

Supercriticality of an annealed approximation of Boolean networks

Thomas Mountford* Daniel Valesin†

Abstract

We consider a model recently proposed by Chatterjee and Durrett [1] as an “annealed approximation” of boolean networks, which are a class of cellular automata on a random graph, as defined by S. Kauffman [5]. The starting point is a random directed graph on n vertices; each vertex has r input vertices pointing to it. For the model of [1], a discrete time threshold contact process is then considered on this graph: at each instant, each vertex has probability q of choosing to receive input; if it does, and if at least one of its input vertices were in state 1 at the previous instant, then it is labelled with a 1; in all other cases, it is labelled with a 0. r and q are kept fixed and n is taken to infinity. Improving one of the results of [1], we show that if $qr > 1$, then the time of persistence of activity of the dynamics is exponential in n .

Keywords: boolean networks; interacting particle systems; metastability.

AMS MSC 2010: 82C22; 92B20; 94C10.

Submitted to ECP on December 4, 2012, final version accepted on April 16, 2013.

1 Introduction

Random boolean networks were introduced by Stuart Kauffman in 1969 [5] as models of gene regulatory networks. A gene regulatory network is a set of genes in a cell that iteratively communicate with each other, using their RNA transcripts as messages, and this communication affects each gene’s activity. They are thus information networks and control systems for the activity of the cell.

Let us define Kauffman’s model. The following definition depends on three parameters: $n, r \in \mathbb{N}$ with $r < n$ and $p \in (0, 1)$ (though Kauffman only considered the case $p = 1/2$). The letters a, b will denote two possible states of a gene. Let $V_n = \{x_1, \dots, x_n\}$ be the set of genes. For each $x \in V_n$, we independently choose:

- a set $y(x) = \{y_1(x), \dots, y_r(x)\} \subset V_n - \{x\}$. The choice is made uniformly among all possibilities. $y(x)$ is called the *influence set* of x . We define the set of directed edges E_n by $E_n = \{(y_i(x), x) : x \in V_n, 1 \leq i \leq r\}$.
- a function $f_x : \{a, b\}^r \rightarrow \{a, b\}$. The values $\{f_x(\omega) : \omega \in \{a, b\}^{y(x)}\}$ are chosen independently, with probability p to be equal to a and $1 - p$ to be equal to b .

Having made all these random choices, we define $\Phi : \{a, b\}^{V_n} \rightarrow \{a, b\}^{V_n}$ by

$$[\Phi(\eta)](x) = f_x(\eta(y_1(x)), \dots, \eta(y_r(x)))$$

*Ecole Polytechnique Fédérale de Lausanne. E-mail: thomas.mountford@epfl.ch

†University of British Columbia. E-mail: valesin@math.ubc.ca

and, given an initial configuration $\eta_0 \in \{a, b\}^{V_n}$, we define a deterministic, discrete time dynamics $(\eta_t)_{t=0,1,\dots}$ by putting $\eta_{t+1} = \Phi(\eta_t)$, $t \geq 0$. The dynamics is explained in words as follows: at each instant, and for each vertex x , we inspect the previous states in the influence set of x and from these, determine the state of x using the function f_x .

A set $\Gamma \subset \{a, b\}^{V_n}$ such that $\Phi(\Gamma) = \Gamma$ and $\Phi(\Gamma') \neq \Gamma'$ for any proper subset Γ' of Γ is called a periodic orbit of Φ . Since the state space is finite, every initial configuration η_0 is in the domain of attraction of a periodic orbit Γ (meaning that, for some t_0 , $\{\eta_t : t \geq t_0\} = \Gamma$). Typical aspects of interest in random boolean networks are the number of these orbits, their stability, periods and the time to reach them. As thoroughly explained in [6], simulations of the model suggested the existence of two regimes, depending on the choice of parameters, in which drastically different behaviours arise. Among other important differences, in the *ordered* (or subcritical) regime, the lengths of the orbits grow slowly with n , whereas in the *chaotic* (or supercritical) regime, they grow rapidly with n .

In [2], Derrida and Pomeau proposed an “annealed approximation” of random boolean networks; in it, the random aspects of the network (namely, the underlying graph and the rules of evolution) are updated at each time step instead of remaining fixed. The process thus obtained is a Markov chain. The simplification destroys important correlations in the system, but allowed the authors to identify (through a not fully rigorous analysis of the transition kernel) a phase transition given by a curve that agrees with simulations, $2rp(1-p) = 1$ (the ordered regime corresponding to $2rp(1-p) < 1$).

In [1], Chatterjee and Durrett proposed a model which was an approximation to the *activity* of Boolean networks. The *activity process* associated to (η_t) is the process $(\bar{\eta}_t)_{t=0,1,\dots}$ with state space $\{0, 1\}^{V_n}$ and given by

$$\bar{\eta}_0 \equiv 1, \quad \bar{\eta}_{t+1}(x) = I_{\{\eta_{t+1}(x) \neq \eta_t(x)\}}, \quad x \in V_n, \quad t \geq 0,$$

where I is the indicator function. The idea in considering $(\bar{\eta}_t)$ rather than (η_t) is the possibility of identifying the phase transition in a process that is in some respects easier to study than the original process. Indeed, (ξ_t) , the proposed approximation to $(\bar{\eta}_t)$ to be defined below, has the more tractable dynamics of a threshold contact process on a random graph (in particular, the graph is sampled only once, and not re-sampled as the dynamics advances). For (ξ_t) , Chatterjee and Durrett proved the phase transition and identified the same critical curve as the one mentioned above, $2rp(1-p) = 1$. Their work allows for insight into this phase transition by an analogy between the flow of information in random boolean networks and the evolution of branching processes.

Let us now define the model of [1]. We start with parameters $n, r \in \mathbb{N}$ with $r < n$ and $q \in (0, 1)$. Define the oriented random graph $G_n = (V_n, E_n)$ exactly as before. We will now define a discrete time Markov chain $(\xi_t)_{t \geq 0}$ with state space $\{0, 1\}^{V_n}$ and initial configuration $\xi_0 \equiv 1$. Its transition kernel is given by

$$p(\xi, \xi') = \left(\prod_{x \in V_n : \sum_i \xi(y_i(x))=0} I_{\{\xi'(x)=0\}} \right) \left(\prod_{x \in V_n : \sum_i \xi(y_i(x))>0} (q \cdot I_{\{\xi'(x)=1\}} + (1-q) \cdot I_{\{\xi'(x)=0\}}) \right),$$

where $\xi, \xi' \in \{0, 1\}^{V_n}$. It will be useful to construct this Markov chain with a set of auxiliary Bernoulli random variables. Let $\{B_t^x : x \in V_n, t \geq 1\}$ be a family of independent Bernoulli random variables with parameter q ; given $\xi_t \in \{0, 1\}^{V_n}$, we put

$$\xi_{t+1}(x) = \begin{cases} 1 & \text{if } B_{t+1}^x = 1 \text{ and } \sum_{i=1}^r \xi_t(y_i(x)) > 0; \\ 0 & \text{otherwise.} \end{cases}$$

When $B_t^x = 1$, we say that x receives *input* at time t ; therefore, a vertex is set to 1 if and only if it receives input at that time and at least one of its input vertices $y_1(x), \dots, y_r(x)$

was set to 1 at the previous time. We sometimes abuse notation and associate $\xi \in \{0, 1\}^{V_n}$ with $\{x \in V_n : \xi(x) = 1\}$.

In the comparison with boolean networks, q plays the role of $2p(1 - p)$, which is the probability that two independent random variables with distribution $p \cdot \delta_{\{a\}} + (1 - p) \cdot \delta_{\{b\}}$ are different. See [1] for a more detailed explanation of the relationship between (ξ_t) and $(\bar{\eta}_t)$.

It is readily seen that the identically zero configuration is absorbing for the chain (ξ_t) and that it is eventually reached with probability 1. In [1], the authors study the time τ_n it takes for this to occur and the typical proportion of sites that are in state 1 at times before τ_n . By a simple comparison between the time dual of the model (as defined below) and a subcritical branching process, it is easy to show that, if $qr < 1$, then τ_n behaves as $\log n$, and this is associated to the ordered regime of random boolean networks. In [1], the following result is shown, characterizing the chaotic regime. Let $\rho = \rho(q, r)$ denote the probability of survival for a branching process in which individuals have probability q of having r children and probability $1 - q$ of having none. Let $|A|$ denote the cardinality of the set A . Finally, let \mathbb{P}_n denote a probability measure both for the choice of G_n and for the family $\{B_t^x\}$ (they are of course taken independently).

Theorem. [1] *If $q(r - 1) > 1$, then for every $\epsilon > 0$ there exists $c > 0$ such that, as $n \rightarrow \infty$,*

$$\inf_{0 \leq t \leq e^{cn}} \mathbb{P}_n \left(\frac{|\xi_t|}{n} \geq \rho - \epsilon \right) \xrightarrow{n \rightarrow \infty} 1.$$

Under the more general hypothesis $qr > 1$, only a weaker result was obtained: the function e^{cn} in the above infimum had to be replaced by a function of the form e^{cn^b} , for $b, c > 0$. The proof of this weaker result was established through a different method than that of the proof of the above theorem. In this paper we give a unified proof that establishes the stronger result

Theorem 1.1. *If $qr > 1$, then there exists $c > 0$ such that, for any $\epsilon > 0$ and any sequence (t_n) with $t_n \rightarrow \infty$ and $t_n \leq e^{cn}$,*

$$\inf_{t_n \leq t \leq e^{cn}} \mathbb{P}_n \left(\rho - \epsilon < \frac{|\xi_t|}{n} < \rho + \epsilon \right) \xrightarrow{n \rightarrow \infty} 1.$$

To explain why this is to be expected and, in particular, the link with the mentioned branching process, we introduce the time dual of the process. Fix a realization of $G_n = (V_n, E_n)$ and $\{B_t^x : x \in V_n, t \geq 1\}$, define \hat{E}_n as the set of directed edges obtained by inverting the edges of E_n and $\hat{G}_n = (V_n, \hat{E}_n)$. Note that

$$\{y_i(x) : 1 \leq i \leq r\} = \{z : (x, z) \in \hat{E}_n\};$$

that is, in \hat{G}_n each vertex “points to” r vertices. Fix $T > 0$ and put $\hat{B}_t^{x,T} = B_{T-t}^x$, for $0 \leq t < T$. Given $A \subset V_n$, define $\hat{\xi}_0^{A,T} = I_A$ and, for $0 \leq t < T$,

$$\hat{\xi}_{t+1}^{A,T}(z) = \begin{cases} 1 & \text{if for some } x, i, \text{ we have } y_i(x) = z, \hat{\xi}_t^{A,T}(x) = 1 \text{ and } \hat{B}_t^{x,T} = 1; \\ 0 & \text{otherwise.} \end{cases} \quad (1.1)$$

When $\hat{\xi}_t^{A,T}(x) = 1$ and $\hat{B}_t^{x,T} = 1$, we say that x gives birth at time t . Let us describe the dual dynamics in words. Given the configuration $\hat{\xi}_t$, we go over every vertex that is in state 1 and determine which of them give birth at time t – for each vertex, this happens with probability q and independently. For each vertex x that gives birth at time t , we set the vertices $y_1(x), \dots, y_r(x)$ to 1 at time $t + 1$. Vertices that are not set to 1 by this procedure are then set to 0. We then have the duality equation

$$\{\xi_T \cap A \neq \emptyset\} = \{\hat{\xi}_T^{A,T} \neq \emptyset\}$$

(recall that we take $\xi_0 \equiv 1$). The above equality holds since both events are equal to

$$\{\exists x_1, \dots, x_T \in V_n : x_1 \in y(x_2), \dots, x_{T-1} \in y(x_T) \text{ and } B_1^{x_1} = B_2^{x_2} = \dots = B_T^{x_T} = 1\}.$$

By taking $A = \{x\}$ for each $x \in V_n$ in the duality equation, we see that, under \mathbb{P}_n , $|\xi_T|$ and $|\{x : \hat{\xi}_T^{\{x\}, T} \neq \emptyset\}|$ have the same distribution.

Since we will mostly work with the dual process, we drop the superscript T and assume that $\hat{\xi}_t^A$ is defined for all $t \geq 0$ with the evolution rule explained above. We also write $\hat{\xi}_t^x$ instead of $\hat{\xi}_t^{\{x\}}$. The convergence in Theorem 1.1 can be re-stated as

$$\inf_{t_n \leq t \leq e^{cn}} \mathbb{P}_n \left(\rho - \epsilon < \frac{|\{x : \hat{\xi}_t^x \neq \emptyset\}|}{n} < \rho + \epsilon \right) \xrightarrow{n \rightarrow \infty} 1. \quad (1.2)$$

Now, assume that n is very large with respect to r . If g is another integer that is much larger than r and much smaller than n , then with high probability, the subgraph of \hat{G}_n with vertex set

$$\{z \in V_n : \text{for some } k \leq g \text{ and } z_1, \dots, z_k \in V_n, \text{ we have } x \rightarrow z_1 \rightarrow \dots \rightarrow z_k \rightarrow z \text{ in } \hat{G}_n\}$$

and edge set equal to the set of edges of \hat{E}_n that start and end at vertices in the above set will simply be a directed tree of degree r rooted in x . Conditioning on the event that this subgraph is indeed a tree, the evolution of $|\hat{\xi}_t^x|$ up to time g will be exactly that of the branching process mentioned before Theorem 1.1. In addition, it is not difficult to see that, without any conditioning, $|\hat{\xi}_t^x|$ is stochastically dominated by such a process. These remarks clarify why the model exhibits two phases in exact correspondence with the branching process. If the expected offspring size $qr < 1$, then $\hat{\xi}_t^x$ dies out faster than the corresponding subcritical branching process, and the primal ξ_t rapidly reaches the zero state. On the other hand, if $qr > 1$, the above theorem states that the system survives for a time that is exponentially large in n , characterizing the supercritical regime.

The structure of our proof is similar to that of [1]. First, using the comparison with the branching process and a second moment argument, we show that with probability tending to 1 as $n \rightarrow \infty$, the set of vertices $S = \{x : |\hat{\xi}_{s_n}^x| > k_n\}$, where $k_n = (\log n)^2$ and $s_n = (\log \log n)^2$, has size close to $\rho \cdot n$ (see Proposition 2.1). Second, in Proposition 2.2, we show that with probability tending to 1 as $n \rightarrow \infty$, the graph \hat{G}_n is “fertile” in the following sense. For any choice of $A \subset V_n$ with $|A| \geq (\log n)^2$, the process $\hat{\xi}_t^A$ defined on \hat{G}_n has probability larger than $1/n^2$ of remaining active up to time e^{cn} , for some fixed constant c . We can then use a simple union bound to argue that with high probability, for every x in S , $(\hat{\xi}_t^x)$ remains active until time e^{cn} .

Our main contribution is Proposition 2.2; let us briefly explain the ideas that go into its proof. Given $A \subset V_n$, suppose we reveal, one by one, the elements of the set $A_1 = \{y_i(x) : 1 \leq i \leq r, x \in A\}$, then $A_2 = \{y_i(x) : 1 \leq i \leq r, x \in A_1\}$, until A_g , for some fixed $g \in \mathbb{N}$. Let $B(A, g)$ be the subgraph of \hat{G}_n with vertex set $A \cup A_1 \cup \dots \cup A_g$ and edge set equal to the edges of \hat{E}_n which start and end at vertices in this set. For most choices of A , $B(A, g)$ is just a disjoint union of $|A|$ directed trees, so that $\{|\hat{\xi}_t^A|\}_{0 \leq t \leq g}$ is exactly a branching process. However, for some choices of A , when revealing A_1, \dots, A_g , we will see some “collisions”, that is, some vertices will be found more than once. We say that A is *expansive* if the number of collisions is not too large, so that $\{|\hat{\xi}_t^A|\}_{0 \leq t \leq g}$ is not too far from the branching process and consequently, $|\hat{\xi}_g^A|$ is very likely to be larger than $|A|$ (see Lemma 2.4). We then show that, with high probability, for some $c > 0$, there is no set $A \subset V_n$ with $(\log n)^2 \leq |A| \leq cn$ that is not expansive (Lemma 2.5). It is then quite easy to put Lemmas 2.4 and 2.5 together to obtain Proposition 2.2.

2 Proof of Theorem 1.1

In this section, we will exclusively work with the dual process. Let us present the notation we will use. For fixed n , \mathcal{P}_n is a probability measure under which the random graph \hat{G}_n is defined; as explained in the Introduction, \hat{G}_n is a directed graph in which each vertex x “points to” r distinct vertices $y_1(x), \dots, y_r(x)$. \hat{G}_n will denote the (finite) set of possible realizations of \hat{G}_n . For a fixed realization of the graph \hat{G}_n , $P_{\hat{G}_n}$ is a probability measure under which independent Bernoulli(q) variables $\{\hat{B}_t^x : t \geq 0, x \in V_n\}$ are defined, and thus the family of processes $\{(\hat{\xi}_t^A)_{t \geq 0} : A \subset V_n\}$ are all defined by the rule (1.1). Finally, \mathbb{P}_n is the annealed probability measure: under \mathbb{P}_n , we first sample the graph \hat{G}_n (with the probability measure \mathcal{P}_n) and then the processes $\{(\hat{\xi}_t^A) : A \subset V_n\}$ on \hat{G}_n (with the probability measure $P_{\hat{G}_n}$).

In all results and proofs that follow, we assume that $qr > 1$. We start with two propositions that together will yield Theorem 1.1. Proposition 2.1 is proved essentially by a repetition of arguments in [1]; we include a proof for completeness.

Proposition 2.1. *Let (a_n) and (b_n) be two sequences of integers satisfying*

$$a_n, b_n \geq 0 \ \forall n, \quad a_n \xrightarrow{n \rightarrow \infty} \infty, \quad \frac{a_n}{\log n} \xrightarrow{n \rightarrow \infty} 0 \quad \text{and} \quad \frac{b_n}{(qr)^{a_n}} \xrightarrow{n \rightarrow \infty} 0.$$

For any $\epsilon > 0$, we have

$$\lim_{n \rightarrow \infty} \mathbb{P}_n \left(\rho - \epsilon < \frac{\left| \left\{ x \in V_n : |\hat{\xi}_{a_n}^x| > b_n \right\} \right|}{n} < \rho + \epsilon \right) = 1.$$

Proof. Let $(Z_t)_{t=0,1,\dots}$ be the branching process with $Z_0 = 1$ and the offspring distribution that gives mass q to r and $1 - q$ to 0 . Also let $H = \{Z_t \neq 0 \ \forall t\}$, $\rho = \mathbb{P}(H)$ and $M_t = \frac{Z_t}{(qr)^t}$. By Theorem 5.3.9 and Exercise 5.3.12 in [3], our assumption that $qr > 1$ implies $\rho > 0$ and the fact that M_t almost surely converges to a limit M , which is strictly positive on H and identically zero on H^c . Let $\rho_n = \mathbb{P}(Z_{a_n} > b_n) = \mathbb{P}(M_{a_n} > b_n / (qr)^{a_n})$. Since $\lim_{n \rightarrow \infty} b_n / (qr)^{a_n} = 0$, we almost surely have $\lim_{n \rightarrow \infty} I_{\{M_{a_n} > \frac{b_n}{(qr)^{a_n}}\}} = I_{\{M > 0\}} = I_H$. Indeed, for almost every $\omega \in H$ we have $\lim_{n \rightarrow \infty} M_{a_n}(\omega) = M(\omega) > 0$ and for almost every $\omega \in H^c$, $M_{a_n}(\omega) = 0$ for n large enough. Consequently, by the dominated convergence theorem,

$$\rho_n \rightarrow \rho. \tag{2.1}$$

For a set of vertices A in the graph \hat{G}_n , let $y^{(0)}(A) = A$, $y^{(1)}(A) = y(A) = \{y_i(x) : x \in A, 1 \leq i \leq r\}$ and $y^{(k+1)}(A) = y(y^{(k)}(A))$ for $k \geq 0$. Given a vertex x and $R \in \mathbb{N}$, we define the ball $B(x, R)$ as the subgraph of \hat{G}_n with vertices $\cup_{k=0}^R y^{(k)}(x)$ and all the edges of \hat{E}_n that start and end at these vertices. Let $F(x, R)$ denote the event that $B(x, R)$ has no cycles and $F(x, y, R)$ the event that $B(x, R)$ and $B(y, R)$ have no cycles and are disjoint. We claim that

$$\lim_{n \rightarrow \infty} \mathcal{P}_n(F(x_1, a_n)) = \lim_{n \rightarrow \infty} \mathcal{P}_n(F(x_1, x_2, a_n)) = 1. \tag{2.2}$$

We will prove only that the first limit is 1, and it should be clear that a similar proof works for the second. We explore the ball $B(x_1, a_n)$ level by level: we reveal the vertices of $y(x_1)$ one by one (in any order we desire), then the vertices of $y^{(2)}(x_1)$ one by one, and so on, and say that a collision occurs if at some point before having revealed all vertices in $B(x_1, a_n)$, we reveal a vertex that had already been revealed at an earlier step; the exploration is then stopped and said to have been unsuccessful. The exploration is thus successful if and only if $F(x_1, a_n)$ occurs. Note that the maximum number

of vertices revealed in the whole exploration is $\sum_{i=0}^{\lceil a_n \rceil} r^i \leq r^{a_n+2}$. Also, at any point in the exploration, there are at least $n - r$ choices for the next vertex (since for any $x \in V_n$, $y_1(x), \dots, y_r(x)$ are necessarily all distinct and different from x), so the probability that the next vertex results in a collision (and thus an unsuccessful exploration) is less than $r^{a_n+2}/(n - r)$. The probability that the exploration is unsuccessful is thus less than $(r^{a_n+2})^2/(n - r)$, which tends to 0 as $n \rightarrow \infty$ since $a_n/\log n \rightarrow 0$.

For $n \geq 1$ and $i \in \{1, \dots, n\}$, let $X_{n,i} = I_{\{|\hat{\xi}_{a_n}^{x_i}| > b_n\}}$. Note that, by symmetry,

$$X_{n,1}, \dots, X_{n,n} \text{ are identically distributed under } \mathbb{P}_n. \tag{2.3}$$

If $F(x_1, a_n)$ occurs, then $(|\hat{\xi}_t^{x_1}|)_{0 \leq t \leq a_n}$ has the same distribution as $(Z_t)_{0 \leq t \leq a_n}$, so that, by (2.1) and (2.2),

$$\mathbb{E}_n[X_{n,1}] = \mathcal{P}_n(F(x_1, a_n)) \cdot \rho_n + \mathbb{E}_n[X_{n,1} \cdot I_{F(x_1, a_n)^c}] \xrightarrow{n \rightarrow \infty} \rho. \tag{2.4}$$

Likewise, if $F(x_1, x_2, a_n)$ occurs, then $(|\hat{\xi}_t^{x_1}|)_{0 \leq t \leq a_n}$, $(|\hat{\xi}_t^{x_2}|)_{0 \leq t \leq a_n}$ are distributed as two independent copies of $(Z_t)_{0 \leq t \leq a_n}$, so that

$$\mathbb{E}_n[X_{n,1} \cdot X_{n,2}] = \mathcal{P}_n(F(x_1, x_2, a_n)) \cdot (\rho_n)^2 + \mathbb{E}_n[X_{n,1} \cdot X_{n,2} \cdot I_{F(x_1, x_2, a_n)^c}] \xrightarrow{n \rightarrow \infty} \rho^2. \tag{2.5}$$

By (2.4) and (2.5) we get

$$\text{Cov}(X_{n,1}, X_{n,2}) \xrightarrow{n \rightarrow \infty} 0. \tag{2.6}$$

Now, (2.3), (2.4), (2.6) and Chebyshev's inequality imply that $\frac{1}{n} \sum_{i=1}^n X_{n,i}$ converges to ρ in probability, as desired. \square

We will write

$$k_n = (\log n)^2, \quad s_n = (\log \log n)^2.$$

For $t > 0$, let us say that a graph $\hat{G}_n \in \mathcal{G}_n$ is t -fertile if

$$\text{for every } A \subset V_n \text{ with } |A| \geq k_n, P_{\hat{G}_n}(\hat{\xi}_t^A = \emptyset) < n^{-2}. \tag{2.7}$$

Let $\mathcal{H}_n(t)$ denote the set of graphs in $\hat{\mathcal{G}}_n$ that are t -fertile.

Proposition 2.2. *There exists $\bar{c} > 0$ such that $\lim_{n \rightarrow \infty} \mathcal{P}_n(\mathcal{H}_n(e^{\bar{c}n})) = 1$.*

Proving this result takes most of our effort. We postpone the proof and first show how the two propositions are used to establish the main theorem.

Proof of Theorem 1.1. We will use the fact that

$$\text{if } t' < t'', \text{ then } \{x \in V_n : \hat{\xi}_{t''}^x \neq \emptyset\} \subset \{x \in V_n : \hat{\xi}_{t'}^x \neq \emptyset\}. \tag{2.8}$$

Let \bar{c} be the constant of Proposition 2.2. Fix $\epsilon > 0$ and a sequence (t_n) as in the statement of the theorem. If $t \leq e^{\bar{c}n}$, by (2.8) we have

$$\begin{aligned} \mathbb{P}_n \left(\frac{|\{x \in V_n : \hat{\xi}_t^x \neq \emptyset\}|}{n} \leq \rho - \epsilon \right) &\leq \mathbb{P}_n \left(\frac{|\{x \in V_n : \hat{\xi}_{e^{\bar{c}n}}^x \neq \emptyset\}|}{n} \leq \rho - \epsilon \right) \\ &\leq \mathbb{P}_n \left(\frac{|\{x \in V_n : |\hat{\xi}_{s_n}^x| > k_n\}|}{n} \leq \rho - \epsilon \right) + \mathbb{P}_n \left(\exists x \in V_n : |\hat{\xi}_{s_n}^x| > k_n, \hat{\xi}_{e^{\bar{c}n}}^x = \emptyset \right) \end{aligned} \tag{2.9}$$

The first term vanishes as $n \rightarrow \infty$ by Proposition 2.1. The second term is less than

$$\begin{aligned} & \mathcal{P}_n((\mathcal{H}_n(e^{\bar{c}n}))^c) + \sum_{\hat{G} \in \mathcal{H}_n(e^{\bar{c}n})} \mathcal{P}_n(\hat{G}) \cdot \sum_{x \in V_n} \sum_{\substack{A \subset V_n: \\ |A| \geq k_n}} P_{\hat{G}}(\hat{\xi}_{s_n}^x = A) \cdot P_{\hat{G}}(\hat{\xi}_{e^{\bar{c}n} - s_n}^A = \emptyset) \\ & \leq \mathcal{P}_n((\mathcal{H}_n(e^{\bar{c}n}))^c) + n^{-2} \sum_{\hat{G} \in \mathcal{H}_n(e^{\bar{c}n})} \mathcal{P}_n(\hat{G}) \cdot \sum_{x \in V_n} P_{\hat{G}_n}(|\hat{\xi}_{s_n}^x| \geq k_n) \\ & \leq \mathcal{P}_n((\mathcal{H}_n(e^{\bar{c}n}))^c) + n^{-1} \cdot \mathcal{P}_n(\mathcal{H}_n(e^{\bar{c}n})) \xrightarrow{n \rightarrow \infty} 0. \end{aligned}$$

This shows that

$$\inf_{t \leq e^{\bar{c}n}} \mathbb{P}_n \left(\frac{|\{x \in V_n : \hat{\xi}_t^x \neq \emptyset\}|}{n} > \rho - \epsilon \right) \xrightarrow{n \rightarrow \infty} 1. \tag{2.10}$$

Now let us consider the reverse inequality. If $t \geq t_n$, by (2.8) we have

$$\mathbb{P}_n \left(\frac{|\{x \in V_n : \hat{\xi}_t^x \neq \emptyset\}|}{n} \geq \rho + \epsilon \right) \leq \mathbb{P}_n \left(\frac{|\{x \in V_n : \hat{\xi}_{\min(t_n, s_n)}^x \neq \emptyset\}|}{n} \geq \rho + \epsilon \right).$$

We can now apply Proposition 2.1 with $a_n = \min(t_n, s_n)$ and $b_n \equiv 0$; the right-hand side thus vanishes as $n \rightarrow \infty$. Thus,

$$\inf_{t \geq t_n} \mathbb{P}_n \left(\frac{|\{x \in V_n : \hat{\xi}_t^x \neq \emptyset\}|}{n} < \rho + \epsilon \right) \xrightarrow{n \rightarrow \infty} 1. \tag{2.11}$$

(2.10) and (2.11) together yield (1.2). □

We now need to prove Proposition 2.2; three preliminary results will be needed: Lemmas 2.3, 2.4 and 2.5.

Once and for all, fix $\tilde{q} < q$, $\delta > 0$ and $g \in \mathbb{N}$ so that

$$\tilde{q}r > 1, \delta < \min((\tilde{q}r - 1), 1) \text{ and } (\tilde{q}r - 1 - \delta)(\tilde{q}r)^{g-1} > 1 + \delta.$$

We now give some definitions and notations.

Given $m \in \mathbb{N}$, let

$$\begin{aligned} T_m^0 &= \{1, \dots, m\}, \\ T_m^i &= \{1, \dots, m\} \times \{1, \dots, r\}^i, \quad 1 \leq i \leq g, \\ T_m &= \cup_{i=0}^g T_m^i. \end{aligned}$$

For $\sigma = (\sigma_0, \dots, \sigma_i), \sigma' = (\sigma'_0, \sigma'_1, \dots, \sigma'_j) \in T_m$, we say $\sigma \prec \sigma'$ either if $i < j$ or if $i = j$ and σ is less than σ' in lexicographic order. With this order, we can take an increasing enumeration

$$T_m = \{\sigma^1, \dots, \sigma^{(1+r+\dots+r^g)m}\} \tag{2.12}$$

Then, $T_m^0 = \{\sigma^1, \dots, \sigma^m\}$ and, for $i \geq 1$, $T_m^i = \{\sigma^{(1+r+\dots+r^{i-1})m+1}, \dots, \sigma^{(1+r+\dots+r^i)m}\}$.

Next, we endow T_m with directed edges by setting

$$\sigma \rightarrow \sigma' \text{ if and only if } \sigma = (\sigma_0, \dots, \sigma_i), \sigma' = (\sigma_0, \dots, \sigma_i, \sigma'_{i+1}) \text{ for some } i.$$

T_m is thus the disjoint union of m rooted, directed trees, each with g generations above the root. If we can go from σ to σ' by following a path of oriented edges of the tree, we say that σ is an *ancestor* of σ' and that σ' is a *descendant* of σ .

The set $\{0, 1\}^{T_m}$ will be called the space of configurations. Given vertex $\sigma \in T_m$ and configuration $\psi \in \{0, 1\}^{T_m}$, $\psi(\sigma) \in \{0, 1\}$ will denote the value of ψ at σ .

Now assume $\hat{G}_n = (V_n, \hat{E}_n)$ is given and $A \subset V_n$ with $|A| = m$. We can enumerate $A = \{x_{j_1}, \dots, x_{j_m}\}$ in the order of the indices of V_n . Given $\sigma = (\sigma_0, \dots, \sigma_i) \in T_m$ with $i > 0$, let $z^\sigma = y_{\sigma_i}(y_{\sigma_{i-1}}(\dots(y_{\sigma_1}(x_{j_{\sigma_0}}))\dots))$. Finally, define

$$\mathcal{A}^\sigma = \{z^{\sigma'} \in T_m : \sigma' \prec \sigma\}.$$

We now present an algorithm to construct a configuration $\psi = \psi(A) \in \{0, 1\}^{T_m}$ from A . The index j in the algorithm follows the enumeration given in (2.12).

```

for  $j = 1$  to  $m$ 
| set  $\psi(\sigma^j) = 0$ ;
for  $j = m + 1$  to  $(1 + r + \dots + r^g)m$ 
| if [ $\psi(\sigma) = 1$  for some  $\sigma$  ancestor of  $\sigma^j$ ] or [ $z^{\sigma^j} \notin \mathcal{A}^{\sigma^j}$ ]
| | then set  $\psi(\sigma^j) = 0$ 
| | else set  $\psi(\sigma^j) = 1$ 
    
```

In words, vertices are inspected in order; the roots are all set to 0 and the other vertices are set to 0 either if one of their ancestors has already been marked with a 1 or if their image under the map $\sigma \mapsto z^\sigma$ has never been seen before; otherwise they are set to 1. Figure 1 presents an example of the effect of the algorithm.

As will become clear in the proof of Lemma 2.4, an essential property of this construction is the fact that $\sigma \mapsto z^\sigma$ injectively maps the set

$$\{\sigma \in T_m : \psi(\sigma) = 0 \text{ and } \psi(\sigma') = 0 \text{ for every ancestor } \sigma' \text{ of } \sigma\}$$

onto the vertex set of $B(A, g)$. Note that this property does not depend on the value of ψ at any vertex σ' such that $\psi(\sigma) = 1$ for some ancestor σ of σ' . On the other hand, we will want to argue that with high probability there are few vertices of T_m where ψ is equal to 1. This is why we set the algorithm to “artificially” set ψ to 0 at all vertices that descend from a vertex σ such that $\psi(\sigma) = 1$; these should be understood as “dummy” 0’s, that is, they have no counterpart in the geometry of $B(A, g)$.

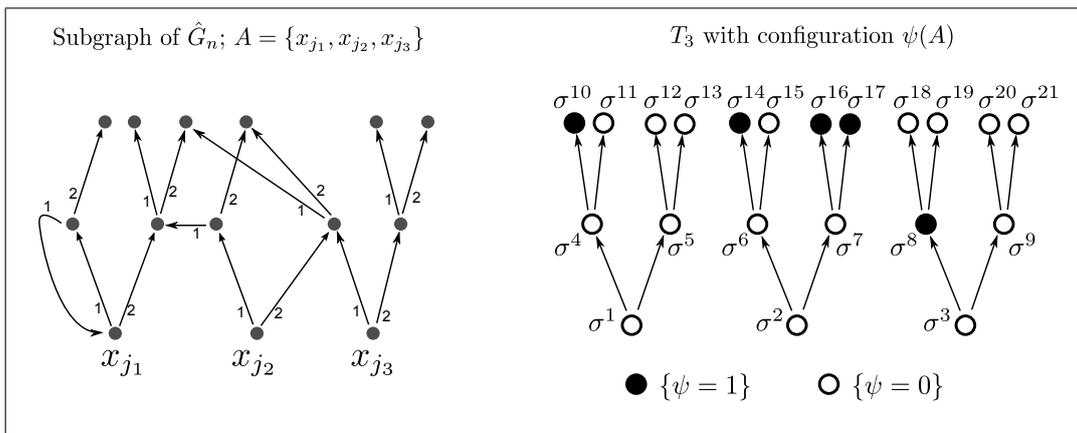


Figure 1: Example of the algorithm. Here $r = 2$, $g = 2$. The numbers in the arrows in the left diagram serve to distinguish $y_1(x)$ and $y_2(x)$ for each vertex x

Lemma 2.3. Given $A \subset V_n$ with $|A| = m$ and $\sigma^{i_1}, \dots, \sigma^{i_k} \in T_m$,

$$\mathcal{P}_n([\psi(A)](\sigma^{i_1}) = \dots = [\psi(A)](\sigma^{i_k}) = 1) \leq \left(\frac{m + rm + \dots + r^g m}{n - r} \right)^k.$$

Proof. There is no loss of generality in assuming that $\sigma^{i_a} \prec \sigma^{i_b}$ when $a < b$. We then have

$$\mathcal{P}_n([\psi(A)](\sigma^{i_k}) = 1 \mid [\psi(A)](\sigma^{i_1}) = \dots = [\psi(A)](\sigma^{i_{k-1}}) = 1) \leq \frac{m + rm + \dots + r^g m}{n - r}.$$

Indeed, let Θ^{i_k} denote the event that none of the ancestors of σ^{i_k} in T_m is marked with a 1 in $\psi(A)$. First note that $\{[\psi(A)](\sigma^{i_k}) = 1\} \subset \Theta^{i_k}$, because the algorithm fills all positions above a 1 with 0's. Next, fix $a_{m+1}, a_{m+2}, \dots, a_{i_k-1} \in V_n$ such that

$$\{z^{\sigma^{m+1}} = a_{m+1}, \dots, z^{\sigma^{i_k-1}} = a_{i_k-1}\} \subset \Theta^{i_k} \cap \{[\psi(A)](\sigma^{i_1}) = \dots = [\psi(A)](\sigma^{i_{k-1}}) = 1\}$$

(we start at $m + 1$ because $z^{\sigma^1}, \dots, z^{\sigma^m}$ are always equal to the points of A). Then, conditioned on $\{z^{\sigma^{m+1}} = a_{m+1}, \dots, z^{\sigma^{i_k-1}} = a_{i_k-1}\}$, there are at least $n - r + 1$ possible positions for $z^{\sigma^{i_k}}$, and $[\psi(A)](z^{\sigma^{i_k}}) = 1$ precisely when $z^{\sigma^{i_k}} \in \mathcal{A}^{\sigma^{i_k}}$, a set of size less than $m + rm + \dots + r^g m$. \square

Given $A \subset V_n$ with $|A| = m$, let

$$d_i(A) = |\{\sigma \in T_m^i : [\psi(A)](\sigma) = 1\}|, \quad d(A) = \sum_{i=1}^g d_i.$$

We say that A is *expansive* if $d(A) \leq (1 + \delta)m$. The next lemma shows the motivation for this definition; see (2.15) in the proof.

Lemma 2.4. There exists $c_1 > 0$ such that, if $A \subset V_n$ is expansive, then

$$P_{\hat{G}_n}(|\hat{\xi}_g^A| < (1 + \delta)|A|) \leq e^{-c_1|A|}.$$

Proof. Let $m = |A|$. If $i < g$ and $B \subset T_m^i$, we will write

$$J(B) = \{\sigma' \in T_m : \sigma \rightarrow \sigma' \text{ for some } \sigma \in B\} \subset T_m^{i+1}.$$

Consider the process $(\hat{\xi}_t^A)_{0 \leq t \leq g}$; define the sets

$$\begin{aligned} B_0 &= \{\sigma \in T_m^0 : z^\sigma \text{ gives birth at time } 0\}; \\ B_i &= \{\sigma \in J(B_{i-1}) \cap \{\psi(A) = 0\} : z^\sigma \text{ gives birth at time } i\}, \quad 1 \leq i < g \end{aligned}$$

The definition of B_0 implies that $\hat{\xi}_1^A \supset \{z^\sigma : \sigma \in J(B_0)\}$. From the construction of $\psi(A)$ we see that $\sigma \mapsto z^\sigma$ is injective on $J(B_0) \cap \{\psi(A) = 0\}$, so we have $|\hat{\xi}_1^A| \geq |J(B_0) \cap \{\psi(A) = 0\}|$. Iterating this argument we get

$$|\hat{\xi}_i^A| \geq |J(B_{i-1}) \cap \{\psi(A) = 0\}|, \quad 1 \leq i \leq g. \quad (2.13)$$

Define the events

$$\begin{aligned} F_0 &= \{|B_0| < \tilde{q}m\}, \\ F_i &= \{|B_i| < \tilde{q} \cdot |J(B_{i-1}) \cap \{\psi(A) = 0\}|\}, \quad 1 \leq i < g. \end{aligned}$$

We now claim that

$$\left(\bigcup_{i=0}^{g-1} F_i \right)^c \subset \left\{ |\hat{\xi}_g^A| \geq (1 + \delta)|A| \right\}. \quad (2.14)$$

Indeed, if none of the F_i occurs, we have

$$\begin{aligned} |B_0| &\geq \tilde{q}m; \\ |J(B_0) \cap \{\psi(A) = 0\}| &\geq r \cdot |B_0| - d_1 \geq \tilde{q}rm - d_1; \\ |B_1| &\geq \tilde{q} \cdot |J(B_0) \cap \{\psi(A) = 0\}| \geq \tilde{q}^2rm - \tilde{q}d_1; \\ |J(B_1) \cap \{\psi(A) = 0\}| &\geq r \cdot |B_1| - d_2 \geq (\tilde{q}r)^2m - \tilde{q}rd_1 - d_2; \\ &\dots \\ |J(B_{i-1}) \cap \{\psi(A) = 0\}| &\geq (\tilde{q}r)^i m - (\tilde{q}r)^{i-1}d_1 - (\tilde{q}r)^{i-2}d_2 - \dots - \tilde{q}rd_{i-1} - d_i \end{aligned}$$

for $i \leq g$. In particular, using $\tilde{q}r > 1$ and the definition of expansiveness, for $0 < i \leq g$ we have

$$|J(B_{i-1}) \cap \{\psi(A) = 0\}| \geq (\tilde{q}r)^i m - (\tilde{q}r)^{i-1}d \geq (\tilde{q}r)^{i-1}(\tilde{q}r - 1 - \delta)m. \tag{2.15}$$

By the choice of g , this gives

$$|J(B_{g-1}) \cap \{\psi(A) = 0\}| \geq (1 + \delta)m.$$

Together with (2.13), this proves (2.14).

The proof of the lemma will thus be complete if we show that, for some $c_1 > 0$,

$$P_{\hat{G}_n} \left(\bigcup_{i=0}^{g-1} F_i \right) \leq e^{-c_1 m}. \tag{2.16}$$

We start by writing

$$P_{\hat{G}_n} \left(\bigcup_{i=0}^{g-1} F_i \right) \leq P_{\hat{G}_n}(F_0) + \sum_{i=1}^{g-1} P_{\hat{G}_n} \left(F_i \mid \bigcap_{j=0}^{i-1} F_j^c \right).$$

In order to bound the terms of this sum, we will need the estimate

$$\mathbb{P}(\text{Bin}(k, p) \leq xkp) \leq \exp\{-\gamma(x)kp\} \quad \text{for all } x \in (0, 1),$$

where $\gamma(x) = x \log x - x + 1$. This follows from Markov's inequality; see Lemma 2.3.3 in [4]. We then have

$$P_{\hat{G}_n}(F_0) = \mathbb{P}(\text{Bin}(m, q) < \tilde{q}m) \leq \exp\{-\gamma(\tilde{q}/q)qm\}$$

Also, on the event $\bigcap_{j=0}^{i-1} F_j^c$, by (2.15) we have $|J(B_{i-1}) \cap \{\psi(A) = 0\}| > (\tilde{q}r - 1 - \delta)(\tilde{q}r)^{i-1}m > (\tilde{q}r - 1 - \delta)m$, so

$$P_{\hat{G}_n} \left(F_i \mid \bigcap_{j=0}^{i-1} F_j^c \right) \leq \exp\{-\gamma(\tilde{q}/q)q(\tilde{q}r - 1 - \delta)m\}.$$

The proof of (2.16) is now complete. □

Lemma 2.5. *There exists $\kappa > 0$ such that, putting $K_n = \kappa \cdot n$,*

$$\mathcal{P}_n(\exists A \subset V_n : k_n \leq |A| \leq K_n, \psi(A) \text{ is not expansive}) \xrightarrow{n \rightarrow \infty} 0.$$

Proof. For fixed m we have

$$\begin{aligned} \mathcal{P}_n(\exists A \subset V_n : |A| = m, \psi(A) \text{ is not expansive}) &\leq \sum_{A:|A|=m} \mathcal{P}_n(\psi(A) \text{ is not expansive}) \\ &\leq \sum_{A:|A|=m} \sum_{d=\lceil(1+\delta)m\rceil}^{(1+r+\dots+r^g)m} \sum_{D \subset T_m:|D|=d} \mathcal{P}_n([\psi(A)](\sigma) = 1 \forall \sigma \in D). \end{aligned}$$

We now bound $|\{D \subset T_m : |D| = d\}|$ by $2^{|T_m|}$ and use Lemma 2.3 to bound the probability; the above is less than

$$\binom{n}{m} (1+r+\dots+r^g) m 2^{(1+r+\dots+r^g)m} \left(\frac{(1+r+\dots+r^g)m}{n-r} \right)^{(1+\delta)m} \quad (2.17)$$

Now we use the bound $\binom{n}{m} \leq n^m/m! \leq (ne/m)^m$ (since $e^m = \sum_{i=0}^{\infty} m^i/i! \geq m^m/m!$); (2.17) is less than

$$\left(\frac{ne}{m} \right)^m C^m \left(\frac{m}{n} \right)^{(1+\delta)m} \left(\frac{n}{n-r} \right)^{(1+\delta)m} \leq \left(C \left(\frac{m}{n} \right)^\delta \right)^m;$$

here C is a constant that only depends on r, g and δ , and whose value has changed in the last inequality. Now choose κ such that $C\kappa^\delta < 1/e$. The probability in the statement of the lemma is then less than

$$\sum_{i=k_n}^{K_n} e^{-i} \leq \kappa n e^{-(\log n)^2} \xrightarrow{n \rightarrow \infty} 0.$$

□

Proof of Proposition 2.2. Assume that n is large enough that $\delta k_n > 1$ and that \hat{G}_n satisfies

$$\text{for every } A \subset V_n \text{ with } k_n \leq |A| \leq K_n, \psi(A) \text{ is expansive.} \quad (2.18)$$

Let $\bar{c} = \frac{c_1 \kappa}{2}$, where c_1 and κ are the constants of the two previous lemmas. We will prove that \hat{G}_n is $e^{\bar{c}n}$ -fertile, that is, we will verify that (2.7) holds with $t = e^{\bar{c}n}$. Together with Lemma 2.5, this will imply the result we need.

We start noting that, if $|A| \geq k_n$, then

$$P_{\hat{G}_n} \left(|\hat{\xi}_g^A| < \min(|A| + 1, K_n) \right) < e^{-c_1 \min(|A|, K_n)}. \quad (2.19)$$

Indeed, if $|A| < K_n$, this follows directly from Lemma 2.4 and $(1+\delta)|A| > |A| + \delta k_n > |A| + 1$. If $|A| \geq K_n$, we can take a subset $A' \subset A$ with $|A'| = \lfloor K_n \rfloor$ and use the previous argument for A' together with the fact that $\hat{\xi}_g^{A'} \subset \hat{\xi}_g^A$.

Using (2.19), we have

$$\begin{aligned} P_{\hat{G}_n} \left(|\hat{\xi}_{j-g}^A| \geq \min(|A| + j, K_n) \text{ for } 1 \leq j \leq e^{\bar{c}n} \right) &\geq 1 - \sum_{j=0}^{e^{\bar{c}n}} e^{-c_1 \min(|A|+j, K_n)} \\ &\geq 1 - \sum_{j=0}^{\lfloor K_n - k_n \rfloor} e^{-c_1(k_n+j)} - \sum_{j=\lfloor K_n - k_n \rfloor + 1}^{e^{\bar{c}n}} e^{-c_1 K_n} \\ &\geq 1 - K_n \cdot e^{-c_1 k_n} - e^{\bar{c}n} \cdot e^{-c_1 K_n} \\ &\geq 1 - \kappa n e^{-c_1(\log n)^2} - e^{-\frac{c_1 \kappa n}{2}} > 1 - n^{-2} \end{aligned}$$

when n is large enough, proving (2.7). □

References

- [1] S. Chatterjee, R. Durrett, *Persistence of Activity in Threshold Contact Processes, an “Annealed Approximation” of Random Boolean Networks*, Random Structures and Algorithms 39 (2011), issue 2, 228 - 246. MR-2850270
- [2] B. Derrida, Y. Pomeau, *Random networks of automata: a simple annealed approximation*, Europhysics Letters 1 (1986), 45-49

- [3] Durrett, R. (2010) *Probability: Theory and Examples*. 4th edition. Cambridge Series in Statistical and Probabilistic Mathematics. MR-2722836
- [4] Durrett, R. (2007) *Random Graph Dynamics*. Cambridge University Press. MR-2656427
- [5] S. Kauffman, *Metabolic stability and epigenesis in randomly constructed genetic nets*. Journal of Theoretical Biology 22 (1969), 437-467
- [6] S. A. Kauffman, *Origins of Order: Self-Organization and Selection in Evolution*. Oxford University Press, 1993.

Acknowledgments. We thank the two anonymous referees of ECP for their helpful suggestions and Johel Beltrán for his help in improving the exposition of the paper.