The Influence of Canalization on the Robustness of Boolean Networks

C. Kadelka[a,b,1,*], J. Kuipers[c], R. Laubenbacher[d,e,1]

[a] *Institute of Medical Virology, University of Zurich, 8006 Zurich, Switzerland*
[b] *Division of Infectious Diseases and Hospital Epidemiology, University Hospital Zurich, 8091 Zurich, Switzerland*
[c] *D-BSSE, ETH Zurich, Mattenstrasse 26, 4058 Basel, Switzerland*
[d] *Center for Quantitative Medicine, University of Connecticut Health Center, Farmington, CT 06030, USA*
[e] *Jackson Laboratory for Genomic Medicine, Farmington, CT 06030, USA*

Abstract

Time- and state-discrete dynamical systems are frequently used to model molecular networks. This paper provides a collection of mathematical and computational tools for the study of robustness in Boolean network models. The focus is on networks governed by k-canalizing functions, a recently introduced class of Boolean functions that contains the well-studied class of nested canalizing functions. The activities and sensitivity of a function quantify the impact of input changes on the function output. This paper generalizes the latter concept to c-sensitivity and provides formulas for the activities and c-sensitivity of general k-canalizing functions as well as canalizing functions with more precisely defined structure. A popular measure for the robustness of a network, the Derrida value, can be expressed as a weighted sum of the c-sensitivities of the governing canalizing functions, and can also be calculated for a stochastic extension of Boolean networks. These findings provide a computationally efficient way to obtain Derrida values of Boolean networks, deterministic or stochastic, that does not involve simulation.

Keywords: k-canalizing function, Derrida value, Boolean network, nested canalizing function, stability

1. Introduction

The robustness of dynamic networks has long been an important topic of investigation in a wide range of contexts, using various definitions of the concept [1, 2]. Due to the important role of stochasticity in the dynamic behavior of biological networks, in particular gene regulatory networks, the concept of robustness has been studied extensively in this context [3]. Since the introduction of Boolean and logical network models to the study of the properties of gene regulatory networks [4, 5], time- and state-discrete models, referred to subsequently as discrete dynamical systems, have become an increasingly popular representation of molecular networks [6, 7, 8]. For the most part, these consist of Boolean networks and various generalizations thereof. Questions regarding the robustness of molecular networks, modeled in the discrete dynamical systems framework, frequently involve the relationship between structural features of the network and its resulting dynamics. One commonly used measure of the robustness of a discrete dynamic network is the so-called Derrida value of the network, a measure of how perturbations propagate through the network [9]. This

*Corresponding author
Email addresses: kadelka.claus@virology.uzh.ch (C. Kadelka), jack.kuipers@bsse.ethz.ch (J. Kuipers), laubenbacher@uchc.edu (R. Laubenbacher)
[1]Supported by NSF Grant CMMI-0908201 and US DoD Grant W911NF-14-1-0486

measure can then be related to network structure, such as the type of logical rules used to represent the regulatory mechanisms for individual network nodes; see, e.g., [10].

A frequently investigated concept related to robustness is that of *canalization* in developmental biology, introduced by Waddington in the 1940s [11]. It was intended to account for the absence of a known mechanism that enables the genetic regulatory protocols driving embryonal development to accurately produce a specific phenotype, even in light of substantial variation in the developing organism's environment. The underlying idea is that phenotypes can be thought of as valleys in a landscape, and canalization is a "force" that channels the developmental trajectory accurately into a particular valley, protecting it from perturbations. Kauffman introduced a version of this concept to Boolean network modeling of gene regulatory networks by studying canalizing functions [12], as well as the special subclass of so-called nested canalizing functions [13]. Several authors recently extended this work by considering canalization as a property of Boolean functions [14, 15]. Generalizing findings from [16], they showed that every Boolean function has a unique algebraic form, characterized by three invariants: canalizing depth, dominance layer numbers and the non-canalizing core-polynomial. The canalizing depth of a function describes its degree of canalization and a k-canalizing function is defined to have canalizing depth of at least k. This extension neatly stratifies the set of all Boolean functions on n variables by their canalizing depth and dominance layer numbers, including so-called nested canalizing functions (n-canalizing), canalizing functions (1-canalizing) and non-canalizing functions (0-canalizing).

A *canalizing* function possesses at least one input variable such that, if this variable takes on a certain "canalizing" value, then the output value is already determined, regardless of the values of the remaining input variables. If this variable takes on another value, and then there is a second variable with this same property, and so on. If k variables follow this pattern, the function is *k-canalizing*, and if all variables follow this pattern, the function is *nested canalizing* (NCF). By definition, any $(k+1)$-canalizing function is also k-canalizing, and the Boolean AND function is an example of an NCF.

The relationship between network stability, frequently measured using Derrida values, and the proportion of canalizing functions and their degree of canalization has received much attention in recent years. Boolean networks governed by canalizing functions are more stable than those constructed using random functions [13, 17, 18]. In general, network stability and the degree of canalization are positively correlated, with networks governed by NCFs exhibiting the most stable dynamics [19, 20, 21, 22]. These findings clearly motivate the study of k-canalizing functions in the context of understanding the regulatory logic of gene networks.

In this paper, we introduce a closed formula for the efficient computation of the Derrida values of a network governed by k-canalizing functions. In detail, the paper is ordered as follows. In Section 2, we formally introduce the computational concept of canalization. To make this paper self-contained, we restate some frequently used definitions and remarks, see, e.g., [13, 15, 16]. The activity of a variable in a function quantifies the influence of that variable on the whole function, while the average sensitivity of a function measures how sensitive a function is to a random input change and can be expressed as a weighted sum of the activities. Both quantities have been extensively studied by the Boolean modeling community, see, e.g., [14, 23, 24, 25]. In Section 3, we derive formulas for the expected activities of any k-canalizing function as well as any canalizing function with known dominance layer numbers. We also introduce the c-sensitivity of a function as a natural generalization of the sensitivity and show how to calculate the average c-sensitivity of a canalizing function from the activities of its variables. In Section 4, we use the normalized average c-sensitivities to derive a formula for the Derrida values of a Boolean network that is governed by k-canalizing function. This greatly simplifies the application of this tool for robustness analyses of networks, which otherwise requires extensive simulations (difficult or infeasible for large networks).

We explore the relationship between the Derrida value, the canalizing depth, the dominance layer numbers and the non-canalizing core-polynomial of a Boolean function. In Section 5, we extend the formula for the Derrida values to the case of stochastic networks, which are an important modeling paradigm for gene regulatory networks. Section 6 concludes this paper with some remarks and possible avenues of future work.

2. The Concept of Canalization

In this section we review some well-known concepts and definitions, mainly from [15], to introduce the computational concept of *canalization*.

Definition 2.1. *A Boolean function $f(x_1, x_2, \ldots, x_n)$ is essential in the variable x_i if there exist $r, s \in \mathbb{F}_2$ and $(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n) \in \mathbb{F}_2^{n-1}$ such that*

$$f(x_1, \ldots, x_{i-1}, r, x_{i+1}, \ldots, x_n) \neq f(x_1, \ldots, x_{i-1}, s, x_{i+1}, \ldots, x_n).$$

Definition 2.2. *A Boolean function $f : \mathbb{F}_2^n \to \mathbb{F}_2$ is canalizing if there exist a variable x_i, a Boolean function $g(x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n)$ and $a, b \in \mathbb{F}_2$ such that*

$$f(x_1, \ldots, x_n) = \begin{cases} b & \text{if } x_i = a \\ g \not\equiv b & \text{if } x_i \neq a, \end{cases}$$

in which case x_i is called a canalizing variable, the input a is the canalizing input, and the output value b when $x_i = a$ is the corresponding canalized output.

Definition 2.3. *A Boolean function $f(x_1, \ldots, x_n)$ is k-canalizing, where $0 \leq k \leq n$, with respect to the permutation $\sigma \in \mathcal{S}_n$, inputs $a_1, \ldots, a_k$ and outputs $b_1, \ldots, b_k$, if*

$$f(x_1, \ldots, x_n) = \begin{cases} b_1 & x_{\sigma(1)} = a_1, \\ b_2 & x_{\sigma(1)} \neq a_1, x_{\sigma(2)} = a_2, \\ b_3 & x_{\sigma(1)} \neq a_1, x_{\sigma(2)} \neq a_2, x_{\sigma(3)} = a_3, \\ \vdots & \vdots \\ b_k & x_{\sigma(1)} \neq a_1, \ldots, x_{\sigma(k-1)} \neq a_{k-1}, x_{\sigma(k)} = a_k, \\ g \not\equiv b_k & x_{\sigma(1)} \neq a_1, \ldots, x_{\sigma(k-1)} \neq a_{k-1}, x_{\sigma(k)} \neq a_k, \end{cases}$$

where $g = g(x_{\sigma(k+1)}, \ldots, x_{\sigma(n)})$ is a Boolean function on $n-k$ variables. When g is not a canalizing function itself, the integer k is the canalizing depth of f (as in [14]), and if g is in addition not constant, it is called the core function of f, denoted by f_C.

Remark 2.4. *Since $g \not\equiv b_k$, a function that is k-canalizing with respect to $\sigma \in \mathcal{S}_n$, inputs a_i and outputs b_i is essential in each $x_{\sigma(i)}$, $1 \leq i \leq k$.*

Remark 2.5. *If we consider the set of all Boolean functions on n variables, then*

(a) *The n-canalizing functions are precisely the well-studied nested canalizing functions.*
(b) *The 1-canalizing functions are precisely the canalizing functions.*
(c) *Every Boolean function is 0-canalizing.*
(d) *Every non-canalizing function has canalizing depth 0.*

3. Activities and Normalized Average Sensitivities

Some variables of a Boolean function have a greater influence over the output of the function than others. The activity of variable x_i in the function $f(x_1, \ldots, x_n)$ is defined as

$$\alpha_i^f = \frac{1}{2^n} \sum_{\mathbf{x} \in \{0,1\}^n} \chi[f(\mathbf{x}) \neq f(\mathbf{x} \oplus e_i)],$$

where χ is an indicator function, $\oplus$ is addition modulo 2 and e_i is the ith unit vector. A change in a highly active variable frequently affects the function f, while a change in a variable with activity 0 never alters f.

Another important quantity, directly related to the activity of a variable and introduced in [26], measures how sensitive the output of a function is to input changes. The sensitivity of a function f on a vector $\mathbf{x}$ is defined as the number of Hamming neighbors of $\mathbf{x}$ (vectors that differ from $\mathbf{x}$ in exactly one bit) with a different function value than $f(\mathbf{x})$. That is,

$$S^f(\mathbf{x}) = \sum_{i=1}^{n} \chi[f(\mathbf{x}) \neq f(\mathbf{x} \oplus e_i)].$$

The average sensitivity S^f is the expected value of $S^f(\mathbf{x})$ under the distribution of $\mathbf{x}$. Under the uniform distribution,

$$S^f = \mathbb{E}[S^f(\mathbf{x})] = \frac{1}{2^n} \sum_{\mathbf{x} \in \{0,1\}^n} \sum_{i=1}^{n} \chi[f(\mathbf{x}) \neq f(\mathbf{x} \oplus e_i)] = \sum_{i=1}^{n} \alpha_i^f$$

In this section, we will generalize the concept of sensitivity to c-sensitivity and calculate the normalized average c-sensitivity for different classes of canalizing functions, which we will then use in the next section for the calculation of stability properties of Boolean networks.

Definition 3.1. *Any vector that differs at exactly c bits from a given vector $\mathbf{x}$ is called a c-Hamming neighbor of $\mathbf{x}$. The c-sensitivity of $f(x_1, \ldots, x_n)$ on $\mathbf{x}$ is defined as the number of c-Hamming neighbors of $\mathbf{x}$ on which the function value is different from its value on $\mathbf{x}$. That is,*

$$S_c^f(\mathbf{x}) = \sum_{\substack{I \subseteq \{1,2,\ldots,n\} \\ |I|=c}} \chi[f(\mathbf{x}) \neq f(\mathbf{x} \oplus e_I)],$$

where χ is an indicator function, $\oplus$ is addition modulo 2 and e_I is a vector with 1 at all indices in I and 0 everywhere else. Assuming a uniform distribution of $\mathbf{x}$,

$$S_c^f = \mathbb{E}[S_c^f(\mathbf{x})] = \frac{1}{2^n} \sum_{\mathbf{x} \in \{0,1\}^n} \sum_{\substack{I \subseteq \{1,2,\ldots,n\} \\ |I|=c}} \chi[f(\mathbf{x}) \neq f(\mathbf{x} \oplus e_I)]$$

is the average c-sensitivity of f. The range of S_c^f is $[0, \binom{n}{c}]$. Let us therefore define the normalized average c-sensitivity of f as

$$q_c^f = \frac{S_c^f}{\binom{n}{c}} \in [0, 1].$$

This definition generalizes the concept of sensitivity in a natural way ($S_1^f = S^f$) and allows the impact of a simultaneous change in more than one input of a function to be studied. We will now derive the expected activities of a k-canalizing function, where the expectation is taken over all k-canalizing function and a uniform distribution is assumed.

Theorem 3.2. *Let f be a k-canalizing function of n variables. By relabeling the variables if necessary, assume that f is k-canalizing in the variable order $x_1, x_2, \ldots, x_k$. The expected activity of x_j in f is*

$$\mathbb{E}[\alpha_j^f] = \begin{cases} \frac{1}{2^j} & \text{if } j < k \\ \frac{1}{2^{k-1}} \frac{2^{2^{n-k}-1}}{2^{2^{n-k}}-1} & \text{if } j = k \\ \frac{1}{2^k} \frac{2^{2^{n-k}-1}}{2^{2^{n-k}}-1} & \text{if } j > k. \end{cases}$$

Proof. See Appendix. □

The expected average c-sensitivity of any k-canalizing function is a weighted sum of the activities of its variables.

Theorem 3.3. *By relabeling the variables if needed, assume that $f(x_1, \ldots, x_n)$ is a k-canalizing function in the variable order $x_1, x_2, \ldots, x_k$. The average c-sensitivity of f is*

$$S_c^f = \sum_{j=1}^{n-c+1} \binom{n-j}{c-1} \alpha_j^f.$$

Proof. See Appendix. □

Corollary 3.4. *The expected activities of the variables $(x_{\sigma(1)}, x_{\sigma(2)}, \ldots, x_{\sigma(n)})$ of an NCF f are*

$$\mathbb{E}[\alpha^f] = \left(\frac{1}{2}, \frac{1}{4}, \ldots, \frac{1}{2^{n-2}}, \frac{1}{2^{n-1}}, \frac{1}{2^{n-1}}\right),$$

and the expected normalized average c-sensitivity is

$$\mathbb{E}[q_c^f] = \frac{c}{2n} {}_2F_1\left[1, c-n; 1-n; \frac{1}{2}\right] = \begin{cases} \frac{1}{n} & \text{if } c = 1 \\ \frac{1}{\binom{n}{c}} \sum_{j=1}^{n-c+1} \binom{n-j}{c-1} \left(\frac{1}{2}\right)^j & \text{if } 1 < c \leq n \end{cases}$$

with ${}_2F_1$ the hypergeometric function.

Theorem 4.6 of [15] shows that any Boolean function can be written in a unique standard monomial form, in which the variables are grouped into different layers based on their dominance (see also [27, 16] for earlier work on this topic for NCFs). Any canalizing variable is part of the first layer. All variables that become canalizing once the variables from the first layer are excluded, are part of the second layer, etc. The number of layers is called the *layer number*, denoted by r. The number of variables in the ith layer is the *dominance number of layer i*, denoted by k_i, and the number of all variables that become eventually canalizing is the *canalizing depth* $k = \sum k_i$. All remaining variables that never become canalizing are part of the *core polynomial*, which is simply an affine transformation of the core function.

Theorem 3.2 yields the expected activities of a function, for which only its minimal canalizing depth is known. If the exact canalizing depth of a function and its dominance layer numbers are known, we can quantify the dynamical properties much more accurately. In [16], the authors have already

computed the activities and average sensitivity of an NCF with known dominance layer numbers. We will now determine the activities of a Boolean function, for which the exact canalizing depth, the dominance layer numbers and the Hamming weight of its core function are known. In this case, we do not require an expected value since all such functions share the same activities.

Theorem 3.5. *Let $f(x_1,\dots,x_n)$ be any Boolean function with canalizing depth $k \geq 0$, r layers of canalization with layer structure $\{k_1,\dots,k_r\}$, where $k_i \geq 1, \sum_{i=1}^{r} k_i = k$, and $v \in \{1,\dots,2^{n-k}\}$ entries $\neq b_k$ in the truth table of its non-canalizing core g. By relabeling the variables if necessary, assume that f is canalizing in the variable order $x_1, x_2, \dots, x_k$. The activity of x_j on f is*

$$\alpha_j^f = \begin{cases} \varphi_{L(j)} + \frac{1}{2^{k-1}}\psi_{L(j)} & \text{if } j \leq k \\ \frac{v(2^{n-k}-v)}{2^{n-1}(2^{n-k}-1)} & \text{if } j > k, \end{cases}$$

where $L(j) \in \{1,\dots,r\}$ denotes the dominance layer number of variable x_j and

$$\varphi_{r+1} = \varphi_r = 0, \varphi_i = \varphi_{i+2} + \sum_{s=0}^{k_{i+1}-1} \left(\frac{1}{2}\right)^{k_1+\dots+k_i+s} \text{ for } i \geq 1$$

$$\psi_r = \frac{v}{2^{n-k}}, \psi_i = 1 - \psi_{i+1} \text{ for } i \geq 1$$

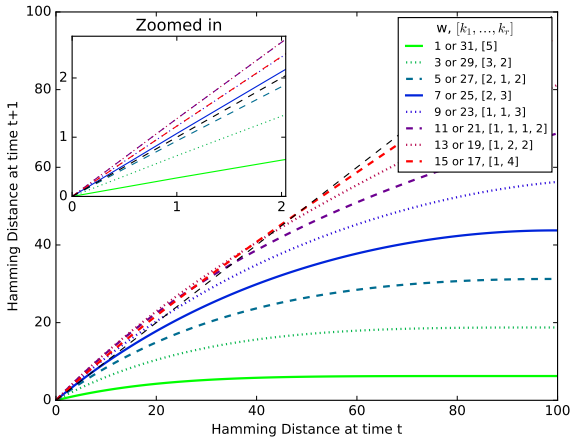
Proof. See Appendix. □

4. Derrida Values of Networks Governed by *k*-canalizing functions

Gene regulatory networks need to be robust to perturbations. The so-called Derrida plot is a common technique to evaluate the robustness of a Boolean discrete dynamical system [9]. It describes how a perturbation of a given size propagates on average over time. If a small perturbation vanishes over time, the system is considered to be in the ordered regime. The network then typically consists of many steady states and short limit cycles. If the perturbation amplifies over time, the system is in the chaotic regime. A chaotic network typically possesses long limit cycles. Lastly, if the perturbation remains of similar size, then the system is in the narrow threshold between these regimes, often called the critical threshold. Many biological systems seem to operate at this "edge of chaos"; they must be robust enough to withstand perturbations caused by environmental changes but also flexible enough to allow adaptation [10, 28].
In this section we formally define the concept of Derrida values in the framework of discrete dynamical systems, using an annealed approximation. Although this approximation corresponds to a system in which interactions are randomly rewired at each time step, it has been shown that its use does not alter the Derrida plot of random Boolean networks [29]. Until now, the calculation of Derrida values has required extensive Monte Carlo simulations [13, 17]. We derive direct formulas for the Derrida values of Boolean networks. Especially for systems with many regulators, this offers a substantial improvement, since the time required to approximate the Derrida plot through simulations increases exponentially in the number of regulators.

Definition 4.1. *Let $F = (f_i)_{i=1}^N$ be a synchronous Boolean network of N nodes. Let $I(f_i) \subseteq \{1,\dots,N\}$ be the set of variables of f_i. Moreover, let $\mathbf{x}, \mathbf{y} \in \{0,1\}^N$ be two system configurations that differ in m coordinates. Lastly, let $J(f_i) = I(f_i) \cap \{j \mid x_j \neq y_j\} \in \{0,\dots,m\}$ be the set of variables of f_i where $\mathbf{x}$ and $\mathbf{y}$ differ. Then, for an initial perturbation of size m, the Derrida value of F is defined as the average size of the perturbation after one update,*

$$D(F,m) = \mathbb{E}\Big[d\big(F(\mathbf{x}), F(\mathbf{y})\big) \;\Big|\; d(\mathbf{x},\mathbf{y}) = m\Big], \tag{4.1}$$



w	|bias|	r	$[k_1,\ldots,k_r]$	D(F,1)	D(F,2)	D(F,3)
1 or 31	0.9375	1	[5]	0.3125	0.6124	0.9000
3 or 29	0.8125	2	[3, 2]	0.6875	1.3548	2.0024
5 or 27	0.6875	3	[2, 1, 2]	0.9375	1.8523	2.7448
7 or 25	0.5625	2	[2, 3]	1.0625	2.1048	3.1273
9 or 23	0.4375	3	[1, 1, 3]	1.1875	2.3548	3.5023
11 or 21	0.3125	4	[1, 1, 1, 2]	1.3125	2.6023	3.8698
13 or 19	0.1875	3	[1, 2, 2]	1.3125	2.6048	3.8774
15 or 17	0.0625	2	[1, 4]	1.1875	2.3624	3.5250
expected value when r = 1				0.3125	0.6124	0.9000
expected value when r = 2				0.9792	1.9407	2.8849
expected value when r = 3				1.1458	2.2706	3.3748
expected value when r = 4				1.3125	2.6023	3.8698
expected value when r and $[k_1,\ldots,k_r]$ are unknown				1	1.9811	2.9436

Figure 1: Derrida plot for networks of $N = 100$ genes governed by NCFs with $n = 5$ regulators and varying layer structure. A black dashed line shows the line $y = x$. The table shows the Derrida values for perturbations of size up to 3 (first 8 rows), as well as average values for cases when only the layer number but not the exact layer structure is known (rows 9-12) and average values when both are unknown (last row).

where $d : \{0,1\}^N \times \{0,1\}^N \to \{0,1,\ldots,N\}$ *is the Hamming distance (the standard* ℓ^1 *metric) and the expected value is taken uniformly over all pairs of configurations with distance* m*.*

Theorem 4.2. *The Derrida value of a synchronous Boolean network* $F = (f_i)_{i=1}^N$ *can be expressed as a weighted sum of the normalized average* c*-sensitivities of its update functions,*

$$D(F,m) = \sum_{i=1}^{N} \mathbb{P}\left(f_i(\mathbf{x}) \neq f_i(\mathbf{y}) \;\middle|\; d(\mathbf{x},\mathbf{y}) = m\right) = \sum_{i=1}^{N}\sum_{c=0}^{m} \mathbb{P}\left(|J(f_i)| = c\right) q_c^{f_i},$$

where $\left|J(f_i)\right|$ *follows a hypergeometric distribution,*

$$\mathbb{P}\big(|J(f_i)| = c\big) = \frac{\binom{m}{c}\binom{N-m}{n_i-c}}{\binom{N}{n_i}} = \frac{\binom{n_i}{c}\binom{N-n_i}{m-c}}{\binom{N}{m}}.$$

Proof. Since $\mathbf{x}$ and $\mathbf{y}$ are synchronously updated, the update of each component is independent from the update of other components. This implies that the Derrida value is simply the sum of the probabilities that $f_i(\mathbf{x})$ and $f_i(\mathbf{y})$ differ after the update, $i = 1,\ldots,N$, which equals the normalized average c-sensitivity of f_i. Conditioning with respect to $|J(f_i)|$ leads to the second equality. $J(f_i)$ is the intersection of two sets so that its magnitude $|J(f_i)|$ follows a hypergeometric distribution. □

We will now use this theorem together with the results from Section 3 to calculate average Derrida values for a multitude of different Boolean networks.
The Hamming weight w of a Boolean function is defined as the number of 1s in its the truth table, and the bias is the probability that a randomly chosen entry in the truth table is 1 minus the probability that it is 0. A Boolean function with equally many 0s and 1s has bias 0 and is called balanced. Constant functions are the most biased with absolute bias 1, and it is easy to see that there is a 1-1 correspondence between the absolute bias and the layer structure of an NCF (see also [16]). Figure 1 depicts the Derrida values for networks of $N = 100$ genes, which are governed by NCFs with $n = 5$ regulators and varying layer structures. The calculation of all the 800 plotted values took less than a second on a regular desktop computer. In networks governed by highly

	k_1	|bias|	r
n = 5	-0.944	-0.892	0.747
n = 7	-0.920	-0.869	0.636
n = 10	-0.912	-0.862	0.528
n = 15	-0.911	-0.862	0.420
n = 20	-0.911	-0.862	0.357

Table 1: For a network governed by NCFs with n regulators, this table shows the Spearman correlations between the Derrida value of a single perturbation $D(F, 1)$, and the number of most dominant variables k_1 (first column), the absolute bias (second column), and the layer number r (last column).

unbalanced NCFs with Hamming weights 1, 3, 29 and 31, small perturbations vanish on average over time; these networks operate in the stable regime. Networks of NCFs with Hamming weights 5, 7, 25 and 27 operate close to the critical threshold, while networks of NCFs with Hamming weights between 11 and 23 operate in the chaotic regime. Surprisingly, networks of NCFs with Hamming weights 11, 13, 19, and 21 are more chaotic than those governed by almost balanced NCFs with Hamming weights 15 and 17. One possible explanation for this observation may be the layer number r. NCFs with Hamming weight 15 or 17 consist of two layers, while NCFs with Hamming weight 13 and 19 (11 and 21) have three (four) different layers. The number of layers is positively correlated with the Derrida value for small perturbations (see rows 9-12 in Figure 1B and Table 1). Similarly, the number of variables in the most dominant layer, k_1, and the absolute bias are negatively correlated. Interestingly, the correlation of the Derrida value for small perturbations with k_1 and with the absolute bias remains high for NCFs with many regulators, whereas the correlation with the layer number decreases with increasing number of regulators.

Table 2 shows the impact of a single perturbation on networks governed by canalizing functions with $n = 7$ regulators, canalizing depth $k = 4$, various layer structures (r and $k_1, \ldots, k_r$) and various numbers of entries $\neq b_k$ in the truth table of the core function (v). As for NCFs, the Derrida value increases with increasing layer number and decreases with increasing number of most dominant variables, k_1. Each combination of layer structure and v yields a canalizing function with a different absolute bias and Figure 2 exhibits the connection between Derrida value and absolute bias. Again, almost balanced functions give rise to more robust networks than functions with intermediate absolute bias, while networks governed by highly biased functions operate in the stable regime. Moreover, networks with a higher proportion of canalizing variables are more robust. The gain of additional dynamical stability decreases however quickly when adding canalizing variables (see [21] for similar findings).

Theorem 3.2 and Theorem 3.5 are used in Table 3 to investigate the difference between k-canalizing functions (i.e., functions with canalizing depth $\geq k$) and functions with exact canalizing depth k. As expected, k-canalizing functions give rise to slightly more stable networks. When the number of non-canalizing variables increases, the difference in robustness vanishes, however, quickly since the vast majority of k-canalizing functions has indeed also exact canalizing depth k when $n \gg k$ [15]. To our knowledge, the number of non-canalizing functions with a given Hamming weight is unknown. For $n - k > 4$, we therefore approximated the distribution by generating 10^7 random non-canalizing functions. For $n - k \leq 4$, we simply used exhaustive enumeration.

In all these analyses, we did not specify the size of the network since we focused on the impact of a single perturbation, for which it does not matter. When considering larger perturbations, the network size has a theoretical impact on the Derrida value. We found this impact, however, to be negligible as long as the proportion of perturbed nodes remains small.

r	$[k_1,\ldots,k_r]$	v = 2 (# = 4)	v = 3 (# = 32)	v = 4 (# = 64)	v = 5 (# = 32)	v = 6 (# = 4)	expected value when v is unknown
1	[4]	0.2054	0.2879	0.3571	0.4129	0.4554	0.3524
2	[3, 1]	0.7679	0.7567	0.7321	0.6942	0.6429	0.7274
2	[2, 2]	1.0804	1.1004	1.1071	1.1004	1.0804	1.1024
3	[2, 1, 1]	0.8929	0.9442	0.9821	1.0067	1.0179	0.9774
3	[1, 1, 2]	1.1429	1.1942	1.2321	1.2567	1.2679	1.2274
4	[1, 1, 1, 1]	1.3304	1.3504	1.3571	1.3504	1.3304	1.3524
2	[1, 3]	1.1429	1.1942	1.2321	1.2567	1.2679	1.2274
3	[1, 2, 1]	1.3304	1.3504	1.3571	1.3504	1.3304	1.3524
expected value when r = 1		0.2054	0.2879	0.3571	0.4129	0.4554	0.3524
expected value when r = 2		0.9970	1.0171	1.0238	1.0171	0.9970	1.0191
expected value when r = 3		1.1220	1.1629	1.1905	1.2046	1.2054	1.1857
expected value when r = 4		1.3304	1.3504	1.3571	1.3504	1.3304	1.3524
expected value when r and $[k_1,\ldots,k_r]$ are unknown		0.9866	1.0223	1.0446	1.0536	1.0491	**1.0399**

Table 2: Impact of a single perturbation ($D(F,1)$) on networks governed by canalizing functions with $n = 7$ regulators, canalizing depth $k = 4$ for varying layer structures $(r, k_1, \ldots, k_r)$ and varying numbers of entries $\neq b_k$ in the truth table of the core function (v).

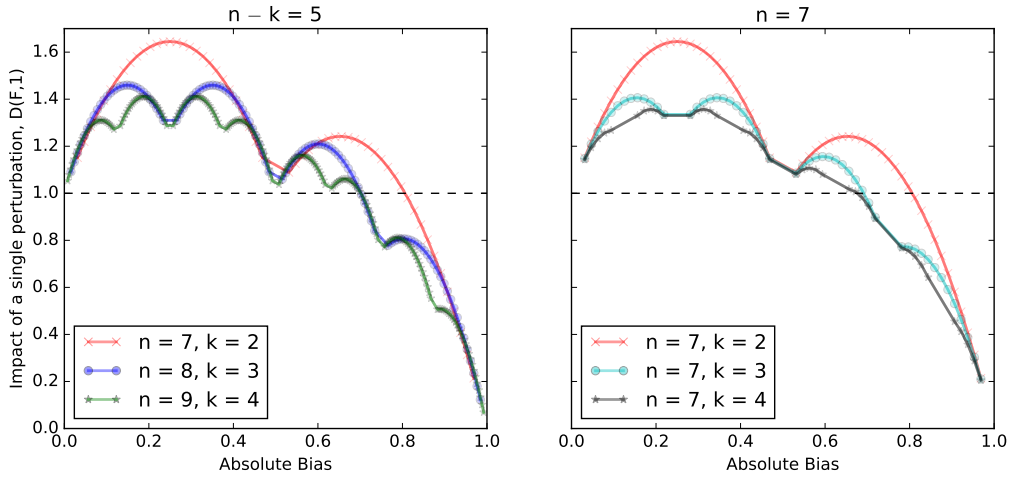


Figure 2: For all different functions of n variables with canalizing depth k, the impact of a single perturbation ($D(F,1)$) is plotted against the absolute bias of the function. In the left plot, the number of non-canalizing variables is constant, while in the right plot, the total number of variables is constant. For visualisation purposes, the scatter points are connected by a line.

		k = n-7	k = n-6	k = n-5	k = n-4	k = n-3	k = n-2	k = n-1	k = n
n = 5	canalizing depth ≥ k	d.n.e.	d.n.e.	2.5	1.50002	1.12745	1.01667	1	1
	canalizing depth = k	d.n.e.	d.n.e.	2.50003*	1.51087	1.15966	1.04167	d.n.e.	1
n = 6	canalizing depth ≥ k	d.n.e.	3	1.75	1.25001	1.06373	1.00833	1	1
	canalizing depth = k	d.n.e.	3.00042*	1.75018*	1.25543	1.07983	1.02083	d.n.e.	1
n = 7	canalizing depth ≥ k	3.5	2	1.375	1.12501	1.03186	1.00417	1	1
	canalizing depth = k	3.50013*	2.00011*	1.37504*	1.12772	1.03992	1.01042	d.n.e.	1

Table 3: Impact of a single perturbation ($D(F,1)$) on networks governed by k-canalizing functions (gray shaded rows) or functions with exact canalizing depth k (white rows) for varying numbers of regulators (n) and varying values of k. Approximated values are marked with *.

5. Derrida Values of Stochastic Discrete Dynamical Systems

Gene regulatory networks are stochastic in nature. A recently introduced generalization of Boolean networks, called Stochastic Discrete Dynamical Systems (SDDS), captures this inherent stochasticity by assigning gene-specific activation probabilities $p_i^{\uparrow}$ and degradation probabilities $p_i^{\downarrow}$, which describe how likely a specific state change happens at a given update step [30]. This framework allows modeling of different time scales as well as stochastic variability, while preserving the simplicity of a Boolean network model. We extend the Derrida value concept (Equation 4.1) to this type of system, and derive a formula for its computation.

Theorem 5.1. *The Derrida value of an SDDS* $F = \left(f_i, p_i^{\uparrow}, p_i^{\downarrow}\right)_{i=1}^{n}$ *is*

$$D(F,m) = \sum_{i=1}^{n}\sum_{c=0}^{m} \mathbb{P}(|J(f_i)| = c)\left(\frac{m}{n}\left[q_c^{f_i}\gamma_1 + \left(1-q_c^{f_i}\right)\gamma_2\right] + \frac{n-m}{n}\left[q_c^{f_i}\gamma_3 + \left(1-q_c^{f_i}\right)\gamma_4\right]\right),$$

where $|J(f_i)|$ *is hypergeometrically distributed as in Theorem 4.2 and*

$$\begin{aligned}
\gamma_1 &= 1 - \frac{1}{2}\left(p_i^{\uparrow} + p_i^{\downarrow}\right) + p_i^{\uparrow}p_i^{\downarrow}\\
\gamma_2 &= 1 - \frac{1}{2}\left(p_i^{\uparrow} + p_i^{\downarrow}\right)\\
\gamma_3 &= \frac{1}{2}\left(p_i^{\uparrow} + p_i^{\downarrow}\right)\\
\gamma_4 &= \frac{1}{2}\left(p_i^{\uparrow} + p_i^{\downarrow}\right) - \frac{1}{2}\left(p_i^{\uparrow}p_i^{\uparrow} + p_i^{\downarrow}p_i^{\downarrow}\right).
\end{aligned}$$

Proof. See Appendix. □

6. Discussion

A characteristic ability of organisms is their capability to operate in highly variable environments, striking a fine balance between the ability to adapt to changing conditions and the robustness to function predictably. This extends to the molecular networks that drive everything from embryonal development to metabolism. Understanding the mechanisms that confer this ability is one of the central challenges in studying biology and its fundamental principles. As in physics, mathematical modeling and analysis is an enabling technology for the drive to connect structural properties of dynamic biological networks to their dynamics. The study of robustness, as instantiated in the concept of canalization, is no exception, and many such studies have been published, see, e.g., [10, 17, 31, 29].

Here, we have focused on one computational instantiation of robustness, that of canalization in the context of Boolean discrete dynamical systems. This includes, as a special case, nested canalization. We have provided a collection of practical and theoretical tools for the analysis of systems governed by k-canalizing functions. In order to study the impact of a simultaneous change in more than one input of a function, we have, at the function level, generalized the concept of sensitivity. At the network level, this allowed us to provide easy-to-use closed form formulas for the Derrida values, a commonly used metric for the stability of networks. We explored the relationship between Derrida value, the canalizing depth, absolute bias, number of layers and number of most dominant variables of a function. In addition, we derived formulas for the Derrida values of a stochastic discrete dynamical system, a modeling framework that can cope with the inherent stochasticity of gene regulatory networks.

The presented formulas significantly simplify the study of robustness via Derrida values. While we started to disentangle the influence of the different parameters of a canalizing function on its robustness, much work remains to be done. Knowledge of the exact number of non-canalizing functions with a certain number of variables and a certain Hamming weight would be helpful in this effort.

References

[1] S. H. Strogatz, Exploring complex networks, Nature 410 (2001) 268–276.

[2] S. Boccaletti, V. Latora, Y. Moreno, M. Chavez, D.-U. Hwang, Complex networks: Structure and dynamics, Physics Reports 424 (2006) 175–308.

[3] H. Kitano, Biological robustness, Nature Reviews Genetics 5 (2004) 826–837.

[4] S. A. Kauffman, Metabolic stability and epigenesis in randomly constructed genetic nets, Journal of Theoretical Biology 22 (1969) 437–467.

[5] R. Thomas, Boolean formalization of genetic control circuits, Journal of Theoretical Biology 42 (1973) 563–585.

[6] R. Albert, H. G. Othmer, The topology of the regulatory interactions predicts the expression pattern of the segment polarity genes in drosophila melanogaster, Journal of Theoretical Biology 223 (2003) 1–18.

[7] F. Li, T. Long, Y. Lu, Q. Ouyang, C. Tang, The yeast cell-cycle network is robustly designed, Proceedings of the National Academy of Sciences 101 (2004) 4781–4786.

[8] M. I. Davidich, S. Bornholdt, Boolean network model predicts cell cycle sequence of fission yeast, PLOS ONE 3 (2008) e1672.

[9] B. Derrida, G. Weisbuch, Evolution of overlaps between configurations in random Boolean networks, Journal de Physique 47 (1986) 1297–1303.

[10] E. Balleza, E. R. Alvarez-Buylla, A. Chaos, S. Kauffman, I. Shmulevich, M. Aldana, Critical dynamics in genetic regulatory networks: examples from four kingdoms, PLOS ONE 3 (2008) e2456.

[11] C. H. Waddington, Canalization of development and the inheritance of acquired characters, Nature 150 (1942) 563–565.

[12] S. Kauffman, The large scale structure and dynamics of gene control circuits: an ensemble approach, Journal of Theoretical Biology 44 (1974) 167–190.

[13] S. Kauffman, C. Peterson, B. Samuelsson, C. Troein, Random Boolean network models and the yeast transcriptional network, Proceedings of the National Academy of Sciences 100 (2003) 14796–14799.

[14] L. Layne, E. Dimitrova, M. Macauley, Nested canalyzing depth and network stability, Bulletin of Mathematical Biology 74 (2012) 422–433.

[15] Q. He, M. Macauley, Stratification and enumeration of Boolean functions by canalizing depth, Physica D: Nonlinear Phenomena 314 (2016) 1–8.

[16] Y. Li, J. O. Adeyeye, D. Murrugarra, B. Aguilar, R. Laubenbacher, Boolean nested canalizing functions: A comprehensive analysis, Theoretical Computer Science 481 (2013) 24–36.

[17] S. Kauffman, C. Peterson, B. Samuelsson, C. Troein, Genetic networks with canalyzing Boolean rules are always stable, Proceedings of the National Academy of Sciences 101 (2004) 17102–17107.

[18] F. Karlsson, M. Hörnquist, Order or chaos in Boolean gene networks depends on the mean fraction of canalizing functions, Physica A: Statistical Mechanics and its Applications 384 (2007) 747–757.

[19] D. Murrugarra, R. Laubenbacher, Regulatory patterns in molecular interaction networks, Journal of Theoretical Biology 288 (2011) 66–72.

[20] N. Kochi, M. T. Matache, Mean-field Boolean network model of a signal transduction network, Biosystems 108 (2012) 14–27.

[21] K. Jansen, M. T. Matache, Phase transition of Boolean networks with partially nested canalizing functions, The European Physical Journal B 86 (2013) 1–11.

[22] E. S. Dimitrova, O. I. Yordanov, M. T. Matache, Difference equation for tracking perturbations in systems of Boolean nested canalyzing functions, Physical Review E 91 (2015) 062812.

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

[24] R. B. Boppana, The average sensitivity of bounded-depth circuits, Information Processing Letters 63 (1997) 257–261.

[25] W. Liu, L. Harri, E. R. Dougherty, I. Shmulevich, et al., Inference of Boolean networks using sensitivity regularization, EURASIP Journal on Bioinformatics and Systems Biology 2008 (2008) 1–12.

[26] S. Cook, C. Dwork, R. Reischuk, Upper and lower time bounds for parallel random access machines without simultaneous writes, SIAM Journal on Computing 15 (1986) 87–97.

[27] A. S. Jarrah, B. Raposa, R. Laubenbacher, Nested canalyzing, unate cascade, and polynomial functions, Physica D: Nonlinear Phenomena 233 (2007) 167–174.

[28] M. Nykter, N. D. Price, M. Aldana, S. A. Ramsey, S. A. Kauffman, L. E. Hood, O. Yli-Harja, I. Shmulevich, Gene expression dynamics in the macrophage exhibit criticality, Proceedings of the National Academy of Sciences 105 (2008) 1897–1900.

[29] C. Fretter, A. Szejka, B. Drossel, Perturbation propagation in random and evolved Boolean networks, New Journal of Physics 11 (2009) 033005.

[30] D. Murrugarra, A. Veliz-Cuba, B. Aguilar, S. Arat, R. Laubenbacher, Modeling stochasticity and variability in gene regulatory networks, EURASIP Journal on Bioinformatics and Systems Biology 2012 (2012) 1–11.

[31] M. Marques-Pita, L. M. Rocha, Canalization and control in automata networks: body segmentation in Drosophila melanogaster, PLOS ONE 8 (2013) e55946.

Appendix

Theorem 3.2

Proof. Let $f(x_1,\dots,x_n)$ be a k-canalizing function with canalizing order $x_1,x_2,\dots,x_k$, inputs a_i and outputs b_i, $1\le i\le k$. We will use a similar argument as in [14] to find the expected activities of f. By definition, the activity of x_j in f is the probability that a change in x_j changes the output of f. If x_j is a canalizing variable (i.e., $j\le k$), a change in x_j can only affect the output of f if none of the variables $x_1,\dots,x_{j-1}$ receive their canalizing input. Thus,

$$\begin{aligned}\alpha_j^f &= \mathbb{P}\big(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_j)\big)\\ &= \mathbb{P}(x_1\neq a_1,\dots,x_{j-1}\neq a_{j-1})\mathbb{P}\big(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_j)\mid x_1\neq a_1,\dots,x_{j-1}\neq a_{j-1}\big).\end{aligned}$$

Since each canalizing variable receives its canalizing input with probability $\frac{1}{2}$,

$$\mathbb{P}(x_1\neq a_1,\dots,x_{j-1}\neq a_{j-1})=\frac{1}{2^{j-1}}$$

For any $j\le k$, the subfunction $f(1-a_1,\dots,1-a_{j-1},x_j,x_{j+1},\dots,x_n)$ is canalizing in x_j and can therefore be written as $(x_j+a_j)\bar{g}(x_{j+1},\dots,x_n)+b_j$ for some polynomial $\bar{g}\not\equiv 0$ as in [15]. Hence,

$$\mathbb{P}\big(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_j)\mid x_1\neq a_1,\dots,x_{j-1}\neq a_{j-1}\big)=\mathbb{P}\big(\bar{g}(x_{j+1},\dots,x_n)=1\big).$$

If $j<k$, $\bar{g}\not\equiv\mathbf{1}$ since a k-canalizing function must be essential in all its canalizing variables (Remark 2.4). Both constant functions are thus excluded from the set of possible choices for $\bar{g}$ so that

$$\mathbb{P}\big(\bar{g}(x_{j+1},\dots,x_n)=1\big)=\frac{1}{2}.$$

If $j=k$, $\bar{g}\equiv\mathbf{1}$ does not cause a contradiction. In this case, there are $2^{2^{n-k}}$ choices of Boolean functions for $\bar{g}$, half of which satisfy $\bar{g}(x_{k+1},\dots,x_n)=1$. Only $\mathbf{0}$ is excluded from the set of choices for $\bar{g}$ so that

$$\mathbb{P}\big(\bar{g}(x_{k+1},\dots,x_n)=1\big)=\frac{\frac{1}{2}2^{2^{n-k}}}{2^{2^{n-k}}-1}=\frac{2^{2^{n-k}-1}}{2^{2^{n-k}}-1} \tag{6.1}$$

On the other hand, if x_j is a non-canalizing variable ($j>k$), then a change in x_j can only affect the output of f if none of the k canalizing variables receive their canalizing input. Thus,

$$\begin{aligned}\alpha_j^f &= \mathbb{P}\big(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_j)\big)\\ &= \mathbb{P}(x_1\neq a_1,\dots,x_k\neq a_k)\mathbb{P}\big(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_j)\mid x_1\neq a_1,\dots,x_k\neq a_k\big)\\ &= \frac{1}{2^k}\mathbb{P}\big(g(x_{k+1},\dots,x_n)\neq g(y_{k+1},\dots,y_n)\big),\end{aligned}$$

where $g\not\equiv b_k$ from Definition 2.3 and $\mathbf{y}=\mathbf{x}\oplus e_j$. Similar arguments as for Equation 6.1 yield

$$\mathbb{P}\big(g(x_{k+1},\dots,x_n)\neq g(y_{k+1},\dots,y_n)\big)=\frac{\frac{1}{2}2^{2^{n-k}}}{2^{2^{n-k}}-1}=\frac{2^{2^{n-k}-1}}{2^{2^{n-k}}-1}. \tag{6.2}$$

□

Theorem 3.3

Proof. Let $\mathbf{x} \in 0,1^n$ be a randomly chosen vector and let $\mathbf{y} = \mathbf{x} \oplus e_I$ be its c-Hamming neighbor, $d(\mathbf{x},\mathbf{y}) = c$. Note that Equation 6.2 is true for any $\mathbf{y}$ that differs from $\mathbf{x}$ in at least one bit $x_j, j > k$, and that the expected activity of all non-canalizing variables is the same. This implies that for any k-canalizing function f, the probability that $f(\mathbf{x})$ and $f(\mathbf{y})$ differ only depends on the first variable where $\mathbf{x}$ and $\mathbf{y}$ differ,

$$\frac{1}{2^n}\sum_{\mathbf{x}\in\{0,1\}^n}\chi\left[f\left(\mathbf{x}\right)\neq f\left(\mathbf{x}\oplus e_I\right)\right]=\mathbb{P}\left(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_I)\right)=\mathbb{P}\left(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_{\min(I)})\right)=\alpha^f_{\min(I)}$$

There are $\binom{n}{c}$ c-subsets in the set $\{1,\ldots,n\}$, $\binom{n-j}{c-1}$ of which contain j as its lowest element, $j = 1,\ldots,n-c+1$. Therefore,

$$S^f_c=\sum_{\substack{I\subseteq\{1,2,\ldots,n\}\\|I|=c}}\frac{1}{2^n}\sum_{\mathbf{x}\in\{0,1\}^n}\chi[f(\mathbf{x})\neq f(\mathbf{x}\oplus e_I)]=\sum_{\substack{I\subseteq\{1,2,\ldots,n\}\\|I|=c}}\alpha^f_{\min(I)}=\sum_{j=1}^{n-c+1}\binom{n-j}{c-1}\alpha^f_j.$$

□

Theorem 3.5

Proof. As in the proof of Theorem 3.2, the activity of any canalizing variable x_j $(j \leq k)$ is

$$\begin{aligned}\alpha^f_j&=\mathbb{P}(x_1\neq a_1,\ldots,x_{j-1}\neq a_{j-1})\mathbb{P}\big(f(\mathbf{x})\neq f(\mathbf{x}\oplus e_j)\mid x_1\neq a_1,\ldots,x_{j-1}\neq a_{j-1}\big)\\&=\frac{1}{2^{j-1}}\mathbb{P}\big(f(1-a_1,\ldots,1-a_j,x_{j+1},\ldots,x_n)\neq b_j\big)\end{aligned}$$

Due to the canalizing nature of f, the probability can be further written as

$$\begin{aligned}&\mathbb{P}\big(f(1-a_1,\ldots,1-a_j,x_{j+1},\ldots,x_n)\neq b_j\big)=\\&=\sum_{i=j+1}^{k}\Big[\mathbb{P}\big(x_{j+1}\neq a_{j+1},\ldots,x_{i-1}\neq a_{i-1},x_i=a_i\big)\cdot\\&\qquad\cdot\mathbb{P}\big(f(1-a_1,\ldots,1-a_{i-1},a_i,x_{i+1},\ldots,x_n)\neq b_j\mid x_{j+1}\neq a_{j+1},\ldots,x_{i-1}\neq a_{i-1},x_i=a_i\big)\Big]+\\&\quad+\mathbb{P}\big(x_{j+1}\neq a_{j+1},\ldots,x_k\neq a_k\big)\mathbb{P}\big(f(1-a_1,\ldots,1-a_k,x_{k+1},\ldots,x_n)\neq b_j\mid x_{j+1}\neq a_{j+1},\ldots,x_k\neq a_k\big)\\&=\sum_{i=j+1}^{k}\frac{1}{2^{i-j}}\chi(b_i\neq b_j)+\frac{1}{2^{k-j}}\mathbb{P}\big(g(x_{k+1},\ldots,x_n)\neq b_j\big).\end{aligned}$$

Let $L := L(j)$ and $L(i)$ be the layer of the variables x_j and x_i, $i \leq k$. The canalized output is equal for all variables of the same layer and alternates among layers. Therefore, $b_i \neq b_j$ if and only if $L(i) - L(j)$ is odd, and the first sum can be rewritten as a sum over all k_{L+2t-1} variables of every second layer $L + 2t - 1 > L$ $(t \geq 1)$,

$$\sum_{i=j+1}^{k}\frac{1}{2^{i-j}}\chi(b_i\neq b_j)=\sum_{t=1}^{\lceil(r-L)/2\rceil}\sum_{s=1}^{k_{L+2t-1}}\left(\frac{1}{2}\right)^{k_1+\ldots+k_{L+2t-2}+s-j}$$

Also, since g contains v entries $\neq b_k$ in its truth table,

$$\psi_l := \mathbb{P}\left(g(x_{k+1},\dots,x_n) \neq b_j\right) = \begin{cases} \mathbb{P}\left(g(x_{k+1},\dots,x_n) \neq b_k\right) & \text{if } r-l \text{ is even} \\ \mathbb{P}\left(g(x_{k+1},\dots,x_n) = b_k\right) & \text{if } r-l \text{ is odd} \end{cases}$$
$$= \begin{cases} \frac{v}{2^{n-k}} & \text{if } r-l \text{ is even} \\ 1-\frac{v}{2^{n-k}} & \text{if } r-l \text{ is odd} \end{cases}$$

Altogether,

$$\alpha_j^f = \frac{1}{2^{j-1}}\left[\sum_{t=1}^{\lceil (r-L)/2\rceil}\sum_{s=1}^{k_{L+2t-1}}\left(\frac{1}{2}\right)^{k_1+\dots+k_{L+2t-2}+s-j} + \frac{1}{2^{k-j}}\psi_l\right]$$
$$= \underbrace{\sum_{t=1}^{\lceil (r-L)/2\rceil}\sum_{s=0}^{k_{L+2t-1}-1}\left(\frac{1}{2}\right)^{k_1+\dots+k_{L+2t-2}+s}}_{:=\varphi_l} + \frac{1}{2^{k-1}}\psi_l,$$

where φ_l and ψ_l can be calculated recursively.
On the other hand, the activity of any non-canalizing variable x_j $(j>k)$ is simply

$$\begin{aligned} \alpha_j^f &= \mathbb{P}(x_1 \neq a_1,\dots,x_k \neq a_k)\mathbb{P}\big(f(\mathbf{x}) \neq f(\mathbf{x}\oplus e_j) \mid x_1 \neq a_1,\dots,x_k \neq a_k\big) \\ &= \frac{1}{2^k}\mathbb{P}\big(g(x_{k+1},\dots,x_n) \neq g(x_{k+1},\dots,x_{j-1},1-x_j,x_{j+1},\dots,x_n)\big) \\ &= \frac{1}{2^k}\frac{v(2^{n-k}-v)}{\binom{2^{n-k}}{2}} \\ &= \frac{v(2^{n-k}-v)}{2^{n-1}(2^{n-k}-1)} \end{aligned}$$

□

Theorem 5.1

Proof. Let $\mathbf{x},\mathbf{y} \in \{0,1\}^n$ be two randomly chosen vectors that differ at m positions. For each node $i \in \{1,\dots,n\}$, define three events

$$A_i = \{x_i \neq y_i\},$$
$$B_i = \left\{f_i(\mathbf{x}) \neq f_i(\mathbf{y}) \text{ before applying } p_i^{\uparrow}, p_i^{\downarrow}\right\},$$
$$C_i = \left\{f_i(\mathbf{x}) \neq f_i(\mathbf{y}) \text{ after applying } p_i^{\uparrow}, p_i^{\downarrow}\right\}.$$

Then, since A_i is independent from B_i, we have

$$\begin{aligned} D(F,m) &= \sum_{i=1}^{n}\mathbb{P}(C_i) \\ &= \sum_{i=1}^{n}\sum_{c=0}^{m}\mathbb{P}(|J(f_i)|=c)\Big(\mathbb{P}(C_i|A_i,B_i)\mathbb{P}(B_i|c)\mathbb{P}(A_i) + \mathbb{P}(C_i|A_i,\neg B_i)\mathbb{P}(\neg B_i|c)\mathbb{P}(A_i) + \\ &\qquad + \mathbb{P}(C_i|\neg A_i,B_i)\mathbb{P}(B_i|c)\mathbb{P}(\neg A_i) + \mathbb{P}(C_i|\neg A_i,\neg B_i)\mathbb{P}(\neg B_i|c)\mathbb{P}(\neg A_i)\Big) \end{aligned}$$

$$= \sum_{i=1}^{n} \sum_{c=0}^{m} \mathbb{P}(|J(f_i)| = c) \Big(\mathbb{P}(A_i) \Big[\mathbb{P}(C_i|A_i, B_i)\mathbb{P}(B_i|c) + \mathbb{P}(C_i|A_i, \neg B_i)\mathbb{P}(\neg B_i|c) \Big] + \\ + \mathbb{P}(\neg A_i) \Big[\mathbb{P}(C_i|\neg A_i, B_i)\mathbb{P}(B_i|c) + \mathbb{P}(C_i|\neg A_i, \neg B_i)\mathbb{P}(\neg B_i|c) \Big] \Big).$$

Since $\mathbf{x}$ and $\mathbf{y}$ differ at m out of n positions, $\mathbb{P}(A_i) = \frac{m}{n}$. $\mathbb{P}(B_i|c)$ is simply the normalized average sensitivity of f. Lastly, the probability that $\mathbf{x}$ and $\mathbf{y}$ differ after applying the propensity probabilities needs to be calculated. If $x_i \neq y_i$ and $f_i(\mathbf{x}) \neq f_i(\mathbf{y})$ before applying $p_i^{\uparrow}, p_i^{\downarrow}$, we can assume that $x_i = 0, y_i = 1$. Then, either $f_i(\mathbf{x}) = 0, f_i(\mathbf{y}) = 1$ or $f_i(\mathbf{x}) = 1, f_i(\mathbf{y}) = 0$, both with probability $\frac{1}{2}$. In the first case, there is no change in values so that the propensity probabilities $p_i^{\uparrow}, p_i^{\downarrow}$ play no role and $f_i(\mathbf{x}) \neq f_i(\mathbf{y})$ after applying $p_i^{\uparrow}, p_i^{\downarrow}$ with probability 1. In the second case, $f_i(\mathbf{x})$ and $f_i(\mathbf{y})$ only differ after applying $p_i^{\uparrow}, p_i^{\downarrow}$, if either both updates happen (probability $p_i^{\uparrow} p_i^{\downarrow}$) or neither update happens (probability $(1 - p_i^{\uparrow})(1 - p_i^{\downarrow})$). That means,

$$\gamma_1 := \mathbb{P}\left(C_i|A_i, B_i\right) = \frac{1}{2} \cdot 1 + \frac{1}{2} \left(p_i^{\uparrow} p_i^{\downarrow} + \left(1 - p_i^{\uparrow}\right) \left(1 - p_i^{\downarrow}\right) \right) = 1 - \frac{1}{2} \left(p_i^{\uparrow} + p_i^{\downarrow} \right) + p_i^{\uparrow} p_i^{\downarrow}.$$

Similarly, we can derive

$$\gamma_2 := \mathbb{P}\left(C_i|A_i, \neg B_i\right) = 1 - \frac{1}{2} \left(p_i^{\uparrow} + p_i^{\downarrow} \right),$$
$$\gamma_3 := \mathbb{P}\left(C_i|\neg A_i, B_i\right) = \frac{1}{2} \left(p_i^{\uparrow} + p_i^{\downarrow} \right),$$
$$\gamma_4 := \mathbb{P}\left(C_i|\neg A_i, \neg B_i\right) = \frac{1}{2} \left(p_i^{\uparrow} + p_i^{\downarrow} \right) - \frac{1}{2} \left(p_i^{\uparrow} p_i^{\uparrow} + p_i^{\downarrow} p_i^{\downarrow} \right).$$

□