2009

On the sensitivity to noise of a Boolean function

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

Valentin Matache
University of Nebraska at Omaha, vmatache@unomaha.edu

ON THE SENSITIVITY TO NOISE OF A BOOLEAN FUNCTION

MIHAELA T. MATACHE* AND VALENTIN MATACHE

Department of Mathematics

University of Nebraska at Omaha

Omaha, NE 68182-0243, USA

*dmatache@mail.unomaha.edu, vmatache@mail.unomaha.edu

Abstract. In this paper we generate upper and lower bounds for the sensitivity to noise of a Boolean function using relaxed assumptions on input choices and noise. The robustness of a Boolean network to noisy inputs is related to the average sensitivity of that function. The average sensitivity measures how sensitive to changes in the inputs the output of the function is. The average sensitivity of Boolean functions can indicate whether a specific random Boolean network constructed from those functions is ordered, chaotic, or in critical phase. We give an exact formula relating the sensitivity to noise and the average sensitivity of a Boolean function. The analytic approach is supplemented by numerical results that illustrate the overall behavior of the sensitivities as various Boolean functions are considered. It is observed that, for certain parameter combinations, the upper estimates in this paper are sharper than other estimates in the literature and that the lower estimates are very close to the actual values of the sensitivity to noise of the selected Boolean functions.

Keywords: sensitivity to noise of a Boolean function, average sensitivity of a Boolean function, lower and upper estimates, generalized elementary cellular automata rules.

1. Introduction

Boolean network models have been used for modelling networks in which the node or cell activity can be described by two states, 1 and 0, ON and OFF, “active and nonactive”, “up-regulated and down-regulated”, and in which each node is updated based on logical relationships with other nodes. The random Boolean networks have been originally developed by Stuart Kauffman as models for genetic regulatory networks [1]. They are referred to as N/K models or Kauffman networks. Boolean networks can model a variety of real or artificial networks including among others: genetic regulatory networks (e.g. Shmulevich et. al. [2], [3]), strongly disordered systems that are common in physics (e.g. Kauffman [1], or Kaufman et.al. [4]), biology (e.g. Klemm and Bornholdt [5], or Raeymaekers [6]), neural networks (e.g. Aldana and Cluzel [7], or Huepe and Aldana [8]), and artificial life (e.g. Wolfram [9]).

As mentioned by Goodrich and Matache in [10], it is known that real networks (biological/genetic, physical, neural, chemical, social etc.) are always subject to disturbances, have the ability to reach functional diversity, and aim to maintain the same state under environmental noise (e.g. food source or energy changes). There are intrinsic or environmental disturbances as well as possible mutations within the network (e.g. genetic mutations). Inducing disturbance in the system by changing the value of certain nodes in the network (according to a deterministic or stochastic rule) is a good model for an environmental or intrinsic type of perturbation. For example, in [10] it is shown that in a Boolean network governed by a specific type of Boolean functions, the introduction of noise can stabilize the system for a wide range of parameters. In Bilke and Sjunnesson [11] a node of the stable core of a

Kauffman network is chosen at random and inverted after the system has reached a limit cycle. The sensitivity of the attractors is investigated. The authors find that the stable core of the network lacks the well-known insensitivity observed in full Kauffman networks. In [12], Beck and Matache apply a particular stochastic noise procedure typical for neural networks, to a Boolean network governed by a certain generalized elementary cellular automata rule. It is shown that there is no critical value of the noise parameter that differentiates between ordered and random behavior of the system.

The study of the robustness of a Boolean network to various types of perturbations is an important aspect of the evolution of systems under Boolean models. These systems have to respond and adapt to interior and exterior disturbances. The interest is in suppressing chaos and bringing the system into an ordered regime. So it is important to understand what the impact of noise is when applied to the input of a Boolean function: does the output change or not? Thus, is the Boolean function sensitive to noise or not? In this paper we answer such questions in a probabilistic setting for certain types of Boolean functions and noise choices.

The average sensitivity of a Boolean function measures how sensitive to changes in the inputs the output of the function is. The average sensitivity has been studied by Friedgut in [13] the conclusion being that Boolean functions with low average sensitivity depend on few coordinates, or Shmulevich and Kauffman in [2] in the context of random Boolean networks. Those authors show that this numerical characteristic can reflect the dynamical behavior of the random Boolean networks constructed from some specific choices of Boolean functions, such as canalizing functions. The average sensitivity can indicate whether a specific random Boolean network is ordered, chaotic, or in critical phase. Intuitively, the concept of sensitivity should be related to the robustness against errors in the inputs. As mentioned by Schober in [14], it is of interest to estimate the probability that a random flip of the value of each input generates a different output of a Boolean function. More precisely, it is of interest to understand what the sensitivity to noise of a Boolean function is and to relate it to the average sensitivity of that function. In [14] this is done in a particular case of input choice and noise application. The author of [14] shows that, the sensitivity to noise is bounded above by the average sensitivity of the function multiplied by a small noise-related parameter. Therefore, if the average sensitivity is small, the noise is not amplified. The noise sensitivity of Boolean functions and its applications to percolation have been studied by Benjamini et.al. in [15]. In the current paper we extend the results of [14] in the context of more relaxed assumptions on the input choices and noise, and determine both upper and lower bounds for the sensitivity to noise. Under certain assumptions, we give an exact formula relating the sensitivity to noise and the average sensitivity of a Boolean function. The analytic approach provided here is supplemented by numerical results that illustrate the overall behavior of the sensitivities as various Boolean functions are considered.

In Section 2 we define the main concepts of sensitivity of a Boolean function and noise and provide upper bounds for the sensitivity to noise of a Boolean function. In Section 3 we focus on special Boolean functions and provide lower and upper bounds for the sensitivity to noise under specific function and parameter choices. In particular we study the elementary cellular automata (ECA) rules 0-256 and generalizations of some of them. We provide numerical results that compare our upper bound to the one in [14] and describe the general behavior of the sensitivity to noise. In Section 4 we relate the sensitivity to noise to the average sensitivity of a Boolean function. Section 5 is dedicated to conclusions and the description of further research directions.

2. Sensitivity and Noise

Let $\Omega = \{0, 1\}$ and consider a probability measure μ on the σ-algebra of all parts of Ω^n.

Definition 1. *For all $j = 1, 2, \ldots, n$ and all $x \in \Omega^n$, recall that the concept of* ***sensitivity of a Boolean function*** $f : \Omega^n \to \Omega$ ***at*** x *is defined as the quantity*

$$s(f, x) := |\{y \in \Omega^n : d(x, y) = 1 \, and \, f(x) \neq f(y)\}| \tag{1}$$

where $|S|$ denotes the cardinality of the set S, and $d(x, y) = \sum_{j=1}^{n} |x_j - y_j|$ is the Hamming distance between the vectors x and y. The ***average sensitivity*** *of f is*

$$avs(f) := \int_{\Omega^n} s(f, x) d\mu(x). \tag{2}$$

Thus, the sensitivity of f at x provides the number of vectors in Ω^n that differ in exactly one coordinate from x (i.e. the Hamming neighbors of x) and generate a flip of the output of the function f. One can regard $s(f, x)$ as output values of a random variable valued in $\{1, 2, \ldots, n\}$ with associated probabilities $\mu(x), x \in \Omega^n$. So the average sensitivity is simply the mean value of this random variable.

Definition 2. *By* ***noise*** *applied to the vector $x \in \Omega^n$ we understand a random variable $N(x) = (N_1(x), N_2(x), \ldots, N_n(x))$ valued in Ω^n that transforms the input vector x into another vector y according to some given rule.*

Assume we work with a given noise–process, (according to Definition 2, noise is actually an Ω_n–indexed, Ω_n–valued stochastic process). The associated noise–operator is defined as follows.

Definition 3. *Given the noise N, the* ***noise operator*** *T_N is the linear operator*

$$(T_N f)(x) := \sum_{y\in\Omega^n} f(y)P(N(x)=y) \qquad x\in\Omega^n,\ f:\Omega^n\to\mathbb{R}.$$

The extent to which a Boolean function f is affected by noise is encoded in the quantity $\rho_f(N)$ defined as follows:

$$\rho_f(N) := \int_{\Omega^n} P(f(N(x))\neq f(x))\,d\mu(x) \qquad f:\Omega^n\to\Omega.$$

Definition 4. *We call $\rho_f(N)$ the* ***sensitivity of the Boolean function*** *f* ***to the noise*** *N.*

The **noise–insensitive** Boolean functions are the Boolean functions f with the property $\rho_f(N) = 0$. There are always at least 2 such functions, namely the constant Boolean functions 0 and 1.

We denote $\chi_f := (-1)^f$. The Hilbert space where our considerations take place is $L^2(\Omega^n,\mathbb{R})$, (that is the space of real–valued functions on Ω^n endowed with the norm $\|f\|_2 = \sqrt{\int_{\Omega^n}|f(x)|^2\,d\mu(x)}\,$) induced by the inner product $\langle f,g\rangle = \int_{\Omega^n} f(x)g(x)d\mu(x)$. The following result in [14] is valid in general, (that is for all noise–operators, not just the particular one considered in [14]). This can be established with exactly the same proof as the one provided in the cited paper.

Proposition 1 ([14, Lemma 2])**.** *The following formula holds*

$$\rho_f(N) = \frac{1}{2}\left(1-\langle T_N\chi_f,\chi_f\rangle\right) \qquad f:\Omega^n\to\Omega. \tag{3}$$

Based on that we prove:

Proposition 2. *Let* λ_{T_N} *denote the least eigenvalue of the real part* $\Re T_N = (T_N^* + T_N)/2$ *of* T_N*. The following estimate holds*

$$\rho_f(N) \leq \frac{1}{2}(1 - \lambda_{T_N}) \qquad f : \Omega^n \to \Omega. \tag{4}$$

Proof. Consider $\Re T_N = (T_N + T_N^*)/2$ acting on $L^2(\Omega^n, \mathbb{C})$. Then, the numerical range $W(\Re T_N) = \{\langle \Re T_N f, f\rangle : \|f\|_2 = 1\}$ equals the line interval $[\lambda_{T_N}, \Lambda_{T_N}]$, where Λ_{T_N} is the largest eigenvalue of $\Re T_N$. This equality combines with formula (3) and the evident fact that $\|\chi_f\|_2 = 1$ into establishing (4), since T_N leaves invariant the subspace $L^2(\Omega^n, \mathbb{R})$ of $L^2(\Omega^n, \mathbb{C})$. □

For more on numerical ranges of operators, one can check the basic reference [16]. It should be added that the argument above also produces the estimate

$$\frac{1}{2}(1 - \Lambda_{T_N}) = \frac{1}{2}(1 - \|\chi_f\|_2^2 \Lambda_{T_N}) \leq \rho_f(N)$$

which is not interesting. Indeed, $\frac{1}{2}(1 - \Lambda_{T_N}) \leq 0$ (since $\Re T_N 1 = 1$), and clearly $0 \leq \rho_f(N)$. The lower bound 0 of $\rho_f(N)$ is attained each time when f is noise–insensitive.

Our next upper bound for ρ_f is not f–independent like the upper bound in (4).

Proposition 3. *For any Boolean function* $f : \Omega^n \to \Omega$ *the following holds*

$$\rho_f(N) \leq \|f\|_2^2(1 - 2\lambda_{T_N}) + \langle T_N f, 1\rangle \tag{5}$$

where 1 *denotes both the scalar* 1 *and the Boolean function constantly equal to* 1.

Proof. Estimate (5) is an immediate consequence of the following considerations. For any Boolean function f, it is obviously true that $\chi_f = 1 - 2f$. Therefore, by (3), one has

$$\rho_f(N) = \frac{1}{2}\langle (I - T_N)(1 - 2f), 1 - 2f\rangle = -\langle (I - T_N)(f), 1 - 2f\rangle$$

because $(I - T_N)(1) = 0$. Of course, I denotes the identity operator $I(f) = f$. One obtains

$$\rho_f(N) = -\langle f, 1\rangle + 2\|f\|_2^2 + \langle T_N f, 1\rangle - 2\langle T_N f, f\rangle = \|f\|_2^2 + \langle T_N f, 1\rangle - 2\langle T_N f, f\rangle$$

because f is a Boolean function. For $f \neq 0$, the quantity $\langle T_N f, f\rangle / \|f\|_2^2$ belongs to $W(\Re T_N) = [\lambda_{T_N}, \Lambda_{T_N}]$, which leads to (5). □

A consequence of one of the equalities in the proof above is:

Proposition 4. *A Boolean function f is noise–insensitive if it is an invariant function of the noise operator, that is, if $T_N f = f$. If $\mu(x) \neq 0$ for all $x \in \Omega^n$, then a Boolean function f is noise–insensitive if and only if it is an invariant function of the noise operator.*

Proof. Indeed, in the previous proof we established formula

$$\rho_f(N) = \|f\|_2^2 + \langle T_N f, 1\rangle - 2\langle T_N f, f\rangle. \tag{6}$$

If $T_N f = f$, one obtains

$$\rho_f(N) = \langle T_N f, 1\rangle - \|f\|_2^2 = 0.$$

For the only if part of the statement in the text of the proposition, observe first that $0 \leq (T_N f)(x) \leq 1$ for all $x \in \Omega^n$. If f, a Boolean function, is noise–insensitive then (6) must hold that is

$$\sum_{x \in \Omega^n} (f(x)(1 - (T_N f)(x)) + (T_N f)(x)(1 - f(x)))\mu(x) = 0$$

hence

$$f(x)(1 - (T_N f)(x)) = (T_N f)(x)(1 - f(x)) = 0 \qquad x \in \Omega^n$$

since $\mu(x) \neq 0$ for all $x \in \Omega^n$.

Thus, for all $x \in \Omega^n$, if $f(x) = 0$ then $(T_N f)(x) = 0$ and if $f(x) = 1$ then $(T_N f)(x) = 1$. It follows that $T_N f = f$. □

It is easy to construct noises for which a preassigned Boolean function f is noise–insensitive. Indeed, given such an f, consider for each $x \in \Omega^n$ a map $N(x) : \Omega^n \to \Omega^n$ with the property $N(x)(\Omega^n) \subseteq f^{-1}(f(x))$. Given this property of the noise N, one has that

$$\{f(N(x)) \neq f(x)\} = \emptyset \qquad x \in \Omega^n,$$

for which reason

$$\rho_f(N) := \int_{\Omega^n} P(f(N(x)) \neq f(x))\, d\mu(x) = 0,$$

that is, f is noise–insensitive.

Estimates for $\rho_f(N)$ are interesting only if f is noise–sensitive. All the noise examples considered in the next sections, have only 2 noise–insensitive Boolean functions: the constant ones.

3. The Case of Product Measures and Examples

Consider a product–probability measure μ on the σ-algebra of all parts of Ω^n. More exactly, consider the measure $\mu = \mu_1 \times \mu_2 \times \cdots \times \mu_n$ where

$$\mu_j(x) = \begin{cases} p_j & \text{if } x_j = 1 \\ 1 - p_j & \text{if } x_j = 0 \end{cases} \qquad x = (x_1, \ldots, x_n) \in \Omega^n \tag{7}$$

and $0 < p_j < 1$ are fixed for $j = 1, 2, \ldots, n$.

Thus each input x_j can be viewed as a Bernoulli random variable with parameter p_j and

$$\mu(x) = \prod_{j=1}^{n} \mu_j(x) = \prod_{\{j:x_j=1\}} p_j \prod_{\{j:x_j=0\}} (1 - p_j). \tag{8}$$

Now let

$$\varphi_j(x) := (-1)^{x_j} \left(\frac{1 - p_j}{p_j}\right)^{\frac{2x_j - 1}{2}} \qquad x = (x_1, \ldots, x_n) \in \Omega^n,\ j = 1, \ldots, n. \tag{9}$$

Consider $\mathcal{B} := \{\phi_u := \prod_{j=1}^{n} \varphi_j^{u_j} : u = (u_1, u_2, \ldots, u_n) \in \Omega^n\}$. Then the following holds.

Remark 1. *$\mathcal{B}$ is a complete orthonormal basis of $L^2(\Omega^n, \mathbb{R})$.*

Indeed, consider $u, v \in \Omega^n$. One gets:

$$\langle \phi_u, \phi_v \rangle = \int_{\Omega^n} \prod_{j=1}^{n} \varphi_j^{u_j + v_j}(x)\, d\mu(x) = \delta_{u,v},$$

since

$$\int_{\Omega^n} \varphi_j(x)\, d\mu(x) = 0 \quad \text{and} \quad \int_{\Omega^n} \varphi_j^2(x)\, d\mu(x) = 1 \qquad j = 1, \ldots, n.$$

The linear dimension of $L^2(\Omega^n, \mathbb{R})$ being $|\Omega^n|$, this establishes our claim.

The complete orthonormal basis $\mathcal{B}$ appears in the particular case $p_1 = \cdots = p_n$ in [17].

Let us consider now a noise example that is a generalization of the noise used in [14]. We will assume that we work with the product measure described above in this section.

Example 1. *For all $j = 1, 2, \dots, n$, let $P(N_j(x) = x_j) = \delta_j$ and $P(N_j(x) = y_j) = 1 - \delta_j$ where y_j is a Bernoulli random variable with parameter p_j. In other words, we assume that each coordinate of the input vector x is unchanged with probability $\delta_j \in [0, 1]$, and with probability $1 - \delta_j$ it is given by a Bernoulli random variable. Denoting the noise by N we get*

$$P(N(x) = y) = \prod_{\{j:x_j=0,y_j=0\}} [\delta_j + (1-\delta_j)(1-p_j)] \cdot \prod_{\{j:x_j=0,y_j=1\}} [(1-\delta_j)p_j] \cdot \prod_{\{j:x_j=1,y_j=0\}} [(1-\delta_j)(1-p_j)] \cdot \prod_{\{j:x_j=1,y_j=1\}} [\delta_j + (1-\delta_j)p_j] \tag{10}$$

for $y \in \Omega^n$.

Observe that for large δ_j values, or large or small p_j values, the sensitivity of the Boolean function f to the noise N is small, so that the values of $\rho_f(N)$ are close to zero.

Consider now the complete orthonormal basis $\mathcal{B} = \{\phi_u : u \in \Omega^n\}$.

Proposition 5. *Given the noise operator of Example 1, the vectors ϕ_u of the orthonormal basis $\mathcal{B}$ are eigenvectors of T_N with corresponding eigenvalues $\prod_{i=1}^n \delta_i^{u_i}$. That is*

$$T_N \phi_u = \left(\prod_{i=1}^n \delta_i^{u_i} \right) \phi_u. \tag{11}$$

Proof: Using the definitions of T_N and ϕ_u we basically need to show

$$\sum_{y\in\Omega^n}\prod_{i=1}^{n}\left((-1)^{y_i}\left(\frac{1-p_i}{p_i}\right)^{\frac{2y_i-1}{2}}\right)^{u_i} P(N(x)=y)=$$

$$=\left(\prod_{i=1}^{n}\delta_i^{u_i}\right)\cdot\prod_{i=1}^{n}\left((-1)^{x_i}\left(\frac{1-p_i}{p_i}\right)^{\frac{2x_i-1}{2}}\right)^{u_i}.$$

We introduce the following notations: $K^{xuy}_{abc}=\{j\in\{1,2,\dots,n\}\,:\,x_j=a,u_j=b,y_j=c\}$ and $K^{xu}_{ab}=\{j\in\{1,2,\dots,n\}\,:\,x_j=a,u_j=b\}$, where $a,b,c\in\Omega$. Using these notations we can rewrite $T_N\phi_u$ as

$$T_N\phi_u(x)=\sum_{y\in\Omega^n}\left[\left(\prod_{i\in K^{xuy}_{010}\cup K^{xuy}_{110}}\sqrt{\frac{p_i}{1-p_i}}\right)\cdot\left(\prod_{i\in K^{xuy}_{011}\cup K^{xuy}_{111}}\left(-\sqrt{\frac{1-p_i}{p_i}}\right)\right)\right.$$

$$\cdot\left(\prod_{i\in K^{xuy}_{000}\cup K^{xuy}_{010}}[\delta_i+(1-\delta_i)(1-p_i)]\right)\cdot\left(\prod_{i\in K^{xuy}_{001}\cup K^{xuy}_{011}}[(1-\delta_i)p_i]\right)$$

$$\left.\cdot\left(\prod_{i\in K^{xuy}_{100}\cup K^{xuy}_{110}}[(1-\delta_i)(1-p_i)]\right)\cdot\left(\prod_{i\in K^{xuy}_{101}\cup K^{xuy}_{111}}[\delta_i+(1-\delta_i)p_i]\right)\right].$$

However

$$K^{xuy}_{000}\cup K^{xuy}_{001}=K^{xu}_{00},K^{xuy}_{010}\cup K^{xuy}_{011}=K^{xu}_{01},K^{xuy}_{100}\cup K^{xuy}_{101}=K^{xu}_{10},K^{xuy}_{110}\cup K^{xuy}_{111}=K^{xu}_{11}.$$

Combining the terms based on these associations between the sets K and separating the sum over $y\in\Omega^n$ accordingly we obtain that $(T_N\phi_u)(x)$ is equal to

$$\left(\sum_{l=0}^{|K^{xu}_{11}|}\sum_{\{i_1<i_2<\cdots<i_l\}\subseteq K^{xu}_{11}}\left[\prod_{j=1}^{l}\sqrt{\frac{p_{i_j}}{1-p_{i_j}}}[(1-\delta_{i_j})(1-p_{i_j})]\cdot\right.\right.$$

$$\left.\left.\cdot\prod_{j\neq i_t,t=1,2,\dots,l}\left(-\sqrt{\frac{1-p_j}{p_j}}\right)[\delta_j+(1-\delta_j)p_j]\right]\right)$$

$$\cdot\left(\sum_{l=0}^{|K^{xu}_{01}|}\sum_{\{i_1<i_2<\cdots<i_l\}\subseteq K^{xu}_{01}}\left[\prod_{j=1}^{l}\sqrt{\frac{p_{i_j}}{1-p_{i_j}}}[\delta_{i_j}+(1-\delta_{i_j})(1-p_{i_j})]\cdot\right.\right.$$

$$\left.\left.\cdot\prod_{j\neq i_t,t=1,2,\dots,l}\left(-\sqrt{\frac{1-p_j}{p_j}}\right)[(1-\delta_j)p_j]\right]\right)$$

$$\cdot\left(\sum_{l=0}^{|K_{00}^{xu}|}\sum_{\{i_1<i_2<\cdots<i_l\}\subseteq K_{00}^{xu}}\left[\prod_{j=1}^{l}[\delta_{i_j}+(1-\delta_{i_j})(1-p_{i_j})]\cdot\prod_{j\neq i_t,t=1,2,\ldots,l}[(1-\delta_j)p_j]\right]\right)$$

$$\cdot\left(\sum_{l=0}^{|K_{10}^{xu}|}\sum_{\{i_1<i_2<\cdots<i_l\}\subseteq K_{10}^{xu}}\left[\prod_{j=1}^{l}[(1-\delta_{i_j})(1-p_{i_j})]\cdot\prod_{j\neq i_t,t=1,2,\ldots,l}[\delta_j+(1-\delta_j)p_j]\right]\right).$$

Note that the last two sums are equal to 1 and thus we finally get

$$(T_N\phi_u)(x)=\prod_{i\in K_{11}^{xu}}\left(\sqrt{\frac{p_i}{1-p_i}}[(1-\delta_i)(1-p_i)]-\sqrt{\frac{1-p_i}{p_i}}[\delta_i+(1-\delta_i)p_i]\right)$$

$$\cdot\prod_{i\in K_{01}^{xu}}\left(\sqrt{\frac{p_i}{1-p_i}}[\delta_i+(1-\delta_i)(1-p_i)]-\sqrt{\frac{1-p_i}{p_i}}[(1-\delta_i)p_i]\right)$$

$$=\prod_{i\in K_{11}^{xu}}\left(-\sqrt{\frac{1-p_i}{p_i}}\right)\delta_i\cdot\prod_{i\in K_{01}^{xu}}\left(\sqrt{\frac{p_i}{1-p_i}}\right)\delta_i$$

$$=\prod_{i\in K_{01}^{xu}}\delta_i\cdot\prod_{i\in K_{11}^{xu}}\delta_i\cdot\prod_{i=1}^{n}\left((-1)^{x_i}\left(\frac{1-p_i}{p_i}\right)^{\frac{2x_i-1}{2}}\right)^{u_i}$$

$$=\left(\prod_{i=1}^{n}\delta_i^{u_i}\right)\cdot\prod_{i=1}^{n}\left((-1)^{x_i}\left(\frac{1-p_i}{p_i}\right)^{\frac{2x_i-1}{2}}\right)^{u_i}.$$

□

Remark 2. *As specified in* [14], *the typical noise example usually assumes that each individual input x_j is flipped with some probability ϵ (not necessarily the same for all $j=1,2,\ldots n$). This would make our analysis a lot harder since the result in the previous proposition would not hold. However, in the special case $p_j=1/2$ and $\delta_j=1-2\epsilon$, $\forall j=1,2,\ldots,n$, the two noise models coincide as shown in* [14].

Now, since the eigenvalues of T_N are $\prod_{i=1}^{n}\delta_i^{u_i}$, $u\in\Omega^n$, it follows that the least eigenvalue is $\lambda_{T_N}=\prod_{i=1}^{n}\delta_i$. Then (4) becomes

$$\rho_f(N)\leq\frac{1}{2}\left(1-\prod_{i=1}^{n}\delta_i\right)\tag{12}$$

while (5) becomes

$$\rho_f(N) \leq \|f\|_2^2 \left(1 - 2\prod_{i=1}^{n} \delta_i\right) + \langle T_N f, 1\rangle. \tag{13}$$

Estimate (12) is sharper than (13) if and only if $Q = \mu(\{f = 1\}) > 1/4$. Indeed, observe that $\langle T_N f, 1\rangle = \langle f, T_N 1\rangle = \langle f, 1\rangle = Q$ and clearly $\|f\|_2^2 = Q$. So the upper estimate (13) can be written

$$\rho_f(N) \leq 2Q(1 - \prod_{i=1}^{n} \delta_i)$$

and obviously one has

$$\frac{1}{2}(1 - \prod_{i=1}^{n} \delta_i) < 2Q(1 - \prod_{i=1}^{n} \delta_i)$$

if and only if $Q > 1/4$ (of course we exclude the trivial case $\delta_i = 1, \forall\, i = 1, 2, \ldots, n$).

Clearly, estimate (13) is an equality if f is the null function, but is this the only Boolean function with that property? The answer is affirmative. We sketch briefly the proof. Estimate (5) is an equality for some nonzero f if and only if $\lambda_{T_N} = \langle T_N f, f\rangle / \|f\|_2^2$. Since (13) is the particular form of (5), in the case of the particular noise operator described above and, since that operator is diagonal with positive diagonal entries, it coincides with its real part. Hence λ_{T_N} is $\prod_{i=1}^{n} \delta_i$ and it is straightforward to see that, (given the diagonal matrix of T_N and the minimality of λ_{T_N}), the equality $\lambda_{T_N} = \langle T_N f, f\rangle / \|f\|_2^2$ can hold if and only if f is a scalar multiple of the basic vector $\phi_{(1,\ldots,1)}$, that is of the vector belonging to $\mathcal{B}$ which is an eigenvector of T_N associated to eigenvalue λ_{T_N}. The only scalar multiple of $\phi_{(1,\ldots,1)}$ which is a Boolean function is 0 since $\phi_{(1,\ldots,1)}$ assumes both positive and negative values.

Given that the operator T_N has a diagonal matrix in the complete orthonormal basis $\mathcal{B}$, one can obtain the following useful lower–bound estimate for $\rho_f(N)$.

Proposition 6. *Let T_N be as above. Then, for all $f : \Omega^n \to \Omega$,*

$$2Q(1-Q)\min\{(1-\prod_{j=1}^{n}\delta_j^{u_j}) : u \neq 0\} \leq \rho_f(N) \tag{14}$$

where, as above, $Q = \mu(\{f = 1\})$.

Proof. The operator $I - T_N$ has diagonal matrix with entries $\{1 - \prod_{j=1}^{n}\delta_j^{u_j}\}$. Thus only the diagonal entry $1 - \prod_{j=1}^{n}\delta_j^{0}$ is null. That entry corresponds to the function $\phi_0 = 1$. Let us denote by $\{\hat{\chi}_f(u)\}$ the Fourier coefficients of χ_f with respect to basis $\mathcal{B}$ and calculate $\hat{\chi}_f(0)$. One gets

$$\hat{\chi}_f(0) = \langle \chi_f, 1\rangle = \langle (-1)^f, 1\rangle = \sum_{f(x)=0}\mu(x) - \sum_{f(x)=1}\mu(x).$$

Since $\|\chi_f\|_2 = \sum_{u\in\Omega^n}\left(\hat{\chi}_f(u)\right)^2 = 1$, formula (3) can be written as follows:

$$\rho_f(N) = \frac{1}{2}\langle (I - T_N)\chi_f, \chi_f\rangle.$$

Thus, given the diagonal form of the matrix of T_N, one has

$$\rho_f(N) = \frac{1}{2}\sum_{u\neq 0}(1-\prod_{j=1}^{n}\delta_j^{u_j})|\hat{\chi}_f(u)|^2 \geq \frac{\min\{(1-\prod_{j=1}^{n}\delta_j^{u_j}) : u\neq 0\}}{2}\sum_{u\neq 0}|\hat{\chi}_f(u)|^2$$
$$= \frac{\min\{(1-\prod_{j=1}^{n}\delta_j^{u_j}) : u\neq 0\}}{2}(1 - |\hat{\chi}_f(0)|^2)$$

which establishes

$$\frac{\min\{(1-\prod_{j=1}^{n}\delta_j^{u_j}) : u\neq 0\}}{2}\left(1 - \left(\sum_{f(x)=1}\mu(x) - \sum_{f(x)=0}\mu(x)\right)^2\right),$$

that is (14) after an elementary computation.

□

Remark 3. *Lower bound (14) is tight. That is, we can find a particular set of parameters and Boolean functions for which the value of $\rho_f(N)$ is equal to that lower bound. Indeed, consider the following special case: $p_j = 1/2, \delta_j = \delta$ for all $j = 1, 2, \ldots, n$, and the "dictatorship" function $f(x) = x_1$. Then it is easy to see that $\mu(x) = 1/2^n$ for all $x \in \Omega^n$ and $Q = 1/2$. Therefore the lower bound (14) becomes $\frac{1-\delta}{2}$ while $\rho_f(N) = \frac{1}{2^n}\sum_{x\in\Omega^n} P(N_1(x) \neq x_1) = \frac{1-\delta}{2}$ (using the total probability formula).*

Similarly one can obtain yet another upper estimate for ρ_f, namely

$$\rho_f(N) \leq \tag{15}$$

$$\frac{\max\{(1-\prod_{j=1}^n \delta_j^{u_j}) : u \neq 0\}}{2}\left(1 - \left(\sum_{f(x)=1}\mu(x) - \sum_{f(x)=0}\mu(x)\right)^2\right).$$

Thus, combining the upper estimates (12), (13), (15) we obtain the upper estimate

$$\min\left\{\frac{1}{2}\left(1-\prod_{i=1}^n \delta_i\right), 2Q\left(1-\prod_{i=1}^n \delta_i\right), 2Q(1-Q)\max\{(1-\prod_{j=1}^n \delta_j^{u_j}) : u \neq 0\}\right\} \tag{16}$$

and we will be using this estimate in the next examples. We focus on some particular Boolean rules with specified parameter combinations and consider the noise operator defined in the previous example. We investigate numerically the accuracy of the estimates for the cases considered.

Let us change the orthonormal basis and work with $\tilde{\mathcal{B}} = \{e_u : u \in \Omega^n\}$ where for all $u \in \Omega^n$, $e_u(x) = \frac{1}{\sqrt{\mu(u)}}\delta_{ux}, x \in \Omega^n$. Checking that $\tilde{\mathcal{B}}$ is a complete orthonormal basis of $L^2(\Omega^n, \mathbb{R})$ is straightforward. Observe that, the matrix of the linear operator T_N with respect to $\tilde{\mathcal{B}}$ is simply given by

$$[T_N] = [a_{xy}]_{x,y\in\Omega^n} \quad \text{where} \quad a_{xy} = P(N(x) = y)\sqrt{\frac{\mu(x)}{\mu(y)}}$$

which is a $2^n \times 2^n$ matrix. We assume that the vectors in Ω^n are ordered, say according to their base 10 representation.

Example 2. *Consider the Boolean rule $f : \Omega^n \to \Omega$ given by*

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} 0 & \textit{if} \quad \sum_{i=1}^n x_i = 0 \textit{ or } \sum_{i=1}^n x_i = n \\ 1 & \textit{otherwise.} \end{cases}$$

This is the extensively studied generalized ECA rule 126 considered for example in [10] in the context of noise driven Boolean networks. The ECA rules have been explained and studied in great detail by Wolfram in [9].

In [14] the following upper bound is found in the particular case $p_j = p, \delta_j = \delta$ for all $j = 1, 2, \ldots, n$.

$$\rho_f \leq \epsilon \cdot avs(f) \qquad \text{where} \quad \epsilon = 2p(1-p)(1-\delta).$$

Using rule 126 of Example 2 we get the equivalent expression $\rho_f \leq 2np(1-p)(1-\delta)(p^{n-1}+(1-p)^{n-1})$ according to the definition of $avs(f)$. We would like to compare this estimate with our bound (16). To do this, we graph the upper estimates against p and δ in Figure 1 (bottom right) with $n = 3$ for simplicity, that is the actual ECA rule 126. We generate a mesh for each of the estimates for a grid of values of $p \in [0, 1/2]$ and $\delta \in [0, 1]$. We also graph the corresponding values of ρ_f to determine the accuracy of the bounds. The two surfaces of the estimates intersect along a curve. We observe that the upper estimate (16) (denoted E_1) is sharper than the estimate established in [14] (denoted E_2) for larger p and smaller δ. At the same time in Figure 2 we graph the lower bound (14) versus ρ_f. We can see that the lower estimate is quite accurate.

Note the general behavior of ρ_f in Figures 1 and 2. The value of ρ_f is close to zero for large $\delta \in [0, 1]$ and small $p \in [0, 0.5]$ as expected. Also, it decreases with increased δ, that is with an increased probability of not changing the input value when applying the noise operator. The values of ρ_f decrease also with decreased p since with probability $1 - \delta$, the noise operator applies a Bernoulli random variable with probability p, under which the nodes become 0 with probability $1 - p$.

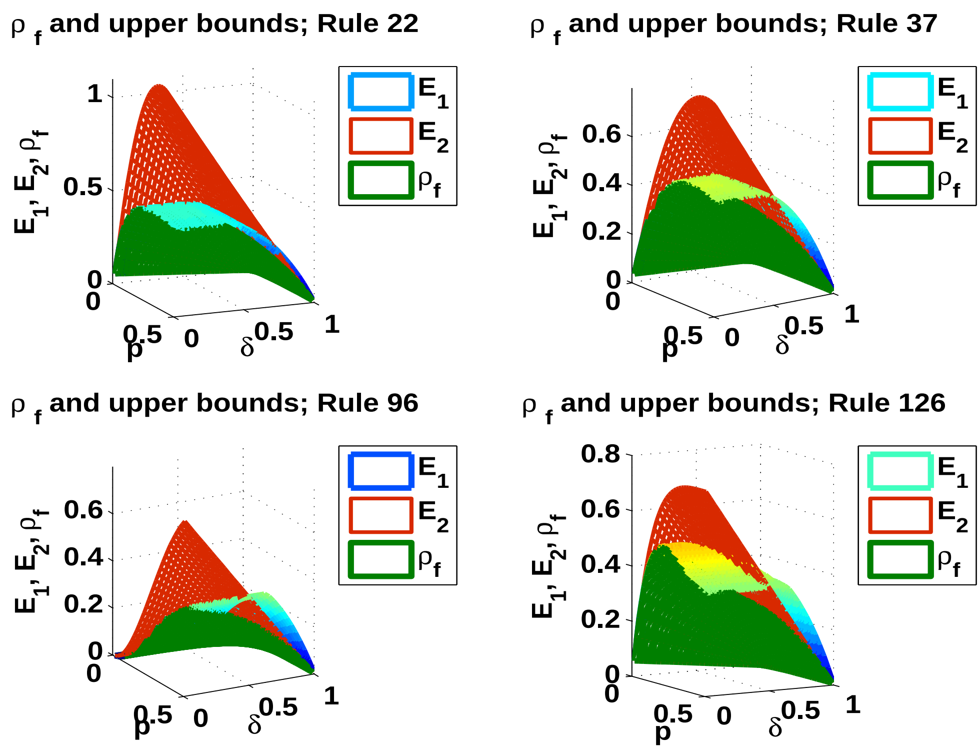

FIGURE 1. (Color online) Plot of ρ_f and its upper estimate (16) labeled E_1, and the estimate in [14], labeled E_2, versus a grid of $p_j = p \in [0, 0.5]$ and $\delta_j = \delta \in [0, 1]$ values. This is done for rules 22, 37, 96, and 126 as typical examples for the ECA rules with $n = 3$. Note that E_1 is sharper than E_2 for larger p and smaller δ.

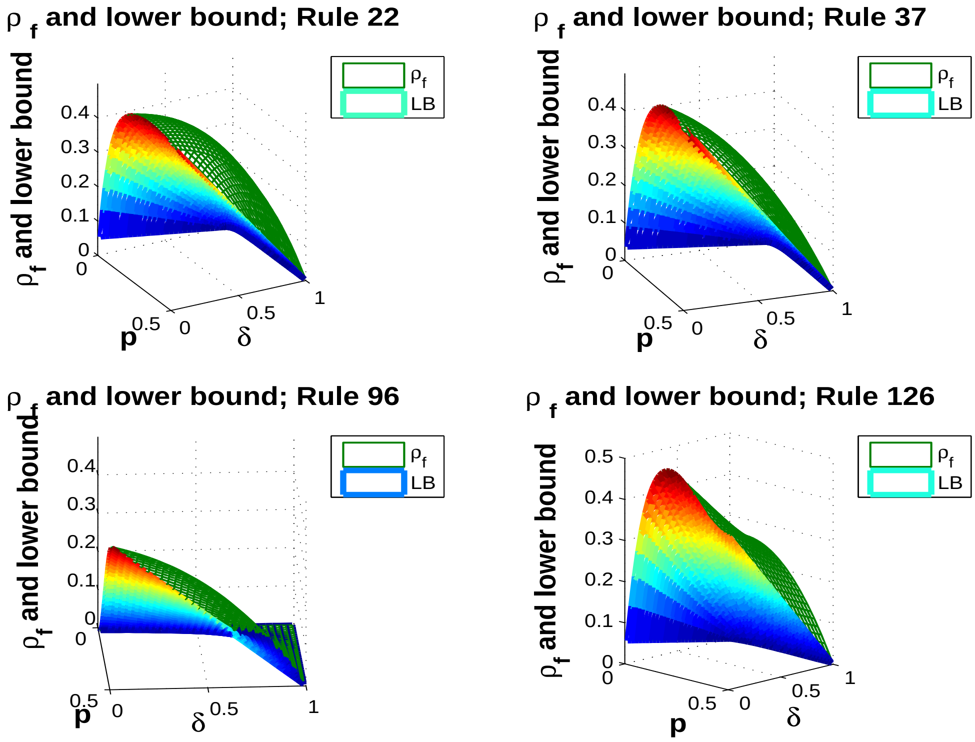

FIGURE 2. (Color online) Plot of ρ_f and its lower estimate (14) labeled LB, versus a grid of $p_j = p \in [0, 0.5]$ and $\delta_j = \delta \in [0, 1]$ values. This is done for rules 22, 37, 96, and 126 as typical examples for the ECA rules with $n = 3$. Note that the lower estimate is quite accurate.

Example 3. *Consider now the Boolean rule $f : \Omega^n \to \Omega$ given by*

$$f(x_1, x_2, \ldots, x_n) = \begin{cases} 1 & \textit{if} \quad d_1 \le \frac{\sum_{i=1}^n x_i}{n} \le d_2 \\ 0 & \textit{if} \quad \textit{otherwise.} \end{cases}$$

Here $0 \le d_1 \le d_2 \le 1$ are fixed parameters. In the context of Boolean networks the rule means that the node under consideration is turned ON if and only if the fraction of 1s is within the given bounds d_1 and d_2. Otherwise the node is turned

OFF. We assume that the node is turned OFF under complete isolation or complete crowding.

This is the generalized ECA rule 22 studied in [12] in the context of synchronous Boolean networks evolving under this rule, whose dynamics are studied also under a stochastic noise procedure. As specified in that paper, this generalization allows insight into Boolean systems governed by rules meaningful for neural networks or biology. For example, if $d_2 = 1$, then the output is 1 (or it fires) if and only if the sum of all the inputs is at least the threshold $0 < d_1 < 1$. Thus we deal with a Boolean linear threshold function that is typical for neural networks. On the other hand, a small value of d_1 implies that fewer active inputs have the property of activating the node under consideration. Hence there is a bias towards the activators of the node. If d_1 is large, then there is a bias towards the inhibitors of the node. In [6], the author indicates that biologically meaningful Boolean functions have input elements that are activators or inhibitors, which can act alone or in conjunction with other activators and/or inhibitors. In a cellular automaton governed by biologically meaningful functions with 3 or 4 inputs, increasing significantly the bias towards the inhibitors or the activators has the effect of decreasing the length of the cycles and of the run-ins, which represent the initial part of trajectories before cycles are reached [6]. We will draw some conclusions on the impact of the bias towards inhibitors or activators on the sensitivity of a Boolean function to noise.

As in Example 2, we graph the estimates and ρ_f against p and δ in Figures 1 and 2 (top left) for $n = 3$ and $d_1 = d_2 = 1/3$, representing exactly the ECA rule 22. We observe again that the estimate generated in this paper (E_1) is significantly sharper than (E_2), the estimate established in [14], for larger p and smaller δ. The actual values of ρ_f tend to E_1 for large p and small δ values.

There are 256 ECA rules ($n = 3$) as described in [9]. One only needs to consider 128 of them, namely rules 0 to 127, since the remaining ones are obtained by symmetry (switching 0 and 1). To have a complete view on the estimates for all ECA rules, we have generated graphs of E_1, E_2, LB, and ρ_f against p and δ for all 128 rules. The results are similar to the ones in Figures 1 and 2 where we show ECA rules 37 and 96 which are typical as well. The general conclusion is that the upper estimate (16) is somewhat less accurate than the one in [14] for small p and large δ, but can be significantly sharper for larger p and smaller δ values. The lower estimate is accurate for all values of the parameters.

The advantage of (16) is that it can be used also in case of varying p and δ values, for which the estimate in [14] is not valid. So, to understand even better the accuracy of E_1 for varying p and δ values, we concentrate now on graphing ρ_f and its estimates for various values of δ. For each δ we plot the results of 8 different runs with selected $p_j, j = 1, 2, \ldots, n$ to obtain a wide range of values of $\|p\| = \sqrt{\sum_{j=1}^{n} p_j^2} \in [0, \sqrt{n}]$. We create three dimensional plots of E_1 versus δ and $\|p\|$ for ECA rules 22, 37, 96 and 126 in Figure 3. In this figure $n = 3$ and $\delta_j = \delta$ for $j = 1, 2, \ldots n$, but the graphs are similar for larger n with generalized ECA rules or varying δ_j values (not shown). Observe that the estimates are sharper for rule 22 and that the shape of the plots exhibits a certain symmetry with respect to $\|p\|$: the values of ρ_f and its estimates tend to be closer to zero for small and large $\|p\|$, that is when the p_j values tend to be all small or all large which induces mostly zeros or ones in the input (here $p \in [0, 1]$ as opposed to $p \in [0, 0.5]$ in previous figures to point out the symmetry). The values of ρ_f are larger for medium values of $\|p\|$.

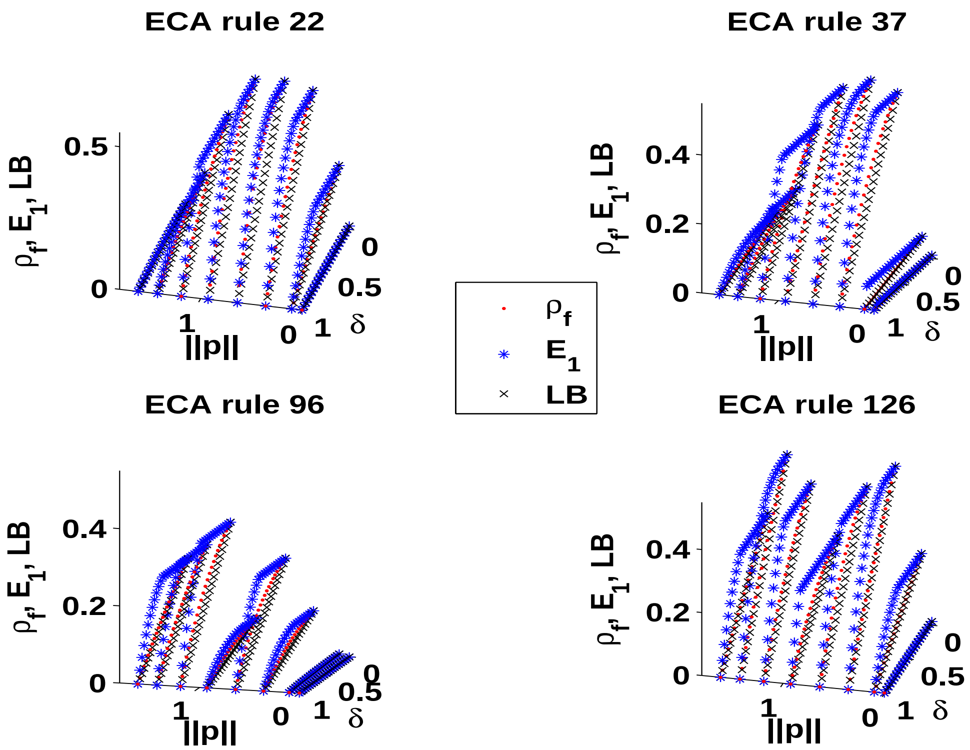

FIGURE 3. (Color online) Plot of ρ_f and its estimates (14) and (16) versus δ and $\|p\|$ for rules 22, 37, 96, and 126. Here $n = 3$ and the values p_j are allowed to vary. Note that rule 22 yields somewhat sharper estimates, and that small or large $\|p\|$ generate smaller values for ρ_f and its estimates, while medium values generate mostly a higher sensitivity to noise.

We note that similar situations are obtained for the generalized rules 22 and 126 with $n > 3$ and various choices of d_1 and d_2 in rule 22. For example, if $d_2 = 1$ and d_1 is allowed to vary, we observe that the estimates are more accurate for smaller d_1, that is under a bias towards the activators. As observed before, the values of the estimates decrease with the increase of d_1, that is with an increased bias towards the inhibitors of the node. However, the general shape of the graphs in all cases is very similar to Figures 1 and 2 with some variation in terms of accuracy of the estimates or ranges of values of the estimates and ρ_f. In all cases, the lower estimate (14) is quite sharp. It is also easy to compute. Note also that although the upper estimate (15) may be less accurate than (16) (the one used in the numerical investigations above), it is actually much easier to compute. So, for large n, one may want to accept a reduced accuracy of the upper estimate to ease the computational burden. In Figure 4 we graph the upper estimate (15) together with the lower estimate (14) for $n = 10$ as an example. We consider the generalized ECA rule 22 with several parameter combinations, as specified in the titles of the subplots. In particular, we fix again d_2 and allow d_1 to vary to account for a bias towards activators of inhibitors. Note that $d_1 = 0$ actually represents exactly the generalized ECA rule 126. The two surfaces have a common boundary for $\delta = 0$ and are at zero for $p = 0$ as expected. The lower bound is zero for $\delta = 1$. Thus, the lower and upper surfaces generate an "envelope" for the values of ρ_f. We can see that, for smaller d_1 values which indicate a bias towards the activators, the estimates take on larger values, while the surfaces tend to flatten out and to approach zero as d_1 increases forcing the bias towards the inhibitors.

4. Average sensitivity

In this section we tie ρ_f to the average sensitivity $avs(f)$ as defined in (2). We do this in the particular case of the noise operator of Example 1 for which we have shown that $T_\delta \phi_u = (\prod_{i=1}^n \delta_i^{u_i})\phi_u$, $u \in \Omega^n$. We also relate our results to the estimate in [14].

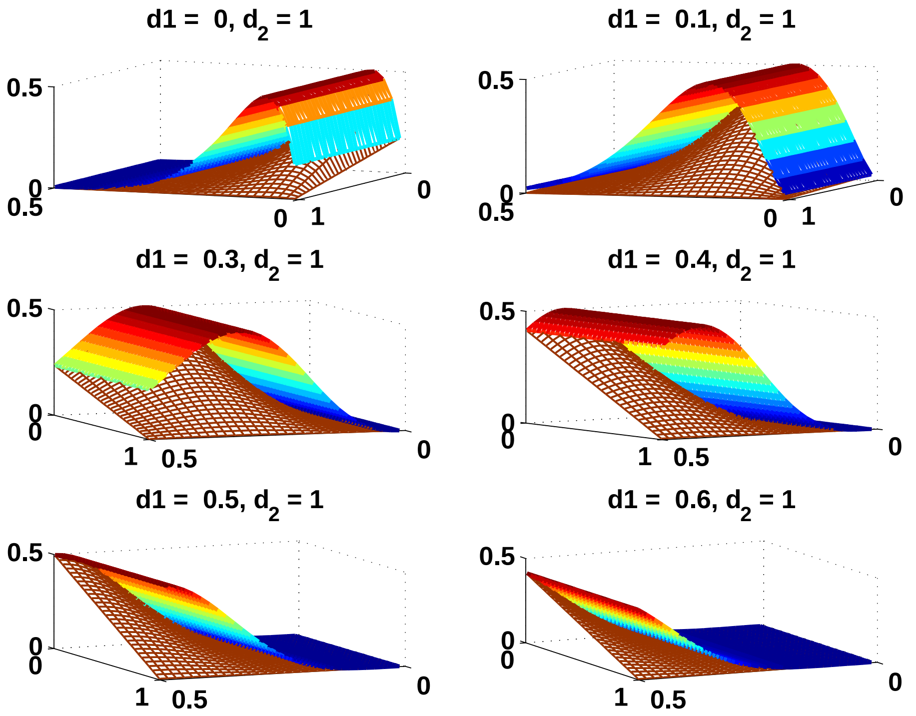

FIGURE 4. (Color online) Plot of the lower estimate (14) and the upper estimate (15) versus a grid of $p_j = p \in [0, 0.5]$ and $\delta_j = \delta \in [0, 1]$ values. This is done for various parameter combinations for the generalized ECA rule 22 specified in the titles. The two meshes provide an "envelope" for the actual values of ρ_f.

Let $f : \Omega^n \to \mathbb{R}$. For all $j = 1, \ldots, n$ and all $x \in \Omega^n$ denote by $x \oplus j$ the vector obtained from x by flipping the entry x_j. The following definition is basically due to [18].

Definition 5. *The* ***influence of the variable*** j ***on the function*** $f : \Omega^n \to \Omega$ *is defined as follows:*

$$I_j(f) := \mu(\exists x \in \Omega^n \ s.t. \ f(x) \neq f(x \oplus j)).$$

Clearly $I_j(f) = \int_{\Omega^n} |f(x) - f(x \oplus j)| \, d\mu(x)$ and therefore $avs(f) = \sum_{j=1}^{n} I_j(f)$. We introduce the following operator, which is essentially due to [17]:

$$\Delta_j f(x) = \begin{cases} (1-p_j)(f(x) - f(x \oplus j))) & \text{if } x_j = 1 \\ p_j(f(x) - f(x \oplus j)) & \text{if } x_j = 0 \end{cases} \qquad f : \Omega^n \to \mathbb{R}. \tag{17}$$

Proposition 7. *The operator Δ_j satisfies the following:*

$$\Delta_j \phi_u(x) = \begin{cases} \phi_u(x) & \textit{if} \quad u_j = 1 \\ 0 & \textit{if} \quad u_j = 0. \end{cases} \tag{18}$$

In other words, Δ_j is the orthogonal projection onto the subspace $\mathcal{L}_j$ of $L^2(\Omega^n, \mu)$ spanned by the functions in $\{\phi_u : u \in \Omega^n, \, u_j = 1\}$.

Proof: Observe that

$$\Delta_j \phi_u(x) = \begin{cases} (1-p_j)(\phi_u(x) - \phi_u(x \oplus j)) & \text{if} \quad x_j = 1 \\ p_j(\phi_u(x) - \phi_u(x \oplus j)) & \text{if} \quad x_j = 0 \end{cases}$$

and that

$$\phi_u(x) - \phi_u(x \oplus j) = \left[\prod_{i=1}^{n} \varphi_i(x)^{(1-\delta_{ij})u_i} \right] \cdot \left[\varphi_j(x)^{u_j} - \varphi_j(x \oplus j)^{u_j} \right].$$

Then it is clear that $u_j = 0$ implies that the quantity above is zero, so $\Delta_j \phi_u(x) = 0$.

On the other hand, if $u_j = 1$ we can see that

$$\phi_u(x) - \phi_u(x \oplus j) = \left[\prod_{i=1}^{n} \varphi_i(x)^{(1-\delta_{ij})u_i} \right] \cdot \left[(-1)^{x_j} \left(\frac{1-p_j}{p_j} \right)^{\frac{2x_j-1}{2}} - (-1)^{1-x_j} \left(\frac{1-p_j}{p_j} \right)^{\frac{1-2x_j}{2}} \right]$$

$$= \begin{cases} \left[\prod_{i=1}^{n} \varphi_i(x)^{(1-\delta_{ij})u_i} \right] \cdot \left(-\frac{1}{\sqrt{p_j(1-p_j)}} \right) & \text{if} \quad x_j = 1 \\ \left[\prod_{i=1}^{n} \varphi_i(x)^{(1-\delta_{ij})u_i} \right] \cdot \left(\frac{1}{\sqrt{p_j(1-p_j)}} \right) & \text{if} \quad x_j = 0 \end{cases}$$

which implies that $\Delta_j \phi_u(x) = \phi_u(x)$ for $u_j = 1$. □

A consequence of this result and of the linearity of Δ_j is the fact that $\Delta_j f(x) = \sum_{u\in\Omega^n} \hat{f}(u)\phi_u(x)u_j$, where $\hat{f}(u)$ are the corresponding Fourier coefficients of f with respect to basis $\mathcal{B}$.

Now one can see that

$$I_j(f) = \frac{1}{4p_j(1-p_j)} \cdot ||\Delta_j \chi_f||_2^2 = \frac{1}{4p_j(1-p_j)} \cdot \sum_{u\in\Omega^n} \hat{\chi}_f(u)^2 u_j.$$

Indeed,

$$||\Delta_j \chi_f||_2^2 = \sum_{x\in\Omega^n} \Delta_j \chi_f(x)^2 \mu(x)$$

$$= \sum_{x\in\Omega^n} \left[(1-p_j)^2(\chi_f(x) - \chi_f(x\oplus j))^2 p_j + p_j^2(\chi_f(x) - \chi_f(x\oplus j))^2(1-p_j)\right] \mu(x)$$

$$= \sum_{x\in\Omega^n} \left[4p_j(1-p_j)^2|f(x) - f(x\oplus j)| + 4p_j^2(1-p_j)|f(x) - f(x\oplus j)|\right] \mu(x)$$

$$= 4p_j(1-p_j) \sum_{x\in\Omega^n} |f(x) - f(x\oplus j)|\, \mu(x) = 4p_j(1-p_j) I_j(f)$$

where we take into account that $P(x_j = 1) = p_j$ and $P(x_j = 0) = 1 - p_j$. Thus, since $avs(f) = \sum_{j=1}^n I_j(f)$ we get

$$avs(f) = \sum_{u\in\Omega^n} \left(\sum_{j=1}^n \frac{u_j}{4p_j(1-p_j)} \right) \hat{\chi}_f(u)^2. \tag{19}$$

Using formula (3) and the Fourier expansion of χ_f in the orthonormal basis $\mathcal{B}$ we get

$$\rho_f = \frac{1}{2}\left(1 - \sum_{u\in\Omega^n} \left(\prod_{i=1}^n \delta_i^{u_i} \right) \hat{\chi}_f(u)^2 \right).$$

But since $u_i = 0$ or 1 we have that $\delta_i^{u_i} = 1 - (1-\delta_i)u_i$ which implies

$$\rho_f = \frac{1}{2}\left(1 - \sum_{u\in\Omega^n} \prod_{i=1}^n (1 - (1-\delta_i)u_i)\, \hat{\chi}_f(u)^2 \right) =$$

$$= \frac{1}{2} \sum_{u\in\Omega^n} \left(\sum_{i=1}^n (1-\delta_i)u_i - \sum_{1\le i<j\le n} (1-\delta_i)u_i(1-\delta_j)u_j + \cdots - (-1)^n \prod_{i=1}^n (1-\delta_i)u_i \right) \hat{\chi}_f(u)^2$$

where we have used the fact that $\sum_{u\in\Omega^n}\hat{\chi}_f(u)^2 = 1$. Now using the notation $\epsilon_i = 2p_i(1-p_i)(1-\delta_i)$ for $i = 1, 2, \ldots, n$, solving for $1-\delta_i$, and replacing in the formula of ρ_f we obtain

$$\rho_f = \sum_{u\in\Omega^n}\left(\sum_{i=1}^{n}\frac{\epsilon_i u_i}{4p_i(1-p_i)} - \right.$$

$$\left. -\frac{1}{2}\sum_{1\le i<j\le n}\frac{\epsilon_i u_i}{2p_i(1-p_i)}\cdot\frac{\epsilon_j u_j}{2p_j(1-p_j)} + \cdots - (-1)^n\frac{1}{2}\prod_{i=1}^{n}\frac{\epsilon_i u_i}{2p_i(1-p_i)}\right)\hat{\chi}_f(u)^2.$$

Observe that in the special case when $\epsilon_i = \epsilon$ and therefore $\delta_i = 1 - \frac{\epsilon}{2p_i(1-p_i)}$, $i = 1, 2, \ldots, n$ we get

$$\rho_f = \epsilon\cdot avs(f) - \tag{20}$$

$$-\sum_{u\in\Omega^n}\left(\frac{\epsilon^2}{2}\sum_{1\le i<j\le n}\frac{u_i}{2p_i(1-p_i)}\cdot\frac{u_j}{2p_j(1-p_j)} + \cdots + (-1)^n\frac{\epsilon^n}{2}\prod_{i=1}^{n}\frac{u_i}{2p_i(1-p_i)}\right)\hat{\chi}_f(u)^2.$$

Thus ρ_f and $avs(f)$ are related by formula (20). Furthermore, if $p_i = p, \forall i = 1, 2, \ldots, n$ as in [14], then

$$\rho_f = \epsilon\cdot avs(f) - \tag{21}$$

$$-\frac{1}{2}\sum_{u\in\Omega^n}\left(\sum_{l=2}^{n}\left[\left(\frac{-\epsilon}{2p(1-p)}\right)^l\cdot\sum_{1\le i_1<i_2<\cdots<i_l\le n}u_{i_1}u_{i_2}\ldots u_{i_l}\right]\right)\hat{\chi}_f(u)^2.$$

In [14], the author obtains in this case the upper estimate $\rho_f \le \epsilon\cdot avs(f)$ which follows as a consequence of (21).

The proof in this section follows closely the lines in [14].

5. Conclusions

In this paper we provide upper and lower bounds for the sensitivity to noise of a Boolean function under certain assumptions on the input generation, noise induction, and choice of Boolean functions. The lower bound is very close to the actual values of the sensitivity to noise, while the upper bound is sharper for certain parameter combinations, and less accurate for others. Under the assumptions used in [14] the upper bound in that paper is shown to be significantly less accurate in case of a larger probability of inducing noise by generating input values based on the flip of a coin. That bound is also deduced from an exact formula relating the sensitivity to noise to the average sensitivity of a Boolean function. Analytic results are supplemented by numerical investigations that illustrate typical overall behavior of the sensitivity to noise of a Boolean function. In general, the bounds take on larger values under a bias towards the activators of a node, while a bias towards the inhibitors has the effect of decreasing the sensitivity to noise.

It would be interesting to extend the current work to multiple iterations of the Boolean functions and to Boolean networks in general. In this context, a natural direction would be to use also other noise operators or underlying Boolean functions, such as threshold functions typical for neural networks (e.g. those used in [8] and [12]). In this case the output of one iteration of the network becomes the input for the next iteration. The nodes of the network could be updated or disturbed say by turning a node on if the concentration of active nodes in its neighborhood reaches a given threshold. This could be done in connection to average sensitivities of higher order. More precisely, the sensitivity of order j at a given input vector x is the number of vectors with j flipped values that generate a change in the output of a given Boolean function (the average sensitivity in this paper corresponds to $j = 1$). Some preliminary work using the higher order sensitivity has been done in [19]. One could use the estimates to generate ranges of parameter values that yield ordered, chaotic or critical dynamical behavior, thus extending the results obtained in [2]. In the context of Boolean networks the number of inputs of each node could vary, the network could be considered synchronous or asynchronous, and the Boolean functions could be chosen deterministically or stochastically from specialized Boolean functions, such as canalizing, or other biologically meaningful functions. Moreover, one could be interested in dealing with multiple Boolean rules, each of them being chosen with a given probability, thus extending the work in this paper to the so called probabilistic Boolean networks studied for example in [3]. That would provide a more realistic approach to, say, biological cellular networks whose update schemes are dependent upon various protein interactions.

References

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

[2] Shmulevich I., Kauffman S.A., *Activities and sensitivities in Boolean network models*, Physical Review Letters, 93(2004), no. 4, 048701(4).

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

[4] Kaufman V., Mihaljev T., Drossel B., *Scaling in critical random Boolean networks*, Phys. Rev. E 72, 046124(2005).

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

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

[7] Aldana M., Cluzel P., *A natural class of robust networks*, PNAS 100 (2003), p. 8710-8714.

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

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

[10] Goodrich C.S., Matache M.T., *The stabilizing effect of noise on the dynamics of a Boolean network*, Physica A 379 (2007) 334-356.

[11] Bilke S. and Sjunnesson F. , *Stability of the Kauffman model*, Physical Review, E 65, 016129, 2001.

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

[13] Friedgut E., *Boolean functions with low average sensitivity depend on few coordinates*, Combinatorica 18(1), 1998, p. 27-35.

[14] Schober S., *About Boolean networks with noisy inputs*, Proceedings of the Fifth International Workshop on Computational Systems Biology, WCSB 2008, p. 173-176.

[15] Benjamini I., Kalai G., Schramm O., *Noise sensitivity of Boolean functions and applications to percolation*, Inst. Hautes Études Sci. Publ. Math. 90, 1999, p. 5-43.

[16] Gustafson K.E., Rao D.K.M., *Numerical range*, Springer-Verlag, New York, Heidelberg, Berlin, 1997.

[17] Talagrand M., *On Russo's approximate zero–one law*, Ann. of Probability, 22(1994), no. 3, 1576–1587.

[18] Ben-Or M., Linial N., *Collective coin flipping*, in "Randomness and Computation" (S. Micali ed.) Academic Press, New York, 1989, pp. 91-115.

[19] Schober S., Bossert M., *Analysis of random Boolean networks using the average sensitivity*, arXiv:0704.0197v1 [nlin.CG].