ANALYZING DYNAMICS OF BOOLEAN NETWORKS WITH CANALYZING FUNCTIONS USING SPECTRAL METHODS

Juha Kesseli, Pauli Rämö and Olli Yli-Harja

Institute of Signal Processing, Tampere University of Technology, P.O.Box 553, 33101 Tampere, Finland
kesseli@cs.tut.fi

ABSTRACT

In this paper we apply spectral methods to the problem of analyzing dynamics in Boolean networks with canalyzing functions. The structure of the Walsh spectrum of these functions is studied and formulas for expected values of spectral parameters are computed for selected distributions of Boolean functions. Using these results we derive the expected Derrida plots for networks generated with the chosen distributions of canalyzing functions. In addition, parameter values at which transition from ordered into chaotic dynamics occurs according to the annealed approximation are shown for networks with a fixed number of input variables for each node and for ones with scale-free topology.

1. INTRODUCTION

Boolean networks are deterministic dynamic systems that have been used e.g. to study the phenomena of phase transition between order and chaos [1]. The goal of many Boolean network studies is to find interesting distributions of functions and to find out the behavior of networks generated with these functions. Some characteristics of the dynamical behavior of networks are e.g. the average number of attractors and the distribution of their lengths, the length of transients and the propagation of small perturbations in the state of the network. Because of the difficulty of studying any particular large complex networks directly the studies are typically concentrated on properties of large networks that have been generated randomly with some fixed distributions and parameters. The connections, in particular, are typically selected so that for each input of a function each node is just as likely to be connected to that input as any other.

The first function classes studied [2] are functions in which the outputs for each input state are selected randomly and independently and the number of input variables is the same fixed value \( K \) for each node in the network. In studies of random networks like these the annealed approximation was presented as a method to determine the chaoticity of networks with different values of parameters for the number of input variables and for the probability of one in the function output [3, 4].

In the annealed approximation it is assumed that the behavior of perturbations in the network can be approximated by reducing the information about the original and the perturbed state to only the distance between the states. The approximation is then performed by using a mapping describing the average distance at timestep \( t + 1 \) for two states at a given distance at time \( t \). This mapping is dependent on the choice, i.e. the distribution, of functions in the network. Even though this approximation neglects correlation effects that can be observed in networks in general, the method can be seen to give correct results for the parameter values at which the phase transition occurs in the case of these random functions.

The mapping of distances from timestep \( t \) to \( t + 1 \) is called the Derrida plot and by finding the fixed point of this mapping the predicted order or chaos of the network can be determined. If the fixed point is at zero all perturbations die out eventually according to the approximation. If this does not occur, the network is chaotic. A simple indicator of the chaoticity of a network can be based on the slope of the Derrida plot at the origin. The slope is equivalent to what is called the average sensitivity of functions [5]. The equivalent of a Derrida plot for single functions is also called the generalized sensitivity [6].

Canalyzing functions are a class of functions that are seen to create order in Boolean networks [7]. Since data collected from papers describing real genetic regulatory networks [8] suggests that canalyzing functions are common in nature these functions are of considerable interest in studies focusing on large-scale properties of complex regulatory networks. A subclass of canalyzing functions, nested canalyzing functions, has also been studied [9]. Importantly for Boolean network studies a function can easily be recognized as canalyzing [10] and functions that have exactly a given number of canalyzing variables can be generated [11].

Harmonic analysis has been used in the study of binary circuits for decades. Transforms of Boolean functions have been of particular interest in e.g. applications of spectral fault detection [12]. Transforms are used in studies of complexity classes of functions [6].

Scale-free networks are a topological class of networks found in numerous places in nature and in human affairs. There have been numerous studies on the mechanisms giving rise to a scale-free network and the properties of networks with a scale-free topology [13]. Scale-free Boolean
networks in particular have been studied in e.g. [14].

In this paper we study the spectrum of canalyzing functions and show how their Derrida plots can be written based on the spectrum. In addition, we study the predictions that the annealed approximation makes on the parameter values at which ordered and chaotic dynamics are obtained. Section 2 contains the basic definitions needed for analysis of Boolean functions, canalyzing and nested canalyzing functions and Derrida plots, a simple measure of chaoticity. Some results presented in [15] that are utilized later are stated. Section 3 has formulas that show the structure of the spectrum of canalyzing and nested canalyzing functions. In Section 4 these formulas are used to compute the expected Derrida plot of a Boolean network with selected distributions of canalyzing functions. These results are also interpreted as predictions of the transition between order and chaos as the parameter values are changed. Section 5 contains discussion on the topics of this paper and further work to be done in this direction.

2. DEFINITIONS

2.1. Analysis of Boolean functions

Denote \( B = \{0, 1\} \). The set of all real functions on the hypercube \( B^K \), \( \mathcal{F} = \{ f : B^K \to \mathbb{R} \} \), is a \( 2^K \)-dimensional real vector space with an inner product defined by

\[
\langle f, g \rangle = \frac{1}{2^K} \sum_{x \in B^K} f(x)g(x).
\]

Denote the \( i \)-th component of vector \( w \) by \( w_i \) and for each \( w \in B^K \) let \( W(w) = \{ i \in \{1, 2, ..., K\} \mid w_i = 1 \} \).

Denote \( |w| = \sum w_i \). For each \( w \in B^K \) a Fourier transform kernel function \( Q_w : B^K \to \{-1, 1\} \) is defined as the parity function over the corresponding subset \( W(w) \) of variables:

\[
Q_w(x) = (-1)^{w^T x} = (-1)^{\sum_{i \in W(w)} x_i},
\]

\( \{Q_w\}_{w \in B^K} \) is an orthonormal basis for \( \mathcal{F} \) [6]. Let \( f : B^K \to \mathbb{B} \). The abstract Fourier transform (in this context also the name Walsh transform is used) of Boolean function \( f \) is the rational valued function \( f^* : B^K \to \mathbb{Q} \) which defines the coordinates of \( f \) with respect to the basis \( \{Q_w\}_{w \in B^K} \), i.e.,

\[
f^*(w) = \langle Q_w, f \rangle = \frac{1}{2^K} \sum_{x \in B^K} Q_w(x)f(x).
\]

\( f \) can then be reconstructed from the coefficients as

\[
f(x) = \sum_w f^*(w)Q_w(x).
\]

We define the Fourier spectrum of function \( f : B^K \to \mathbb{B} \) as

\[
a_f(i) = 4 \sum_{|w|=i} f^*(w)^2, i = 1, \ldots, K.
\]

The fast Walsh transform can be used to calculate the spectral coefficients and thus also the Fourier spectrum efficiently [16].

2.2. Canalyzing and nested canalyzing functions

A canalyzing function is defined as a function \( f : B^K \to \mathbb{B} \) such that there is at least one input variable \( k \) with the property that \( \exists \exists x \in B^K : x_k = i \implies f(x) = j. i \) is called the canalyzing value of input \( k \) of \( f \) and \( j \) the canalyzing value of \( f \). Note that in a function there can be inputs with different canalyzing values but there can not be two variables that would canalyze the output to a different value. The identity function has both values zero and one as canalyzing values. We denote by \( F_{f,i} \) the set of all input variables of \( f \) with canalyzing value \( i \). An arbitrary canalyzing function that has at least one input with canalyzing value \( i \) can then be written as

\[
f(x) = \begin{cases} j, & \exists k \in F_{f,i} : x_k = i \\ f_R(x_R), & \text{otherwise}, \end{cases}
\]

where \( x_R \) contains the values of inputs corresponding to inputs of \( f \) that are not canalyzing with canalyzing value \( i \) and \( f_R : B^{K-\{F_{f,i}\}} \to \mathbb{B} \) is what we call the reduced function of \( f \). Nested canalyzing functions are functions \( f \) in which this reduction can be repeated, i.e. \( f_R \) is a canalyzing function itself and so on.

A nested canalyzing function can be described as a sequence of decisions based on individual bits of the input. Chain functions are a special case of nested canalyzing functions presented in [17]. In chain functions the canalyzing value is always zero.

2.3. Derrida plots

In this section we define Derrida plots and state some results presented in [15] that will be used in what follows. Let \( P_{\rho,K} = \{ y \in B^K : |y| = \rho K \} \). For simplicity of notation let \( \binom{N}{K} = 0 \) if \( K < 0 \). The Derrida plot of a Boolean function \( f : B^K \to \mathbb{B} \) is defined as

\[
d_f(\rho) = \frac{1}{2^K (\rho K)} \sum_{x \in B^K} \sum_{y \in P_{\rho,K}} f(x) \oplus f(x \oplus y),
\]

where \( \rho = 0, \frac{1}{K}, \frac{2}{K}, \ldots, \frac{K}{K}. \)

Boolean network with \( N \) nodes can be viewed as a mapping \( F : B^N \to B^N \) in which each node, i.e. variable, is assigned a binary output according to a Boolean function selected for the node. The output value of the variable in question is computed by taking the Boolean function assigned to the variable and computing it with inputs taken from the inputs determined by the directed edges in the graph describing the connections of the network. The network nodes are updated synchronously. The Derrida plot of a Boolean network \( F : B^N \to B^N \) is defined as

\[
d_{F}(\rho) = \frac{1}{2^N (\rho N)} \sum_{x \in B^N} \sum_{y \in P_{\rho,N}} \frac{1}{N} |F(x) \oplus F(x \oplus y)|,
\]

where \( \rho = 0, \frac{1}{N}, \frac{2}{N}, \ldots, \frac{N}{N}. \)

The Derrida plot of Boolean function \( f : B^K \to \mathbb{B} \) with Fourier spectrum \( a_f \) can be written as [6]:

\[
d_f(\rho) = \sum_i a_f(i)p_{i,\rho}(\rho),
\]
where

\[ p_i(K) = \sum_{j=0}^{\lfloor \frac{K-1}{2} \rfloor} \binom{K-i}{2j+1} (\rho K - 2j - 1) \]

is the Derrida plot of a parity function with \( i \) inputs and with added redundant variables so that the total number of variables is \( K \). As redundant variables, i.e. variables not affecting the output in any way, are added to the function so that \( K \to \infty \) the Derrida plot of a parity function with \( i \) input variables approaches [15]

\[ p_i(\rho) = \frac{1}{2} - 2^{i-1}(\frac{1}{2} - \rho)^i \]

as a pointwise limit.

The Derrida plot of a Boolean network \( F : \mathbb{B}^N \to \mathbb{B}^N \) generated randomly from a given distribution of functions \( f \) approaches

\[ d_F(\rho) = \sum_i E[a_f(i)]p_i(\rho), \tag{1} \]

as \( N \to \infty \) [15].

3. SPECTRAL PROPERTIES OF CANALYZING AND NESTED CANALYZING FUNCTIONS

There are four possible types of canalyzing inputs in a function. We denote by \( C_{i,j} \) the set of functions that are canalyzing with canalyzed value \( j \) and that have at least one input with canalyzing value \( i \). For each type of a function a formula holds that can be used to compute the spectrum with ease based on the spectrum of the reduced function \( f_R \). It should be noted that these equations can alternatively be obtained by an application of known results for composition of Boolean functions [6].

Let \( f \) be a canalyzing function with at least one input with canalyzing value \( i \). If the set of canalyzing input variables of \( f \) with canalyzing value \( i \) is denoted by \( F_{f,i} \) we can define

\[ c_{f,i}(x) = \begin{cases} 1, & \text{if } \exists k \in F_{f,i} : x_k = i \\ 0, & \text{otherwise.} \end{cases} \]

That is, \( c_{f,i}(x) \) is one if and only if the output of function \( f \) is canalyzed to its value with canalyzing value \( i \) when the inputs have value \( x \). Also denote by \( w_C \) the bits of \( w \) corresponding to canalyzing inputs and by \( w_R \) the bits corresponding to the inputs of the reduced function. Let \( x_C \) and \( x_R \) denote the values of the variables in question. If the variables are ordered e.g. so that the canalyzing variables are first, this means that we can write \( w = [w_R^T, w_C^T]^T \) and \( x = [x_R^T, x_C^T]^T \).

First, let \( f : \mathbb{B}^K \to \mathbb{B} \in C_{1,1} \).

\[ f^*(w) = \frac{1}{2^K} \sum_x Q_w(x)f(x) \]

\[ = \frac{1}{2^K} \sum_{x: c_{f,i}(x) = 1} (-1)^{w_C^T x_C + w_R^T x_R} f(x) \]

\[ + \frac{1}{2^K} \sum_{x: c_{f,i}(x) = 0} (-1)^{w_C^T x_C + w_R^T x_R} f(x) \]

\[ = \frac{1}{2^K} \sum_{x: c_{f,i}(x) = 1} (-1)^{w_C^T x_C + w_R^T x_R} \delta(|w_R|) \]

\[ + \frac{1}{2^K} \sum_{x: c_{f,i}(x) = 0} Q_w(x)f_R(x) \]

\[ = \frac{1}{2^K} \left[ \{2^{F_{f,i}} \delta(|w_C|) - 1 \} \delta(|w_R|) + f_R(w_R) \right] \]

\[ = \delta(|w|) + \frac{1}{2^{F_{f,i}}} \left( f_R^*(w_R) - \delta(|w_R|) \right), \]

where

\[ \delta(x) = \begin{cases} 1, & \text{if } x = 0 \\ 0, & \text{otherwise.} \end{cases} \]

For \( f : \mathbb{B}^K \to \mathbb{B} \in C_{0,0} \) we have

\[ f^*(w) = \frac{1}{2^K} \sum_x Q_w(x)f(x) \]

\[ = \frac{1}{2^K} \sum_{x: c_{f,i}(x) = 1} (-1)^{w_C^T x_C + w_R^T x_R} f(x) \]

\[ + \frac{1}{2^K} \sum_{x: c_{f,i}(x) = 0} Q_w(x)f_R(x) \]

\[ = \frac{1}{2^K} \left[ \{2^{F_{f,i}} \delta(|w_C|) - 1 \} \delta(|w_R|) + f_R^*(w_R) \right] \]

\[ = \delta(|w|) + \frac{1}{2^{F_{f,i}}} \left( f_R^*(w_R) - \delta(|w_R|) \right). \]
For a canalyzing function with inputs with canalyzing values of both zero and one the spectrum can be computed by applying a suitable formula for both types of canalyzing inputs separately. It can be seen that if the function $f$ is a nested canalyzing function, we can again apply the formulas for the function $f_R$, which is in this case canalyzing as well. The entire spectrum of an arbitrary nested canalyzing function can thus be computed using these formulas recursively.

4. APPLICATIONS

4.1. Derrida plots of networks with canalyzing functions

In this section we will show how the formulas of the previous section can be used to obtain the expected values of the spectrum, $E[a_f(i)]$, for functions chosen from a selected distribution of canalyzing functions. This can, in turn, be used to compute the Derrida plot for networks containing the same distribution of canalyzing functions.

We compute the expected spectrum for canalyzing functions by specifying the canalyzing and canalyzed values of $f: \mathbb{B}^K \rightarrow \mathbb{B}$, the number $|F_f|$ of canalyzing inputs of $f$ with canalyzing value $1$ and the bias parameter $b$. For simplicity we study functions in $C_{1,0}$ in this section and can thus write $F_f = F_{f,i}$. The bias parameter $b$ determines the distribution of the reduced function $f_R: \mathbb{B}^{K-|F_f|} \rightarrow \mathbb{B}$ of $f$ such that for each $x$ the output $f_R(x) \in \mathbb{B}$ is selected randomly with $E[f_R(x)] = b$.

First, we need to know the expected value $E[f_R(w_R)^2]$ for a reduced function chosen from the distribution described. Denoting $K_R = K - |F_f|$ this can be computed [6] from

$$E[f_R(w_R)^2] = E\left[\frac{1}{2^{K_R}} \sum_{x \in \mathbb{B}^{K_R}} f_R(x)Q_{w_R}(x)\right]$$

$$= \frac{1}{2^{K_R}} \sum_{x \in \mathbb{B}^{K_R}} \sum_{y \in \mathbb{B}^{K_R}} Q_{w_R}(x)Q_{w_R}(y)E[f_R(x)f_R(y)].$$

Since $E[f_R(x)] = b$ and $E[f_R(x)f_R(y)] = b^2$ for $x \neq y$, we have

$$E[f_R(w_R)^2] = \frac{1}{2^{K_R}} \left[ (b - b^2) \sum_{x \in \mathbb{B}^{K_R}} Q_{w_R}(x)Q_{w_R}(x) \right] + b^2 \sum_{x \in \mathbb{B}^{K_R}} Q_{w_R}(x) \sum_{y \in \mathbb{B}^{K_R}} Q_{w_R}(y).$$

Since

$$\sum_{x \in \mathbb{B}^{K_R}} Q_{w_R}(x) = 2^{K_R} \delta(|w_R|)$$

and the basis $\{Q_{w_R}\}$ is orthonormal this can be written simply as

$$E[f_R(w_R)^2] = \frac{b(1 - b)}{2^{K_R}} + b^2 \delta(|w_R|).$$

We can now compute the expected squares of spectral coefficients for functions $f: \mathbb{B}^K \rightarrow \mathbb{B} \subset C_{1,0}$. In this case

$$E[f^*(w)^2] = E\left[\left(\frac{1}{2^{K_R}} f_R(w_R)^2\right)^2\right] = \frac{1}{2^{2K_R}} E[f_R(w_R)^2]^2$$

$$= \frac{b(1 - b)}{2^{K_R+|F_f|}} + \frac{b^2 \delta(|w_R|)}{2^{2|F_f|}}.$$

Now for these functions it can be seen that

$$E[a_f(i)] = 4 \sum_{w:|w|=i} E[f^*(w)^2]$$

$$= 4\left( \frac{|F_f|}{K} \right)^2 \frac{b(1 - b)}{2^{K+|F_f|}} + \frac{b^2}{2^{2|F_f|}}.$$

We can now derive the Derrida plot for a Boolean network generated from random functions $f: \mathbb{B}^K \rightarrow \mathbb{B}$ by substituting this into Eq. (1),

$$d_f(\rho) = \sum_i a_f(i)\left(\frac{1}{2} - 2^{-i-1}\left(\frac{1}{2} - \rho\right)^i\right)$$

$$= 4\left( \frac{|F_f|}{K} \right)^2 \sum_i \left( \frac{K}{i} \right) \left(\frac{1}{2} - 2^{-i-1}\left(\frac{1}{2} - \rho\right)^i\right) + 4\frac{b^2}{2^{2|F_f|}} \sum_i \left( \frac{|F_f|}{i} \right) \left(\frac{1}{2} - 2^{-i-1}\left(\frac{1}{2} - \rho\right)^i\right)$$

$$= \frac{1}{2^{2|F_f|}} \left[ (b(1 - b)(1 - (1 - \rho)^K) + b^2(1 - (1 - \rho)^{|F_f|})\right].$$

It can be seen that for symmetry reasons the Derrida plot is the same for functions taken from $C_{1,0}$. For the functions with canalyzed value one, i.e. for functions in $C_{0,1}$ and $C_{1,1}$, $b$ will have to be replaced with $1 - b$ in the above formula. Mixtures of canalyzing functions with different numbers of input variables and canalyzing variables can be studied by weighing the spectra of different types with their probability and then applying Eq. (1).

In Fig. 1 the Derrida plots for networks with functions of three inputs and zero, one and two canalyzing inputs are shown. The Derrida plots obtained by sampling the state space of a network of size $N = 100$, shown in dotted line, are close to the expected ones computed with Eq. (2). $N$ is thus already so large that the approximation error in Eq. (1) can be seen to be of no significance in practice. 100 pairs of points were sampled at each value of $\rho$. 
which networks with canalyzing functions are predicted. An approximation we can find the values of parameters for.

If the slope is less than one, there will be a stable fixed point at zero and that the network is chaotic. If the slope is larger than one, the annealed approximation predicts that the perturbation size has a fixed point greater than zero, and thus the network is chaotic. If the slope is equal to one, we have a fixed point at zero. If the slope is greater than one, the network is chaotic.

In [5] average sensitivities are computed for canalyzing functions selected from a distribution like the one used here, except for the fact that a special case with only one canalyzing variable and \( b = \frac{1}{2} \) is studied. Average sensitivity is a measure of chaoticity of Boolean networks that can be seen to be equivalent to the slope of the Derrida plot at the origin [15]. This enables us to compute the average sensitivities in an alternative way using the preceding formulas. We denote the average sensitivity by \( s_f \).

Since

\[
\frac{d}{d\rho} d_f(\rho) = \frac{1}{2^{|F_f|}-1} \left[ b(1-b)K(1-\rho)^{K-1} + b^2 |F_f|(1-\rho)^{|F_f|-1} \right]
\]

we have

\[
s_f = \left. \frac{d}{d\rho} d_f(\rho) \right|_{\rho=0} = \frac{1}{2^{|F_f|}-1} \left[ b(1-b)K + b^2 |F_f| \right]. \tag{3}
\]

This formula generalizes the study in [5] for canalyzing functions with an arbitrary number of canalyzing variables and an arbitrary \( b \) for the reduced function.

If the slope of the Derrida plot, i.e. the average sensitivity, is larger than one, the annealed approximation predicts that the perturbation size has a fixed point greater than zero and that the network is chaotic. If the slope is less than one, there will be a stable fixed point at zero and the network is ordered. Based on this use of annealed approximation we can find the values of parameters for which networks with canalyzing functions are predicted to be chaotic.

Figure 2 shows the predicted phase transition between order and chaos as curves in the parameter space for networks with functions taken from \( C_{0,0} \) and \( C_{1,0} \) with the same number of inputs for all the nodes in the network. Each curve corresponds to networks with a given constant number of canalyzing inputs. In studying the figure it should be noted that the parameter \( b \) on the horizontal axis is the bias parameter of the reduced function. This means that for different numbers of canalyzing inputs \(|F_f|\) the expected bias \( E[\sum x f(x)] \) is different at same \( b \). As canalyzing variables are added the amount of order is increased if the expected bias of the functions is kept constant.

In order to study scale-free networks we choose the in-degrees of nodes according to the distribution

\[
P(K) = \frac{1}{\zeta(\gamma)} K^{-\gamma},
\]

where

\[
\zeta(\gamma) = \sum_{K=1}^{\infty} K^{-\gamma}.
\]

\( \zeta(\gamma) \) is the Riemann zeta function and \( \gamma \) is the distribution parameter. The slope of the Derrida plot can now be computed using Eq. (3),

\[
s_f = \frac{1}{\zeta(\gamma)} \sum_{K=1}^{\infty} K^{-\gamma} \left[ \frac{1}{2^{|F_f|-1}} (b(1-b)K + b^2 |F_f|) \right] = \frac{b^2 |F_f|}{2^{|F_f|-1}} + \frac{b(1-b)}{2^{|F_f|-1}} \frac{\zeta(\gamma-1)}{\zeta(\gamma)}.
\]

By setting \( s_f = 1 \), we can again obtain curves of transition from order to chaos in the parameter space. These curves are shown in Fig. 3 for different numbers of canalyzing input variables. This time the functions are selected with canalyzing value one so that \( b \) in the above formula is replaced with \( 1-b \). The previous comments on parameter \( b \) apply to this figure as well.

5. DISCUSSION

The formulas and methods presented make studies on properties of canalyzing functions easier by enabling simple results on Derrida plots to be obtained and giving theoretical tools for further work. Not only Derrida plots but also other characteristics of the functions can be studied using the spectrum. The Derrida plots that can be drawn using the results presented can be used as a first approximation of the behavior of Boolean networks with canalyzing functions.

As has already been noted in the literature [7], this approximation leaves a lot to be desired. In particular, networks constructed e.g. with only functions from \( C_{0,0} \) are much more ordered than would be predicted by the annealed approximation as it is used here. The so-called forcing effect is significant in freezing nodes in these networks and is not taken into account in this study.

The results obtained here are thus only applicable as measuring chaoticity properties realistically to cases where the forcing effect is not significant. It can be observed by
numerical experiments with randomly generated networks that networks with canalyzing functions in which the forcing effect is not significant are at least ones in which there are either only canalyzing functions from $C_{0,0}$ and $C_{1,0}$ or else there are approximately as many functions from $C_{0,0}$ as there are from $C_{1,1}$.

Since genetic regulatory networks seem to have a distribution in which there are lots of functions with canalyzed value 0 and inputs with canalyzing value 0, this means that in the main application of Boolean networks the results presented here are not applicable as such. Further work should thus concentrate on extending the current techniques so that better measures of chaoticity can be achieved.

6. REFERENCES


