fig. 3. This GP map shows simplicity bias. **(a)** and **(b)** show two cyclic attractors, of length $L = 2$ states and $L = 6$ states respectively, for a GRN with $N = 10$ genes. In both panels, every column represents the binary state of a node, with yellow for `1` and blue for `0`. **(c)** and **(d)** show this GP map is biased towards low-complexity phenotypes. Both measures show the number of GRNs mapping to a phenotype is bounded by an exponential decay in its complexity, as measured by their attractor length and Kolmogorov complexity respectively. Finally, **(e)** shows a heat map for all $N = 4$ GRNs, comparing attractor cycle length to attractor complexity, indicating that more GRNs produce phenotypes of lower complexity, even when comparing phenotypes of the same cycle length. Darker colours represent higher phenotype frequency.

the details of every GP map or input-output map, as long as a phenotype can be represented a binary string, its Kolmogorov complexity can be approximated using the Lempel-Ziv algorithm [30], as described by Dingle et al. [16]

Figures 3a and b respectively show examples of simple and complex phenotypes, for a GRN with $N = 10$ genes. Both panels represent cyclic attractors, with every column representing the binary state of a node and every row indicating a state in the cyclic attractor, with yellow for `1` and blue for `0`; for example, Fig. 3a indicates the 2-state cycle `1000110001` → `1110001110`, while Fig. 3b indicates a 6-state cycle starting with `0000001010` → `1001110111`.

Even though the length of a cyclic attractor in a BTN can be used as a proxy for its complexity [36], cycle length alone does not explain the variation of phenotype frequency that also happens among the cycles of the same equal length, as shown above in Figs. 2b and c.

Since two cycles with the same length can represent two different patterns of gene activation, an ideal measure of phenotype complexity should take these differences into account. Here we quantify these differences in complexity between different attractors by "stacking" all states of the cycle in a single binary string, such that a sequence such as `0110` → `1111` cycle would be represented as `01101111`. We then estimate the Kolmogorov complexity of the resulting string using the canonical Lempel-Ziv method [16,30]. The result of this analysis, presented in Fig. 2d, shows the same trend revealed by looking at cycle length in Fig. 2c: the number of genotypes mapping to a phenotype is bounded by a function that decreases exponentially with complexity – a hallmark of simplicity bias, as discussed by Dingle et al. [16].

Given that both cycle length and Kolmogorov complexity point towards the presence of simplicity bias in this GP map, as shown by Figs. 2c and d respectively, it could be the case that both measures are simply revealing the bias towards shorter attractors. This is, however, not the case. Figure 2e shows a heatmap where every column represents attractors of the same cycle length, such as the ones grouped in the same subplots in Fig. 2e. In the figure, the darker shades of red are always in the lower, low-complexity side of the red-shaded part of every column. Since darker colours represent more genotypes, this plot indicates that more genotypes map to phenotypes of lower complexity, even when comparing phenotypes of the same attractor length.

Even though defining the phenotype of a network as its main attractor in its state space is a practical way to model the behaviour of a GRN, this phenotype definition discards information about which initial conditions lead to that main attractor, while also ignoring other attractors present in the GRN state space. In a case where all biologically relevant initial conditions lead to the same attractors, such as the yeast cell cycle studied by Li et al. [31], these limitations are not a problem, but this definition of phenotype is not suited to represent a bistable system, or a cell that, provided the right initial conditions, might differentiate into multiple cell types [29].

With that in mind, in the next section we study a fine-grained GP map for GRNs, defining a phenotype taking into account the attractor corresponding to every possible initial condition in state space. Or, in biological terms, which gene expression patterns lead to which cellular behaviours.

4 Results for Phenotype 2: Multiple Attractors

For this analysis, we define the phenotype of a GRN with N genes as a 2^N-long string of the attractors corresponding to each one of its 2^N possible states. For example, for $N = 4$, a phenotype would be a string such as `AABACBDBCCCDDCCD`, indicating that that states `0000` and `0001` map to attractor A, state `0010` maps to attractor B, and so on.

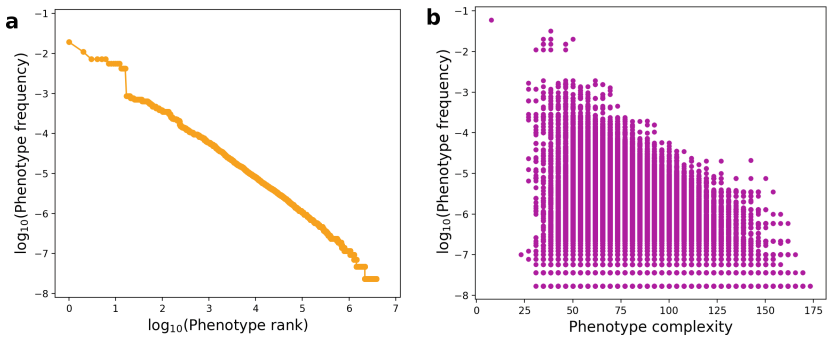

Fig. 4. (a) Rank plot for $N = 4$ GRNs showing a very biased distribution of phenotype frequency, for the phenotype defined as the list of attractors. **(b)** Scatter plot of phenotype frequency versus phenotype complexity, measured using the Lempel-Ziv algorithms, showing a strong bias towards simple phenotypes.

Having the phenotype written down as a string also allows us to calculate its Kolmogorov complexity using the Lempel-Ziv method. We do this by first iterating over all $3^{N\times N}$ possible genotypes for a given number of genes, and producing their corresponding strings of attractors. We then translate those phenotype strings into the shortest binary strings that can accommodate the range of attractors presented by this GP map. In the case of $N = 4$, the whole genotype space produces a total of 8172 different attractors, meaning that the index corresponding to each attractor can be represented in $\lceil \log_2(8172) \rceil = 13$ bits, from `0000000000001` to 1111111101100, and the whole phenotype string made of the $2^N = 16$ attractors corresponding to every initial state can be represented as a string with $16 \times 13 = 208$ bits, and its complexity can then be measured using the Lempel-Ziv method.

4.1 Redundancy and Simplicity Bias

As this GP map uses a finer-grained definition of phenotype, it produces a larger number of phenotypes than the more coarse-grained GP map presented in the previous section. Figure 4a shows a rank plot of the frequency of the 3.8×10^6 phenotypes produced by this GP map for $N = 4$ GRNs. This GP map also shows simplicity bias, as indicated in Fig. 4b. The scatter plot exhibits the characteristic

triangular shape shown for the previous GP map in Figs. 3c and d, with the most frequent phenotype being when all initial conditions lead to 0000 as a fixed point, representing when all genes turn inactive.

5 Robustness and Evolvability

In addition to having a highly biased distribution of neutral network sizes and a strong bias towards simple outputs, both GP maps studied in this paper show evidence of a very correlated neutral network landscape. This can be measured by comparing the frequency f_P of a given phenotype P, i.e. the fraction of the genotype space that maps to P, with the phenotype's mutational robustness ρ_P, which is defined as the chance that a mutation to one of those genotypes which map to P produces a phenotype which still maps to P. In other words, it is the chance that a mutation to one of those phenotypes is a *neutral mutation*, that is, a mutation that leads to the same phenotype [1,5,11,12,24,41].

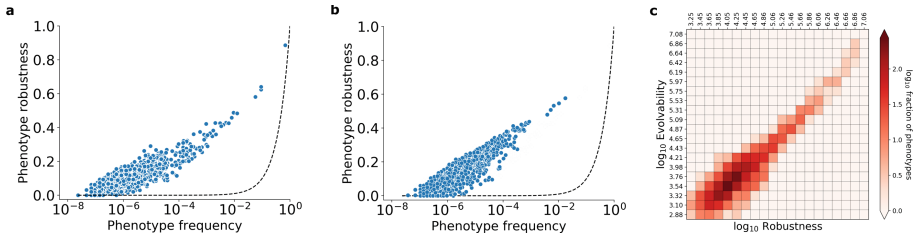

Fig. 5. Robustness and evolvability. Panels **(a)** and **(b)** show phenotype frequency versus phenotype robustness for phenotypes 1 and 2 respectively. In both cases, phenotype robustness grows proportionally to the logarithm of phenotype frequency. The dashed black line represents the random GP map expectation where $\rho_p = f_p$. Panel **(c)** shows a heatmap comparing phenotype robustness and evolvability for phenotype 2, showing that the most robust phenotypes are also the most evolvable ones.

For a random GP map where there is no correlation between neighbour, the chance of finding P in the neighbourhood of P's neutral network should be equal to the chance of finding P anywhere else in genotype space. Put simply, $\rho_P = f_P$. This is indicated by the black dashed lines in Figs. 5a and b. The GP maps studied here do not show this pattern: instead, phenotype robustness to mutations scales linearly with the logarithm of phenotype frequency, for both phenotype definitions. Finally, these GP maps also have a high correlation between phenotype robustness and phenotype evolvability, defined as the number of different phenotypes within one point mutation of a given phenotype [45]. This is illustrated for phenotype 2 in Fig. 5c. This behaviour echoes what has been observed for other GP maps in the literature, which often show correlated landscapes, where robust phenotypes are also the most evolvable [1,24,39].

6 Conclusion

In this paper, we have studied the map from a Boolean threshold network to its dynamic behaviour as a GP map, and showed that this map presents a series of properties observed in the GP maps literature, such as high levels of redundancy, a wide distribution of neutral network sizes, high robustness and evolvability, and most notably a bias towards low-complexity phenotypes.

Although these results are interesting, they only concern the space of GRN topologies. One could ask if we would obtain the same results if we had instead fixed the network topology and varied the strength of gene-gene interactions, or varied topology and interactions – both interesting research directions.

More broadly, the presence of a strong bias towards simplicity raises many questions about the systems modelled by Boolean threshold networks. GRNs are behind cell signalling, differentiation, embryonic development, and other complex biological patterns, but BTNs are also the building block of neural network models in machine learning [35]. Following the analogy between both types of networks, the evolutionary search for a gene network with a given set of biological properties can be compared to the machine learning task of finding a set of parameters that minimises a cost function. The parallels between evolution and computation, in particular between evolution and (machine) learning, are discussed in a large body of literature [43,44,47]. We believe that the study of GP maps might contribute to this literature. Based on the work presented in this paper, one could expect that phenomena such as simplicity bias, and even more biologically relevant properties such as robustness and evolvability, might have a computational or learning equivalent. Needless to say, the translation of concepts from one field to another leads to more questions than answers, and plenty of interesting work ahead.

References

1. Ahnert, S.E.: Structural properties of genotype–phenotype maps. J. R. Soc. Interface **14**(132), 20170275 (2017). https://doi.org/10.1098/rsif.2017.0275
2. Albert, R., Othmer, H.G.: The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in Drosophila melanogaster. J. Theor. Biol. **223**(1), 1–18 (2003)
3. Aral, N., Kabakçıoğlu, A.: Coherent regulation in yeast's cell-cycle network. Phys. Biol. **12**(3), 036002 (2015)
4. Aral, N., Kabakçıoğlu, A.: Coherent organization in gene regulation: a study on six networks. Phys. Biol. **13**(2), 026006 (2016)
5. Azevedo, R.B., Lohaus, R., Srinivasan, S., Dang, K.K., Burch, C.L.: Sexual reproduction selects for robustness and negative epistasis in artificial gene networks. Nature **440**(7080), 87 (2006)
6. Bergman, A., Siegal, M.L.: Evolutionary capacitance as a general feature of complex gene networks. Nature **424**(6948), 549 (2003)
7. Boldhaus, G., Bertschinger, N., Rauh, J., Olbrich, E., Klemm, K.: Robustness of Boolean dynamics under knockouts. Phys. Rev. E **82**(2), 021916 (2010)

8. Boldhaus, G., Klemm, K.: Regulatory networks and connected components of the neutral space. Eur. Phys. J. B-Condens. Matter Complex Syst. **77**(2), 233–237 (2010)
9. Borenstein, E., Krakauer, D.C.: An end to endless forms: epistasis, phenotype distribution bias, and nonuniform evolution. PLoS Comput. Biol. **4**(10), e1000202 (2008)
10. Chen, H., Wang, G., Simha, R., Du, C., Zeng, C.: Boolean models of biological processes explain cascade-like behavior. Sci. Rep. **7** (2016). Article number 20067. https://www.nature.com/articles/srep20067
11. Ciliberti, S., Martin, O.C., Wagner, A.: Innovation and robustness in complex regulatory gene networks. Proc. Natl. Acad. Sci. **104**(34), 13591–13596 (2007)
12. Ciliberti, S., Martin, O.C., Wagner, A.: Robustness can evolve gradually in complex regulatory gene networks with varying topology. PLoS Comput. Biol. **3**(2), e15 (2007)
13. Cowperthwaite, M.C., Economo, E.P., Harcombe, W.R., Miller, E.L., Meyers, L.A.: The ascent of the abundant: how mutational networks constrain evolution. PLoS Comput. Biol. **4**(7), e1000110 (2008)
14. Davidich, M.I., Bornholdt, S.: Boolean network model predicts cell cycle sequence of fission yeast. PLoS One **3**(2), e1672 (2008)
15. Davidich, M.I., Bornholdt, S.: Boolean network model predicts knockout mutant phenotypes of fission yeast. PLoS One **8**(9), e71786 (2013)
16. Dingle, K., Camargo, C.Q., Louis, A.A.: Input-output maps are strongly biased towards simple outputs. Nat. Commun. **9**(1), 761 (2018)
17. Dingle, K., Schaper, S., Louis, A.A.: The structure of the genotype-phenotype map strongly constrains the evolution of non-coding RNA. Interface Focus **5**(6), 20150053 (2015)
18. Espinosa-Soto, C., Padilla-Longoria, P., Alvarez-Buylla, E.R.: A gene regulatory network model for cell-fate determination during Arabidopsis thaliana flower development that is robust and recovers experimental gene expression profiles. Plant Cell **16**(11), 2923–2939 (2004)
19. Fauré, A., Naldi, A., Chaouiya, C., Thieffry, D.: Dynamical analysis of a generic Boolean model for the control of the mammalian cell cycle. Bioinformatics **22**(14), e124–e131 (2006)
20. Giacomantonio, C.E., Goodhill, G.J.: A Boolean model of the gene regulatory network underlying mammalian cortical area development. PLoS Comput. Biol. **6**(9), e1000936 (2010)
21. Glass, L., Mackey, M.C.: From Clocks to Chaos: The Rhythms of Life. Princeton University Press, Princeton (1988)
22. Greenbury, S.F., Ahnert, S.E.: The organization of biological sequences into constrained and unconstrained parts determines fundamental properties of genotype-phenotype maps. J. R. Soc. Interface **12**(113), 20150724 (2015)
23. Greenbury, S.F., Johnston, I.G., Louis, A.A., Ahnert, S.E.: A tractable genotype-phenotype map modelling the self-assembly of protein quaternary structure. J. R. Soc. Interface **11**(95), 20140249 (2014)
24. Greenbury, S.F., Schaper, S., Ahnert, S.E., Louis, A.A.: Genetic correlations greatly increase mutational robustness and can both reduce and enhance evolvability. PLoS Comput. Biol. **12**(3), e1004773 (2016)
25. Helikar, T., Konvalina, J., Heidel, J., Rogers, J.A.: Emergent decision-making in biological signal transduction networks. Proc. Natl. Acad. Sci. **105**(6), 1913–1918 (2008)

26. Hopfield, J.J., Tank, D.W.: Collective computation with continuous variables. In: Disordered Systems and Biological Organization, pp. 155–170. Springer (1986)
27. Hopfield, J.J., Tank, D.W.: Computing with neural circuits - a model. Science **233**(4764), 625–633 (1986)
28. Kauffman, S.A.: Homeostasis and differentiation in random genetic control networks. Nature **224**(5215), 177–178 (1969)
29. Kauffman, S.A.: The Origins of Order: Self-Organization and Selection in Evolution. Oxford University Press, Oxford (1993)
30. Lempel, A., Ziv, J.: On the complexity of finite sequences. IEEE Trans. Inf. Theory **22**(1), 75–81 (1976)
31. Li, F., Long, T., Lu, Y., Ouyang, Q., Tang, C.: The yeast cell-cycle network is robustly designed. Proc. Natl. Acad. Sci. U.S.A. **101**(14), 4781–4786 (2004)
32. Li, H., Helling, R., Tang, C., Wingreen, N.: Emergence of preferred structures in a simple model of protein folding. Science **273**(5275), 666 (1996)
33. Li, S., Assmann, S.M., Albert, R.: Predicting essential components of signal transduction networks: a dynamic model of guard cell abscisic acid signaling. PLoS Biol. **4**(10), e312 (2006)
34. May, R.M.: Models for two interacting populations. In: Theoretical Ecology: Principles and Applications, pp. 49–70 (1976)
35. McCulloch, W.S., Pitts, W.: A logical calculus of the ideas immanent in nervous activity. Bull. Math. Biophys. **5**(4), 115–133 (1943)
36. Nochomovitz, Y.D., Li, H.: Highly designable phenotypes and mutational buffers emerge from a systematic mapping between network topology and dynamic output. Proc. Natl. Acad. Sci. U.S.A. **103**(11), 4180–4185 (2006)
37. Raman, K., Wagner, A.: The evolvability of programmable hardware. J. Roy. Soc. Interface **8**(55), rsif20100212 (2010). https://royalsocietypublishing.org/doi/full/10.1098/rsif.2010.0212#ref-list-1
38. 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: Transactions on Computational Systems Biology VII, pp. 56–72. Springer (2006)
39. Schaper, S., Louis, A.A.: The arrival of the frequent: how bias in genotype-phenotype maps can steer populations to local optima. PLoS One **9**(2), e86635 (2014)
40. Schuster, P., Fontana, W., Stadler, P.F., Hofacker, I.L.: From sequences to shapes and back: a case study in RNA secondary structures. Proc. Roy. Soc. London B: Biol. Sci. **255**(1344), 279–284 (1994)
41. Steiner, C.F.: Environmental noise, genetic diversity and the evolution of evolvability and robustness in model gene networks. PLoS One **7**(12), e52204 (2012)
42. Tarjan, R.: Depth-first search and linear graph algorithms. SIAM J. Comput. **1**(2), 146–160 (1972)
43. Valiant, L.: Probably Approximately Correct: Nature's Algorithms for Learning and Prospering in a Complex World. Basic Books, New York (2013)
44. Valle-Pérez, G., Camargo, C.Q., Louis, A.A.: Deep learning generalizes because the parameter-function map is biased towards simple functions. In: International Conference on Learning Representations (ICLR) (2019)
45. Wagner, A.: Robustness and evolvability: a paradox resolved. Proc. Roy. Soc. London B: Biol. Sci. **275**(1630), 91–100 (2008)
46. Wagner, A.: Robustness and Evolvability in Living Systems. Princeton University Press, Princeton (2013)

47. Watson, R.A., Szathmáry, E.: How can evolution learn? Trends Ecol. Evol. **31**(2), 147–157 (2016)
48. Zañudo, J.G., Aldana, M., Martínez-Mekler, G.: Boolean threshold networks: virtues and limitations for biological modeling. In: Information Processing and Biological Systems, pp. 113–151 (2011)