Nat Comput (2017) 16:427–439
DOI 10.1007/s11047-016-9584-z

Activity in Boolean networks

Abhijin Adiga[1] · Hilton Galyean[2] · Chris J. Kuhlman[1] · Michael Levet[4] · Henning S. Mortveit[1,3] · Sichao Wu[1]

Published online: 14 October 2016

Abstract In this paper we extend the notion of activity for Boolean networks introduced by Shmulevich and Kauffman (Phys Rev Lett 93(4):48701:1–4, 2004). In contrast to existing theory, we take into account the actual graph structure of the Boolean network. The notion of activity measures the probability that a perturbation in an initial state produces a different successor state than that of the original unperturbed state. It captures the notion of sensitive dependence on initial conditions, and provides a way to rank vertices in terms of how they may impact predictions. We give basic results that aid in the computation of activity and apply this to Boolean networks with threshold functions and nor functions for elementary cellular automata, *d*-regular trees, square lattices, triangular lattices, and the Erdős–Renyi random graph model. We conclude with some open questions and thoughts on directions for future research related to activity, including long-term activity.

Keywords Boolean networks · Finite dynamical system · Activity · Sensitivity · Network · Sensitive dependence on initial conditions

✉ Henning S. Mortveit
henning.mortveit@vt.edu

Abhijin Adiga
adiga@vbi.vt.edu

Hilton Galyean
hgalyean@vt.edu

Chris J. Kuhlman
ckuhlman@vbi.vt.edu

Michael Levet
mlevet@vt.edu

Sichao Wu
sichao@vbi.vt.edu

1 Network Dynamics and Simulation Science Laboratory, VBI, Virginia Tech, 1015 Life Science Circle, Blacksburg, VA 24061, USA

2 Department of Physics, Virginia Tech, Blacksburg, VA 24061, USA

3 Department of Mathematics, Virginia Tech, Blacksburg, VA 24061, USA

4 Department of Computer Science, University of South Carolina, Columbia, SC, USA

1 Introduction

A *Boolean network (BN)* is a map of the form

$$F = (f_1, \ldots, f_n) : \{0,1\}^n \longrightarrow \{0,1\}^n \ . \qquad (1)$$

BNs were originally proposed as a model for many biological phenomena (Kauffman 1969, 1993), but have now been used to capture and analyze a range of complex systems and their dynamics (Aldana et al. 2003). Associated to F we have the *dependency graph of* F whose vertex set is $\{1, 2, \ldots, n\}$ and with edges all (i, j) for which the function f_i depends non-trivially on the variable x_j. See (Goles and Martinez 1990; Robert 1986; Mortveit and Reidys 2007) for more general discussions of maps F.

The study of stability and the response to perturbations of Boolean networks is central to increased understanding of their dynamical properties. Perturbations may take many forms, with examples including perturbations of the dependency graph (Adiga et al. 2013; Luo and Turner 2012; Pomerance et al. 2009), the vertex states (Shmulevich and Kauffman 2004; Ghanbarnejad and Klemm 2012; Serra et al. 2010), the vertex functions (Shmulevich et al. 2003; Xiao and Dougherty 2007), or combinations of these.

This paper is concerned with noise applied to vertex states. Specifically, we want to know the following:

What is the probability that $F(x)$ and $F(x+e_i)$ are different?

Here i is a vertex while e_i is the ith unit vector with the usual addition modulo 2. In Shmulevich and Kauffman (2004), Shmulevich and Kauffman considered Boolean networks over regular graphs with $f_j = f$ for all vertices j, that is, those induced by a common function. They defined the notion of *activity* of f with respect to its ith argument as the expected value of the Boolean derivative of f with respect to its ith variable. Under their assumptions, this may give a reasonable indication of the expected impact of perturbations to the ith variable under the evolution of F. However, this approach does not consider the impact of the dependency graph structure. We remark that Layne et al. (2012) compute the activities of nested canalyzing functions given their canalyzing depth, extending results in Shmulevich and Kauffman (2004) on canalyzing functions.

This question of sensitivity has also been studied when x is restricted to attractors in order to assess stability of long-term dynamics under state noise. The notion of threshold ergodic sets (TESs) is introduced in Ribeiro and Kauffman (2007) and studied further in Luo and Turner (2012), and Kuhlman and Mortveit (2014). The structure of TESs capture long-term stability under state perturbations of periodic orbits and the resulting mixing between attractors that may happen as a result.

Other tools for analyzing sensitivity of vertex noise includes Lyapunov exponents, see for example (Baetens and Baets 2010; Baetens et al. 2012), although this is perhaps mostly relevant or suited for the infinite case such as cellular automata over (infinite) regular lattices. Also, the notion of *Derrida diagrams* has been used to quantify how *Hamming classes* of states separate on average (Derrida and Pomeau 1986; Fretter et al. 2009) after one or ℓ transitions under F. Derrida diagrams, however, are mainly analyzed through numerical experiments via sampling. Moreover, analyzing how Hamming classes of large distance separate under F may not be so insightful—it seems more relevant to limit oneself to the case of nearby classes of vertex states.

Returning to the original question, we note that $F(x)$ and $F(x+e_i)$ may only differ in the components j for which f_j depends on x_i. The answer to the question therefore depends on the vertex functions in the 1-neighborhood of i in the dependency graph X, and therefore the structure of the induced subgraph of the 2-neighborhood of i in X with the omission of edges connecting pairs of vertices both of distance 2 from i. We denote this subgraph by $X(i; 2)$, see Fig. 1.

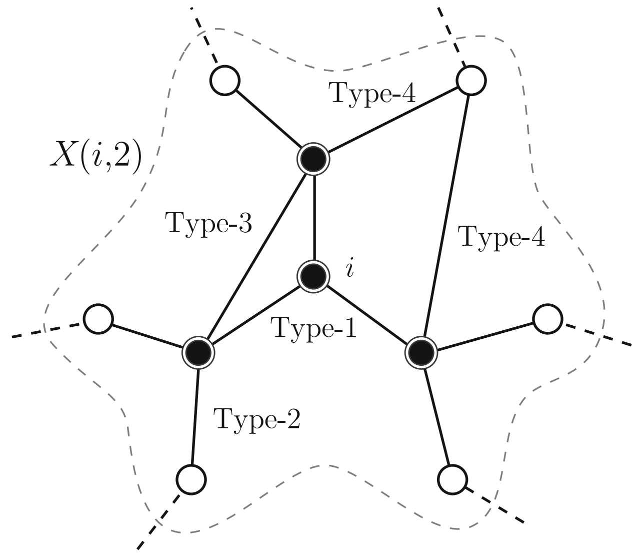

Fig. 1 The subgraph $X(i; 2)$ of X induced by vertex i and its distance ≤ 2 neighbors. Vertices belonging to the 1-neighborhood of i (*note* that i is included) are marked using *filled circles*, remaining vertices are marked using *open circles*. Edges connecting vertices of distance 2 from i in X do not belong to $X(i; 2)$. Type-1 edges are edges incident to i; Type-2 edges are edges of $X(i; 2)$ not incident to i and not part of any 3- or 4-cycle. Type-3 (resp. type-4) edges are edges in 3-cycles (resp. 4-cycles) containing i that are not incident to i. For additional terminology we refer to Sect. 2

To determine the activity of vertex i one will in general have to evaluate F over all possible states of $X(i; 2)$, a problem which, in the general case, is computationally intractable. We will address three cases in this paper: the first case is when $X(i; 2)$ is a tree, the second case is that of elementary cellular automata (a special case of the former case), and the case where X is a regular, square, 2-dimensional lattice. Work for other graph classes is in progress and we comment on some of the challenges in the Summary section. Here we remark that a major source of challenges for analytic computations is the introduction of *type-3* and *type-4* edges as illustrated in Fig. 1. Lack of symmetry adds additional challenges. The activity of a vertex can be evaluated analytically using the inclusion–exclusion principle, and type-3 and type-4 edges impact the complexity of the combinatorics.

A goal of the work on activity started here is as follows: when given a network X and vertex functions $(f_i)_i$, we would like to rank the vertices by activity in decreasing order. Just being able to identify for example the ten (say) vertices of highest activity would also be very useful. This information would allow one to identify the vertices for which state perturbations are most likely to produce different outcomes, at least in the short term. In a biological experiment for which there is a BN model of the form (1), one may then be able to allocate more resources to the measurement of such states, or perhaps ensure that these states are carefully controlled and locked at their intended values.

1.1 Paper organization

After basic definitions and terminology in Sect. 2 we carefully define a new notion of activity denoted by $\bar{\alpha}_{F,i}$. Basic results to help in analytic evaluations of $\bar{\alpha}_{F,i}$ are given in Sect. 3 followed by specific results for the elementary cellular automata in Sect. 4, d-regular trees in Sect. 5, and then nor- and threshold-BNs over square and triangular lattices in Sect. 6. In Sect. 8 we consider activity for the $G_{n,p}$ random graph model. A central goal is to relate the structure of $X(i;\,2)$ and the functions $(f_j)_j$ to $\bar{\alpha}_{F,i}$. We conclude with open questions and possible directions for followup work in Sect. 9. We remark that this paper is an extended version of work that we presented at Automata 2015 and that appeared in its proceedings.

2 Background, definitions and terminology

In this paper we consider the discrete dynamical systems of the form (1) where each map f_i is of the form $f_i : \{0,1\}^n \longrightarrow \{0,1\}$. However, f_i will in general depend nontrivially only on some subset of the variables $x_1, x_2, \ldots, x_n$, a fact that is captured by the dependency graph defined in the introduction and denoted by X_F or simply X when F is implied. The graph X is generally directed and will contain a loop at every vertex. We will, however, limit ourselves to undirected graphs. A particular graph that we will refer to is the *circle graph on n vertices* denoted by Circle_n and defined by vertex and edge sets

$$\mathrm{v}[\mathrm{Circle}_n] = \{1,2,\ldots,n\}$$
$$\mathrm{e}[\mathrm{Circle}_n] = \{\{i,i+1\} \mid 1 \leqslant i \leqslant n\}\ ,$$

respectively, and where indices are taken modulo 2.

Each vertex i has a *vertex state* $x_i \in K = \{0,1\}$ and a *vertex function* of the form $f_i : \{0,1\}^{d(i)+1} \longrightarrow \{0,1\}$ taking as arguments the states of vertices in the 1-neighborhood of i in X. Here $d(i)$ is the degree of vertex i. We write $n[i]$ for the ordered sequence of vertices contained in the 1-neighborhood of i with i included and set $n(i) = n[i] \setminus \{i\}$. Similarly, we write $x[i]$ for the subsequence of vertex states corresponding to $n[i]$. Finally, we denote the *system state* by $x = (x_1,\ldots,x_n) \in \{0,1\}^n$.

We will write the evaluation of F in (1) as

$$F(x) = (f_1(x[1]), f_2(x[2]), \ldots, f_n(x[n]))\ .$$

The phase space of the map F in (1) is the directed graph $\Gamma(F)$ with vertex set $\{0,1\}^n$ and directed edges all pairs $(x, F(x))$. A state on a cycle in $\Gamma(F)$ is called a *periodic point* and a state on a cycle of length one is a *fixed point*. The sets of all such points are denoted by $\mathrm{Per}(F)$ and $\mathrm{Fix}(F)$ respectively. All other states are *transient states*. Since $\{0,1\}^n$ is finite, the phase space of F consists of a collection of oriented cycles (called periodic orbits), possibly with directed trees attached at states contained on cycles.

In this paper we analyze *short-term stability of dynamics* through the function $\alpha_{F,i} : K^n \longrightarrow \{0,1\}$ defined by

$$\alpha_{F,i}(x) = \mathbb{I}[F(x+e_i) \neq F(x)] \tag{2}$$

where $\mathbb{I}$ is the indicator function and e_i is the ith unit vector with $1 \leqslant i \leqslant n$. In other words, $\alpha_{F,i}(x)$ measures whether perturbing x by e_i results in a different successor state under F than $F(x)$.

Definition 1 The *activity of F with respect to vertex i* is the expected value of $\alpha_{F,i}$ using the uniform measure on K^n:

$$\bar{\alpha}_{F,i} = \mathbb{E}[\alpha_{F,i}]\ . \tag{3}$$

The *activity of F* is the vector

$$\bar{\alpha}_F = (\bar{\alpha}_{F,1}, \bar{\alpha}_{F,2}, \ldots, \bar{\alpha}_{F,n})\ , \tag{4}$$

while the *sensitivity of F* is the average activity $\bar{\alpha} = \sum_{i=1}^n \bar{\alpha}_{F,i}/n$.

For a randomly chosen state $x \in K^n$, the value $\bar{\alpha}_{F,i}$ may be interpreted as the probability that perturbing x_i will cause $F(x+e_i) \neq F(x)$ to hold. This activity notion may naturally be regarded as a measure of sensitivity with respect to initial conditions.

From (2) it is clear that $\bar{\alpha}_{F,i}$ depends on the functions f_j with $j \in n[i]$ and the structure of the distance-2 subgraph $X(i;\,2)$, see Fig. 1. The literature (see, e.g. Shmulevich and Kauffman 2004; Layne et al. 2012) has focused on a very special case when considering activity. Rather than considering the general case and Eq. (2), they have focused on the case where X is a regular graph where each vertex has degree d and all vertex functions are induced by a common function $f : K^{d+1} \longrightarrow K$ through $f_v(x[v]) = f(x[v])$. In this setting, activity is defined with respect to f and its ith argument, that is, as the expectation value of the function $\mathbb{I}[f(x+e_i) \neq f(x)]$ where $x \in K^{d+1}$. Clearly, this measure of activity is always less than or equal to $\mathbb{E}[\alpha_{F,i}]$. Again, we note that this simpler notion of activity does not account for the network structure of $X(i;\,2)$.

3 Preliminary results

For the evaluation of $\bar{\alpha}_{F,i}$ we introduce some notation. In the following we set $K = \{0,1\}$, write N_i for the size of $X(i;\,2)$, and $K(i) = K^{N_i}$ for the projection of K^n onto the set of vertex states associated to $X(i;\,2)$. For $j \in n[i]$, define the sets $A_j(i) \subset K(i)$ by

$$A_j(i) = \{x \in K(i) \mid F(x+e_i)_j \neq F(x)_j\} \; .$$

These sets appear in the evaluation of $\bar{\alpha}_{F,i}$, see Proposition 1. For convenience, we also set

$$A_j^m(i) = \{x \in A_j(i) \mid x_i = m\} \; ,$$

for $m = 0, 1$. We write $\bar{A}_j(i) = A_j^0(i)$ and $\binom{n}{k}$ for binomial coefficients using the convention that it evaluates to zero if either $k < 0$ or $n - k < 0$.

The following proposition provides a somewhat simplified approach for evaluating $\bar{\alpha}_{F,i}$ in the general case.

Proposition 1 *Let X be a graph and F a map over X as in (1). The activity of F with respect to vertex i is*

$$\bar{\alpha}_{F,i} = \Pr\left(\bigcup_{j \in n[i]} A_j \mid x_i = 0\right) = \Pr\left(\bigcup_{j \in n[i]} \bar{A}_j\right) . \tag{5}$$

Proof As stated earlier, we can write

$$\bar{\alpha}_{F,i} = \mathbb{E}[\alpha_{F,i}] = \sum_{x \in K^n} \alpha_{F,i}(x)\Pr(x) = \sum_{x \in K(i)} \alpha_{F,i}(x)\Pr(x)$$
$$= \Pr\left[\bigcup_{j \in n[i]} A_j\right] ,$$

where probabilities in the first and second sum are in K^n and $K(i)$, respectively. This follows by considering the relevant cylinder sets. Equation (5) follows by conditioning on the possible states for x_i. Since we are in the Boolean case, we have a bijection between each pair of sets A_j^0 and A_j^1 for $j \in n[i]$ which immediately allows us to deduce Eq. (5). □

As an example, consider the complete graph $X = K_n$ with threshold functions at every vertex. Recall that the standard Boolean *threshold function*, denoted by $\tau_{k,n} : K^n \longrightarrow K$, is defined by

$$\tau_{k,n}(x_1, \ldots, x_n) = \begin{cases} 1, & \text{if } \sum_{j=1}^n x_j \geqslant k, \\ 0, & \text{otherwise,} \end{cases} \tag{6}$$

where k is the threshold. Here we have

$$\bar{\alpha}_{F,i} = \binom{n-1}{k-1} \Big/ 2^{n-1} \; .$$

To see this, note first that all the sets $\bar{A}_j(i)$ are identical. For a state x with $x_i = 0$ to satisfy $\tau_k(x) \neq \tau_k(x + e_i)$ it is necessary and sufficient that x belong to Hamming class $k - 1$. Since $x_i = 0$, it follows from Proposition 1 that $\Pr(\bar{A}_j(i)) = \frac{1}{2^{n-1}} |\bar{A}_j| = \binom{n-1}{k-1} / 2^{n-1}$ as stated.

Next, consider the Boolean nor-function $\mathrm{nor}_m : K^m \longrightarrow K$ defined by

$$\mathrm{nor}_m(x_1, \ldots, x_m) = (1 + x_1) \cdots (1 + x_m) \; , \tag{7}$$

with arithmetic operations modulo 2. If we use the nor-function over K_n we obtain

$$\bar{\alpha}_{F,i} = 1/2^{n-1} \; .$$

Again, $\bar{A}_j = \bar{A}_k$ for all $j, k \in n[i]$ and we have $F(x + e_i)_i \neq F(x)_i$ precisely when $x_j = 0$ for all $j \in n(i)$ leading to $\Pr(\bar{A}_i) = \frac{1}{2^{n-1}}$. We record the previous two results as a proposition:

Proposition 2 *If F is the Boolean network induced by the* nor*-function over K_n, then*

$$\bar{\alpha}_{F,i} = \frac{1}{2^{n-1}} ,$$

and if F is induced by the k-threshold function over K_n, then

$$\bar{\alpha}_{F,i} = \binom{n-1}{k-1} \Big/ 2^{n-1} \; .$$

In the computations to follow, we will frequently need to evaluate the probability of the union of the A_j's. For this, let B denote the union of A_j's for all $j \neq i$. We then have

$$\begin{aligned} \Pr\left[\bigcup_{j \in n[i]} A_j\right] &= \Pr(A_i \cup B) = \Pr(B) + \Pr(A_i \cap B^c) \\ &= 1 - \Pr(B^c) + \Pr(A_i \cap B^c) \; , \end{aligned} \tag{8}$$

where B^c denotes the complement of B.

4 Activity of elementary cellular automata

The evaluation of Eq. (5) can often be done through the inclusion–exclusion principle. We demonstrate this in the context of elementary cellular automata (ECA). This also makes it clear how the structure of X comes into play for the evaluation of $\bar{\alpha}_{F,i}$.

Let F be the ECA map over $X = \mathrm{Circle}_n$ where each vertex function is given by $f : \{0,1\}^3 \longrightarrow \{0,1\}$. Here we will assume that $n \geqslant 5$; the case $n = 3$ corresponds to the complete graph K_3 and the case $n = 4$ can be handled quite easily. Here we have

$$\bar{\alpha}_{F,i} = \Pr(\bar{A}_{i-1}(i) \cup \bar{A}_i(i) \cup \bar{A}_{i+1}(i)) \; .$$

Applying the definitions,

$$\begin{aligned} \bar{A}_j(i) = \{x = (&x_{i-2}, x_{i-1}, x_i = 0, x_{i+1}, x_{i+2}) \mid \\ &f(x[j]) \neq f((x + e_i)[j])\} \end{aligned} \tag{9}$$

where $j \in n[i] = \{i-1, i, i+1\}$ and with indices modulo n.

Proposition 3 *The activity for a k-threshold ECA is*

$$\bar{\alpha}_{F,i} = \begin{cases} 0, & \text{if } k = 0 \text{ or } k > 3\ , \\ 1/2, & \text{if } k = 1 \text{ or } k = 3\ , \\ 7/8, & \text{if } k = 2\ . \end{cases}$$

Proof Clearly, for $k = 0$ and $k > 3$ we always have $F(x + e_i) = F(x)$ so in these cases it follows that $\bar{\alpha}_{F,i} = 0$. The cases $k = 1$ and $k = 3$ are symmetric (rule r and $255 - r$ have identical activity, see below), and, using $k = 1$, we have

$$\bar{A}_{i-1} = \{(0,0,0,x_{i+1},x_{i+2})\}, \quad \bar{A}_i = \{(x_{i-2},0,0,0,x_{i+2})\}, \\ \text{and } \bar{A}_{i+1} = \{(x_{i-2},x_{i-1},0,0,0)\}\ .$$

By the inclusion–exclusion principle, it follows that

$$\begin{aligned} |\bar{A}_{i-1} \cup \bar{A}_i \cup \bar{A}_{i+1}| &= |\bar{A}_{i-1}| + |\bar{A}_i| + |\bar{A}_{i+1}| \\ &\quad - |\bar{A}_{i-1} \cap \bar{A}_i| - |\bar{A}_{i-1} \cap \bar{A}_{i+1}| - |\bar{A}_i \cap \bar{A}_{i+1}| + |\bar{A}_{i-1} \cap \bar{A}_i \cap \bar{A}_{i+1}| \\ &= 3 \times 4 - 2 - 1 - 2 + 1 = 8\ . \end{aligned}$$

This yields $\bar{\alpha}_{F,i} = 8/2^4 = 1/2$. The proof for the case $k = 2$ is similar to that of $k = 1$ so we leave this to the reader, but see also Table 1. □

Remark 1 For the ECA induced by the nor-function, $\bar{\alpha}_{F,i} = 1/2$ for $n \geqslant 5$.

More generally, an ECA rule $f : \{0,1\}^3 \longrightarrow \{0,1\}$ may be represented by an integer $0 \leqslant r(f) \leqslant 255$, or simply r. Each triple (x_{i-1}, x_i, x_{i+1}) can be viewed as a binary number $0 \leqslant j \leqslant 7$ where x_{i-1} is taken as the most significant digit. Let x^j denote the triple corresponding to the integer and set $a_j = f(x^j)j$. We can then represent f as the 8-tuple $a = (a_0, a_1, \ldots, a_7)$. The associated rule number is $r(f) = \sum_{0 \leqslant j \leqslant 7} a_j \cdot 2^j$. From this and through either inclusion-exclusion or exhaustive enumeration (perhaps easier), one can obtain the activity value of all ECA through conditions on the a_i values. The frequency distribution of activity values for the ECA is derived from the data in Table 1 and is shown in Fig. 2.

Remark 2 In Shmulevich and Kauffman (2004), activity is defined with respect to the function f instead of F and its ith argument. With this context,

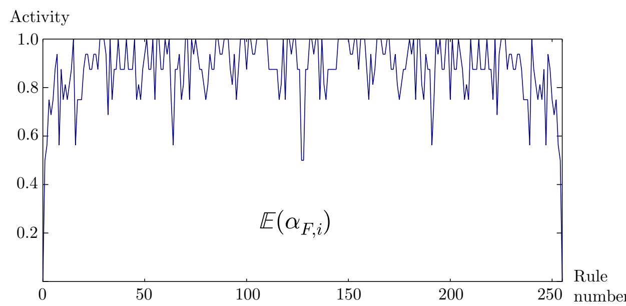

Fig. 2 Activity of ECA by rule number. With the exception of the constant rules of activity 0, all rules have activity ≥1/2 with most rules exceeding 3/4

Table 1 The activity of all ECA rules for $n \geqslant 5$

0	0	1	8	2	9	3	12	4	11	5	12	6	14	7	15
8	9	9	14	10	12	11	13	12	12	13	13	14	14	15	16
16	9	17	12	18	12	19	12	20	14	21	15	22	15	23	14
24	14	25	15	26	15	27	14	28	16	29	16	30	16	31	15
32	11	33	16	34	12	35	14	36	14	37	16	38	14	39	14
40	14	41	16	42	14	43	14	44	14	45	16	46	12	47	13
48	12	49	14	50	15	51	16	52	14	53	14	54	16	55	12
56	16	57	16	58	14	59	14	60	16	61	15	62	16	63	12
64	9	65	14	66	14	67	15	68	12	69	13	70	16	71	16
72	12	73	16	74	15	75	16	76	15	77	14	78	14	79	13
80	12	81	13	82	15	83	14	84	14	85	16	86	16	87	15
88	15	89	16	90	16	91	16	92	14	93	13	94	15	95	12
96	14	97	16	98	16	99	16	100	14	101	16	102	16	103	15
104	15	105	16	106	16	107	16	108	16	109	16	110	16	111	14
112	14	113	14	114	14	115	14	116	12	117	13	118	16	119	12
120	16	121	16	122	15	123	16	124	16	125	14	126	14	127	8

In the table, rule number and activity number alternate. For example, in the lower right corner, "127 8" encodes rule 127 and activity number 8. Here, all activity numbers must be divided by 16 to obtain the actual activity value. Rules in the range $128 \leqslant r \leqslant 255$ are omitted since the activity of rule r and rule $255 - r$ coincide. The reader will observe that the possible activity numbers are $\{0, 8, 9, 11, 12, 13, 14, 15, 16\}$

$\bar{\alpha}_{f,i}(x) = \mathbb{E}[\mathbb{1}[f(x + e_i) \neq f(x)]]$.

Using their definition of activity for ECA with a function of three Boolean variables, the only possible values for activity of ECA are 0, 1/4, 1/2, 3/4, and 1. The fact that our new notion of activity gives $\bar{\alpha} = 7/8$ for threshold-2 Boolean networks demonstrates that network structure does indeed impact results.

Additionally, we always have $\bar{\alpha}_{f,i} \leqslant \bar{\alpha}_{F,i}$. As an example, for a circle graph of girth $\geqslant 5$ and threshold-1 functions (i.e., $k = 1$) it can be verified that $\bar{\alpha}_{f,i} = 1/4$. It was shown above that in this case $\bar{\alpha}_{F,i} = 1/2$.

5 Activity over d-regular trees

A natural starting point for analyzing activity is the case where the graph X is a d-regular tree. The reason is that in this case the sets A_j (or more precisely, the sets $n[j]$) with $j \neq i$ only have vertex i in common. As a result, when we condition on $x_i = m$, the resulting sets are independent. This fact will hold if the girth of the graph is at least 5, so the results we give are applicable to this broader graph class. We remark that a d-regular tree is an infinite tree where all vertices have degree d. Alternatively, one may consider this to be finite trees where all vertices either have degree d or degree 1. In the latter case, our result only applies to vertices whose neighbors all have degree d.

The Boolean bi-threshold function $\tau_{i,k_{01},k_{10},n} : K^n \longrightarrow K$ generalizes standard threshold functions and is defined by

$$\tau_{i,k_{01},k_{10},n}(x_1, \ldots, x_n) = \begin{cases} 1, & \text{if } x_i = 0 \text{ and } \sum_{j=1}^{n} x_j \geqslant k_{01} \\ 0, & \text{if } x_i = 1 \text{ and } \sum_{j=1}^{n} x_j < k_{10} \\ x_i, & \text{otherwise,} \end{cases} \tag{10}$$

where the integers k_{01} and k_{10} are the up- and down-thresholds, respectively. Here i is a designated vertex – it will be the index of a vertex function. If the up- and down-thresholds are the same we get the standard threshold systems.

Proposition 4 *Let X be a d-regular graph of girth ≥ 5, and F the Boolean network over X with the bi-threshold vertex functions as in* Eq. (10). *Then the activity of F with respect to vertex i is given by*

$$\bar{\alpha}_{F,i} = 1 - \frac{1}{2^{d^2}}\left[2^d - \binom{d-1}{k_{01}-1} - \binom{d-1}{k_{10}-2}\right]^d + \frac{1}{2^{d^2+1}} \sum_{k=k_{01},k_{10}} \binom{d}{k-1}\left[2^{d-1} - \binom{d-1}{k_{10}-2}\right]^{k-1} \left[2^{d-1} - \binom{d-1}{k_{01}-1}\right]^{d-(k-1)}. \tag{11}$$

Proof Let B denote the union of A_j's for all $j \neq i$ as in Eq. (8) and note that B^c is the event that none of the A_j's occur for $j \neq i$. Using girth $\geqslant 5$ and the resulting independence from conditioning on x_i we can write

$$\Pr(B^c) = \Pr\left(\bigcap_{j\in n(i)} A_j^c \mid x_i = 0\right) = \prod_{j\in n(i)} \Pr\left(A_j^c \mid x_i = 0\right). \tag{12}$$

Next we have

$$\begin{aligned}\Pr\left(A_j^c \mid x_i = 0\right) &= \frac{1}{2}\sum_{r\in\{0,1\}} \Pr\left(A_j^c \mid x_i = 0 \,\mathrm{and}\, \mathrm{x_j} = \mathrm{r}\right) \\ &= \frac{1}{2}\cdot\frac{2^{d-1} - \binom{d-1}{k_{01}-1}}{2^{d-1}} + \frac{1}{2}\cdot\frac{2^{d-1} - \binom{d-1}{k_{10}-2}}{2^{d-1}} \\ &= \left[2^d - \binom{d-1}{k_{01}-1} - \binom{d-1}{k_{10}-2}\right]\Big/2^d,\end{aligned}$$

which substituted into Eq. (12) yields

$$\Pr(B^c) = \left[2^d - \binom{d-1}{k_{01}-1} - \binom{d-1}{k_{10}-2}\right]^d \Big/ 2^{d^2}. \tag{13}$$

Next, we compute the probability of $A_i \cap B^c$ as:

$$\begin{aligned} 2^{d^2+1} \times \Pr(A_i \cap B^c) &= 2^{d^2} \sum_{r=0,1} \Pr(A_i \cap B^c \mid x_i = r) \\ &= 2^{d^2} \sum_{r=0,1} \Pr(A_i \mid x_i = r)\cdot\Pr(B^c \mid A_i \,\mathrm{and}\, \mathrm{x_i} = \mathrm{r}) \\ &= \binom{d}{k_{01}-1}\left[2^{d-1} - \binom{d-1}{k_{10}-2}\right]^{k_{01}-1} \left[2^{d-1} - \binom{d-1}{k_{01}-1}\right]^{d-(k_{01}-1)} + \binom{d}{k_{10}-1} \\ &\quad \left[2^{d-1} - \binom{d-1}{k_{10}-2}\right]^{k_{10}-1}\left[2^{d-1} - \binom{d-1}{k_{01}-1}\right]^{d-(k_{10}-1)} \\ &= \sum_{k=k_{01},k_{10}} \binom{d}{k-1}\left[2^{d-1} - \binom{d-1}{k_{10}-2}\right]^{k-1} \\ &\quad \times \left[2^{d-1} - \binom{d-1}{k_{01}-1}\right]^{d-(k-1)} \end{aligned} \tag{14}$$

Substituting Eqs. (13) and (14) into Eq. (8) leads to Eq. (11), which ends the proof. □

Since the standard threshold function is a special case of the bi-threshold function when k_{01} and k_{10} coincide, it is straightforward to obtain the following corollary.

Corollary 1 *If F is the Boolean network with k-threshold vertex functions over a d-regular graph of girth ⩾5, then the activity of F with respect to vertex i is*

$$\bar{\alpha}_{F,i} = 1 - \left[1 - \binom{d}{k-1}\Big/2^d\right]^d + \frac{\binom{d}{k-1}}{2^d}\left[1 - \binom{d-1}{k-2}\Big/2^{d-1}\right]^{k-1} \times \left[1 - \binom{d-1}{k-1}\Big/2^{d-1}\right]^{d-(k-1)} . \qquad (15)$$

We next consider the nor-function.

Proposition 5 *Let X be a d-regular graph of girth ⩾5 and F the Boolean network over X induced by the* nor*-function. Then the activity of F with respect to i is given by*

$$\bar{\alpha}_{F,i} = 1 - \left(1 - \frac{1}{2^d}\right)^d + \left(\frac{1}{2} - \frac{1}{2^d}\right)^d . \qquad (16)$$

Proof Conditioning on $x_i = 0$ and using independence we have

$$\Pr(B^c) = \Pr\left(\bigcap_{j\in n(i)} A_j^c\right) = \prod_{j\in n(i)} \Pr\left(A_j^c\right) . \qquad (17)$$

For a state x with $x_i = 0$ to be in $\bar{A}_j$, all remaining d states of $\bar{A}_j$ must be zero, leading to $\Pr\left(\bar{A}_j^c\right) = 1 - \frac{1}{2^d}$, and

$$\Pr(\bar{B}^c) = \left(1 - \frac{1}{2^d}\right)^d . \qquad (18)$$

In order to calculate $\Pr(A_i \cap B^c)$ we note that $\Pr(A_i \cap B^c) = \Pr(A_i)\Pr(B^c|A_i)$ and again use independence to obtain

$$\Pr(\bar{B}^c|\bar{A}_i) = \prod_{j\in n(i)} \Pr\left(\bar{A}_j^c|\bar{A}_i\right) . \qquad (19)$$

Note that $\Pr(\bar{A}_j|\bar{A}_i) = 1/2^{d-1}$ so that $\Pr\left(\bar{A}_j^c|\bar{A}_i\right) = 1 - \Pr\left(\bar{A}_j|\bar{A}_i\right) = 1 - 1/2^{d-1}$ which substituted into (19) gives

$$\Pr(\bar{B}^c|\bar{A}_i) = \left(1 - \frac{1}{2^{d-1}}\right)^d .$$

Noting that $\Pr(\bar{A}_i) = 1/2^d$ we obtain the third term in (15), finishing the proof. □

6 Activity over square lattices

In this section we consider graphs with girth-4 edges or girth 4. Here the 1-neighborhoods $n[j]$ with $j \neq i$ may intersect, the key aspect we want to address here. As an example, the reader may verify that for threshold-2 functions and Circle$_4$, the activity of any vertex is 3/4 and not 7/8 as when $n \geqslant 5$ in Proposition 3 (Fig. 3).

To be specific we take the graph X to be a regular, square 2-dimensional lattice, see Fig. 4. It may be either infinite or with periodic boundary conditions. In the latter case, we assume for simplicity that its two dimensions are at least 5. This graph differs from the 4-regular tree by the introduction of type-4 edges: sets A_j and A_{j+1} of $n(i)$, when conditioned on the state of vertex i, are no longer independent. The graph has cycles of size 4 containing i. We first illustrate this case using nor-functions as these allow for a somewhat simplified evaluation as compared to what happens for threshold functions.

Proposition 6 Let X be the 2-dimensional lattice as above where every vertex has degree 4, and let F be the Boolean network over X induced by nor-functions. Then the activity of F for every vertex i is $\bar{\alpha}_{F,i} = 1040/2^{12} = 0.254$.

Proof The neighborhoods of i involved are illustrated in Fig. 4. We rewrite Eq. (8) as follows

$$\Pr\left(\bigcup_{j\in n[i]} A_j\right) = \Pr(B) + \Pr(A_i)(1 - \Pr(B|A_i)) . \qquad (20)$$

To determine $|B|$, refer to Fig. 4. With indices $p, q, r, s \in \{1, 2, 3, 4\}$ we have

$$|A_1 \cup A_2 \cup A_3 \cup A_4| = \sum_{p=1}^{4} A_p - \sum_{p<q} |A_p \cap A_q| + \sum_{p<q<r} |A_p \cap A_q \cap A_r| - \sum_{p<q<r<s} |A_p \cap A_q \cap A_r \cap A_s| , \qquad (21)$$

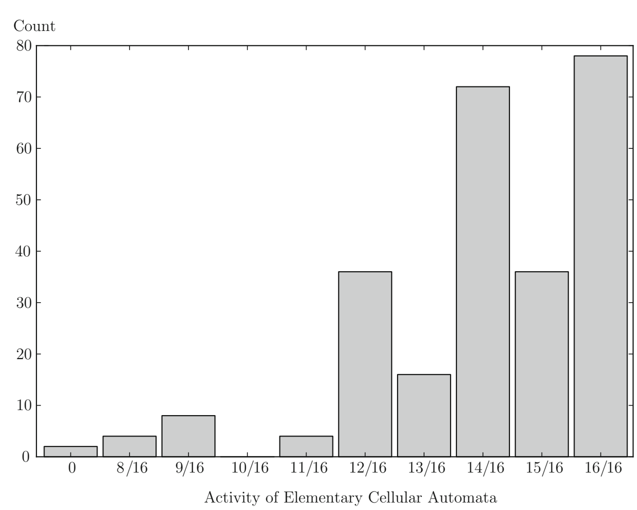

Fig. 3 A diagram showing the possible activity values of ECA (*horizontal axis*) and their frequency (*vertical axis*)

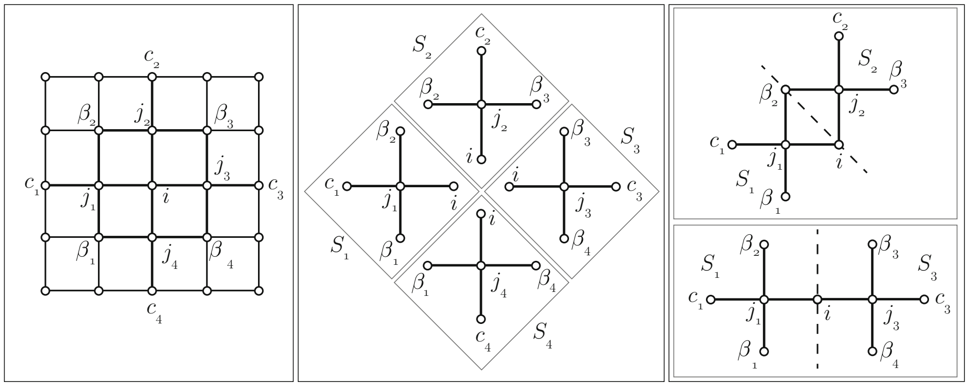

Fig. 4 A subgraph of the square lattice including $X(i;2)$ at center (*left*). The four adjacent sets of A_{j_ℓ} correspond to the subgraphs S_ℓ that are illustrated in the middle as templates. Finally, on the *right*, the two possible intersections of 2 adjacent sets/templates are illustrated. For one set/template (middle) or the intersection of 3 sets (not shown), there is only one possible configuration modulo rotational symmetry

which, along with Eq. (20), is also valid if we condition on $x_i = 0$ in every set. We take $i \neq j_\ell$ with $\ell \in \{1,2,3,4\}$ and let j_ℓ denote the vertex at the center of $n[j_\ell]$ and i the vertex at the center of $n[i]$. Here $|X(i;2)| = 13$. Using the symmetry of the sets $\bar{A}_{j_\ell}$ and the fact that a state x is in the set $\bar{A}_{j_\ell}$ precisely when $x[j_\ell] = 0$ leads to

$$\sum_{\ell=1}^{4} |\bar{A}_{j_\ell}| = 4 \cdot 2^8 \ .$$

For the second term on the right in (21) we have two cases for intersections: (*i*) two adjacent vertices $j_\ell, j_{\ell+1} \in n(i)$ with all indices modulo 4, of which there are four instances, and (*ii*) two non-adjacent vertices $j_\ell, j_{\ell+2} \in n(i)$, of which there are two instances. In the first case we obtain $|\bar{A}_{j_\ell} \cap \bar{A}_{j_{\ell+1}}| = 2^5$, while in the second case we have $|\bar{A}_{j_\ell} \cap \bar{A}_{j_{\ell+2}}| = 2^4$, making the second term on the right in (21)

$$\sum_{p<q} |\bar{A}_{j_p} \cap \bar{A}_{j_q}| = 4 \cdot 2^5 + 2 \cdot 2^4 \ . \tag{22}$$

The third sum appearing on the right in (21) contains four terms all corresponding to an intersection of three sets A_{j_ℓ}, and we obtain

$$\sum_{p<q<r} |\bar{A}_{j_p} \cap \bar{A}_{j_q} \cap \bar{A}_{j_r}| = 4 \cdot 2^2 \ . \tag{23}$$

Finally, the intersection of all four sets $\bar{A}_{j_\ell}$ has size 1, and we get

$$\Pr(\bar{B}) = |\bar{B}|/2^{12} = \left[4 \cdot 2^8 - (4 \cdot 2^5 + 2 \cdot 2^4) + 4 \cdot 2^2 - 1\right]/2^{12} \ . \tag{24}$$

We next determine $\Pr(B|A_i)$. For a state x to be in A_i we must have $x[i] = 0$ so that $|A_i| = 2^8$. Conditioning on this, we can calculate $|\bigcup_{\ell=1}^{4} A'_{j_\ell}|$ as above. The reader can verify that

$$\begin{aligned} \sum_{\ell=1}^{4} |A'_{j_\ell}| &= 4 \cdot 2^5 \ , \\ \sum_{p<q} |A'_{j_p} \cap A'_{j_q}| &= 4 \cdot 2^3 + 2 \cdot 2^2 \ , \\ \sum_{p<q<r} |A'_{j_p} \cap A'_{j_q} \cap A'_{j_r}| &= 4 \cdot 2 \ , \\ |A'_{j_1} \cap A'_{j_2} \cap A'_{j_3} \cap A'_{j_4}| &= 1 \ , \end{aligned} \tag{25}$$

leading to the expression

$$\begin{aligned} \Pr(\bar{A}_i \cap \bar{B}^c) &= \Pr(\bar{A}_i)(1 - \Pr(\bar{B}|\bar{A}_i)) \\ &= \frac{1}{2^4}\left(1 - \frac{1}{2^8}\left[2^7 - 2^5 - 2^3 + 2^3 - 1\right]\right) \ , \end{aligned}$$

which, together with (24) in (20) gives the stated result of $\bar{\alpha}_{F,i} = 1040/2^{12}$. □

6.1 Activity of threshold functions over square lattices

We next consider the square lattice where each vertex function is the k-threshold function. For this case, the analytic derivations resemble those of the nor-case just covered, but the expressions appearing are more involved.

Proposition 7 *Let X be a rectangular grid (torus) in which every vertex has degree 4, and F the map over X with the threshold vertex functions. Then the activity of F with respect to vertex i is given by*

$$\alpha_{F,i} = \left(\frac{1}{2^{12}}\right)[\gamma_1 - \gamma_2 + \gamma_3 - \gamma_4 + \gamma_5] \ , \tag{26}$$

where

$$\gamma_1 = 2^{10} \cdot \binom{4}{k-1} \ , \tag{27}$$

$$\gamma_2 = 2^7\left[\binom{3}{k-1}^2 + \binom{3}{k-2}^2\right] + 2^5\binom{4}{k-1}^2, \tag{28}$$

$$\gamma_3 = 2^4\left[\binom{3}{k-1}^2\binom{2}{k-1} + 2\binom{3}{k-2}\binom{2}{k-2}\binom{3}{k-1} + \binom{3}{k-2}^2\binom{2}{k-3}\right], \tag{29}$$

$$\gamma_4 = \left[\binom{2}{k-1}^4 + 4\binom{2}{k-1}^2\binom{2}{k-2}^2 + 4\binom{2}{k-1}\binom{2}{k-2}^2\binom{2}{k-3} + 2\binom{2}{k-2}^4 + 4\binom{2}{k-2}^2\binom{2}{k-3}^2 + \binom{2}{k-3}^4\right], \tag{30}$$

and

$$\gamma_5 = \sum_{x\in M} \mathbb{I}\left[x_{j_\ell} + x_{\beta_\ell} + x_{\beta_{\ell+1}} + x_{c\ell} \neq k-1\right], \tag{31}$$

where M is the subset of states associated to $X(i;\,2)$ for which the Hamming norm of its elements projected to the components j_ℓ with $1 \leqslant \ell \leqslant 4$ is $k-1$, that is, $|(x_{j_1}, x_{j_2}, x_{j_3}, x_{j_4})|_H = k-1$, and where x_{j_ℓ}, x_{β_ℓ} and x_{c_ℓ} are the states of the vertices j_ℓ, β_ℓ and c_ℓ as given in Fig. 4.

Proof We begin with Eq. (8), and by the symmetry of thresholds, we can condition on $x_i = 0$, to write

$$\bar{\alpha}_{F,i} = \Pr(B \mid x_i = 0) + \Pr(A_i \cap B^c \mid x_i = 0)\ . \tag{32}$$

Here, B is the union of A_{j_ℓ}s, and we take $i \neq j_\ell$ for all $\ell \in \{1,2,3,4\}$. We have A_{j_ℓ} as a particular A_j. Let j_ℓ denote the vertex at the "center" of subgraph S_ℓ with $d = 4$ (per Fig. 4), as vertex i is the "center" of S_i, and let x_{j_ℓ} denote the state of j_ℓ. There is a one-to-one correspondence between S_ℓ and A_{j_ℓ} since the vertices associated with each are the same subset of those of $n[j_\ell]$. We evaluate $\Pr(B \mid x_i = 0)$ and $\Pr(A_i \cap B^c \mid x_i = 0)$ in turn, using the "templates" formed from the S_ℓ.

The first four terms γ_r with $1 \leq r \leq 4$ in Eq. (26), come from the first term on the right hand side (RHS) of Eq. (32). We have, by symmetry in the $0 \to 1$ and $1 \to 0$ transitions for x_i,

$$\Pr(B \mid x_i = 0) = \Pr\left[\bigcup_{\substack{j_\ell \in n[i] \\ j_\ell \neq i}} (A_{j_\ell} \mid x_i = 0)\right] = |A_1 \cup A_2 \cup A_3 \cup A_4|/2^{12}\ ,$$

where the 12 vertices in the template of Fig. 4 can be in either state 0 or 1 ($x_i = 0$ is fixed).

The inclusion/exclusion principle can be used to evaluate the numerator in the RHS of this last expression, which is independent of the girth of the graph. We also note that the last expression of the union of the A_{j_ℓ}s is subject to the condition $x_i = 0$. We use the templates in Fig. 4, without loss of generality, to compute the value of the last equation. Each of the expressions for γ_1 through γ_4 take the same form, which is a product of the multiplicity of the particular template in $X(i;\,2)$; the number z of the 12 vertices whose states can be either zero or one because they do not play a role in the particular template being evaluated (giving rise to a term 2^z); and an expression in binomial coefficients that represent the number of sets of vertex states in the template whose value must be 1. In this case, the first summation on the RHS of (21) evaluates to

$$\gamma_1 = \sum_{p=1}^{4} |A_p| = 4 \cdot 2^{12-4} \cdot \binom{4}{k-1}, \tag{33}$$

using symmetry across A_{j_ℓ} with $\ell \in \{1,2,3,4\}$ of Fig. 4, center diagram. The multiplicity is 4 because there are four A_{j_ℓ} instances. The 2^{12-4} comes from the fact that there are 12 vertices in $\bar{K}^{N_i}$ whose states can be chosen, and we specify 4 vertex states in each template S_ℓ. The 4 in the binomial comes from the fact that there are four vertices in each S_ℓ that can be assigned a value in $\{0,1\}$ independently of the state assignments to the other vertices (we call these free vertices), noting $x_i = 0$, and $k-1$ comes from the fact that when x_i changes from state 0 to 1 (i.e., when its threshold is satisfied), A_{j_ℓ} is satisfied.

For the second term on the RHS of Eq. (21), we have two cases (see the right-most drawings of Fig. 4): (*i*) four instances with A_{j_ℓ} and $A_{j_{\ell+1}}$ (all indices are modulo 4) with a total of 7 vertices in each instance that can take states 0 and 1 (so that $z = 12 - 7$); and (*ii*) two instances with A_{j_ℓ} and $A_{j_{\ell+2}}$ where each instance has 8 vertices that can take states 0 and 1, giving $z = 12 - 8$. Hence, the β_ℓ are not free vertices for case (*i*), but they are for case (*ii*). In case (*i*), there is a common vertex, $\beta_{\ell+1}$, and hence there are three free vertices in each S_ℓ. Fixing $x_{\beta_{\ell+1}}$ in turn, to be 0 and then 1, we get $|A_{j_\ell} \cap A_{j_{\ell+1}}| = 4 \cdot 2^{12-7}\left[\binom{3}{k-1}^2 + \binom{3}{k-2}^2\right]$. For case (*ii*), there is no common vertex beyond i, so we have four free vertices in each of A_{j_ℓ} and $A_{j_{\ell+2}}$. We get $|A_{j_\ell} \cap A_{j_{\ell+2}}| = 2 \cdot 2^{12-8}\binom{4}{k-1}^2$. In total, the second summation on the RHS of Eq. (21) is $\sum_{p<q} |A_p \cap A_q| = \gamma_2$, with γ_2 given in Eq. (28).

The third summation in the RHS of Eq. (21) has a single case: the intersection of three of the four A_{j_ℓ}s. There are four instances, with each instance composed of 10 vertices to which states may be assigned, giving a multiplier of $4 \cdot 2^{12-10} = 2^4$ in γ_3 in Eq. (29). Without loss of generality, we focus on A_{j_ℓ}, $\ell \in \{1,2,3\}$ of Fig. 4. Here, A_{j_1} has three free vertices (β_{j_2} is common to A_{j_1} and A_{j_2}); A_{j_2} has two free vertices (β_{j_2} and β_{j_3} are common to A_{j_1} and A_{j_3}, respectively); and A_{j_3} has three free vertices. The three terms in $[\cdot]$ in Eq. (29) are produced by assigning the state pairs $(x_{\beta_2}, x_{\beta_3})$ equal to, in turn, (*a*) (0,0); (*b*) (0,1) and then (1,0); and (*c*) (1,1).

The final summation in the RHS of Eq. (21) also has a single case: the intersection of all four A_{j_ℓ}s. There is one instance. All four β_ℓ are considered, so we take as subcases the number of β_ℓ where $x_{\beta_\ell} = 1$, is, in turn, 0, 1, 2, 3, and 4. It can be shown, following analogous procedures to those for the expression for γ_3, that $\sum_{p<q<r<s} |A_p \cap A_q \cap A_r \cap A_s| = \gamma_4$; see Eq. (30).

To evaluate $\Pr(A_i \cap B^c \mid x_i = 0)$ in Eq. (32), we note that $B^c = \left[\bigcap_{j_\ell \in n(i)} A_{j_\ell}^c \right]$. The states contained in $A_i \cap B^c$ are therefore those contained in A_i but none of the four sets A_{j_ℓ}.

For a given k, consider the state x. We see that $x \in A_i$ if and only if exactly $k-1$ of the adjacent vertices of i have state $x_{j_\ell} = 1$ and the remaining $4-(k-1)$ adjacent states satisfy $x_{j_\ell} = 0$. For $x \notin A_{j_\ell}$ to hold, it is necessary and sufficient that

$$x_{j_\ell} + x_{\beta_\ell} + x_{\beta_{\ell+1}} + x_{c\ell} \neq k-1\ , \tag{34}$$

for $1 \leqslant \ell \leqslant 4$. We can partition these cases as follows: First, consider all 16 elements $x_\beta = (x_{\beta_1}, x_{\beta_2}, x_{\beta_3}, x_{\beta_4}) \in \{0,1\}^4$ (the four overlapping vertices $\beta = (\beta_1, \beta_2, \beta_3, \beta_4)$ in Fig. 4). For each x_β, consider all $\hat{x}_j = (x_{j_1}, x_{j_2}, x_{j_3}, x_{j_4}) \in \{0,1\}^4$ of adjacent vertex states of i of Hamming norm $|\hat{x}_j| = k-1$. Finally, we consider all $x_c = (x_{c_1}, x_{c_2}, x_{c_3}, x_{c_4}) \in \{0,1\}^4$. We thus count the the number of elements $x \in M$ for which Eq. (34) holds, producing γ_5 in Eq. (31) and completing the proof. □

We have shown the graph of the activity as a function of the threshold value k in Fig. 5. Note that for thresholds in the range $1<k<5$, the activity for the threshold function is considerably larger than that for the nor-function given immediately above (Table 2; Fig. 5).

Table 2 Activity value (rounded to four decimal places) as a function of the threshold k for the 4-regular square lattice and the 4-regular tree. Thresholds 0 and 6 omitted

Graph k	1	2	3	4	5
4-lattice	0.2539	0.7095	0.9009	0.7095	0.2539
4-regular tree	0.2641	0.7370	0.9046	0.7370	0.2641

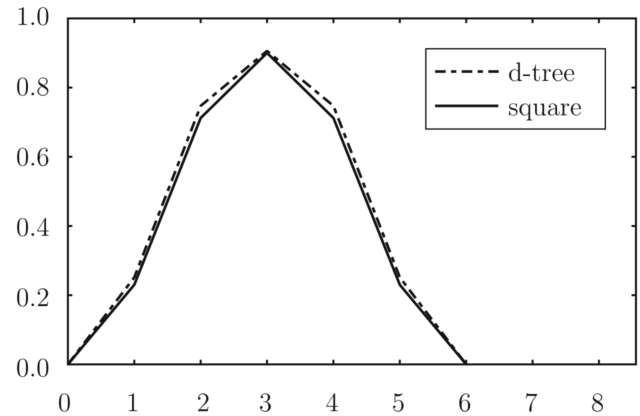

Fig. 5 Activity as a function of the threshold value k for the square lattice of Fig. 4 (*solid line*). For comparison, the activity for the 4-regular tree is included (*dashed line*)

7 Activity over triangular lattices

A triangular lattice is a regular graph with type-3 edges and a girth of 3 as indicated in the left of Fig. 6. In this case, the structure of the intersections among sets A_j (with $j \neq i$) is more involved than for the square lattice case, and type-3 edges are introduced. In principle, the derivations are similar to those of the square lattice case. Rather than giving a similar argument to the proof of Proposition 7, we limit ourselves to showing the graph of the activity as a function of the threshold value k, see Fig. 6 right side, but see also Table 3 for the approximate values.

8 Activity over the Erdős–Rényi random graphs

Here, we provide an upper bound for the expected activity in $G_{n,p}$ for the uniform k-threshold function using the union bound. While bound obtained this way may not turn out to strong, the work in this section takes the first steps towards analyzing activity over ER random graphs.

First, let $B[n, p]$ denote the binomial random variable and let

$$f(t;n,p) = \Pr(B[n,p] = t) = \binom{n}{t} p^t (1-p)^{n-t}\ .$$

Since we are considering a family of graphs, we find it necessary to redefine some of the terminology developed earlier. Let i be the vertex under consideration. Let $\bar{\alpha}_{F,i}(G)$ denote the activity for $G \in G_{n,p}$. The expected activity $\mathbb{E}[\bar{\alpha}_{F,i}] := \mathbb{E}_{G(n,p)}[\bar{\alpha}_{F,i}]$.

Note that, by symmetry $\bar{\alpha}_{F,i}$ is the same for all $i \in \{1,\ldots,n\}$. Next, recall the definition of $\bar{A}_j(i)$ (or $\bar{A}_j$ in short) from Section 3. Here, we modify it as $\bar{A}_j(G)$ to denote the set of states x which satisfy $x_i = 0$ and $F(x + e_i)_j \neq F(x)_j$ in G.

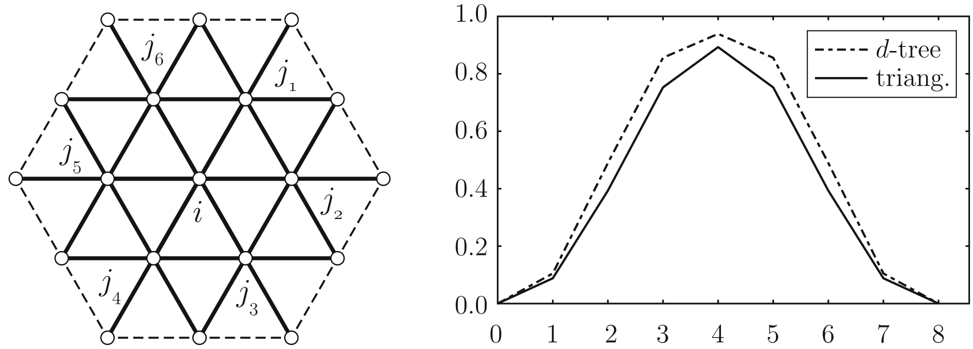

Fig. 6 *Left*: the triangular lattice. *Right*: activity as a function of threshold for the triangular lattice (*solid line*). For comparison, the activity for the 6-regular tree is included (*dashed line*)

Table 3 Activity value (rounded to four decimal places) as a function of the threshold k for the triangular lattice. Thresholds 0 and 8 omitted

Graph k	1	2	3	4	5	6	7
Triangular lattice	0.0881	0.3963	0.7603	0.9019	0.7603	0.3963	0.0881

Proposition 8 If F is the Boolean network induced by the k-threshold function over $G(n, p)$, then $\mathbb{E}[\bar{\alpha}_{F,i}] \leqslant f(k-1; n-1, p/2) + \frac{(n-1)p}{2}(f(k-1; n-1, p/2) + f(k-2; n-2, p/2))$.

Proof Let $n(i, G)$ and $n[i, G]$ denote the open and closed neighborhoods of i in G.

Using the union bound on (5), we have

$$\mathbb{E}[\bar{\alpha}_{F,i}] = \sum_{G \in G_{n,p}} \Pr(G) \Pr\left(\bigcup_{j \in n[i,G]} \bar{A}_j(G)\right) \tag{35}$$

$$\leqslant \sum_{G \in G_{n,p}} \Pr(G) \Pr(\bar{A}_i(G)) + \sum_{G \in G_{n,p}} \Pr(G) \Pr\left(\bigcup_{j \in n(i,G)} \bar{A}_j(G)\right) \tag{36}$$

$$\leqslant \sum_{G \in G_{n,p}} \Pr(G) \Pr(\bar{A}_i(G)) + \sum_{G \in G_{n,p}} \Pr(G) \sum_{j \in n(i,G)} \Pr(\bar{A}_j(G)) . \tag{37}$$

Now we will compute the two summands separately. First, we note that the term $\sum_{G \in G_{n,p}} \Pr(G) \Pr(\bar{A}_i(G))$ is precisely the probability that a randomly chosen pair (x, G) satisfies $x \in \bar{A}_i(G)$. By randomly chosen, we mean that (x, G) is sampled by choosing independently x from K^n and G from $G_{n,p}$. This probability can be computed easily as follows. For any $x \in \bar{A}_i(G)$, i must have exactly $k-1$ neighbors in state 1. For any vertex $j \neq i$, the probability that it is a neighbor of i and is in state 1 is $p/2$. Therefore, the number of neighbors of i in state 1 is binomially distributed with parameters $n-1$ and $p/2$, i.e., $B[n-1, p/2]$. Hence,

$$\sum_{G \in G_{n,p}} \Pr(G) \Pr(\bar{A}_i(G)) = \Pr(B[n-1, p/2] = k-1)$$
$$= f(k-1; n-1, p/2) . \tag{38}$$

Next, we will compute $\sum_{G \in G_{n,p}} \Pr(G) \sum_{j \in n(i,G)} \Pr(\bar{A}_j(G))$. Rearranging the summation,

$$\sum_{G \in G_{n,p}} \Pr(G) \sum_{j \in n(i,G)} \Pr(\bar{A}_j(G)) = \sum_{j \neq i} \sum_{G \in G_{n,p}} \Pr(G) \Pr(\bar{A}_j(G)) .$$

As in the previous case, the term $\sum_{G \in G_{n,p}} \Pr(G) \Pr(\bar{A}_j(G))$ corresponds to the probability that a randomly chosen pair (x, G) satisfies $x \in \bar{A}_j(G)$. This probability can be computed as follows: For (x, G) to satisfy the above condition, j must be a neighbor of i and satisfy one of these conditions: (i) if $x_j = 0$, then, j must have exactly $k-1$ neighbors excluding i in state 1, and (ii) if $x_j = 1$, then, j must have exactly $k-2$ neighbors, excluding i, in state 1. Let Y_j^1 be the number of neighbors of j excluding i at state 1 in x. Note that this is binomially distributed with parameters $n-2$ and $p/2$. Therefore,

$$\sum_{G \in G_{n,p}} \Pr(G) \sum_{j \in n(i,G)} \Pr(\bar{A}_j(G))$$
$$\big(\Pr(x_j = 0) \Pr(x \in \bar{A}_j \mid x_j = 0)$$
$$+ \Pr(x_j = 1) \Pr(x \in \bar{A}_j \mid x_j = 1)\big)$$
$$= \frac{p}{2}\Big(\Pr(Y_j^1 = k-1) + \Pr(Y_j^1 = k-2)\Big)$$
$$= \frac{p}{2}(f(k-1; n-2, p/2) + f(k-2; n-2, p/2)) \tag{39}$$

Finally, applying (38) and (39) in (37), we have

$$\mathbb{E}[\bar{\alpha}_{F,i}] \leqslant f(k-1;n-1,p/2) + \sum_{j\neq i}\frac{p}{2}(f(k-1;n-1,p/2) + f(k-2;n-2,p/2)) = f(k-1;n-1,p/2) + \frac{(n-1)p}{2}(f(k-1;n-1,p/2) + f(k-2;n-2,p/2)).$$

Hence proved. □

9 Summary and research directions

In this extended version of Adiga et al. (2015) we have incorporated new and additional results. Our notion of activity extends what was proposed by Shmulevich and Kauffman Shmulevich and Kauffman (2004). This extension takes into account the impact of the network structure when studying $\mathbb{E}\left[\mathbb{I}[F(x) \neq F(x+e_i)]\right]$, which estimates how likely the perturbation e_i will cause successor states to diverge after one time step.

Naturally, orbits that initially separate may later converge, reflecting that x and $x+e_i$ may belong to the same attractor basin. Nonetheless, this notion of activity provides a measure for sensitivity with respect to initial conditions that accounts for network structure. Investigating *long-term activity*, that is, the probability that perturbing a state in the ith coordinate will cause x and $x+e_i$ to have different ω-limit sets, is an interesting direction for future research. Involving the attractors of F, we naturally expect this to be challenging work, even in most special cases.

In this paper we have started analysis for activity over random graphs. Future research could include extensions of this work to provide stronger bounds. Analysis of dynamics of Boolean networks, as well as the more general class of graph dynamical systems, in the setting of random graphs also seem to hold much potential for interesting insight. Our results also show that the activity for the d-regular trees using threshold functions is larger than those of the corresponding square and triangular grids. Clearly, the combinatorial arguments become more involved with the presence of type-3 and type-4 edges. In general, what can be said about the impact of these classes of edges on activity?

For efficient computations of activity one may consider using model counting SAT solvers. This was not explored by us, however, it was commented by a reviewer. For networks where $n_i := |X(i;2)|$ is "not too large", one may consider developing a tool to map from vertex functions and the expressions involved in the definitions of the sets A_k to suitable formats for this class of SAT solvers. For additional directions for future work we refer to Adiga et al. (2015).

Acknowledgments We thank our external collaborators and members of the Network Dynamics and Simulation Science Laboratory (NDSSL) for their suggestions and comments. We also thank two anonymous reviewers for valuable comments. This work has been partially supported by DTRA Grant HDTRA1-11-1-0016.

References

Adiga A, Kuhlman CJ, Mortveit HS, Vullikanti AKS (2013) Sensitivity of diffusion dynamics to network uncertainty. In: Proceedings of the twenty-seventh AAAI conference on artificial intelligence (AAAI-13), July 14–18, 2013. Bellevue, Washington, USA, pp 2–8

Adiga A, Galyean H, Kuhlman CJ, Levet M, Mortveit HS, Wu S (2015) Network structure and activity in Boolean networks. In: Kari J (ed) Cellular automata and discrete complex systems: proceedings of AUTOMATA 2015, Turku, Finland, June 8–10, 2015, Lecture Notes in Computer Science, vol 9099, pp 210–223, doi:10.1007/978-3-662-47221-7_16

Aldana M, Coppersmith S, Kadanoff LP (2003) Boolean dynamics with random couplings. In: Perspectives and problems in nonlinear science, Springer, pp 23–89

Baetens JM, De Baets B (2010) Phenomenological study of irregular cellular automata based on Lyapunov exponents and Jacobians. Chaos 20:1–15. doi:10.1063/1.3460362

Baetens JM, Van der Weeën P, De Baets B (2012) Effect of asynchronous updating on the stability of cellular automata. Chaos Solitons Fractals 45:383–394. doi:10.1016/j.chaos.2012.01.002

Derrida B, Pomeau Y (1986) Random networks of automata: a simple annealed approximation. Europhys Lett 1:45–49

Fretter C, Szejka A, Drossel B (2009) Perturbation propagation in random and evolved Boolean networks. N J Phys 11:1–13. doi:10.1088/1367-2630/11/3/033005

Ghanbarnejad F, Klemm K (2012) Impact of individual nodes in Boolean network dynamics. EPL (Europhys Lett) 99(5):58,006

Goles E, Martinez S (1990) Neural and automata networks: dynamical behaviour and applications. Kluwer Academic Publishers, Berlin

Kauffman SA (1969) Metabolic stability and epigenesis in randomly constructed genetic nets. J Theor Biol 22:437–467

Kauffman SA (1993) The origins of order: self-organization and selection in evolution. Oxford University Press, Oxford

Kuhlman CJ, Mortveit HS (2014) Attractor stability in nonuniform Boolean networks. Theor Comput Sci 559:20–33. doi:10.1016/j.tcs.2014.08.010 special volume: Non-uniform Cellular Automata

Layne L, Dimitrova E, Matthew M (2012) Nested canalyzing depth and network stability. Bull Math Biol. doi:10.1007/s11538-011-9692-y

Luo JX, Turner MS (2012) Evolving sensitivity balances Boolean networks. PLoS One 7(e36):010. doi:10.1371/journal.pone.0036010

Mortveit HS, Reidys CM (2007) An introduction to sequential dynamical systems. Universitext, Springer. doi:10.1007/978-0-387-49879-9

Pomerance A, Ott E, Girvan M, Losert W (2009) The effect of network topology on the stability of discrete state models of genetic control. Proc Nat Acad Sci 106(20):8209–8214

Ribeiro AS, Kauffman SA (2007) Noisy attractors and ergodic sets in models of gene regulatory networks. J Theor Biol 247:743–755

Robert F (1986) Discrete iterations. A Metric Study. No. 6 in Springer Series in Computational Mathematics, Springer

Serra R, Villani M, Barbieri A, Kauffman S, Colacci A (2010) On the dynamics of random Boolean networks subject to noise: attractors, ergodic sets and cell types. J Theor Biol 265(2):185–193

Shmulevich I, Kauffman SA (2004) Activities and sensitivities in Boolean network models. Phys Rev Lett 93(4):048701:1–4

Shmulevich I, Lähdesmäki H, Dougherty ER, Astola J, Zhang W (2003) The role of certain post classes in Boolean network models of genetic networks. Proc Nat Acad Sci 100(19):10,734–10,739

Xiao Y, Dougherty ER (2007) The impact of function perturbations in Boolean networks. Bioinformatics 23(10):1265–1273