Maximum-entropy large-scale structures of Boolean networks optimized for criticality

2015 New J. Phys. 17 043021

(http://iopscience.iop.org/1367-2630/17/4/043021)

New Journal of Physics
The open access journal at the forefront of physics

Deutsche Physikalische Gesellschaft DPG
IOP Institute of Physics

Published in partnership with: Deutsche Physikalische Gesellschaft and the Institute of Physics

OPEN ACCESS

RECEIVED
18 November 2014

REVISED
4 February 2015

ACCEPTED FOR PUBLICATION
4 March 2015

PUBLISHED
14 April 2015

PAPER

Maximum-entropy large-scale structures of Boolean networks optimized for criticality

Marco Möller[1] **and Tiago P Peixoto**[2]

[1] Institute for condensed matter physics, TU Darmstadt, Hochschulstrasse 6, D-64289 Darmstadt, Germany
[2] Institute for Theoretical Physics, University of Bremen, Hochschulring 18, D-28359 Bremen, Germany

E-mail: tiago@itp.uni-bremen.de

Keywords: Boolean networks, complex networks, evolved networks, criticality

Abstract

We construct statistical ensembles of modular Boolean networks that are constrained to lie at the critical line between frozen and chaotic dynamic regimes. The ensembles are maximally random given the imposed constraints, and thus represent *null models* of critical networks. By varying the network density and the entropic cost associated with biased Boolean functions, the ensembles undergo several phase transitions. The observed structures range from fully random to several ordered ones, including a prominent core–periphery-like structure, and an 'attenuated' two-group structure, where the network is divided in two groups of nodes, and one of them has Boolean functions with very low sensitivity. This shows that such simple large-scale structures are the most likely to occur when optimizing for criticality, in the absence of any other constraint or competing optimization criteria.

1. Introduction

Boolean networks are often used as generic models for the dynamics of complex systems of interacting entities, such as social and economic networks, neural networks, and gene or protein interaction networks [1, 2]. Whenever the states of a system can be reduced to being either 'on' or 'off' without loss of important information, a Boolean approximation captures many features of the dynamics of real networks [3]. One of such features is the transition between two dynamical regimes: a 'frozen' phase, where small perturbations of the dynamics vanish after some time, and a 'chaotic' phase, where localized perturbations grow exponentially fast, and disturb the entire system [2]. It is often posited that many real systems such as gene regulatory networks [4] or the brain [5] possess features similar to networks which are at the critical line between these two phases, and hence share features inherent to both of them. A central question has been how such systems are capable of self-organizing in this critical state [6–8].

In this work, we tackle this question from a different point of view. Instead of describing how a specific dynamics or evolutionary process can drive the structure of the system towards criticality, we focus on the minimal ingredients necessary for a system to be critical under general constraints. In particular, we consider *null models* of critical Boolean networks that possess the necessary topological and functional characteristics for criticality, but are otherwise maximally random. We are interested both in the large-scale structure of such networks, as well as the choice of Boolean functions. With this in mind, we parametrize general ensembles of functional networks using the stochastic block model [9], and obtain configurations which maximize its entropy [10] under the constraint that the dynamics lie in the critical line. By varying the density of the network, and the entropic cost of choosing Boolean functions with specific sensitivities, we observe topological phase transitions from a fully random configuration, to many structured ones. The most prominent structured topological phases we find are: (1) a core–periphery-like structure [11, 12], which achieves criticality by restricting the regulation of most of the network by a few 'core' nodes, and (2) a two-group structure where one of the node groups possess Boolean functions with very low sensitivity. We also show that topologies formed by only two distinct groups of nodes are sufficient to achieve criticality given the general constraints considered. This suggests that more elaborate large-scale structures and choice of Boolean functions do not necessarily arise directly out of an

optimization towards criticality, and may have other direct causes, such as being the outcome of a growth process or some other non-equilibrium dynamics.

This paper is divided as follows. In section 2 we present our model. In section 3 we describe the optimized ensemble of critical networks. In section 4 we present the results of the optimization, and the phase diagram. We then conclude in section 5 with a general discussion.

2. The Model

Our objective is to investigate the large-scale topology of networks which are forced to be critical under global constraints. We parametrize the possible structures as a general directed stochastic block model [9, 10], where the nodes are divided into B groups, where a given group r has size n_r, and e_{rs} is the number of edges randomly placed from group s to group r. We also ascribe to each node in the network a Boolean function, chosen randomly from the set of all possible functions with a given bias p_r (i.e. the fraction of input combinations that correspond to output 1)[3]. On a given realization of this network ensemble, we consider a synchronous Boolean dynamics, where at each discrete time step the Boolean values $\{\sigma_i\}$ of all nodes are updated as

$$\sigma_i(t+1) = f_i\Big(\big\{\sigma_j(t)\big\}\Big), \tag{1}$$

where f_i is the Boolean function of node i, and $\{\sigma_j(t)\}$ is the set of inputs of i. We are interested in describing the time evolution of perturbations of the dynamics, where a single Boolean value is flipped $\sigma_i \to 1-\sigma_i$, and the following cascade of flips is measured. More precisely, we are interested in the Hamming distance between two identical copies of the same network, but where one of them is unperturbed,

$$h(t) = \frac{1}{N}\sum_i \left|\sigma_i(t) - \sigma_i'(t)\right|, \tag{2}$$

where $\{\sigma_i(t)\}$ and $\{\sigma_i'(t)\}$ are the states of the original and perturbed copies, respectively. According to the order parameter $h(t)$ we can distinguish between two phases in the thermodynamic limit $N \gg 1$: a frozen dynamics where $\lim_{t\to\infty} h(t) = 0$, and a 'chaotic' one for $\lim_{t\to\infty} h(t) > 0$. In terms of our imposed block structure, we have

$$h(t) = \frac{1}{N}\sum_r n_r h_r(t), \tag{3}$$

where $h_r = \sum_{i\in r}\left|\sigma_i(t) - \sigma_i'(t)\right| / n_r$ is the contribution to $h(t)$ from the nodes belonging to group r. Instead of investigating the microscopic dynamics of equation (1) directly, here we make use of the annealed approximation [13], and consider that at each time step the edges of the network are re-sampled from the same ensemble. By restricting ourselves to the early times after the perturbation, such that $h_r(t) \ll 1$, we can neglect the probability that more than one input flips simultaneously, which is of order $O(h_r(t)^2)$. This allows us to write the time evolution of the individual $h_r(t)$ values as

$$h_r(t+1) \simeq b_r \sum_s w_{rs} h_s(t), \tag{4}$$

with $w_{rs} = e_{rs}/n_r$ being the fraction of inputs of group r that belong to group s, and $b_r = 2p_r(1-p_r)$ is the probability that if an input of a node belonging to group r flips its output will also flip [14–16]. The general solution of this linear system is

$$\vec{h}(t) = M^t \vec{h}(0) \tag{5}$$

with $\vec{h}(t) = \{h_r(t)\}$ and

$$M_{rs} = b_r w_{rs} = 2p_r\left(1-p_r\right)e_{rs}/n_r. \tag{6}$$

The matrix M is non-negative and non-symmetric, and hence it has at least one purely real non-negative leading eigenvalue λ, with a non-negative eigenvector $\vec{x}$. Henceforth we further assume that M is strictly positive[4], so that this eigenvalue is positive and unique, and then for a sufficiently large t we can write

$$\vec{h}(t) \simeq \lambda^t\left(\vec{h}(0), \vec{x}\right)\vec{x}. \tag{7}$$

[3] For nodes which do not receive any input, we impose that they receive constant functions, with a randomly chosen output value.

[4] In some of the following results we will observe values $M_{rs} \to 0$. These should simply be interpreted as asymptotic limits where the values become arbitrarily small but always strictly positive, and hence do not invalidate equation (7).

From this we have that $\lambda < 1$ corresponds to the frozen phase, since the magnitude of the perturbation will decrease exponentially, and conversely $\lambda > 1$ corresponds to the 'chaotic' phase, since it will increase exponentially. Hence the special value $\lambda = 1$ marks the critical line between the two phases. Therefore, given some parameter choice for $\{n_r\}$, $\{e_{rs}\}$ and $\{p_r^k\}$, we can decide in which phase the corresponding dynamics will lie by computing the largest eigenvalue of the $B \times B$ matrix M. Note that from this criterion we recover trivially the critical value for biased random Boolean networks with $B = 1$, $\lambda = 2p_r^k(1 - p_r^k)\langle k\rangle$ [2].

3. Optimized ensembles

We are interested in generating ensembles of Boolean networks which lie in the critical line $\lambda = 1$. More specifically, we want *null models* of critical networks where—in addition to being critical—the topology and choice of functions is maximally random within the imposed constraints. We achieve this by maximizing the entropy of stochastic block model ensemble [10]

$$\mathcal{S}_{\mathrm{S}}\left(\{n_r\}, \{e_{rs}\}\right) = \ln \Omega\left(\{n_r\}, \{e_{rs}\}\right) \tag{8}$$

with $\Omega(\{n_r\}, \{e_{rs}\})$ being the total number of network realizations, as well as the entropy $\mathcal{S}_{\mathrm{F}}(\{n_r\}, \{p_r^k\})$ of the distribution of Boolean functions, which we describe below. The maximization is performed in a constrained fashion, by imposing a critical sensitivity $\lambda(\{n_r\}, \{e_{rs}\}, \{p_r^k\}) = 1$, and by adjusting the relative entropic cost of modifying the structure of the network (the parameters $\{n_r\}$, $\{e_{rs}\}$) and the Boolean functions (the parameters $\{p_r^k\}$). This is achieved by finding the saddle point of the Lagrangian function

$$\Lambda = (1 - \mu)\mathcal{S}_{\mathrm{S}}\left(\{n_r\}, \{e_{rs}\}\right) + \mu\mathcal{S}_{\mathrm{F}}\left(\{n_r\}, \left\{p_r^k\right\}\right) - c\left[\lambda\left(\{n_r\}, \{e_{rs}\}, \left\{p_r^k\right\}\right) - 1\right], \tag{9}$$

where c is a Lagrange multiplier which imposes $\lambda = 1$, and μ controls the relative entropic cost between the edge placements and the choice of functions.

For the structural entropy $\mathcal{S}_{\mathrm{S}}(\{n_r\}, \{e_{rs}\})$ we have simply [10]

$$\mathcal{S}_{\mathrm{S}}\left(\{n_r\}, \{e_{rs}\}\right) \simeq E - \sum_{rs} e_{rs} \ln\left(\frac{e_{rs}}{n_r n_s}\right), \tag{10}$$

where $E = \sum_{rs} e_{rs}$ is the total number of edges, and the limit of sparse networks was assumed, i.e. $e_{rs} \ll n_r n_s$.

For $\mathcal{S}_{\mathrm{F}}(\{n_r\}, \{p_r^k\})$ we have considered different choices. A first option would be to enumerate all possible Boolean functions of k inputs with a given bias, $\Omega(p, k)$, and compute $\mathcal{S}_{\mathrm{F}}(\{n_r\}, \{p_r^k\}) = \sum_{r,k} n_r p_k^r \ln \Omega(p_r^k, k)$ with p_k^r being the in-degree distribution of group r, which is a Poisson with average $\left\langle k_r\right\rangle = \sum_s e_{sr}/n_r$. For each choice of k there are 2^k possible input combinations, $2^k p$ of which will have output 1 and $2^k(1 - p)$ will have output zero. The total number of Boolean functions is therefore

$$\Omega(p, k) = \binom{2^k}{2^k p}, \tag{11}$$

from which we obtain using Stirling's approximation for $2^k \gg 1$

$$\ln \Omega(p, k) \simeq 2^k H_{\mathrm{b}}(p), \tag{12}$$

where $H_{\mathrm{b}}(x) = -x \ln x - (1 - x)\ln(1 - x)$ is the binary entropy function. As equation (12) shows, this choice will result in an entropy function which grows exponentially with the number of inputs k. This means that, for any choice of $\mu > 0$ in equation (9) above, a trivial maximization of the overall entropy can be achieved by sufficiently increasing the average number of inputs of a vanishingly small fraction of the nodes, since the remaining entropy term of equation (10) will only depend log-linearly on the size and density of the groups. In other words, the functional entropy will exponentially dominate the structural one for many parameter choices, $\mathcal{S}_{\mathrm{F}}(\{n_r\}, \{p_r^k\}) \gg \mathcal{S}_{\mathrm{S}}(\{n_r\}, \{e_{rs}\})$. Hence, the outcome would correspond always to fully random networks, with a uniform bias p selected to enforce $\lambda = 1$. However, the choice of equation (12) corresponds to an unrealistic situation where all entries of the truth table of the Boolean functions with the same bias are equally accessible evolutionary. As equation (12) itself shows, the number of such functions increases comparably fast to the number of all Boolean functions with k inputs, 2^{2^k}. It is known, however, that the vast majority of Boolean functions are not realizable biochemically, and those that are empirically observed often fall within very narrow classes, such as nested canalizing functions [17], which scale in number as $\sim k!2^k$ [18]; far slower than the biased functions above. Furthermore, since we are only considering networks that are forced to lie at the critical line, the vast majority of input combinations of the biased functions considered above are not dynamically accessible. Hence, for the purposes of the actual dynamics, the enumeration done in equation (11) is largely immaterial. In view of this, here we consider instead the modified situation where the input combinations not present during

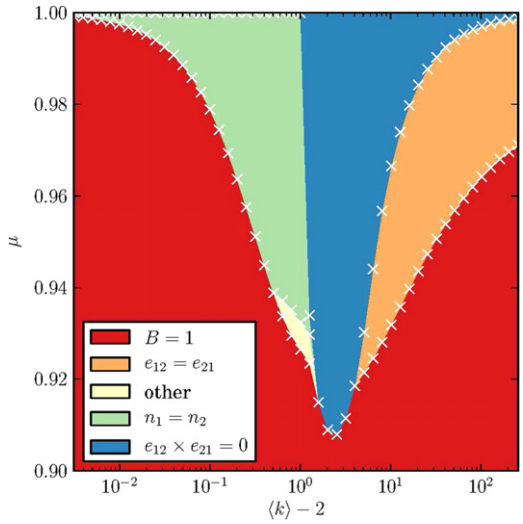

Figure 1. Phase diagram for the critical networks as a function of the average in-degree $\langle k\rangle$ and functional entropy bias μ. The different phases are color-coded according to the legend, and described in more detail in figure 3.

typical dynamical trajectories of equation (1) are not wastefully encoded in the truth table (which can be imagined to contain constant output values instead). We compute functional entropy approximately via the sensitivity to perturbation itself, $2p(1-p)$, which corresponds to the probability that the output of the function changes if one input is perturbed. With this we can write,

$$\mathcal{S}_{\mathrm{F}}\left(\{n_r\},\left\{\underline{p}_r\right\}\right)=\sum_r n_r H_{\mathrm{b}}\left(2\underline{p}_r\left(1-\underline{p}_r\right)\right), \tag{13}$$

which is more well behaved than equation (12)[5].

Although we can compute both $\mathcal{S}_{\mathrm{S}}(\{n_r\},\{e_{rs}\})$ and $\mathcal{S}_{\mathrm{F}}(\{n_r\},\{\underline{p}_r\})$, the value of λ can only be obtained analytically for very small number of groups B. Therefore, for larger values of B we are forced to perform the maximization numerically. For details of the numerical methods we refer to appendix.

4. Numerical results

We obtained the values of $\{n_r\}$, $\{e_{rs}\}$ and $\{\underline{p}_r\}$, characterizing the network structure and Boolean functions, which maximize $(1-\mu)\mathcal{S}_{\mathrm{S}}(\{n_r\},\{e_{rs}\})+\mu\mathcal{S}_{\mathrm{F}}(\{n_r\},\{\underline{p}_r\})$, subject to the constraint that the dynamics lies exactly on the critical line, $\lambda=1$. We obtained results for different values of the average in-degree per node $\langle k\rangle = E/N$, as well as different values of μ, which regulates the relative trade-offs between the structural and functional entropies. We consider only the thermodynamic limit with $N\gg 1$.

We have investigated ensembles with different number of groups B. However, we found that in all cases the obtained structures could be fully equivalently formulated as a $B=2$ structure, where one or more groups could be merged together, resulting in the exact same network ensemble, with the same structural and functional entropies. Hence we concluded that a number of $B=2$ groups is sufficient to describe the obtained topologies for all parameter choices, similarly to previous work focusing on optimization against noise [19], and structural stability [20]. Thus, we focus on the $B=2$ case from now on.

Varying both $\langle k\rangle$ and μ we obtain a variety of structural phases, characterized by distinct large-scale structures, as well as several phase transitions between them, as can be seen in the phase diagram of figure 1. We have identified phases by searching for discontinuities and abrupt changes in $\mathcal{S}_{\mathrm{S}}$, $\mathcal{S}_{\mathrm{F}}$, $\{n_r\}$, $\{e_{rs}\}$ and $\{\underline{p}_r\}$, which correspond to qualitatively distinct structural patterns. An example for this can be found in figure 2, which shows the values of both entropies in the phase diagram, where the phase boundaries can be identified.

Overall, for a functional entropy bias μ becoming sufficiently small, the observed topology is a fully random one with $B=1$, and a functional bias p chosen so that $\lambda=2p(1-p)\langle k\rangle=1$. These correspond to the fully random biased Boolean networks often considered in the literature [2]. As soon as μ increases sufficiently, and hence also the entropic cost of choosing biased functions, the network rewires itself in specific configurations, depending on how dense it is. For $\langle k\rangle$ approaching 2, the network becomes fully random again, since a fully

[5] A similar alternative would be to use the directly the bias p_r instead of $2\underline{p}_r(1-\underline{p}_r)$ in equation (13). We investigated this variant as well, and obtained results qualitatively equivalent to those presented here using equation (13).

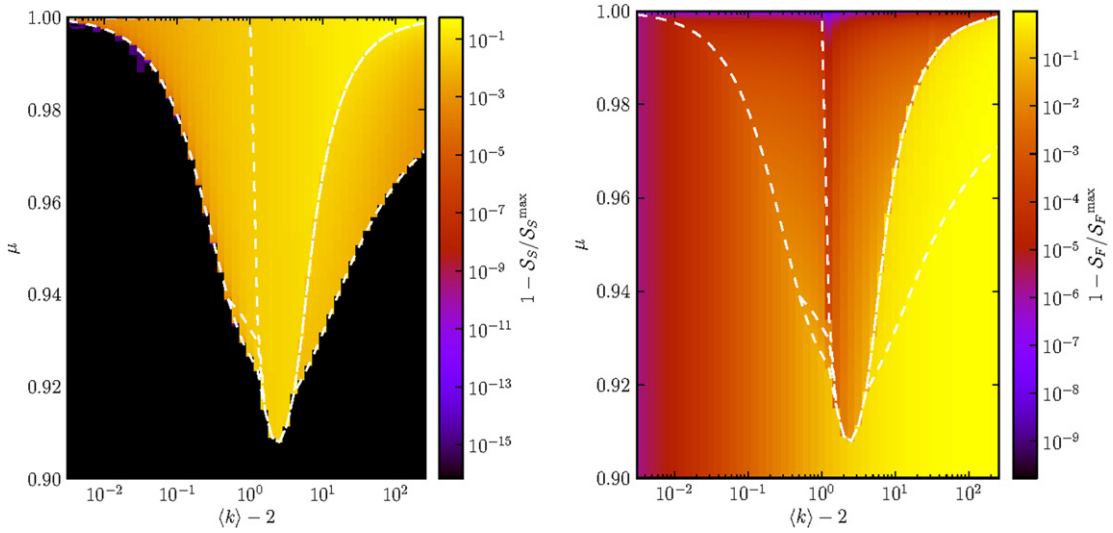

Figure 2. Relative distance of the structural entropy S_S (left) and functional entropy S_F (right) from their maximal value, as a function of $\langle k \rangle$ and μ. The abrupt changes (dotted lines) mark the phase boundaries in figure 1.

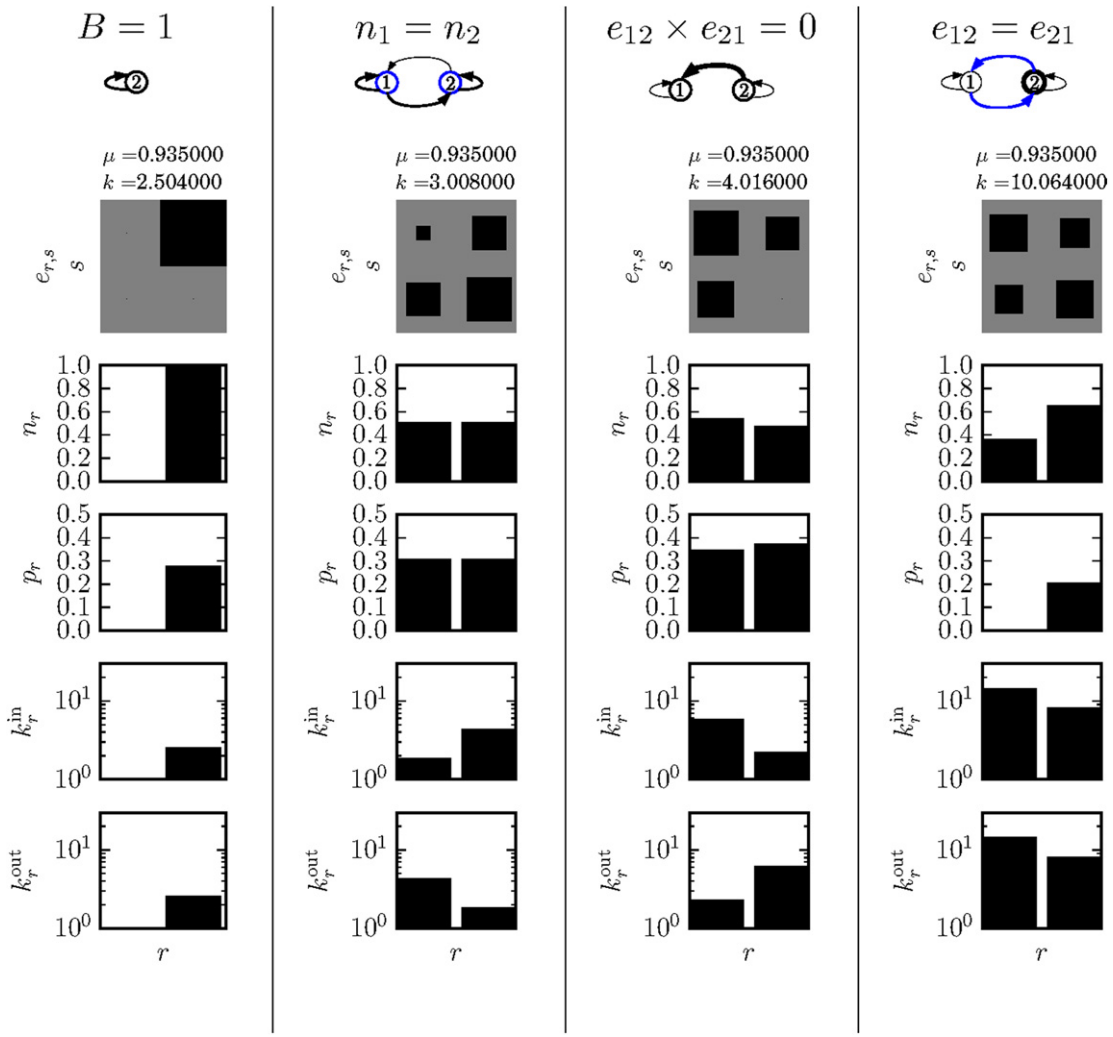

Figure 3. Sample block model parametrizations belonging to the different phases in figure 1. At the top are graphical representations of the matrix e_{rs} (with line thickness corresponding the magnitude), followed the group sizes $\{n_r\}$, function biases $\{p_r\}$, and average in- and out-degree per group $\langle k \rangle_r^{\text{in}}$ and $\langle k \rangle_r^{\text{out}}$, respectively.

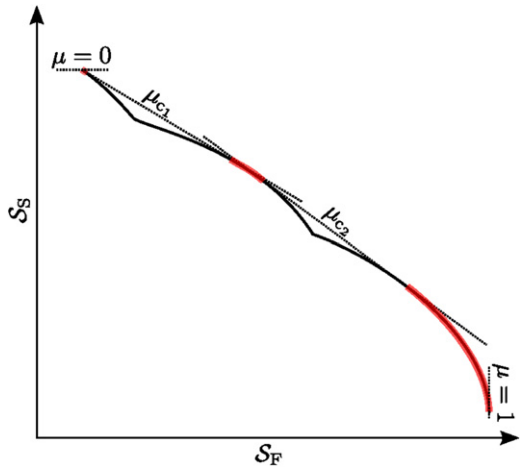

Figure 4. Qualitative example of a Pareto front of the joint optimization of $\mathcal{S}_S$ and $\mathcal{S}_F$ (black). The line segments highlighted in red are the convex hull of the pareto front, which intersects some curve with slope $\mu/(\mu - 1)$ furthest away from the origin (dotted lines).

random Boolean network with $p = 1/2$ is critical precisely at $\langle k \rangle = 2$ [1]. For increasing values of $\langle k \rangle$ a fully random network would be super-critical, and hence other structures arise in order to achieve criticality. For a region around $2 < \langle k \rangle \lesssim 3$ the system may find itself divided into two groups of identical sizes and functional biases ($n_1 = n_2$, $p_1 = p_2$), but which are asymmetrically connected, such that one group receives more links than the other. We observe also a very narrow phase for smaller μ values ('other'), for which our numerical precision did not allow an accurate characterization, but it most likely corresponds to an extension of the $n_1 = n_2$ phase, as figure 2 seems to imply. For even larger values of $\langle k \rangle$ the system moves towards two other phases, the first one being a core–periphery-like ($e_{12} \times e_{21} = 0$), where one of the groups (the periphery) is predominantly regulated by the other group (the core), which regulates itself. Because of this, the core is also more densely connected than the periphery, however it also has a smaller bias. For even larger $\langle k \rangle$, the system transits to yet another phase, where the number of edges between both groups is the same ($e_{12} = e_{21}$), but the groups have different sizes, and hence different in-degrees. The group with a higher in-degree has a much smaller bias p_r, such that its sensitivity to perturbations is significantly lower than the other group. Overall, the core–periphery-like structure is preferred for high μ and large k, in comparison to the other phases.

4.1. On the nature of the phase transitions

In order to understand the nature of the transitions between the phases we have described, it is useful to consider the Pareto front of the joint optimization of both entropy functions $\mathcal{S}_S$ and $\mathcal{S}_F$. The Pareto front is a line in the $(\mathcal{S}_S, \mathcal{S}_F)$ plane where one could not increase one of the values without decreasing the other. Since we are linearly combining both values as $\Lambda = (1 - \mu)\mathcal{S}_S + \mu\mathcal{S}_F$, each line with slope $\mu/(\mu - 1)$ in this plane will have the same value of Λ. The maximization procedure consists in, for a given value of μ, finding the intersection of the slope $\mu/(\mu - 1)$ with the Pareto front which is furthest away from the origin (see figure 4). Note that not all points in the Pareto front will correspond to this intersection with this slope, only those which belong to its convex hull. Since these segments of the Pareto front are often disconnected, a jump from one solution set to the other will correspond to a (first-order) phase transition. The actual value of μ where such a transition occurs depends on $\langle k \rangle$. The only places where second-order phase transitions are observed are at the boundaries between more than two phases, such as between $e_{12} = e_{21}$, $e_{12} \times e_{21} = 0$ and $B = 1$, and at the onset of the $n_1 = n_2$ phase at $\langle k \rangle = 2$.

We also found a relatively simple explanation for the phase transition and its approximate position between the fully random $B = 1$ phase and the other structured phases. For any given value of B, both entropy values will be bounded within some interval depending on μ, i.e. $\mathcal{S}_S \in [\mathcal{S}_S^{\min}, \mathcal{S}_S^{\max}]$ and $\mathcal{S}_F \in [\mathcal{S}_F^{\min}, \mathcal{S}_F^{\max}]$. Using this, we can rewrite our objective function as

$$\Lambda = (1 - \mu)\left(\mathcal{S}_S - \mathcal{S}_S^{\min}\right) + \mu\left(\mathcal{S}_F - \mathcal{S}_F^{\min}\right) + \text{const}, \tag{14}$$

up to some unimportant constant. For small values of μ the value of Λ is dominated by the structural entropy $\mathcal{S}_S$, and for larger values of μ by functional entropy $\mathcal{S}_F$. A plausible hypothesis for the position μ_t where the transition occurs is when both values have an equal contribution

$$\left(1 - \mu_c\right)\Delta\mathcal{S}_S = \mu_c\Delta\mathcal{S}_F, \tag{15}$$

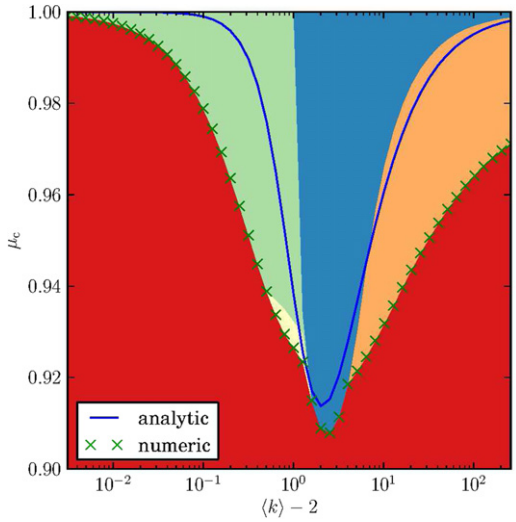

Figure 5. Critical point μ_t marking the transition between a fully random $B = 1$ and one of the structured phases in figure 1. Blue: analytic values corresponding to equation (16). Green: values obtained from the numeric optimization.

so that we have

$$\mu_c = \frac{\Delta \mathcal{S}_{\mathrm{S}}}{\Delta \mathcal{S}_{\mathrm{S}} + \Delta \mathcal{S}_{\mathrm{F}}}. \tag{16}$$

The values of $\Delta \mathcal{S}_{\mathrm{S}} = \mathcal{S}_{\mathrm{S}}^{\max} - \mathcal{S}_{\mathrm{S}}^{\min}$ and $\Delta \mathcal{S}_{\mathrm{F}} = \mathcal{S}_{\mathrm{F}}^{\max} - \mathcal{S}_{\mathrm{F}}^{\min}$ can also be obtained with simple arguments. The maximal entropy will be obtained for the fully random case with $B = 1$ and $p_r = 1/2 - \sqrt{1/4 - 1/2\langle k \rangle}$ [2]. The minimum value we assume to correspond to the core–periphery-like topology $e_{12} \times e_{21} = 0$ with representative values $n_1 = n_2 = N/2$, $e_{21} = 0$, $e_{11} = e_{22} = N$, $e_{12} = E - 2N$ and $p_1 = p_2 = 1/2$. As shown in figure 5, the value of μ_t obtained in this manner corresponds well to the numerical results obtained.

5. Conclusion

We have shown that maximum-entropy ensembles of modular Boolean networks posed at the critical line exhibit structural phase transitions from a fully random topology to several structured ones. The phases occur according to globally imposed constraints, such as the average in-degree and the relative entropic cost of modifying Boolean functions versus the network topology itself. In the limit where non-random Boolean functions possess a very high entropic cost, the emerging topology is a two-group core–periphery-like structure, where a small fraction of the nodes is responsible for the regulation of the entire system, which is formed predominately of a non-regulating majority. If the entropic cost of adapting the Boolean functions diminish, this core–periphery-like structure is replaced by a tiered one, where the network is divided into two groups of comparable size and bidirectional connections, and one of them has a sensitivity much smaller than the other. Finally, for even smaller functional entropic cost, the emerging topology is a fully random network, with uniformly sampled Boolean functions with the necessary critical sensitivity.

The emerging core–periphery-like topology is very similar to the one arising out of optimization against stochastic fluctuations [19] (as well as structural robustness against failure [20]), and is qualitatively similar to what is observed in real gene regulatory networks, where only a minority of genes (transcription factors) are in fact responsible for regulation. Hence our results suggest not only a possible explanation for this property, but also that such a core–periphery-like organization may provide fitness according to multiple criteria.

The ensembles considered in this work are optimized according to a single criterion, and are subject to no constraints other than the total number of inputs per node. Furthermore our approach does not take into account topological features which arise out of non-equilibrium processes such as gene duplication [21–23] or frozen accidents, or even other specific parametrizations of Boolean functions such as canalizing [24, 25] or nested canalizing functions [26, 27]. Nevertheless, this type of analysis can provide insight into the topological features which are necessary outcomes of a given optimization procedure, and allows one to rule out others. For instance, although it is known that gene duplication can result in broad degree distributions [21–23] and assortative modular structure [28], these features are not present in our ensembles, although they are not

forbidden. This strongly suggests that these features are not strictly necessary ingredients of robust systems, and are simply the byproduct of non-equilibrium growth processes.

Although we have forced the networks in the ensemble to lie at the critical line, there are associated properties of critical networks such as the scaling of the 'frozen core' [29] and the number of attractors [30] that were not *a priori* imposed. It remains to be seen how the obtained topologies affect these characteristics, which we leave to future work.

Acknowledgments

The authors would like to thank Barbara Drossel for insightful conversations. This work was partially funded by the University of Bremen, zentrale Forschunsfrderung Linie 04.

Appendix. Numerical methods

Here we describe in more detail the numerical methods used for the optimization described in the main text. The objective is to maximize

$$(1-\mu)\mathcal{S}_{\mathrm{S}}\left(\{n_r\},\{e_{rs}\}\right)+\mu\mathcal{S}_{\mathrm{F}}\left(\{n_r\},\left\{p_r\right\}\right), \tag{A.1}$$

with respect to $\{n_r\}$, $\{e_{rs}\}$ and $\{p_r\}$, subject to the constraint

$$\lambda\left(\{n_r\},\{e_{rs}\},\left\{p_r\right\}\right)=1, \tag{A.2}$$

where $\lambda(\{n_r\},\{e_{rs}\},\{p_r\})$ is the leading eigenvalue of the matrix $\boldsymbol{M}$ in equation (6). The approach we take is to maximize the unconstrained objective function

$$\Lambda=(1-\mu)\mathcal{S}_{\mathrm{S}}+\mu\mathcal{S}_{\mathrm{F}}-\nu(\lambda-1)^2, \tag{A.3}$$

where ν is determines the penalty for deviating from $\lambda=1$. By maxing $\nu\to\infty$ we recover the original constrained optimization. In practice, we make ν sufficiently large, so that the results no longer depend on it (a choice of $\nu\sim 10^7$ was enough for our purposes). For the computation of $\mathcal{S}_{\mathrm{S}}$, $\mathcal{S}_{\mathrm{F}}$ and λ, we used the relative densities $w_r=n_r/N$, $m_{rs}=e_{rs}/N$, and $\langle k\rangle=E/N$ so that the value of our objective function no longer depends on the absolute number of nodes and edges, N and E, except for unimportant multiplicative and additive constants which do not affect the optimization.

We took some steps to reduce the degrees of freedom of the optimization by using some intrinsic constraints. Since we have that $\sum_r w_r=1$ and $\sum_{rs}m_{rs}=1$, one of each variable set can be eliminated and expressed as a function of the remaining sum. Furthermore, we know that λ is a eigenvalue of $\boldsymbol{M}$ which leads to

$$\det(\boldsymbol{M}-\lambda)=0. \tag{A.4}$$

For the special case of $B=2$ and the constraint $\lambda=1$, we can use this to the express the sensitivity of one group as a function of the other

$$2p_2\left(1-p_2\right)=w_2\left(m_{22}-\frac{m_{21}m_{12}}{m_{11}-\frac{w_1}{w_2}}\right)^{-1}. \tag{A.5}$$

We can use this to decrease the number of degrees of freedom, but we still need the dependence on ν in equation (A.3) to enforce that $\lambda=1$ is the *largest* eigenvalue of $\boldsymbol{M}$.

For the actual optimization of Λ we used a mixture of different standard optimizers. First we have used S-metric selection EMOA [31] maximizing two independent objective functions $\Lambda_1=\mu\mathcal{S}_{\mathrm{S}}-\nu(\lambda-1)^2$ and $\Lambda_2=(1-\mu)\mathcal{S}_{\mathrm{F}}-\nu(\lambda-1)^2$, which return the Pareto front of the corresponding multi-objective optimization. The resulting set of solutions for a given value of μ was used starting as points for the direct optimization of Λ. The algorithm used was a combination of Powell's method [32] followed by a downhill simplex heuristic [33], in addition to a simple heuristic of starting from the neighborhood of previously obtained solutions.

References

[1] Kauffman S A 1969 Metabolic stability and epigenesis in randomly constructed genetic nets *J. Theor. Biol.* **22** 437–67
[2] Drossel B 2008 Random boolean networks *Reviews of Nonlinear Dynamics and Complexity* ed H G Schuster vol 1 (New York: Wiley)
[3] Bornholdt S 2005 Less is more in modeling large genetic networks *Science* **310** 449–51
[4] Kauffman S A 1993 *The Origins of Order: Self-Organization and Selection in Evolution* 1st edn (New York: Oxford University Press)

[5] Beggs J M and Plenz D 2003 Neuronal avalanches in neocortical circuits *J. Neurosci.* **23** 11167–77
[6] Bornholdt S and Rohlf T 2000 Topological evolution of dynamical networks: global criticality from local dynamics *Phys. Rev. Lett.* **84** 6114–7
[7] Bornholdt S and Röhl T 2003 Self-organized critical neural networks *Phys. Rev.* E **67** 066118
[8] Rybarsch M and Bornholdt S 2014 Avalanches in self-organized critical neural networks: a minimal model for the neural SOC universality class *PLoS One* **9** e93090
[9] Holland P W, Laskey K B and Leinhardt S 1983 Stochastic blockmodels: first steps *Soc. Netw.* **5** 109–37
[10] Peixoto T P 2012 Entropy of stochastic blockmodel ensembles *Phys. Rev.* E **85** 056122
[11] Holme P 2005 Core–periphery organization of complex networks *Phys. Rev.* E **72** 046111
[12] Rombach M, Porter M, Fowler J and Mucha P 2014 Core–periphery structure in networks *SIAM J. Appl. Math.* **74** 167–90
[13] Derrida B and Pomeau Y 1986 Random networks of automata: a simple annealed approximation *Europhys. Lett.* **1** 45–49
[14] Shmulevich I and Kauffman S A 2004 Activities and sensitivities in boolean network models *Phys. Rev. Lett.* **93** 048701
[15] Ghanbarnejad F and Klemm K 2011 Stability of boolean and continuous dynamics *Phys. Rev. Lett.* **107** 188701
[16] Langton C G 1990 Computation at the edge of chaos: phase transitions and emergent computation *Physica* D **42** 12–37
[17] Harris S E, Sawhill B K, Wuensche A and Kauffman S 2002 A model of transcriptional regulatory networks based on biases in the observed regulation rules *Complexity* **7** 23–40
[18] Bender E A and Butler J T 1978 Asymptotic approximations for the number of fanout-free functions *IEEE Trans. Comput.* **C-27** 1180–3
[19] Peixoto T P 2012 Emergence of robustness against noise: a structural phase transition in evolved models of gene regulatory networks *Phys. Rev.* E **85** 041908
[20] Peixoto T P and Bornholdt S 2012 Evolution of robust network topologies: emergence of central backbones *Phys. Rev. Lett.* **109** 118703
[21] Pastor-Satorras R, Smith E and Solé R V 2003 Evolving protein interaction networks through gene duplication *J. Theor. Biol.* **222** 199–210
[22] Ispolatov I, Krapivsky P L and Yuryev A 2005 Duplication-divergence model of protein interaction network *Phys. Rev.* E **71** 061911
[23] Chung F, Lu L, Dewey T G and Galas D J 2003 Duplication models for biological networks *J. Comput. Biol.* **10** 677–87
[24] Shmulevich I, Lesm L, Dougherty E R, Astola J and Zhang W 2003 The role of certain post classes in boolean network models of genetic networks *Proc. Natl Acad. Sci. USA* **100** 10734–9
[25] Moreira A A and Amaral L N 2005 Canalizing kauffman networks: nonergodicity and its effect on their critical behavior *Phys. Rev. Lett.* **94** 218702
[26] Kauffman S, Peterson C, Samuelsson B and Troein C 2004 Genetic networks with canalyzing boolean rules are always stable *Proc. Natl Acad. Sci. USA* **101** 17102–7
[27] Peixoto T P 2010 The phase diagram of random boolean networks with nested canalizing functions *Eur. Phys. J.* B **78** 187–92
[28] Enemark J and Sneppen K 2007 Gene duplication models for directed networks with limits on growth *J. Stat. Mech.* **11** 7
[29] Mihaljev T and Drossel B 2006 Scaling in a general class of critical random boolean networks *Phys. Rev.* E **74** 046101
[30] Samuelsson B and Troein C 2003 Superpolynomial growth in the number of attractors in kauffman networks *Phys. Rev. Lett.* **90** 098701
[31] Beume N, Naujoks B and Emmerich M 2007 SMS-EMOA: multiobjective selection based on dominated hypervolume *Eur. J. Oper. Res.* **181** 1653–69
[32] Powell M J D 1994 A direct search optimization method that models the objective and constraint functions by linear interpolation *Adv. Optim. Numer. Anal.* **7** 51–67
[33] Nelder J A and Mead R 1965 A simplex method for function minimization *Comput. J.* **7** 308–13