7-2013

Phase transition of Boolean networks with partially nested canalizing functions

Kayse Jansen
University of Nebraska at Omaha

Mihaela Teodora Matache
University of Nebraska at Omaha, dvelcsov@unomaha.edu

Phase transition of Boolean networks with partially nested canalizing functions

Kayse Jansen, Mihaela Teodora Matache*

University of Nebraska at Omaha, Mathematics
Durham Science Center 237, Omaha, NE, 68182, USA
*corresponding author, dmatache@unomaha.edu

Abstract: We generate the critical condition for the phase transition of a Boolean network governed by partially nested canalizing functions for which a fraction of the inputs are canalizing, while the remaining non-canalizing inputs obey a complementary threshold Boolean function. Past studies have considered the stability of fully or partially nested canalizing functions paired with random choices of the complementary function. In some of those studies conflicting results were found with regard to the presence of chaotic behavior. Moreover, those studies focus mostly on ergodic networks in which initial states are assumed equally likely. We relax that assumption and find the critical condition for the sensitivity of the network under a non-ergodic scenario. We use the proposed mathematical model to determine parameter values for which phase transitions from order to chaos occur. We generate Derrida plots to show that the mathematical model matches the actual network dynamics. The phase transition diagrams indicate that both order and chaos can occur, and that certain parameters induce a larger range of values leading to order versus chaos. The edge-of-chaos curves are identified analytically and numerically. It is shown that the depth of canalization does not cause major dynamical changes once certain thresholds are reached; these thresholds are fairly small in comparison to the connectivity of the nodes.

1. Introduction

Boolean Networks (BN) are used for modeling networks in which the node activity, or state of the cell, can be described as a binary value: on-off, active-non active, 1-0, etc. This type of network model has been used to examine the connections among diverse physical and engineered networks such as genetic regulatory or signal transduction networks (Kauffman [1], Shmulevich et.al. [2], [3], [4], Helikar et.al. [5], Kochi and Matache [6]), biological networks (Klemm and Bornholdt [7], Raeymaekers [8]), social networks (Flache and Hegselmann [9], Green et.al. [10], Moreira et.al. [11]), economic/prediction market

networks (Jumadinova et.al. [12]), neural networks (Huepe and Aldana [13]), complex networks in general (Wolfram [14]), and more. Studying these network representations leads to predictive models of real occurrences. For example, specific biological problems studied include cell differentiation, immune response, regulatory networks and neural networks. For cell differentiation and immune response, the basic binary element might be a chemical compound, while in neural networks it might be the state of firing of a neuron.

Recently, there has been an interest in understanding the structure and dynamics of Boolean networks governed by canalizing/nested canalizing rules in which at least one of the inputs can determine the output regardless of the values of the other variables. These types of networks are encountered in many biological/genetic systems (Kauffman [1]). For example, in (Kauffman et.al. [15]) it is shown that stability prevails in genetic networks with nested canalizing Boolean rules. Similar results are obtained for other types of biological networks in (Nikolajewa et.al. [16], Rämö et.al. [17]). As indicated in (Just et.al. [18]) canalizing functions also play an important role in the study of phase transitions in random Boolean networks (Kauffman [1], Shmulevich et.al. [4]). Conflicting results are discovered by Peixoto in (Peixoto [19]), whose phase diagram for nested canalizing functions (NCF) shows large ranges of parameter values where the system appears to be in the chaotic phase. A study of the literature on transcriptional regulation in eukaryotes demonstrates a bias towards canalizing rules (Harris et.al. [20]). As shown in (Layne et.al. [21]) NCFs as typically used in the literature can be somewhat artificial, since biologically relevant rules do not necessarily obey a fully canalized structure. Then it becomes important to understand the dynamics of partially nested canalizing functions (PNCF) that are more realistic models for a variety of real networks. In particular, PNCFs have been observed in (Kochi and Matache [6]) where the Boolean functions corresponding to the signal transduction network of a generic fibroblast cell developed in (Helikar et. al. [5]) are grouped in eleven classes. Three of those classes represent PNCFs, while a fourth class which incorporates a number of functions not identified in detail in that paper are an additional source of PNCFs. Those individual classes are shown to have distinct impacts on the overall activity level of the network, but that they mainly lead to an ordered behavior of the signal transduction network. On the other hand, the class of canalizing functions with exactly two canalizing inputs such that if at least one of the inputs is active/on than the output turns on, is shown to lead to chaos that can be stabilized under certain mutations. This can be of importance in identifying the types of Boolean functions that could be targeted in drug therapies in order to shift the dynamics of a disease from one dynamical regime to another. Although

for NCFs it has been shown that order may prevail, PNCFs could lead to a more complex evolution of the system. Thus, the study of PNCFs can improve our means for finding the effect of altering certain types of nodes that could be essential for the functionality of a biological network.

In this paper, we extend some of the results in (Layne et.al. [21]) under the assumption of non-ergodicity, by taking into account the long run activity of the network in establishing the likelihood of the states of the network. We find the critical condition that separates order from chaos using an approach similar to (Peixoto [19]).

As mentioned in (Layne et.al. [21]), NCFs are very restrictive in structure. It is possible that some nodes do not exhibit the canalizing behavior at all, thus a need to relax the structure is necessary. The authors consider functions that have a partially nested canalizing structure rather than a fully nested canalizing structure. They define the nested canalizing depth as the degree to which a function exhibits a canalizing structure in comparison to its number of inputs. The NCFs are a special case of PNCFs when all inputs are canalizing. It is shown that the average sensitivity to small perturbations of a PNCF increases as the canalizing depth increases; however the difference in sensitivity between PNCFs of sufficient depth and NCFs is very slight. Additionally, it is shown that the dynamics of networks with PNCFs rapidly approach the critical regime, whereas networks with functions of relatively few nested canalizing variables can remain in the chaotic phase as was found in (Peixoto [19]). In (Layne et.al. [21]), the average sensitivity is computed assuming ergodicity of the network, that is all inputs can arise with the same probability during evolution, and the time average over the states visited by the network yields the same result as the average over the whole phase space. This is an implausible assumption that is unlikely to hold for the dynamics of arbitrary BNs as noted in (Moreira and Amaral [22]). In general, the dynamics of BNs converge to limiting cycles that occupy only a fraction of the entire phase space. In this work we take into account the states composing the limiting cycles of PNCF networks and find the average sensitivity and the corresponding phase transitions using a complementary threshold function for the PNCFs. Our analytic results are supported by the numerical simulations.

The paper is structured as follows. In Section 2 we present the partially nested canalizing function model with detailed computations for significant quantities in Subsection 2.1, followed by numerical results in Subsection 2.2. Once our model is explained, we focus on

sensitivity of the network in Section 3, by finding the analytic results for the average sensitivity and the critical condition in Subsection 3.1, followed by phase diagram simulations in Subsection 3.2. We end this work with conclusions and further directions in Section 4.

2. Boolean Network Model with Partially Nested Canalizing Functions

2.1. **The Model.** In this section we construct the PNCF network model in order to analyze the parameters that cause phase transitions from stability to chaos. The procedure used here is based on some ideas from (Peixoto [19]) and (Layne [21]).

Consider a random Boolean network under the ensemble $\mathcal{E}$ of PNCFs given as follows. Each function $F : \{0,1\}^k \to \{0,1\}$ in $\mathcal{E}$ has the following formula

$$(1) \qquad F(\sigma_1, \sigma_2, \ldots, \sigma_u, \ldots \sigma_k) = \begin{cases} s_1 \quad \text{if} \quad \sigma_1 = c_1 \\ s_2 \quad \text{if} \quad \sigma_1 = 1 - c_1, \sigma_2 = c_2 \\ s_3 \quad \text{if} \quad \sigma_1 = 1 - c_1, \sigma_2 = 1 - c_2, \sigma_3 = c_3 \\ \ldots\ldots\ldots\ldots\ldots\ldots \\ s_u \quad \text{if} \quad \sigma_1 = 1 - c_1, \ldots, \sigma_{u-1} = 1 - c_{u-1}, \sigma_u = c_u \\ G(\sigma_{u+1}, \ldots, \sigma_k) \quad \text{otherwise} \end{cases}$$

where $s_i, c_i \in \{0,1\}, \forall i = 1, 2, \ldots u$. Here c_i represents the canalizing value of the i-th input, and s_i the corresponding canalized output value. If $s_i = 1$, then σ_i is called an activator. If $s_i = 0$ then σ_i is a deactivator or inhibitor. An input at its canalizing value is called canalized. If none of the canalizing inputs are at their respective canalizing value, the output is given by a default totalistic Boolean function G with $k - u$ input, for which the output only depends on the sum of the values of the inputs, not on the individual inputs; thus it is the aggregation of the inputs that governs the node evolution. If $u = k$, we obtain the classical NCF case studied in (Kauffman et.al. [23]) or (Peixoto [19]). On the other hand, in (Layne [21]), some properties of PNCFs are analyzed and an algorithm for identifying u from the truth table of a Boolean function is generated, while the choice of G in simulations is given by a random bit generator.

In Figure 1 we show sample pattern formation plots and corresponding densities of ones (activity level) or the fraction of active nodes for three networks with the parameters indicated in the figure and explained below. The bottom graphs correspond to a NCF in which all inputs are canalizing. The dynamics reach stability rather quickly. We note here that

the density of ones is an estimate for the probability of finding a node in an active state at any given time.

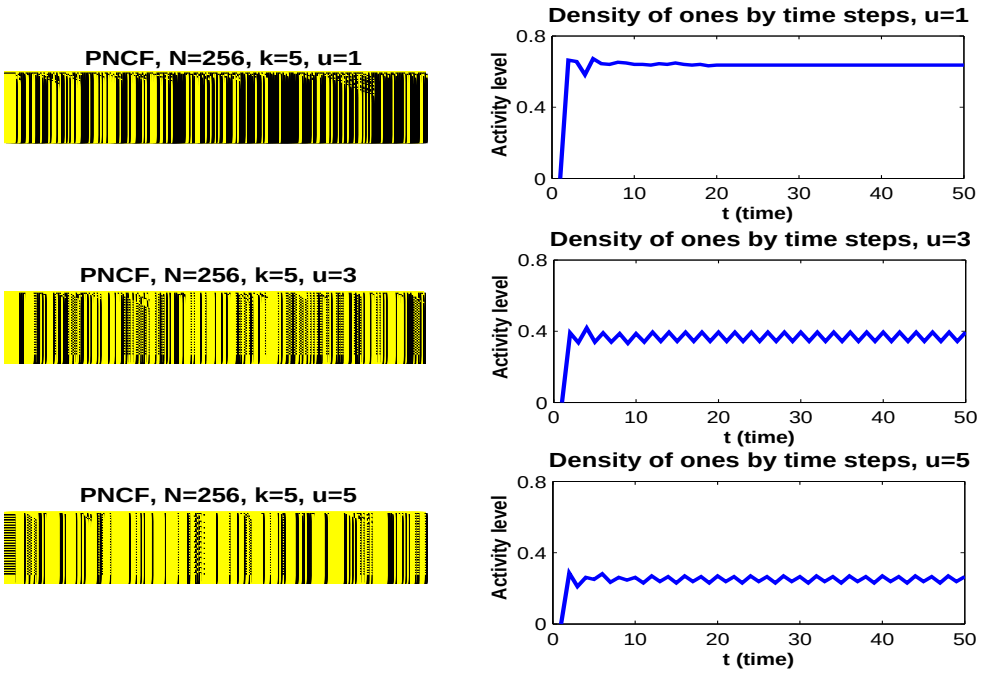

Figure 1. Pattern formation plots and density of ones for PNCF (top and middle) and NCF (bottom) networks where the nodes are ordered horizontally in the left figures, and each node has $k = 5$ inputs: itself and the two nearest neighbors on each side, so it is a cellular automaton. There are $N = 256$ cells, each obeying a PNCF with the indicated canalization depths. The function G that is applied when the canalizing inputs are not on their canalizing values is a generalized elementary cellular automata rule 126: if all inputs are either zero or one, the node becomes zero (yellow) at the next time step, otherwise it becomes a one (black). Only one of the 256 cells is initially black, representing an active node. The automata are evolved 50 time steps (downward) in the left column; the time evolves horizontally in the right column with the corresponding activity levels. Notice that all three networks reach stability with a fixed state for the PNCF with canalization depth $u = 1$ and periodicity for the the others. Also note that the cases $u = 3, u = 5$ are fairly similar with a small decrease in the overall activity level, so the canalization depth does not have a significant impact once it reaches a threshold value (3 in this case).

The parameters and quantities that determine the distribution of the ensemble of PNCFs from which we choose our functions are as follows:

$$a = P(s_i = 1) = \textit{the probability that an arbitrary input is an activator}$$

$$c = \textit{the probability that the canalizing value of an input is } 1$$

$$p_t = \textit{the probability that the function } G \textit{ takes on value } 1$$

b_t = *the density of ones for the network*

We start by deriving a formula for b_t using a mean-field approximation, in which correlations among inputs are ignored. For large networks this assumption has almost no bearing on the overall dynamics. So we consider the function

$$\begin{aligned}\gamma(b_t) &= P(\text{an arbitrary input for an arbitrary node of the network is at its canalizing value})\\ &= P(\text{the input is canalizing}) \cdot P(\text{its canalizing value is } 1 \,|\, \text{the input is canalizing})\\ &\quad \cdot P(\text{the input } = 1 \,|\, \text{its canalizing value is 1 and the input is canalizing})\\ &\quad + P(\text{the input is canalizing}) \cdot P(\text{its canalizing value is } 0 \,|\, \text{the input is canalizing})\\ &\quad \cdot P(\text{the input } = 0 \,|\, \text{its canalizing value is 0 and the input is canalizing})\\ &= \frac{u}{k} \cdot c \cdot b_t + \frac{u}{k} \cdot (1-c) \cdot (1-b_t) = \frac{u}{k}[cb_t + (1-c)(1-b_t)].\end{aligned}$$

We obtain the density of ones at time $t+1$ as the probability that at least one canalizing input is canalized and is an activator, or no canalizing input is canalized and the output of G is 1. More precisely,

$$\begin{aligned}b_{t+1} &= P(\text{at least one input is canalized}) \cdot P(\text{the canalizing input is an activator})\\ &\quad + P(\text{no canalizing input is canalized}) \cdot P(\text{the output of } G = 1)\\ &= [1-(1-\gamma(b_t))^u] \cdot a + (1-\gamma(b_t))^u \cdot p_t = a + (p_t - a)(1-\gamma(b_t))^u.\end{aligned}$$

Observe that for $k \to \infty$, which implies $u \to \infty$, we have that $b_t \to a$.

Now, of all the Boolean functions one could consider for G, let us focus on threshold functions which are typical for neural or genetic networks, and which have been studied for example in (Anthony [24], Beck and Matache [25]), Raeymaekers [8]). We define

$$G(\sigma_{u+1}, \ldots, \sigma_k) = \begin{cases} 1 & \text{if} \quad d_1 \le \frac{1}{k-u} \sum_{i=u+1}^{k} \sigma_i \le d_2 \\ 0 & \text{if} \quad \text{otherwise} \end{cases} \tag{2}$$

where $0 \le d_1 \le d_2 \le 1$ and k are fixed parameters. The probability that the output of (2) is 1 at time $t+1$, denoted p_{t+1}, is obtained under a mean-field approximation (Beck and Matache [25]). Then

$$\begin{aligned}p_{t+1} &= P(G(\sigma_{u+1}(t), \ldots, \sigma_k(t)) = 1) = P\left(d_1 \le \frac{1}{k-u} \sum_{i=u+1}^{k} \sigma_i(t) \le d_2\right)\\ &= \sum_{\substack{d_1(k-u) \le s \le d_2(k-u)\\ s \in \mathbf{N}}} \binom{k-u}{s} b_t^s (1-b_t)^{k-u-s},\end{aligned}$$

since b_t is an estimate for the probability of finding a node in state 1 at time t.

Now, by combining the two formulas found in this section and replacing $\gamma(b_t)$, we obtain:

$$b_{t+1} = a + (p_t - a)\left[1 - \frac{u}{k}[cb_t + (1-c)(1-b_t)]\right]^u \tag{3}$$

and

$$p_{t+1} = \sum_{\substack{d_1(k-u)\le s\le d_2(k-u)\\ s\in\mathbf{N}}} \binom{k-u}{s} b_t^s (1-b_t)^{k-u-s}. \tag{4}$$

By solving the system $b_{t+1} = b_t, p_{t+1} = p_t$ we get the fixed points which indicate the long-term dynamics of our model. We will denote by b^*, p^* the fixed points, or the average of the fixed points or limiting cycles as in (Peixoto [19]).

2.2. **Numerical Results.** In this section we explore the influence of parameters on the dynamics of b_t and p_t using bifurcation diagrams, to have a graphical view of the estimates b^*, p^*. After considering a variety of parameter combinations for simulations, we select some typical graphs that basically clarify the dynamics of the two dimensional map given by equations (3) and (4). All numerical investigations in this paper are performed with MATLAB. In Figure 2 we plot bifurcation diagrams along the probability a that an arbitrary input is an activator (horizontal axis) for $k = 6$, $u = 2, 3, 5$, $c = 0.5, 1$, $d_1 = 0, 0.3, 0.5, 0.8$, and $d_2 = 1$ (corresponding to a simple threshold function typical for neural networks). This way we can assess the impacts of u, c, d_1 as well. Plots for other parameter combinations when $k = 6$ yield similar results. The plots for p_t and b_t are graphed on the same figure for an easy comparison. The diagrams for $p_t \in [0, 1]$ are plotted in $[0, 1]$, while those for $b_t \in [0, 1]$ are presented in the interval $[1, 2]$ above p_t. The canalization depth has little impact once a certain value of u is reached: the cases $u = 1, 2$ are similar, and a more significant modification occurs at $u = 3$, while the cases $u = 3, 4, 5$ are rather similar. This situation has been analyzed also in (Layne et.al. [21]), and it is to be expected since every extra (nested) canalizing input "freezes" at least half of the remaining truth table of the Boolean function.

Based on a variety of simulations that yield more or less the same type of behavior seen in Figure 2, we conclude that the two-dimensional map (3)-(4) exhibits mostly stability with periodic orbits, and period doubling/halving bifurcations may occur. The canalization depth has no impact once it reaches a certain threshold value. The period of orbits may change with increased bias towards inhibition. Thus the density of ones for the network

and the probability that the function G takes on value 1 generate either unique fixed points b^*, p^*, or averages over periodic orbits.

We note here that it is possible that other parameter combinations may yield chaos for p_t or b_t, whose bifurcation diagrams are quite similar as seen in Figure 2. However, we have not observed that phenomenon in our simulations. Regardless, $p*, b*$ are computed as the long term average of the p_t, b_t values, respectively. Although chaos is not observed in simulations for the two-dimensional map p_t, b_t, we will observe chaotic behavior for the network dynamics in what follows.

We now turn to the computation of the sensitivity of the network to perturbations and critical conditions for a phase transition from order to chaos in the network.

3. Sensitivity and Criticality

3.1. **Theoretical Results.** Recall from (Moreira and Amaral [22], Peixoto [19], Shmulevich and Kauffman [2]) that the average influence of the variables of a Boolean function $F : \{0,1\}^k \to \{0,1\}$ belonging to an ensemble of Boolean functions $\mathcal{E}$ is

$$I(F) = \frac{1}{k}\sum_{i=1}^{k} I_i(F), \qquad I_i(F) = P(F \text{ changes value when input } i \text{ is changed}) \tag{5}$$

The quantity $I_i(F)$ is the influence of the i-th variable on F. By averaging this quantity over the entire ensemble $\mathcal{E}$ one obtains the average influence of $\mathcal{E}$, denoted $I(\mathcal{E})$. The critical condition that separates order from chaos is

$$kI(\mathcal{E}) = 1. \tag{6}$$

Note that $I(\mathcal{E})$ depends on k as well and that the value of k for which this condition holds is usually denoted by k_c, the critical connectivity. The quantity $\lambda = kI(\mathcal{E})$ is called the network sensitivity.

In order to determine the critical condition for the PNCF network we will consider a generic PNCF as described in Section 2.1. Observe that the influence of the i-th variable depends on its canalizing versus non-canalizing property, therefore let us look at the possible cases which are explained in sufficient detail.

Case 1. $i \leq u$, meaning that the input σ_i is canalizing.

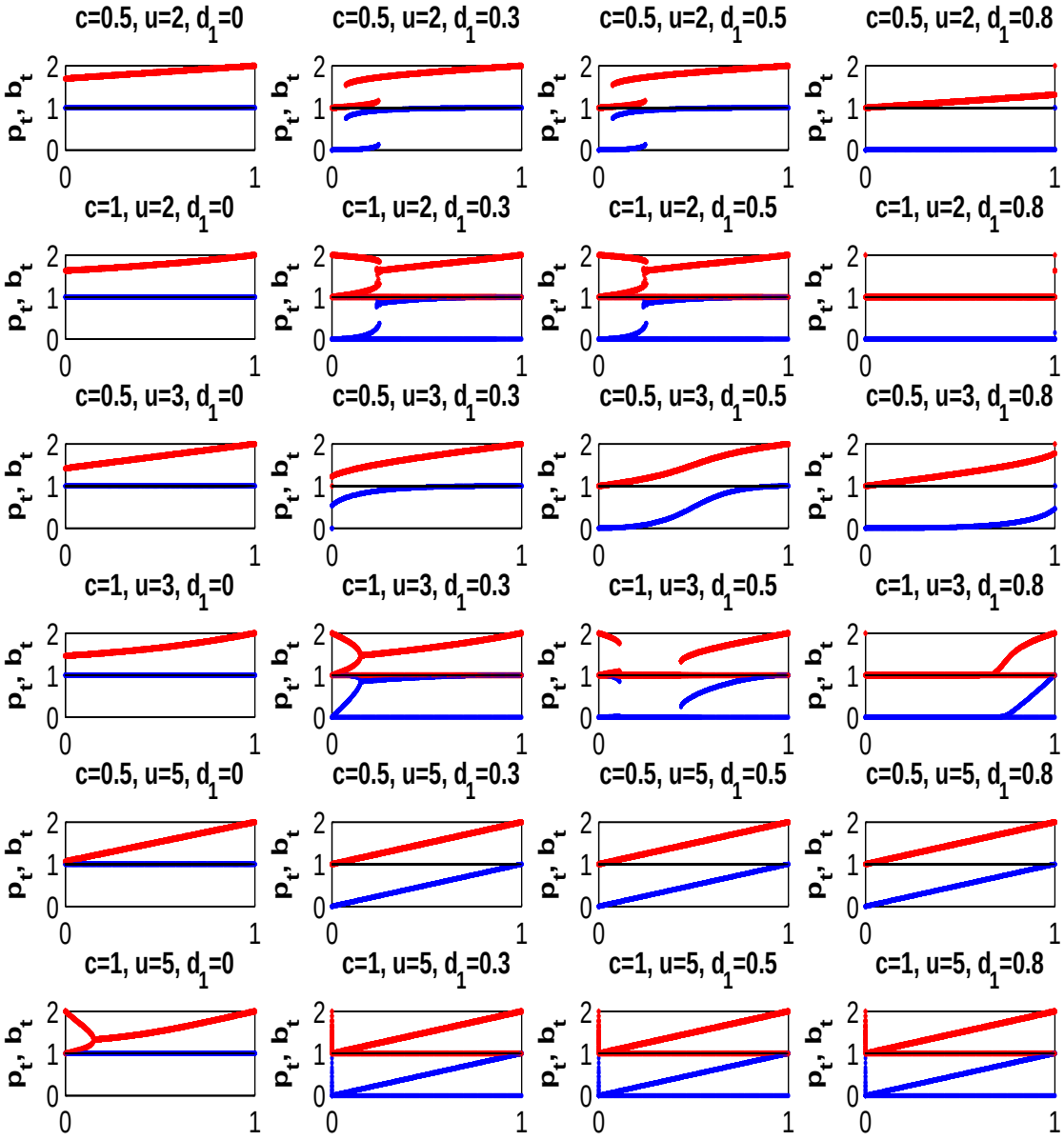

Figure 2. Bifurcation diagrams for p_t, the probability that the function G takes on value 1 (presented with blue in the interval $[0, 1]$ on the y-axis), and b_t, the density of ones of the network (presented with red in the interval $[1, 2]$ on the y-axis), along the probability a that an arbitrary input is an activator (x-axis in all plots, not labeled). The parameters are: connectivity $k = 6$, upper threshold for the function G set to $d_2 = 1$, and combinations of the other parameters as specified in the titles. The parameters $c = 0.5, 1$ (the probability that the canalizing value of an input is 1) and the canalization depth $u = 2, 3, 5$ change by rows (1 and 2, 3 and 4, 5 and 6, respectively), while the lower threshold of the function G takes on values $d_1 = 0, 0.3, 0.5, 0.8$ and changes by columns. Observe stability for all parameter values with fixed points or periodic orbits. As expected, an increased canalization depth does not produce significant modifications.

Observe that in this case, the first $i-1$ inputs cannot be on their canalizing input values, since those would fix the output of F regardless of the i-th input value, so no changes can occur when this one is flipped. Now $P(\text{the first } i-1 \text{ inputs are not on their canalizing values}) = [1-\gamma(b^*)]^{i-1}$. Then there are two ways to get a flip in the output of F when the i-th input is flipped. More precisely, we have two sub-cases.

Case 1(a). None of the remaining canalizing inputs is on its canalizing value. The probability of this happening is given by $[1-\gamma(b^*)]^{u-i}$. In this case the function G will determine the output of F. Thus we want the output of G to be different than the canalized output of σ_i. The probability of this event is $1-\eta_0$, where $\eta_0 = P(G(\sigma_{u+1},\ldots,\sigma_k) = s_i)$. Thus, we obtain the final probability as follows:

$$[1-\gamma(b^*)]^{i-1} \cdot [1-\gamma(b^*)]^{u-i} \cdot (1-\eta_0).$$

To determine η_0, we take into account if the canalized output is a 1 or a 0, and the probability that G will produce the same output (in the long term). More precisely,

$$\eta_0 = P(s_i=1)P(G(\sigma_{u+1},\ldots,\sigma_k)=1) + P(s_i=0)P(G(\sigma_{u+1},\ldots,\sigma_k)=0)$$

$$= ap^* + (1-a)(1-p^*).$$

Case 1(b). At least one of the remaining canalizing inputs $j > i$ is on its canalizing value. The probability of this happening is given by $1-[1-\gamma(b^*)]^{u-i}$. If this is the case, then a flip in the output of F occurs when the canalized output of j is different than the canalized output of i. This happens with probability $1-\eta$ where $\eta = P(\text{any two canalizing inputs have the same canalized output value})$. Thus, we obtain the final probability as follows:

$$[1-\gamma(b^*)]^{i-1} \cdot (1-[1-\gamma(b^*)]^{u-i}) \cdot (1-\eta).$$

Here

$$\eta = P(s_i=1)P(s_j=1) + P(s_i=0)P(s_j=0) = a^2 + (1-a)^2$$

accounting for the fact that the two inputs could be both 1 or 0.

In conclusion, **Case 1** leads to the following final formula:

$$I_i(F) = [1-\gamma(b^*)]^{i-1}\left\{[1-\gamma(b^*)]^{u-i}(1-\eta_0) + (1-[1-\gamma(b^*)]^{u-i})(1-\eta)\right\}. \tag{7}$$

Observe that if $i = u$ the term obtained from **Case 1(b)** does not exist anymore. In this case the formula is $I_u(F) = [1-\gamma(b^*)]^{u-1}(1-\eta_0)$.

Case 2. $i > u$, that is σ_i is not a canalizing input.

In this case none of the canalizing inputs should be on its canalizing value in order to have a possible change in the output. This happens with probability $[1-\gamma(b^*)]^u$. Then a flip in the output of F occurs if the output of G is flipped. If τ is the probability that the output of G is flipped when input i is flipped, the formula is:

$$I_i(F) = \tau[1-\gamma(b^*)]^u. \tag{8}$$

To find τ we use conditional probability on the actual value of node σ_i, namely

$$\begin{aligned}\tau = &\\ &= P(\text{output of } G \text{ is flipped from 0 to 1 when } \sigma_i \text{ is flipped from 0 to 1} | \sigma_i = 0) \cdot P(\sigma_i = 0)\\ &+P(\text{output of } G \text{ is flipped from 1 to 0 when } \sigma_i \text{ is flipped from 0 to 1} | \sigma_i = 0) \cdot P(\sigma_i = 0)\\ &+P(\text{output of } G \text{ is flipped from 0 to 1 when } \sigma_i \text{ is flipped from 1 to 0} | \sigma_i = 1) \cdot P(\sigma_i = 1)\\ &+P(\text{output of } G \text{ is flipped from 1 to 0 when } \sigma_i \text{ is flipped from 1 to 0} | \sigma_i = 1) \cdot P(\sigma_i = 1).\end{aligned}$$

Here we need to consider several cases, depending on the quantities $d_1(k-u), d_2(k-u)$ being integer values or not, since this has influence on the output of G according to the formula (2).

Case 2(a) $0 < [d_1(k-u)] < d_1(k-u) < [d_2(k-u)] \leq d_2(k-u) < k-u$, where $[d]$ represents the integer part of d. This means that $d_1(k-u) \notin \mathbf{N}$ and $d_2 < 1$. Let us denote $s = \sum_{j=u+1}^{k} \sigma_j$ where σ_i is at its original value before being flipped. Then

$$\begin{aligned}&P(\text{output of } G \text{ is flipped from 0 to 1 when } \sigma_i \text{ is flipped from 0 to 1} | \sigma_i = 0) \cdot P(\sigma_i = 0)\\ &= P(s < d_1(k-u), d_1(k-u) \leq s+1 \leq d_2(k-u)) \cdot (1-b^*)\\ &= P(s = [d_1(k-u)])(1-b^*) = \binom{k-u-1}{[d_1(k-u)]} (b^*)^{[d_1(k-u)]} (1-b^*)^{k-u-[d_1(k-u)]}\end{aligned}$$

since a change of σ_i from 0 to 1 will add one unit to s.

Similarly,

$$\begin{aligned}&P(\text{output of } G \text{ is flipped from 1 to 0 when } \sigma_i \text{ is flipped from 0 to 1} | \sigma_i = 0) \cdot P(\sigma_i = 0)\\ &= P(d_1(k-u) \leq s \leq d_2(k-u)), d_2(k-u) < s+1) \cdot (1-b^*)\\ &= P(s = [d_2(k-u)])(1-b^*) = \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]} (1-b^*)^{k-u-[d_2(k-u)]};\end{aligned}$$

$$\begin{aligned}&P(\text{output of } G \text{ is flipped from 0 to 1 when } \sigma_i \text{ is flipped from 1 to 0} | \sigma_i = 1) \cdot P(\sigma_i = 1)\\ &= P(d_1(k-u) \leq s-1 \leq d_2(k-u)), d_2(k-u) < s) \cdot (b^*)\\ &= P(s = [d_2(k-u)]+1)(b^*) = \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]+1} (1-b^*)^{k-u-1-[d_2(k-u)]};\end{aligned}$$

where in the binomial coefficient we use $[d_2(k-u)]$ since $\sigma_i = 1$ and only the other nodes need to be considered. Finally,

$$P(\text{output of } G \text{ is flipped from 1 to 0 when } \sigma_i \text{ is flipped from 1 to } 0 | \sigma_i = 1) \cdot P(\sigma_i = 1)$$

$$= P(s-1 < d_1(k-u), d_1(k-u) \leq s \leq d_2(k-u)) \cdot (b^*)$$

$$= P(s = [d_1(k-u)] + 1)(b^*) = \binom{k-u-1}{[d_1(k-u)]} (b^*)^{[d_1(k-u)]+1} (1-b^*)^{k-u-1-[d_1(k-u)]}.$$

Then τ is the sum of these individual probabilities, so that

$$\begin{aligned}
\tau = & \binom{k-u-1}{[d_1(k-u)]} (b^*)^{[d_1(k-u)]} (1-b^*)^{k-u-[d_1(k-u)]} \\
& + \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]} (1-b^*)^{k-u-[d_2(k-u)]} \\
& + \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]+1} (1-b^*)^{k-u-1-[d_2(k-u)]} \\
& + \binom{k-u-1}{[d_1(k-u)]} (b^*)^{[d_1(k-u)]+1} (1-b^*)^{k-u-1-[d_1(k-u)]}
\end{aligned}$$

which leads to

$$\begin{aligned}
\tau = & \binom{k-u-1}{[d_1(k-u)]} (b^*)^{[d_1(k-u)]} (1-b^*)^{k-u-1-[d_1(k-u)]} \\
& + \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]} (1-b^*)^{k-u-1-[d_2(k-u)]}
\end{aligned} \tag{9}$$

Case 2(b) $0 < [d_1(k-u)] = d_1(k-u) \leq [d_2(k-u)] \leq d_2(k-u) < k-u$. Here $d_1(k-u) \in \mathbf{N}$ and $d_2 < 1$. Using a similar procedure as in **Case 2(a)**, we obtain

$$\begin{aligned}
\tau = & P(s = [d_1(k-u)] - 1)(1-b^*) + P(s = [d_2(k-u)])(1-b^*) \\
& + P(s = [d_2(k-u)] + 1)(b^*) + P(s = [d_1(k-u)])(b^*) \\
= & \binom{k-u-1}{[d_1(k-u)]-1} (b^*)^{[d_1(k-u)]-1} (1-b^*)^{k-u-[d_1(k-u)]+1} \\
& + \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]} (1-b^*)^{k-u-[d_2(k-u)]} \\
& + \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]+1} (1-b^*)^{k-u-1-[d_2(k-u)]} \\
& + \binom{k-u-1}{[d_1(k-u)]-1} (b^*)^{[d_1(k-u)]} (1-b^*)^{k-u-[d_1(k-u)]}
\end{aligned}$$

which leads to

$$\begin{aligned}
\tau = & \binom{k-u-1}{[d_1(k-u)]-1} (b^*)^{[d_1(k-u)]-1} (1-b^*)^{k-u-[d_1(k-u)]} \\
& + \binom{k-u-1}{[d_2(k-u)]} (b^*)^{[d_2(k-u)]} (1-b^*)^{k-u-1-[d_2(k-u)]}
\end{aligned} \tag{10}$$

Observe that the second term of the sum is the same in both cases, since the fact that $d_2(k-u)$ is natural or not does not make a difference in the formula.

Case 2(c) $0 \leq [d_1(k-u)] < d_1(k-u) < [d_2(k-u)] = k-u$, so that $d_2 = 1$ and $d_1(k-u) \notin \mathbf{N}$. Then

$$\tau = P(s = [d_1(k-u)])(1-b^*) + P(s = [d_1(k-u)+1])(b^*)$$

$$= \binom{k-u-1}{[d_1(k-u)]} (b^*)^{[d_1(k-u)]} (1-b^*)^{k-u-1-[d_1(k-u)]} \tag{11}$$

since two of the four possibilities in the sums of the previous two cases are not valid anymore.

Case 2(d) $0 < [d_1(k-u)] = d_1(k-u) < [d_2(k-u)] = k-u$, so that $d_2 = 1$ and $d_1(k-u) \in \mathbf{N}$. Then

$$\tau = P(s = [d_1(k-u)]-1)(1-b^*) + P(s = [d_1(k-u)])(b^*)$$

$$= \binom{k-u-1}{[d_1(k-u)]-1} (b^*)^{[d_1(k-u)]-1} (1-b^*)^{k-u-[d_1(k-u)]} \tag{12}$$

Case 2(e) $0 = [d_1(k-u)] = d_1(k-u) < [d_2(k-u)] = k-u$, so that $d_1 = 0, d_2 = 1$. Then the function G is constant equal to 1 so $\tau = 0$.

This covers all the possibilities for τ and we can now put together formulas (7) and (8) to obtain the influence of the (generic) i-th variable:

$$I_i(F) = \left(\frac{u}{k}\right) [1-\gamma(b^*)]^{i-1} \left\{ [1-\gamma(b^*)]^{u-i}(1-\eta_0) + (1-[1-\gamma(b^*)]^{u-i})(1-\eta) \right\} + \left(\frac{k-u}{k}\right) \tau [1-\gamma(b^*)]^u. \tag{13}$$

Then the average influence of the generic function F (which is equivalent to $I(\epsilon)$) is given by

$$\lambda = kI(F) = \sum_{i=1}^{k} I_i(F) =$$

$$\sum_{i=1}^{k} \left(\frac{u}{k}\right) [1-\gamma(b^*)]^{i-1} \left\{ [1-\gamma(b^*)]^{u-i}(1-\eta_0) + (1-[1-\gamma(b^*)]^{u-i})(1-\eta) \right\} + \sum_{i=1}^{k} \left(\frac{k-u}{k}\right) \tau [1-\gamma(b^*)]^u$$

which leads to the critical condition for sensitivity, $\lambda = kI(F) = 1$, as follows:

$$(14)\qquad \binom{u}{k}(1-\eta)\frac{1-[1-\gamma(b^*)]^k}{\gamma(b^*)}+u(\eta-\eta_0)[1-\gamma(b^*)]^{u-1}+(k-u)\tau[1-\gamma(b^*)]^u=1$$

Since λ represents the response to small perturbations, if $\lambda<1$ the network is in a stable phase. On the other hand, if $\lambda>1$ the network is in a chaotic phase.

Now, let us denote by α the fraction of canalizing nodes, that is $\alpha=u/k$. Then equation (14) becomes

$$(15)\qquad \lambda=\alpha(1-\eta)\frac{1-[1-\gamma(b^*)]^k}{\gamma(b^*)}+k\alpha(\eta-\eta_0)[1-\gamma(b^*)]^{u-1}+k(1-\alpha)\tau[1-\gamma(b^*)]^u\tau.$$

We note here that when we restrict these formulas to the case of a NCF network, that is when $G(\sigma_{u+1},\dots,\sigma_k)=s_d$ with $s_d\in\{0,1\}$ being a default fixed output, we obtain the results of (Peixoto [19]).

3.2. **Numerical Results.** First let us summarize the information needed to compute the sensitivity of the network and the critical condition in the form of an algorithm. Although we are repeating previous formulas, it is useful to have an overview of the procedure and the necessary parameters.

Algorithm 1:

(i) Initialize parameter values a,c,k,u,d_1,d_2.

(ii) Find the fixed points b^*,p^* (or average values of fixed points if more than one) of the following system:

$$\begin{cases} b=a+(p-a)(1-\gamma(b))^u \\ p=\sum_{s\in\mathbf{N}}\binom{k-u}{s}b^s(1-b)^{k-u-s}, \qquad d_1(k-u)\le s\le d_2(k-u)\end{cases}$$

where

$$\gamma(b)=\frac{u}{k}[cb+(1-c)(1-b)].$$

(iii) Compute

$$\eta_0=ap^*+(1-a)(1-p^*)\qquad\text{and}\qquad \eta=a^2+(1-a)^2$$

(iv) Compute τ according to one of the five cases:

(a) $0<[d_1(k-u)]<d_1(k-u)<[d_2(k-u)]\le d_2(k-u)<k-u$

$$\tau=\binom{k-u-1}{[d_1(k-u)]}(b^*)^{[d_1(k-u)]}(1-b^*)^{k-u-1-[d_1(k-u)]}$$

$$+\binom{k-u-1}{[d_2(k-u)]}(b^*)^{[d_2(k-u)]}(1-b^*)^{k-u-1-[d_2(k-u)]}$$

(b) $0<[d_1(k-u)]=d_1(k-u)\le[d_2(k-u)]\le d_2(k-u)<k-u$

$$\tau=\binom{k-u-1}{[d_1(k-u)]-1}(b^*)^{[d_1(k-u)]-1}(1-b^*)^{k-u-[d_1(k-u)]}$$

$$+\binom{k-u-1}{[d_2(k-u)]}(b^*)^{[d_2(k-u)]}(1-b^*)^{k-u-1-[d_2(k-u)]}$$

(c) $0 \leq [d_1(k-u)] < d_1(k-u) < [d_2(k-u)] = k-u$

$$\tau = \binom{k-u-1}{[d_1(k-u)]}(b^*)^{[d_1(k-u)]}(1-b^*)^{k-u-1-[d_1(k-u)]}$$

(d) $0 < [d_1(k-u)] = d_1(k-u) < [d_2(k-u)] = k-u$

$$\tau = \binom{k-u-1}{[d_1(k-u)]-1}(b^*)^{[d_1(k-u)]-1}(1-b^*)^{k-u-[d_1(k-u)]}$$

(e) $0 = [d_1(k-u)] = d_1(k-u) < [d_2(k-u)] = k-u$

$$\tau = 0$$

(v) Find the sensitivity as in (14)

$$\lambda = kI(F) =$$

$$\left(\frac{u}{k}\right)(1-\eta)\frac{1-[1-\gamma(b^*)]^k}{\gamma(b^*)} + u(\eta-\eta_0)[1-\gamma(b^*)]^{u-1} + (k-u)\tau[1-\gamma(b^*)]^u.$$

(vi) Repeat this procedure for different parameter values.

Using this algorithm we can generate phase transition diagrams to identify the stable phase, the chaotic phase, and the critical transition in terms of parameter values. However, before we do that, let us analyze the accuracy of our model in comparison to an actual network obeying the PNCF scenario. To do that, we compare the formula for sensitivity (15) to the outcome of applying perturbations to an actual PNCF network.

More precisely, we can construct Derrida plots (Derrida and Pomeau [26], Kauffman [1], Shmulevich and Kauffman [2]) which map the average Hamming distance at time $t+1$, $H(t+1)$, against the average Hamming distance at time t, where $H(t) = \frac{1}{k}\sum_{i=1}^{k}|\sigma_i(t) - \sigma_i'(t+1)|$, and $(\sigma_1, \sigma_2, \ldots, \sigma_k)$ and $(\sigma_1', \sigma_2', \ldots, \sigma_k')$ are two states of the network. The result of plotting these values as t increases to $t+1$ and averaging over many initial states and networks is the so-called the Derrida curve. If the curve is above the main diagonal, for which $H(t) = H(t+1)$, it reflects instability in the sense that a small disturbance tends to increase during the next time steps, so that the network is sensitive to initial conditions. Derrida curves below the main diagonal indicate the tendency to overcome the disturbance and correspond to stability. Derrida curves along the diagonal indicate the most complex behavior at the edge-of-chaos, where the system is flexible in face of perturbations.

We can compare the Derrida plots for some given parameters with a line having slope λ obtained from (15) with the same parameters. For small initial perturbations the plots should be similar. We show a few examples in Figure 3. Note that the Derrida plots (with dots or markers) follow the line plots of our model fairly closely especially for small

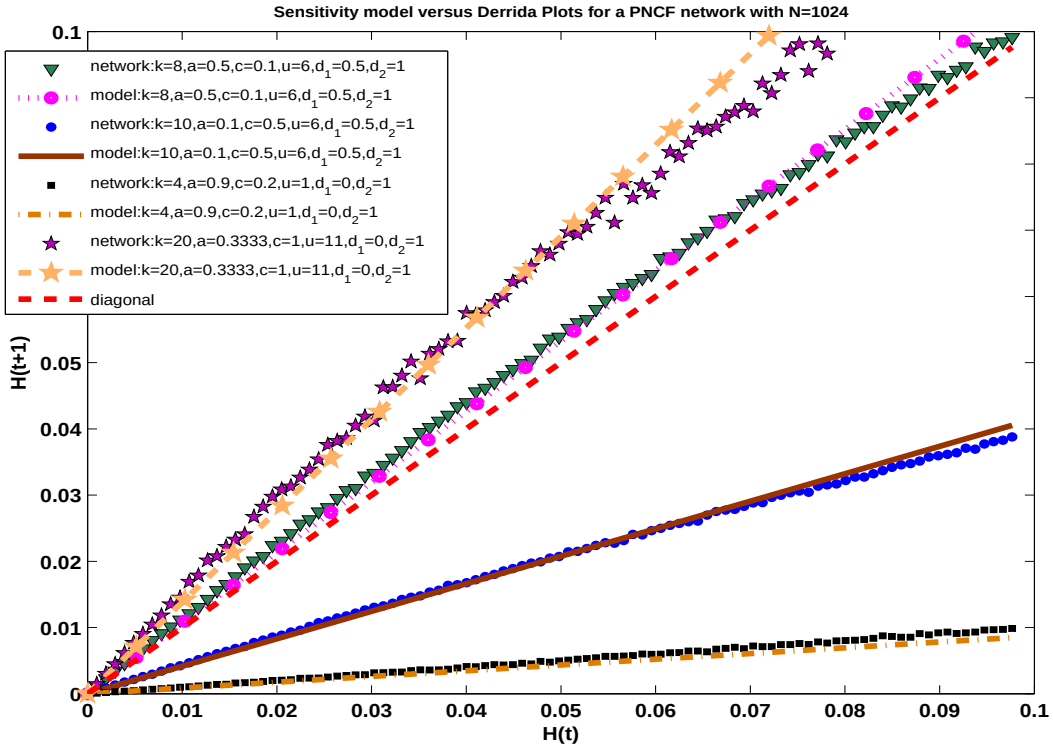

Figure 3. Derrida plots and corresponding model results for various parameter combinations. Our model should match the Derrida plot for small perturbations ($H(t)$ close to zero), while for bigger perturbations ($H(t)$ away from zero) the curves could separate. However, given that λ corresponds to changes when a single input is flipped, we only need to take into account small values of $H(t)$. The parameter combinations are specified in the legend, and their values range as follows: connectivity $k = 4, 8, 10, 20$, the probability that an arbitrary input is an activator $a = 1/10, 1/3, 1/2, 9/10$, the probability that the canalizing value of an input is 1 takes on values $c = 0.1, 0.2, 0.5, 1$, the lower threshold of the function G is $d_1 = 0, 0.5$ and the upper threshold is $d_2 = 1$, while the canalization depth takes on values $u = 1, 6, 11$. These parameters are chosen in order to present curves that lie below, as well as above the main diagonal. Other parameter combinations lead to similar graphs.

values of $H(t)$, representing small perturbations. The graphs are quite similar for various parameter combinations and therefore are not included here. The matches improve for larger networks and for an increased number of initial states and randomly chosen PNCFs for the computation of the average Hamming distance. We conclude that we can use our model to predict the behavior of the system in a variety of parameter settings.

Now, we use our model to generate phase transition diagrams in order to understand the impact of the parameters on the sensitivity of the network to small perturbations. We choose to generate three dimensional plots of λ against parameters a and k which are varied freely. These parameters will show us the impact the number of inputs, k, has on the

stability of the network and also how the probability a canalizing input is an activator, a, impacts stability. We also plot the plane at $\lambda = 1$ which separates order from chaos. The intersection of this plane with the mesh generated with our model is the critical condition, or the edge of chaos. In Figures 4-5 we show several typical graphs that contain some of the parameter combinations shown in the Derrida plots. We can see the parameter values that lead to order versus chaos, as well as the intersection of the two meshes which is the critical condition. In all cases we note again that an increase in the number of canalizing inputs does not generate significant qualitative modifications in the overall shapes of the graphs. We include four different values for $\alpha = u/k$ in each figure. Note that for $d_1 = 0, d_2 = 1$ the graphs indicate stability for a wide range of values, mostly for large values of a. This is expected since lower values of a indicate a higher probability of inhibition, which means more 0 outputs. However, when $d_1 = 0, d_2 = 1$, the probability of activation increases for G. Therefore the increased inhibition corresponding to small a ends up conflicting with the large activation likelihood of G. Additionally, we note in simulations that as c increases, the chaos range of a is shifted towards left so that the chaos region becomes smaller.

A similar impact of c is noted for the case $d_1 = 0.5, d_2 = 1$, however the graphs indicate larger parameter ranges that lead to chaos. At the same time, the graphs are more visually complex, and the case $\alpha = 0.25$ is clearly different than the other values of $\alpha > 0.25$ in each case considered in simulations. Thus we can see again that once a certain threshold of α is reached, the depth of canalization does not induce major changes. This threshold is fairly small in comparison to the connectivity level of the nodes, and we note that it slightly increases as the total number of inputs, k, increases.

Other cases considered in simulations have a similar behavior, although the exact shape of the meshes may differ. A comprehensive account of the impact of all parameters is subject for further research, together with analytical and numerical investigations of the threshold of α over which the dynamics are fairly similar.

4. Conclusions and Directions for Future Work

In this paper we consider Boolean networks governed by PNCFs. Given a certain canalizing depth u we complete the PNCF using a simple threshold function. We construct a model for finding the sensitivity of the network to perturbations and identify the critical condition between order and chaos. We provide Derrida plots indicating that the model matches the network behavior for a wide range of parameter values. The match is improved with larger networks and increased number of PNCFs and initial conditions. The phase

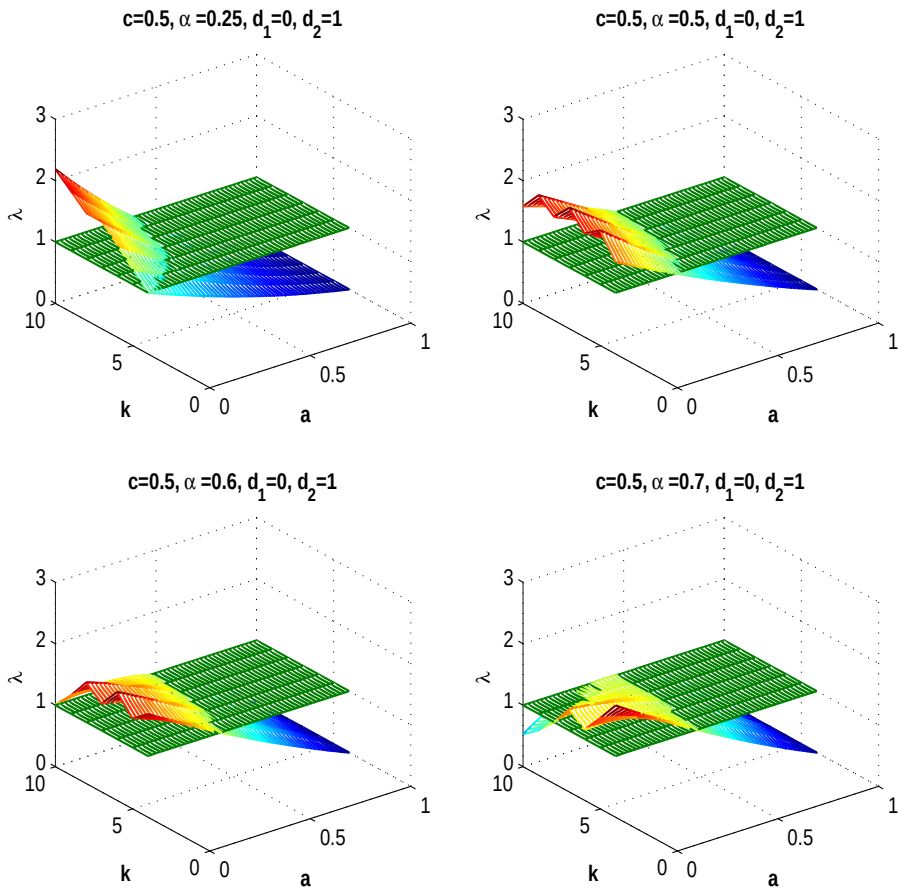

Figure 4. Phase diagram and critical plane for the sensitivity λ versus the probability a that an arbitrary input is an activator and the connectivity k. The other parameters are fixed as follows: the probability that the canalizing value of an input is 1 is set to $c = 0.5$, the thresholds of the function G are $d_1 = 0, d_2 = 1$, and the canalization depth u expressed as a fraction of the connectivity is set to $\alpha = u/k = 0.25, 0.5, 0.6, 0.7$ corresponding to each subplot in the figure. Small values of a induce chaos, while larger values lead to a transition to order. An increase in α does not generate significant qualitative changes in the graphs once a threshold is reached. Chaos can occur for mostly low values of a. An increase in α induces more stability.

transition diagrams show both stability and chaotic behavior for varying parameter values. Our model confirms that the canalizing depth seems to have little impact after a certain value of u is reached; this threshold seems to be fairly low in comparison to the connectivity, k. The threshold value of u slightly increases as the total number of inputs, k, increases.

Areas for future research include providing an in-depth numerical analysis of the parameters that generate a phase transition from order to chaos and generating analytical results regarding the importance of the canalizing depth, as well as its threshold value. Furthermore, it would be interesting to combine these types of Boolean functions with other types encountered in real applications and generate a heterogeneous model of mixed functions

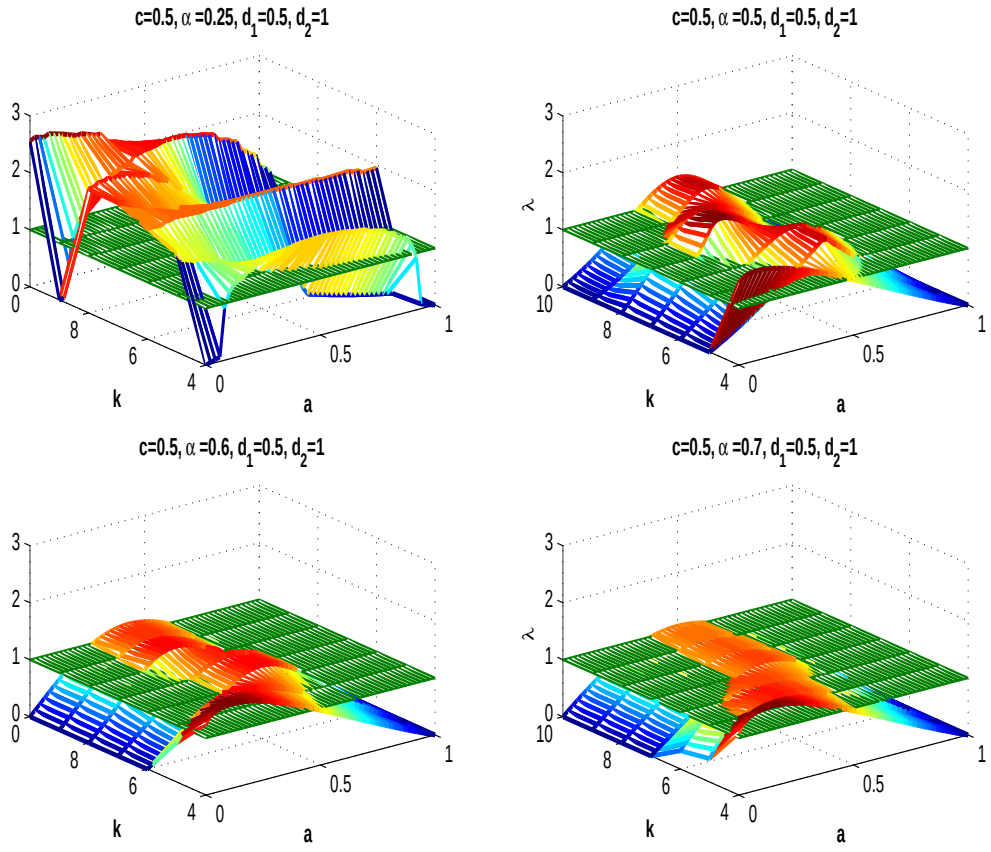

Figure 5. Phase diagram and critical plane for the sensitivity λ versus the probability a that an arbitrary input is an activator and the connectivity k. The other parameters are fixed as follows: the probability that the canalizing value of an input is 1 is set to $c = 0.5$, the thresholds of the function G are $d_1 = 0.5, d_2 = 1$, and the canalization depth u expressed as a fraction of the connectivity is set to $\alpha = u/k = 0.25, 0.5, 0.6, 0.7$ corresponding to each subplot in the figure. This figure is the analog of Figure 4 for the indicated parameters. An increase in α induces more stability, and a wider range of values of λ close to the phase transition.

that can be used in applications such as neurological and biological networks. In particular, as mentioned in the introduction, in (Kochi and Matache [6]) the authors identify several different classes of Boolean functions in a signal transduction network of a fibroblast cell. Some of those classes encompass PNCFs and are shown to have a significant impact on the dynamics of the network. At the same time we note that in (Szejka et.al. [27]), the authors offer a somewhat similar study as in this paper for random threshold networks with a particular type of updating rule that does not change the state of a node when the sum of its inputs gives exactly the threshold value. In that paper it is found that threshold values

close to zero can lead to chaos or criticality, while threshold values away from zero lead mostly to order. Criticality is reached for only one small connectivity value. In comparison, PNCFs seem to exhibit a wider range of parameter values that lead to criticality or chaos. The authors of (Szejka et.al. [27]) also explore the potential differences in phase diagrams obtained with annealed networks versus quenched networks; noticeable differences are obtained for integer thresholds. We plan on aggregating PNCFs, NCFs, threshold functions, as well as simple biased functions in a heterogeneous network whose phase diagram would depend on a variety of parameters. Such a model could serve as a way of identifying the effect of changes/mutations within one class of Boolean functions on the overall network dynamics. In particular, using the information in (Kochi and Matache [6]) and (Helikar et.al. [5]) we could identify various types of proteins that could be targeted in drug therapies for a most effective impact on the network behavior.

5. Acknowledgments

This study is partially supported by a 2012-2013 Fund for Investing in the Research Enterprise (FIRE) grant at the University of Nebraska at Omaha.

References

[1] Kauffman S.A., *The origins of order*, Oxford University Press, 1993, 173-235.

[2] Shmulevich I., Kauffman S.A., *Activities and sensitivities in Boolean network models*, Physical Review Letters 93(4), 2004, 048701.

[3] Shmulevich I., Dougherty E.R., Zhang W., *From Boolean to probabilistic Boolean networks as models for genetic regulatory networks*, Proceedings of the IEEE 90(11), 2002, 1778-1792.

[4] Shmulevich I., Lähdesmäki H., Dougherty E.R., Astola J., Zhang W., *The role of certain Post classes in Boolean network models of genetic networks*, Proceedings of the National Academy of Sciences of the United States of America 100(19), 2003, 10734-10739.

[5] Helikar T., Konvalina J., Heidel J., Rogers J.A., *Emergent decision-making in biological signal transduction networks*, Proceedings of the National Academy of Sciences of the United States of America 105(6), 2008, 1913-1918.

[6] Kochi N., Matache M.T., *Mean-Field Boolean Network Model of a Signal Transduction Network*, Biosystems 108(1-3), 2012, 14-27.

[7] Klemm K., Bornholdt S., *Stable and unstable attractors in Boolean networks*, Phys. Rev. E 72, 2000, 055101.

[8] Raeymaekers L., *Dynamics of Boolean networks controlled by biologically meaningful functions*, Journal of Theoretical Biology 218, 2002, 331-341.

[9] Hegslelmann R., Flache R. , *Understanding Complex Social Dynamics: A Plea For Cellular Automata Based Modelling*, Journal of Artificial Societies and Social Simulation 1(3), 1998.

[10] Green D.G., Leisman T.G., Sadedin S., *The Emergence of Social consensus in Boolean Networks*, Proceedings of IEEE Symposium on Artificial Life, Honolulu, Hawaii, 2007.

[11] Moreira A., Mathur A., Diermeier D., Amaral L., Karp, R., *Efficient System-Wide Coordination in Noisy Environments*, Proceedings of the National Academy of Sciences of the United States of America 101(33), 2004, 12085-12090.

[12] Jumadinova J., Matache M.T., Dasgupta P., *A Multi-Agent Prediction Market Based on Boolean Network Evolution*, Proceedings of the 2011 IEEE/WIC/ACM International Conference on Web Intelligence and Intelligent Agent Technology (WI-IAT), 2011, 171-179.

[13] Huepe C., Aldana-González M., *Dynamical phase transition in a neural network model with noise: an exact solution*, Journal of Statistical Physics 108(3-4), 2002, 527-540.

[14] Wolfram S., *A new kind of science*, Wolfram Media, Champaign, 2002.

[15] Kauffman S., Peterson C., Samuelsson B., Troein C., *Genetic networks with canalyzing Boolean rules are always stable*, Proceedings of the National Academy of Sciences of the United States of America 101(49), 2004, 17102-17107.

[16] Nikolajewa S., Friedel M., Wilhelm T., *Boolean networks with biologically relevant rules show ordered behavior*, Biosystems 90(1), 2007, 40-47.

[17] Rämö P., Kesseli J., Yli-Harja O., *Stability of functions in Boolean models of gene regulatory networks*, Chaos 15(3), 2005, 34101.

[18] Just W., Shmulevich I., Konvalina J., *The number and probability of canalizing functions*, Physica D 197, 2004, 211-221.

[19] Peixoto T.P., *The phase diagram of random Boolean networks with nested canalizing functions*, The European Physical Journal B 78(2), 2010, 187-192.

[20] Harris S.E., Sawhill B.K., Wuensche A., Kauffman S., *A model of transcriptional regulatory networks based on biases in the observed regulation rules*, Complexity 7(4), 2002, 23-40.

[21] Layne L., Dimitrova E., and Macauley M., *Nested Canalyzing Depth and Network Stability*, Bulletin of Mathematical Biology 74(2), 2012, 422-433.

[22] Moreira A.A., Amaral L.A.N., *Canalizing Kauffman networks: nonergodicity and its effect on their critical behavior*, Physical Review Letters 94, 2005, 218702.

[23] Kauffman S., Peterson C., Samuelsson B., Troein C., *Random Boolean network models and the yeast transcriptional network*, Proceedings of the National Academy of Sciences of the United States of America 100(25), 2003, 14796-14799.

[24] Anthony M., *Accuracy of Classification by Iterative Linear Thresholding*, Proceedings of the Workshop on Discrete Mathematics and Data Mining, 3rd SIAM International Conference on Data Mining, San Francisco, 2003.

[25] Beck G.L., Matache M.T., *Dynamical behavior and influence of stochastic noise on certain generalized Boolean networks*, Physica A 387(19-20), 2008, 4947-4958.

[26] Derrida B., Pomeau Y., *Random networks of automata: a simple annealed approximation*, Europhys. Lett. 1, 1986, 45-49.

[27] Szejka A., Mihaljev T., Drossel B., *The phase diagram of random threshold networks*, New J. Phys. 10, 2008, 063009.