

# On slowly percolating sets of minimal size in bootstrap percolation

Fabricio Benevides\*

Departamento de Matemática  
Universidade Federal do Ceará  
Fortaleza, CE 60455-760, Brazil.

`fabricio@mat.ufc.br`

Michał Przykucki

Department of Pure Mathematics and Mathematical Statistics  
University of Cambridge  
Cambridge, CB3 0WB, England, UK

`m.j.przykucki@dpmms.cam.ac.uk`

Submitted: Jul 12, 2012; Accepted: May 26, 2013; Published: Jun 7, 2013

Mathematics Subject Classifications: 05D99, 60K35

## Abstract

Bootstrap percolation, one of the simplest cellular automata, can be seen as a model of the spread of infection. In  $r$ -neighbour bootstrap percolation on a graph  $G$  we assign a state, infected or healthy, to every vertex of  $G$  and then update these states in successive rounds, according to the following simple local update rule: infected vertices of  $G$  remain infected forever and a healthy vertex becomes infected if it has at least  $r$  already infected neighbours. We say that percolation occurs if eventually every vertex of  $G$  becomes infected.

A well known and celebrated fact about the classical model of 2-neighbour bootstrap percolation on the  $n \times n$  square grid is that the smallest size of an initially infected set which percolates in this process is  $n$ . In this paper we consider the problem of finding the maximum time a 2-neighbour bootstrap process on  $[n]^2$  with  $n$  initially infected vertices can take to eventually infect the entire vertex set. Answering a question posed by Bollobás we compute the exact value for this maximum showing that, for  $n \geq 4$ , it is equal to the integer nearest to  $(5n^2 - 2n)/8$ .

---

\*Supported by CNPq and FUNCAP.

# 1 Introduction

Bootstrap percolation is a particular type of cellular automaton, a concept studied, for example, by von Neumann [10]. Given a graph  $G$ , bootstrap percolation on  $G$  is a class of models that describe the spread of ‘infection’ over the set of vertices. Before describing the model let us fix some notation. In the context of percolation, the vertices of  $G$  are usually called *sites* and the edges of  $G$  *bonds*. For each  $v \in V(G)$ , we denote by  $N(v)$  the set of neighbours of  $v$ . Each site  $v \in V(G)$  is in one of the two *states*, say *healthy* or *infected*; we write  $A$  for the set of sites whose initial state is ‘infected’ and call  $A$  the set of *initially infected sites*.

In this paper we consider a particular family of bootstrap percolation models called *r-neighbour bootstrap percolation*. It consists of the process defined as follows: set  $A_0 = A$  and, thinking of  $A_t$  as the set of sites infected at time  $t$ , for  $t \in \mathbb{N}$  set

$$A_t = A_{t-1} \cup \{v \in V(G) : |N(v) \cap A_{t-1}| \geq r\}. \quad (1)$$

This means that sites of  $G$  become infected if they have at least  $r$  infected neighbours. Note that in bootstrap percolation once a site is infected it never becomes healthy.

The closure of  $A \subset V(G)$  is the set  $\langle A \rangle = \bigcup_{t=0}^{\infty} A_t$  of all sites that are eventually infected. We say that a set  $A$  *percolates* if all sites are eventually infected, that is, if  $\langle A \rangle = V(G)$ . Furthermore, we say that  $A$  takes time  $T$  to percolate if  $\langle A \rangle = V(G)$  and  $T$  is the smallest natural number such that  $A_T = V(G)$ .

Most work in bootstrap percolation has been focused on a particular case where, for some natural numbers  $n$  and  $d$ , the graph  $G$  above is the  $d$ -dimensional grid  $[n]^d$  defined as follows: the set of sites of  $G$  is

$$V(G) = \{(i_1, i_2, \dots, i_d) : 1 \leq i_j \leq n \text{ for all } 1 \leq j \leq d\}$$

and two sites  $v, w \in V(G)$  are neighbours if  $\|v - w\|_1 = 1$ , that is,  $v$  and  $w$  differ in exactly one coordinate and by one unit. For  $d = 2$  we can represent  $G = [n]^2$  by an  $n$  by  $n$  square-grid where each site  $(i, j)$  is a unit square whose center has coordinates  $(i - 1/2, j - 1/2)$  in the Cartesian plane. Note that two sites are adjacent if the corresponding squares share an edge. This particular model was introduced in 1979 by Chalupa, Leath and Reich [7], together with bootstrap percolation on regular trees. In this paper we consider some extremal properties of bootstrap percolation on  $[n]^2$  but the first questions asked about this model were mostly of probabilistic nature. Aizenman and Lebowitz [1] considered the problem where the set of initially infected sites is chosen by selecting sites independently at random with uniform probability  $p$ . They tried to determine for what values of  $p$  an initial set  $A$  chosen this way percolates with high probability. The first sharp result was given by Holroyd [8] in 2003: he determined the asymptotic value of the critical probability in 2-neighbour bootstrap percolation on the square grid. Balogh, Bollobás, Duminil-Copin and Morris [2] proved the corresponding result in all dimensions  $d$  and for all values of the infection threshold  $r$ .

Turning to extremal problems, the size of the smallest percolating sets in  $[n]^d$  was studied by Pete and a summary of his results can be found in [3]. However, the case

$d = r = 2$ , which is now a famous coffee-time problem and a puzzle for high-school students, was answered very early. In this case the answer is  $n$  with a diagonal of the  $n \times n$  square being an obvious example of such a percolating set. This little result turned out to be very useful in proving some early probabilistic results in bootstrap percolation which motivated further research on extremal bootstrap percolation problems. Smallest percolating sets for  $G = [n]^2$  and  $r = 2$  have also been considered by Shapiro and Stephens [13]. A famous and still wide open problem in this area is determining the size of the smallest percolating sets for  $G$  being the  $d$ -dimensional hypercube ( $n = 2$ ) and  $r = 3$ .

Another family of extremal questions that have been considered, posed by Bollobás, are those about the maximum size of minimal percolating sets, i.e., sets that percolate but no proper subsets of which do. This problem was studied by Morris [9] for  $G = [n]^2$  and  $r = 2$  and by Riedl [12] for 2-neighbour bootstrap percolation on a hypercube.

In this paper we answer another question posed by Bollobás, that of bounding the time that a percolating subset  $A$  of the set of vertices of  $G = [n]^2$ , such that  $|A| = n$ , can take to percolate under 2-neighbour bootstrap percolation. Time-related questions have recently attracted a lot of attention also in the probabilistic setup, with some sharp results obtained by Bollobás, Holmgren, Smith and Uzzell [5] and by Bollobás, Smith and Uzzell [6].

For small values of  $n$  it is easy to answer our question computationally by an exhaustive search. But as a main result of this paper we prove the following theorem. Let  $T(A)$  denote the time that  $A$  takes to percolate. Moreover, let

$$M(n) = \max\{T(A) : \langle A \rangle = [n]^2 \text{ and } |A| = n\}.$$

**Theorem 1.** *For every  $n \geq 4$ ,*

$$M(n) = \left\lfloor \frac{5n^2 - 2n}{8} \right\rfloor, \tag{2}$$

where  $\lfloor x \rfloor$  denotes the integer nearest to  $x$ .

In a forthcoming paper [4] we consider a similar problem: finding the maximum percolation time in 2-neighbour bootstrap percolation on  $[n]^2$  among all percolating sets of initially infected sites, i.e., not only those of size  $n$ . For that problem we determine an asymptotic answer showing that the maximum time is  $\frac{13}{18}n^2 + O(n)$ , which is interesting as it shows that taking larger initially infected sets might extend the percolation time significantly. For  $G$  being an  $n$ -dimensional hypercube and  $r = 2$  an analogous question was recently answered by Przykucki [11].

## 2 Preliminaries

Given natural numbers  $k$  and  $\ell$ , a  $k$  by  $\ell$  rectangle is a subset of  $\mathbb{Z}^2$  of the form  $\{a, a + 1, \dots, a + k - 1\} \times \{b, b + 1, \dots, b + \ell - 1\}$  for some choice of  $a$  and  $b$ . Let  $\text{Rec}(k, \ell)$  denote the set of all  $k$  by  $\ell$  rectangles in  $[n]^2$ . We say that a rectangle  $R$  is *internally spanned* by a given set of infected sites  $A$  if  $\langle A \cap R \rangle = R$ .

Given a finite set  $A \subset \mathbb{Z}^2$ , we represent a site  $(i, j) \in A$  as a shaded unit square on the grid so that its center has coordinates  $(i - 1/2, j - 1/2)$  in  $\mathbb{R}^2$ . We define the *boundary* of  $A$  as the set of bonds of  $\mathbb{Z}^2$  having exactly one endpoint in  $A$ ; in our pictures this corresponds to a side shared by a shaded and a non-shaded unit square. The *perimeter* of  $A$  is the number of bonds in its boundary. Its *semi-perimeter* is half of the perimeter; we denote it by  $\Phi(A)$ . In particular, if  $R \in \text{Rec}(k, \ell)$  is a  $k$  by  $\ell$  rectangle then its semi-perimeter is  $\Phi(R) = k + \ell$ .

In our proofs we shall talk about distances between sites and rectangles. The distance we use is given by the usual  $L_1$  norm, i.e., the distance between a pair of sites, say  $(i_1, j_1)$  and  $(i_2, j_2)$ , is  $|i_1 - i_2| + |j_1 - j_2|$ . The distance between two rectangles  $R'$  and  $R''$  is the minimum distance between a site in  $R'$  and a site in  $R''$ ; it is denoted by  $\text{dist}(R', R'')$ .

**Remark.** *This definition of distance coincides with the length of the shortest path from  $R'$  to  $R''$  when viewing  $\mathbb{Z}^2$  as a graph. Note that two rectangles are at distance zero from each other if and only if they intersect; and at distance one if and only if they are disjoint but their boundaries share at least one edge.*

**Fact 2.** *For any two finite sets  $A, B \subset \mathbb{Z}^2$  we have  $\Phi(A) + \Phi(B) \geq \Phi(A \cup B)$ . Equality occurs if and only if  $\text{dist}(A, B) \geq 2$ , that is, if  $A$  and  $B$  do not intersect and have disjoint boundaries.*

From now on let us consider 2-neighbour bootstrap percolation on  $[n]^2$  only. Let us start with the following standard and simple observation which follows immediately from the fact that the perimeter of the infected set cannot grow when a new site becomes infected.

**Observation 3.** *Let  $A$  be a set of infected sites. Then  $\Phi(\langle A \rangle) \leq \Phi(A)$ .*

**Corollary 4.** *Given  $k, \ell \in \mathbb{N}$  and a rectangle  $R \in \text{Rec}(k, \ell)$ , if  $A \subset R$  is a set that internally spans  $R$  then  $|A| \geq \lceil \Phi(R)/2 \rceil = \lceil \frac{k+\ell}{2} \rceil$ . In particular, if  $n \in \mathbb{N}$  and  $A \subset [n]^2$  percolates, then  $|A| \geq n$ .*

It is easy to show that the lower bounds in Corollary 4 are sharp. For example, a diagonal is a percolating set of size  $n$  in  $[n]^2$ .

As we mentioned before, we are interested in finding sets of size  $n$  in  $[n]^2$  that do percolate but do so in the maximum possible time  $M(n)$ . To do this we build a family of sets that percolate in a particular way. In order to do so we shall need to use induction on the size of the underlying graph. Hence it is natural to extend the definition of  $M(n)$  to percolation on rectangles.

Given natural numbers  $k$  and  $\ell$  such that  $k + \ell$  is even (the reason why we only look at even values of  $k + \ell$  will become clear in our proof) let  $T(A)$  again denote the time that  $A$  takes to percolate. We define  $M(k, \ell)$  by

$$M(k, \ell) = \max\{T(A) : \langle A \rangle = [k] \times [\ell] \text{ and } |A| = (k + \ell)/2\}.$$

For a rectangle  $R \in \text{Rec}(k, \ell)$  define  $M(R)$  to be the maximum time in which some set of order  $\Phi(R)/2$  internally spans  $R$ . Of course,  $M(R)$  is just another notation for  $M(k, \ell)$ .

Before trying to compute bounds on  $M(n)$  we should also understand how the infection spreads on a broader scale. The first simple but important observation is the following.

**Fact 5.** *Given any set  $A$  of infected sites,  $\langle A \rangle$  is a union of rectangles such that any distinct two of them are at distance at least 3.*

This fact clearly follows from the following argument. The set  $A$  can be viewed as a union of 1 by 1 rectangles and any two fully infected rectangles within distance at most 2 do span a larger rectangle containing both. The next proposition from Holroyd [8] is a much more precise result in this direction.

**Proposition 6.** *Let  $R$  be a rectangle with area at least 2. Suppose that  $R$  is internally spanned by a set of sites  $A$ . Then there exist disjoint subsets of  $A$ , say  $A'$  and  $A''$ , and rectangles  $R'$  and  $R''$  such that:*

1.  $R' \subsetneq R$  and  $R'' \subsetneq R$ ,
2.  $R'$  is internally spanned by sites in  $A'$  and  $R''$  is internally spanned by sites in  $A''$ ,
3.  $\langle R' \cup R'' \rangle = R$ ; in particular,  $\text{dist}(R', R'') \leq 2$ .

**Remark.** *Note that in Proposition 6 we cannot require the rectangles  $R'$  and  $R''$  to be disjoint (see Figure 1).*

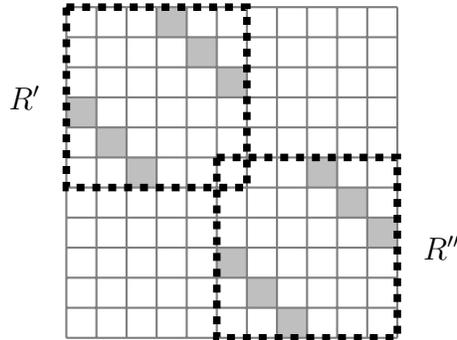


Figure 1: An example where rectangles  $R'$  and  $R''$  are uniquely determined by the initially infected sites and do overlap.

**Remark.** *Although Proposition 6 is sharp, it does not describe the percolation process in a step by step fashion (i.e., as the time  $t$  increases by one). In fact, it may happen that some sites in  $R \setminus (R' \cup R'')$  become infected while some of  $R' \cup R''$  are still healthy. Even though the problem we study is intrinsically time related, we shall be able to make heavy use of Proposition 6.*

### 3 Slowly percolating sets with the minimal number of sites

In this section our aim is to compute the exact value of  $M(n)$  for every  $n \in \mathbb{N}$ . Let us start by giving some intuitions about the solution to this problem. First, we clearly have  $M(n) \leq n^2 - n$ , since at each time step we need to infect at least one of the initially healthy sites to continue the process. Also, without too much effort one can show that  $M(n) \geq \frac{n^2}{2} + O(n)$ . For example, consider the set of initially infected sites of the  $[7]^2$  grid in Figure 2, which generalizes in a self-explanatory way to the  $[n]^2$  grid. It is clear that with this particular starting set at each time step, except the first one, at most two new sites become infected.

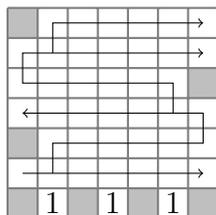


Figure 2: An initial set showing that  $M(n) \geq \frac{n^2}{2} + O(n)$ .

This shows that  $M(n) = \Theta(n^2)$ . As a main result of this paper we prove that the structure of sets maximizing percolation time is actually more sophisticated. To be more precise, we show that to infect a  $k \times \ell$  rectangle  $R$  in the maximum time we should use an initially infected set  $A = A' \cup \{v\}$  such that the set  $A'$  first internally spans either a  $(k - 1) \times (\ell - 1)$ ,  $(k - 2) \times \ell$  or  $k \times (\ell - 2)$  rectangle in the maximum possible time, and then using 'help' from the site  $v$  finishes the infection of  $R$ . As a consequence, when  $k = \ell = n$ , the structure of a set  $A(n)$  with  $|A(n)| = n$  maximizing percolation time can be described as follows. We have  $A(n) = B(n) \cup C(n) \cup D(n)$ , where  $|B(n)| \approx n/4$ ,  $|C(n)| \approx n/2$  and  $|D(n)| \approx n/4$ , and such that

1. Set  $B(n)$  internally spans a rectangle of size roughly  $\frac{n}{2} \times 2$  in time  $\frac{3n}{4} + O(1)$ ,
2. Set  $C(n)$  extends the infected area to a rectangle of size roughly  $n \times \frac{n}{2}$  in time  $\frac{3n^2}{8} + O(n)$ ,
3. Set  $D(n)$  finishes the infection of the  $n \times n$  grid in time  $\frac{n^2}{4} + O(n)$ .

Thus the set  $A(n)$  percolates in time  $5n^2/8 + O(n)$ , significantly beating the simple construction presented in Figure 2.

Now, let us return to Proposition 6. Given  $A \subset [n]^2$ , consider the sets  $A'$ ,  $A''$ ,  $R'$  and  $R''$ , given by Proposition 6, and assume that  $A'$  takes at least as many time steps to internally span  $R'$  as  $A''$  takes to internally span  $R''$ . Then clearly we can bound from above the time that  $A$  takes to percolate  $R$  by the time  $A'$  takes to infect  $R'$  plus the time to grow from  $R' \cup R''$  to  $R$ , that is, to infect all sites in  $R \setminus (R' \cup R'')$  given that all

sites in  $R'$  and  $R''$  are infected. Intuitively, the time to grow from  $R' \cup R''$  to  $R$  does not change much if we only slightly change the sizes of  $R'$  and  $R''$  while the infection time of  $R'$  might grow a lot if we increase the side lengths of  $R'$  even by some small quantities (this follows from our intuitions about the quadratic growth of  $M$ ). It is then intuitive that, to maximize the time that  $A$  takes to percolate,  $R''$  should probably be as small as possible, maybe even a single site. Let us now make our arguments formal.

First, we consider a family of sets of initially infected sites that internally span a rectangle in a particular way. The following definition is the most important concept of this paper.

**Definition 1.** Let  $k$  and  $\ell$ , with  $k + \ell$  even, be given natural numbers. We say that a set  $A$  of initially infected sites is  $(k, \ell)$ -good if it has cardinality  $(k + \ell)/2$  and the 2-neighbour bootstrap percolation process starting from  $A$  can be described as follows. There exists a nested sequence of rectangles  $P_0 \subset P_1 \subset \dots \subset P_r \in \text{Rec}(k, \ell)$ , such that  $P_i \in \text{Rec}(s_i, t_i)$  where  $s_i, t_i$  satisfy the following properties:

1. either  $s_0 \leq 2$  or  $t_0 \leq 2$  or  $s_0 = t_0 = 3$ ; and  $s_1, t_1 \geq 3$  and  $(s_1, t_1) \neq (3, 3)$ ,
2. for every  $1 \leq i \leq r$ , the rectangle  $P_i$  is in
 
$$\text{Rec}(s_{i-1} + 1, t_{i-1} + 1) \cup \text{Rec}(s_{i-1} + 2, t_i) \cup \text{Rec}(s_i, t_{i-1} + 2),$$
3. for all  $0 \leq i \leq r$ , the rectangles  $P_i$  are internally spanned by  $A \cap P_i$  in the maximal possible time, that is, in time  $M(P_i)$ ,
4. for every  $0 \leq i \leq r$ , if the rectangle  $P_i$  has no side of length 1 then among the sites which become infected last in  $P_i$  there is at least one of its corner sites,
5. for every  $0 \leq i < r$ , there exists a site  $v_i \in A$  such that  $P_i \cup \{v_i\}$  internally spans  $P_{i+1}$  and  $v_i$  is at distance exactly 2 from one of the last sites to become infected in  $P_i$  and at distance at least 3 from any other site in  $P_i$  (see Figure 3).

**Definition 2.** If  $A$  is  $(k, \ell)$ -good we say that  $P_0 \subset P_1 \subset \dots \subset P_r \in \text{Rec}(k, \ell)$  is a good sequence of rectangles associated with  $A$  if it satisfies conditions (1)-(5) above.

From condition (2) it follows that for every  $0 \leq i \leq r - 1$  we have  $\Phi(R_{i+1}) = \Phi(R_i) + 2$ . From condition (3), taking  $i = r$ , it follows that any  $(k, \ell)$ -good set infects a  $k \times \ell$  rectangle in the maximum possible time. We shall show that for every  $n \geq 4$  there exists an  $(n, n)$ -good set  $A$ .

For a  $(k, \ell)$ -good set  $A \subset [k] \times [\ell]$  and a good sequence  $P_0 \subset P_1 \subset \dots \subset P_r = [k] \times [\ell]$  associated with it, we say that we use *Move 1* at moment  $i$  (to construct  $P_i$  from  $P_{i-1}$ ) if  $P_i \in \text{Rec}(s_{i-1} + 1, t_{i-1} + 1)$ , that we use *Move 2* at moment  $i$  if  $P_i \in \text{Rec}(s_{i-1} + 2, t_{i-1})$  and that we use *Move 3* at moment  $i$  if  $P_i \in \text{Rec}(s_{i-1}, t_{i-1} + 2)$ .

We shall prove a recursive formula for  $M(k, \ell)$  that works for all values of  $k$  and  $\ell$  such that  $k + \ell$  is even. The reader should keep in mind the description of  $(k, \ell)$ -good initial sets as we are going to build such a set in our proof. In the next two lemmas we deal with some small cases which we will later use as base cases for the recursion. Since  $M(k, \ell) = M(\ell, k)$ , we shall omit some cases where  $k < \ell$ .

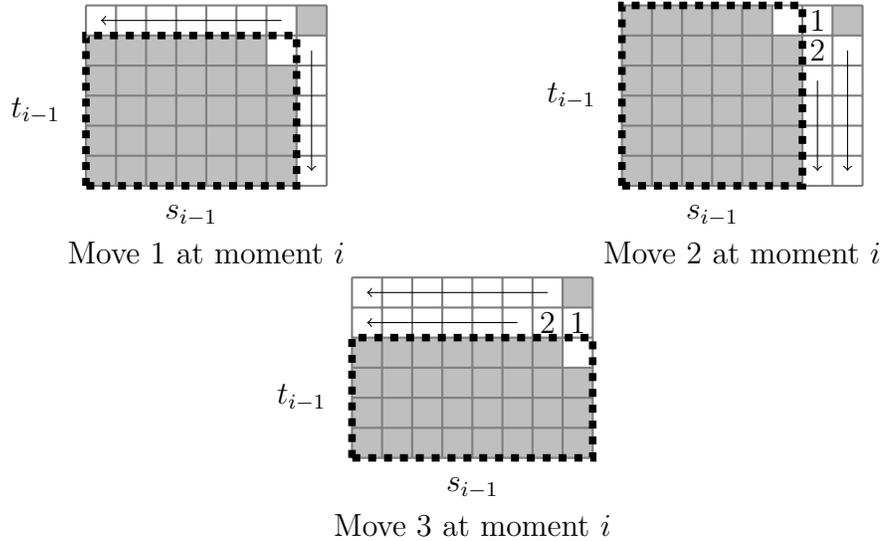


Figure 3: Moves 1, 2 and 3.

**Lemma 7.** *We have  $M(1, 1) = 0$ ;  $M(k, 1) = 1$  for all odd  $k \geq 3$ ; and  $M(3, 3) = 4$ . Furthermore, in all these cases there exist  $(k, \ell)$ -good sets.*

*Proof.* The proof of this lemma is easy and we leave it as an exercise to the reader.  $\square$

**Lemma 8.** *For any even  $k$  we have  $M(k, 2) = (3k - 4)/2$ . Furthermore, there is a  $(k, 2)$ -good set,  $A^0(k, 2)$ , which percolates  $[k] \times [2]$  in time  $M(k, 2)$ .*

*Proof.* We define  $A^0(k, 2)$  to be the set of shaded sites in Figure 4. Clearly  $|A^0(k, 2)| = (k + 2)/2$  and  $A^0(k, 2)$  percolates  $[k] \times [2]$  in time  $(3k - 4)/2$ . Thus we have  $M(k, 2) \geq (3k - 4)/2$  for any  $k$  even. Note that, setting  $P_0 = [k] \times [2]$ , to prove that  $A^0(k, 2)$  is a  $(k, 2)$ -good set we only need to show that in fact  $M(k, 2) = (3k - 4)/2$ .

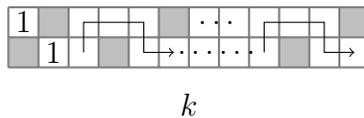


Figure 4: A set of initially infected sites which gives the maximum percolation time on  $[k] \times [2]$  when  $k$  is even.

Now we prove by induction on  $k$  that for any  $k$  even we have  $M(k, 2) \leq (3k - 4)/2$ . Clearly,  $M(2, 2) = 1$ . Assume that we are given some even  $k \geq 4$  and that  $M(k - 2, 2) = (3k - 10)/2$ . Let  $A$ , with  $|A| = (k + 2)/2$ , be any set that percolates  $[k] \times [2]$ .

Since  $A$  percolates, any two consecutive columns of  $[k] \times [2]$  contain at least one site of  $A$ . In particular, each of the 2 by 2 squares of the form  $\{2i - 1, 2i\} \times \{1, 2\}$ ,  $1 \leq i \leq k/2$ , must contain at least one site of  $A$ . So only one of such squares can contain two sites of

A. Therefore, either  $\{1, 2\} \times \{1, 2\}$  or  $\{k - 1, k\} \times \{1, 2\}$  contains exactly one site of  $A$ . Assume without loss of generality that the latter holds. Since  $A$  percolates, either  $(k, 1)$  or  $(k, 2)$  must be an initially infected site. Again without loss of generality we may assume that the latter holds. In this setting it is trivial to check that  $A$  must internally span  $[k - 2] \times [2]$ . Therefore  $A$  takes time at most  $M(k - 2, 2) + 3 = (3k - 4)/2$  to percolate. This completes the proof.  $\square$

Now, we state a lemma giving a recursive formula for  $M(k, \ell)$ . Let us note that in the formula we are about to prove the sum of the parameters that the function  $M(\cdot, \cdot)$  depends on at each recursive step decreases by two. This is why, being interested in the value of  $M(n, n)$ , we only need to look at values of  $k$  and  $\ell$  with even  $k + \ell$ .

**Lemma 9.** *For  $k, \ell \geq 3$  such that  $(k, \ell) \neq (3, 3)$  and  $k + \ell$  is even, we have*

$$M(k, \ell) = \max \begin{cases} M(k - 1, \ell - 1) + \max\{k, \ell\} - 1, \\ M(k - 2, \ell) + \ell + 1, \\ M(k, \ell - 2) + k + 1. \end{cases} \quad (3)$$

Furthermore, for all such  $k$  and  $\ell$  there exists a  $(k, \ell)$ -good set.

*Proof.* We prove Lemma 9 by induction on  $k + \ell$ . Assume that we are given  $k, \ell \geq 3$  such that  $(k, \ell) \neq (3, 3)$  and  $k + \ell$  is even. Our induction hypothesis is that for any  $k', \ell'$  such that  $k' + \ell'$  is even and  $k' + \ell' < k + \ell$  there exists a  $(k', \ell')$ -good set  $A^0(k', \ell')$  which percolates in time  $M(k', \ell')$ , as in the statement of Lemma 9.

The fact that  $k, \ell \geq 3$  guarantees that we have  $k - 1, \ell - 1 \geq 2$ . This will be important for us as in the constructions below we shall use property (4) of Definition 1 of  $(k, \ell)$ -good sets a lot. We shall first prove that the following inequality holds for  $k$  and  $\ell$  as above.

$$M(k, \ell) \geq \max \begin{cases} M(k - 1, \ell - 1) + \max\{k, \ell\} - 1, \\ M(k - 2, \ell) + \ell + 1, \\ M(k, \ell - 2) + k + 1. \end{cases} \quad (4)$$

Consider the following three particular ways of infecting  $[k] \times [\ell]$  (see Figure 3).

- a. By the induction hypothesis there exists a  $(k - 1, \ell - 1)$ -good set  $A^0(k - 1, \ell - 1)$  which internally spans the rectangle  $[k - 1] \times [\ell - 1]$  in time  $M(k - 1, \ell - 1)$ . Without loss of generality, since  $k - 1, \ell - 1 \geq 2$ , we may assume that the site  $(k - 1, \ell - 1)$  becomes infected at time  $M(k - 1, \ell - 1)$ . Let  $A_1(k, \ell) = A^0(k - 1, \ell - 1) \cup \{(k, \ell)\}$ . Then the infection of sites in  $([k] \times [\ell]) \setminus ([k - 1] \times [\ell - 1])$  starts only after  $(k - 1, \ell - 1)$  is infected and so  $A_1(k, \ell)$  takes time  $M(k - 1, \ell - 1) + \max\{k, \ell\} - 1$  to internally span  $[k] \times [\ell]$ . In addition, note that at least one of the corner sites  $(k, 1), (1, \ell)$  becomes infected at time  $M(k - 1, \ell - 1) + \max\{k, \ell\} - 1$ .
- b. When  $k \geq 4$ , by the induction hypothesis there exists a  $(k - 2, \ell)$ -good set  $A^0(k - 2, \ell)$ , internally spanning the rectangle  $[k - 2] \times [\ell]$  in time  $M(k - 2, \ell)$ , which infects the

site  $(k-2, \ell)$  at time  $M(k-2, \ell)$  (this follows from the fact that  $k-2, \ell \geq 2$ ). Let  $A_2(k, \ell) = A^0(k-2, \ell) \cup \{(k, \ell)\}$ . Then the infection of sites in  $([k] \times [\ell]) \setminus ([k-2] \times [\ell])$  starts only after  $(k-2, \ell)$  is infected and so  $A_2(k, \ell)$  takes time  $M(k-2, \ell) + \ell + 1$  to internally span  $[k] \times [\ell]$ . In addition, note that the corner site  $(k, 1)$  becomes infected at time  $M(k-2, \ell) + \ell + 1$ .

- c. When  $\ell \geq 4$ , analogously to case b), by the induction hypothesis there exists a  $(k, \ell-2)$ -good set  $A^0(k, \ell-2)$ , internally spanning the rectangle  $[k] \times [\ell-2]$  in time  $M(k, \ell-2)$ , which infects the site  $(k, \ell-2)$  at time  $M(k, \ell-2)$ . Then the set  $A_3(k, \ell) = A^0(k, \ell-2) \cup \{(k, \ell)\}$  internally spans  $[k] \times [\ell]$  in time  $M(k, \ell-2) + k + 1$ , with the corner site  $(1, \ell)$  becoming infected at the last time step.

The above constructions show that inequality (4) holds when  $k, \ell \geq 4$ . It remains to check that it also holds for  $k = 3$  and  $\ell \geq 5$ , and for  $\ell = 3$  and  $k \geq 5$  (recall that  $k + \ell$  is even so, e.g., when  $k = 3$  we have  $\ell \neq 4$ ). These cases are clearly symmetric so let us just show that  $M(k, 3) \geq M(k, 1) + k + 1$  for  $k \geq 5$ . This is immediate as  $M(k, 1) + k + 1 = k + 2$  and we already know that

$$M(k, 3) \geq M(k-1, 2) + k - 1 \geq M(4, 2) + k - 1 \geq k + 3.$$

Thus the lower bound on  $M(k, \ell)$  is proved.

Let us now show that the set  $A_1(k, \ell)$  defined above satisfies all but possibly condition (3) of a  $(k, \ell)$ -good set. Showing that the same holds for the sets  $A_2(k, \ell)$  and  $A_3(k, \ell)$  is analogous.

Thus, assume that  $A^0(k-1, \ell-1)$  is a  $(k-1, \ell-1)$ -good set with a good sequence of rectangles  $P_0 \subset P_1 \subset \dots \subset P_r = [k-1] \times [\ell-1]$  associated with it. First, clearly  $|A_1(k, \ell)| = |A^0(k-1, \ell-1)| + 1 = (k+\ell-2)/2 + 1 = (k+\ell)/2$ . Now consider the sequence  $P_0 \subset P_1 \subset \dots \subset P_r = [k-1] \times [\ell-1] \subset P_{r+1} = [k] \times [\ell]$  which describes the infection of  $[k] \times [\ell]$  with  $A_1(k, \ell)$  as the set of initially infected sites. This sequence clearly satisfies properties (1) and (2) of  $(k, \ell)$ -good sets since it is obtained from a  $(k-1, \ell-1)$ -good set. It satisfies property (4) for  $i = r+1$  since, as we noticed, in the infection started from  $A_1(k, \ell)$  at least one of the corner sites  $(k, 1), (1, \ell)$  becomes infected at the last time step. It satisfies property (5) for  $i = r+1$  since  $(k, \ell)$  is at distance 2 from  $(k-1, \ell-1)$  (which is infected last) and at distance at least 3 from any other site in  $P_r = [k-1] \times [\ell-1]$ . Properties (4) and (5) for  $i \leq r$  are satisfied for this sequence since it is obtained from one associated with a  $(k-1, \ell-1)$ -good set.

We shall show that at least one of the sets  $A_1(k, \ell), A_2(k, \ell)$  and  $A_3(k, \ell)$  is  $(k, \ell)$ -good by proving an upper bound on  $M(k, \ell)$  analogous to inequality (4), that is,

$$M(k, \ell) \leq \max \begin{cases} M(k-1, \ell-1) + \max\{k, \ell\} - 1, \\ M(k-2, \ell) + \ell + 1, \\ M(k, \ell-2) + k + 1. \end{cases} \quad (5)$$

This will mean that at least one of these sets satisfies the missing property (3) of a  $(k, \ell)$ -good set.

Consider any set  $A$  which internally spans the rectangle  $R = [k] \times [\ell]$  in time  $M(k, \ell)$  and is such that  $|A| = (k + \ell)/2$ . By Proposition 6, there exist disjoint subsets of  $A$ , say  $A'$  and  $A''$ , and two rectangles  $R'$  and  $R''$  satisfying conditions (1)–(3) of Proposition 6. By Proposition 3 and condition (3) of Proposition 6, we have that

$$\Phi(R' \cup R'') \geq \Phi(\langle R' \cup R'' \rangle) = \Phi(R) = k + \ell.$$

By Fact 2, condition (2) of Proposition 6 and Corollary 4,

$$\Phi(R' \cup R'') \leq \Phi(R') + \Phi(R'') \leq 2|A'| + 2|A''| \leq 2|A| = k + \ell.$$

Therefore, each of the above inequalities must be an equality. In particular, we have  $\Phi(R' \cup R'') = \Phi(R') + \Phi(R'')$ . Fact 2 implies that  $\text{dist}(R', R'') \geq 2$ , which together with condition (3) of Proposition 6 gives that  $R'$  and  $R''$  must be at distance exactly 2. Also, we must have  $\Phi(R') = 2|A'|$  and  $\Phi(R'') = 2|A''|$ , therefore, both  $\Phi(R')$  and  $\Phi(R'')$  are even.

Let  $s_1, t_1, s_2, t_2 \geq 1$  be such that  $R' \in \text{Rec}(s_1, t_1)$  and  $R'' \in \text{Rec}(s_2, t_2)$ . We have  $\Phi(R') + \Phi(R'') = \Phi(R)$ , so  $s_1 + s_2 + t_1 + t_2 = k + \ell$ . Since  $R'$  and  $R''$  must be at distance exactly 2, the values of  $s_1, t_1, s_2, t_2$  and the positions of  $R'$  and  $R''$  inside  $R$ , must satisfy exactly one of the following conditions (to avoid redundancy we do not list cases analogous to Conditions (a), (b) and (c) when the alignment of  $R'$  and  $R''$  in  $R$  is a rotation by 90 degrees of the one we consider here).

Condition (a): rectangles  $R'$  and  $R''$  align like in Figure 5 (a) with  $s_1 + s_2 = k - 1$  and  $t_1 + t_2 = \ell + 1$ .

Condition (b): rectangles  $R'$  and  $R''$  align like in Figure 5 (b) with  $s_1 + s_2 = k$ ,  $t_1 + t_2 = \ell$ .

Condition (c): there is an  $0 \leq m \leq t_1 - t_2$  so that the rectangles  $R'$  and  $R''$  align as in Figure 5 (c) with  $s_1 + s_2 = k - 1$ ,  $t_1 = \ell$  and  $t_2 = 1$ .

Additionally, the rectangles  $R'$  and  $R''$  are nonempty and internally spanned by  $\frac{s_1+t_1}{2}$  and  $\frac{s_2+t_2}{2}$  sites respectively.

Note now that no matter which of the Conditions (a), (b) or (c) holds, if at least one of  $s_1, t_1, s_2, t_2$  equals 1 (which for Condition (c) is true by definition with  $t_2 = 1$ ) then, just by possibly moving sites from  $A'$  to  $A''$  or the other way, we can find a partition of  $A$ , say into sets  $\tilde{A}'$  and  $\tilde{A}''$ , such that  $\langle \tilde{A}' \rangle = \tilde{R}'$  is a rectangle,  $\langle \tilde{A}'' \rangle = \tilde{R}''$  is a single site and  $\langle \tilde{R}' \cup \tilde{R}'' \rangle = R$ . Now, returning to the intuitions we gave at the beginning of this section, we can bound from above the time that  $A$  takes to percolate  $[k] \times [\ell]$  by the larger of the maximal times needed to internally span  $R'$  or  $R''$ , plus the time to grow from  $R' \cup R''$  to  $R$ , that is, to infect all sites in  $R \setminus (R' \cup R'')$  given that all sites in  $R'$  and  $R''$  are infected. So if such  $\tilde{R}''$  consisting of a single site can be found then the percolation time clearly cannot be greater than the lower bound given by inequality (4), in which case we are done. Assume therefore this is not the case which allows us to ignore Condition

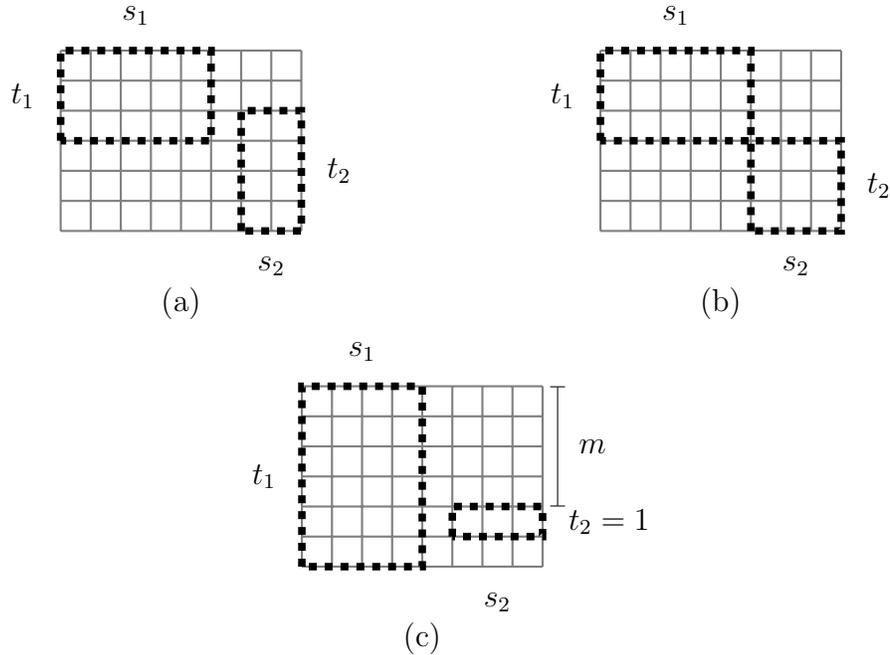


Figure 5: Three possible alignments of rectangles  $R'$  and  $R''$ .

(c). Thus we only need to consider Conditions (a) and (b) with  $s_1, t_1, s_2, t_2 \geq 2$ . For these Conditions we are also free to assume that  $M(R') \geq M(R'')$ .

Therefore, the time  $A$  takes to percolate is at most

$$\begin{cases} M(s_1, t_1) + \max\{s_1 + t_2, s_2 + t_1\}, & \text{if Condition (a) holds,} \\ M(s_1, t_1) + \max\{s_1 + t_2 - 1, s_2 + t_1 - 1\}, & \text{if Condition (b) holds.} \end{cases} \quad (6)$$

From (4) and small case analysis when  $s$  or  $t$  equals 2, we have that the bound  $M(s, t) \geq M(s - 1, t - 1) + \max\{s, t\} - 1$  holds for all  $2 \leq s \leq k$ ,  $2 \leq t \leq \ell$ ,  $s + t$  even. Also,  $M(\cdot, \cdot)$  is clearly non-decreasing in both parameters.

If Condition (a) holds then since  $s_1, t_1, s_2, t_2 \geq 2$  we also have  $s_1, s_2 \leq k - 3$  and  $t_1, t_2 \leq \ell - 1$ . Then

$$\begin{aligned} M(s_1, t_1) + \max\{s_1 + t_2, s_2 + t_1\} &\leq M(k - 3, \ell - 1) + k + \ell - 4 \\ &\leq M(k - 2, \ell) + k + \ell - 4 - (\max\{k - 2, \ell\} - 1) \\ &\leq M(k - 2, \ell) + \min\{\ell - 1, k - 3\} \\ &< M(k - 2, \ell) + \ell + 1. \end{aligned}$$

In case the rectangles  $R'$  and  $R''$  satisfy an analogous condition obtained by rotating Condition (a) by 90 degrees, we get an analogous bound  $M(k, \ell - 2) + k + 1$  for the percolation time of  $A$ .

If Condition (b) holds then since  $s_1, t_1, s_2, t_2 \geq 2$  we also have  $s_1, s_2 \leq k - 2$  and

$t_1, t_2 \leq \ell - 2$ . Then

$$\begin{aligned}
 M(s_1, t_1) + \max\{s_1 + t_2, s_2 + t_1\} - 1 &\leq M(k - 2, \ell - 2) + k + \ell - 5 \\
 &\leq M(k - 1, \ell - 1) + k + \ell - 5 - (\max\{k, \ell\} - 2) \\
 &\leq M(k - 1, \ell - 1) + \min\{\ell, k\} - 3 \\
 &< M(k - 1, \ell - 1) + \max\{k, \ell\} - 1.
 \end{aligned}$$

Thus we conclude that the weakest upper bound on percolation time of  $A$ , equal to

$$\max\{M(k - 1, \ell - 1) + \max\{k, \ell\} - 1, M(k - 2, \ell) + \ell + 1, M(k, \ell - 2) + k + 1\},$$

is obtained when one of  $R'$  or  $R''$  is a single site. Since  $A$  was arbitrary with  $|A| = (k + \ell)/2$  and  $T(A) = M(k, \ell)$ , this is an upper bound on  $M(k, \ell)$  and so (5) is proved. Since this upper bound matches the percolation time of at least one of the sets  $A_1(k, \ell)$ ,  $A_2(k, \ell)$ ,  $A_3(k, \ell)$  constructed in the proof of the lower bound on  $M(k, \ell)$ , we see that at least one of them percolates in time  $M(k, \ell)$ . This was the last step needed to show that one of them is a  $(k, \ell)$ -good set. This completes the proof of Lemma 9.  $\square$

By Lemma 9 for every  $n \geq 4$  there exists an  $(n, n)$ -good set which percolates  $[n]^2$  in the maximal time  $M(n)$ . So, it is enough to determine  $s_0, t_0$  and the sequence of moves 1, 2 and 3 which takes the longest time to percolate. In the next lemma we treat a number of small cases to exclude some, a priori possible, values for the numbers  $s_0$  and  $t_0$ .

**Lemma 10.** *Let  $k$  and  $\ell$  be such that  $k \geq 4, \ell \geq 2$  and  $k + \ell$  is even. Then there exists a  $(k, \ell)$ -good set  $A^0(k, \ell)$  with the sequence  $P_0 \subset P_1 \subset \dots \subset P_r = [k] \times [\ell]$  of rectangles associated with it, with  $P_0 \in \text{Rec}(s, 2) \cup \text{Rec}(2, s)$  for some even  $s \geq 4$ .*

*Proof.* Given  $k, \ell$ , with  $k \geq 4, \ell \geq 2$  and  $k + \ell$  is even, consider any  $(k, \ell)$ -good set,  $A^0(k, \ell)$ , and its associated good sequence of rectangles  $P_0 \subset P_1 \subset \dots \subset P_r = [k] \times [\ell]$ . If  $\ell = 2$  then we have  $r = 0$  and the lemma is trivial. Thus assume that  $\ell \geq 3$ . Then since  $k \geq 4$  we also have  $r \geq 1$ .

Suppose for a contradiction that  $P_0 \in \text{Rec}(s, 1)$ , for some odd  $s$ . By the definition of a  $(k, \ell)$ -good set we have  $P_1 \in \text{Rec}(s_1, t_1)$  with  $s_1, t_1 \geq 3$  and  $\max\{s_1, t_1\} \geq 4$ . The only move we can apply to  $P_0$  to satisfy this is Move 3, so we must have  $P_1 \in \text{Rec}(s, 3)$  with  $s \geq 5$  (recall that  $s_1 + t_1$  is even). This implies  $s - 1 \geq 4$  and so  $M(s - 1, 2) \geq 4$ . By inequality (4) we obtain  $M(P_1) = M(s, 3) \geq M(s - 1, 2) + s - 1 \geq s + 3$ . However, if we apply Move 3 to  $P_0 \in \text{Rec}(s, 1)$  we will percolate  $P_1 \in \text{Rec}(s, 3)$  in time at most  $M(P_0) + s + 1 = s + 2$ . This contradicts the fact that  $A^0(k, \ell)$  is  $(k, \ell)$ -good (more precisely, property (3) of Definition 1 will not hold for  $P_1$ ). We deal with the case  $P_0 \in \text{Rec}(1, s)$  analogously.

Suppose now that  $P_0 \in \text{Rec}(3, 3)$ . We can assume that either  $P_1 \in \text{Rec}(4, 4)$  (if we use Move 1 at moment 1) or  $P_1 \in \text{Rec}(5, 3)$  (if we use Move 2), as the case  $P_1 \in \text{Rec}(3, 5)$  (where we use Move 3) is analogous. In the first case,  $M(P_0) = M(3) = 4$  and it takes 3 time steps to finish the infection of  $P_1$  after  $P_0$  has been fully infected. Thus  $P_1$  becomes fully infected after at most  $4 + 3 = 7$  time steps. However, by inequality (4) we know

that  $M(P_1) = M(4) \geq M(4, 2) + 4 + 1 = 9$ . So, like in the previous paragraph, we have a contradiction to  $A^0(k, \ell)$  being  $(k, \ell)$ -good. In the second case, where  $P_1 \in \text{Rec}(5, 3)$ , it takes 4 time steps to apply Move 2 to  $P_0$  and finish the infection of  $P_1$  after  $P_0$  is fully infected. Thus  $P_1$  is fully infected at time  $M(P_0) + 4 = 8$ . However, starting from  $P'_0 \in \text{Rec}(4, 2)$  and using Move 1 at moment 1 we obtain infection time of  $P_1$  equal to  $M(4, 2) + 4 = 8$ . This does not contradict the  $(k, \ell)$ -goodness of  $A^0(k, \ell)$  but shows that there exists a  $(k, \ell)$ -good set  $A'(k, \ell)$  with the sequence  $P'_0 \subset P_1 \subset \dots \subset P_r = [k] \times [\ell]$  of rectangles associated with it. This completes the proof of Lemma 10.  $\square$

Let  $k$  and  $\ell$  be such that  $k \geq 4, \ell \geq 2$  and  $k + \ell$  is even. By Lemma 10 we know that there exists an  $(k, \ell)$ -good set  $A^0(k, \ell)$  with the sequence  $P_0 \subset P_1 \subset \dots \subset P_r = [k] \times [\ell]$  of rectangles associated with it, with  $P_0 \in \text{Rec}(s, 2) \cup \text{Rec}(2, s)$  for some even  $s \geq 4$ . Note that for each  $i \geq 1$  the infection of the sites in  $P_i \setminus P_{i-1}$  starts only after all sites in  $P_{i-1}$  are infected, which by the definition of  $(k, \ell)$ -good sets happens at time  $M(P_{i-1})$ . The following two observations are crucial to determine the precise value of  $M(n)$ . In fact, with those observations and equation (3) we shall be able to find an  $(n, n)$ -good percolating set, i.e., a set which takes time exactly  $M(n)$  to percolate.

**Observation 11.** *For any  $i \geq 1$ , no matter which of moves 1, 2 or 3 is used at moment  $i$  to extend the rectangle  $P_{i-1}$  to  $P_i$ , at most two new sites become infected at each time step between  $M(P_{i-1}) + 1$  and  $M(P_i)$ .*

By Observation 11, having fixed  $P_0$  and remembering that in our problem the number of initially infected sites is fixed, a sequence of moves 1, 2 and 3 that maximizes the time to infect a rectangle  $R$  must also maximize the number of time steps after  $M(P_0)$  at which only one new site of  $R \setminus A$  becomes infected. This observation also allows us to change the way we think about maximizing percolation time. Instead of thinking of the exact time it takes to apply a particular Move  $j$  at step  $i$  we shall think of a *score* of such move which is equal to the number of time steps at which exactly one new site becomes infected when we use Move  $j$ . Then our task becomes to maximize the cumulative score of our sequence of moves.

**Observation 12.** *For any  $i \geq 1$  the following statements hold.*

1. *If Move 1 is used at moment  $i$  to extend the rectangle  $P_{i-1} \in \text{Rec}(s_{i-1}, t_{i-1})$  to  $P_i \in \text{Rec}(s_{i-1} + 1, t_{i-1} + 1)$  then only one new site becomes infected at exactly  $|s_{i-1} - t_{i-1}|$  time steps between  $M(P_{i-1}) + 1$  and  $M(P_i)$ , i.e., at  $M(P_i) - |s_{i-1} - t_{i-1}| + 1, M(P_i) - |s_{i-1} - t_{i-1}| + 2, \dots, M(P_i) - 1$  and  $M(P_i)$ .*
2. *If Move 2 or Move 3 is used at moment  $i$  to extend the rectangle  $P_{i-1}$  to  $P_i$  then only one new site becomes infected at exactly 3 time steps between  $M(P_{i-1}) + 1$  and  $M(P_i)$ , i.e., at  $M(P_{i-1}) + 1, M(P_{i-1}) + 2$  and  $M(P_i)$ .*

Using these observations we get the next important claim. To talk about sequences of moves we shall use the following notation similar to that of regular expressions. We say that a finite (possibly empty) sequence is of the form  $[a_1|a_2| \dots |a_r]^*$  if all its terms belong

to  $\{a_1, \dots, a_r\} \subset \{1, 2, 3\}$ . We concatenate these expressions to create more general ones which describe the corresponding sets of concatenated sequences. For example, each of the sequences 22133232, 112333, 121233 is of the form  $[1]^*[2]^*[1]^*[2][3]^*$ , but 122331 is not.

**Claim 13.** *For  $k \geq 4, \ell \geq 2$ , there exists a  $(k, \ell)$ -good set  $A$  internally spanning the rectangle  $R \in \text{Rec}(k, \ell)$ , with a good sequence  $P_0 \subset P_1 \subset \dots \subset P_r = R$  associated with it, with  $P_0 \in \text{Rec}(s, 2) \cup \text{Rec}(2, s)$  for some  $s \geq 4$ , such that the sequence of moves  $(m_1, m_2, \dots, m_r)$  used to fully infect  $P_r$  from  $P_0$  is of the form  $[2]^*[1]^*[3]^*$  or of the form  $[3]^*[1]^*[2]^*$ .*

*Proof.* Let us fix an even  $s \geq 4$  and assume that  $P_0 \in \text{Rec}(s, 2) \cup \text{Rec}(2, s)$ . Note that this uniquely defines  $r = n - s/2 - 1$ , which is also the number of initially infected sites outside  $P_0$ . By Observation 12 we immediately see that in such a sequence we should apply moves 1 to rectangles  $P_i \in \text{Rec}(s_i, t_i)$  with as large as possible difference  $|s_i - t_i|$  between the length of the longer side and the length of the shorter side of  $P_i$ . We also note that whenever Move 1 is applied, say to obtain  $P_{i+1} \in \text{Rec}(s_{i+1}, t_{i+1})$  from  $P_i \in \text{Rec}(s_i, t_i)$ , then this difference does not change, i.e.,  $|s_{i+1} - t_{i+1}| = |s_i + 1 - (t_i + 1)| = |s_i - t_i|$ .

If Move 1 does not occur in  $(m_1, m_2, \dots, m_r)$  then every move in the sequence has a constant score 3 depending neither on the step at which it is applied nor on the dimensions of the rectangle it is applied to. Thus every permutation of  $(m_1, m_2, \dots, m_r)$  has the same score and we can clearly rearrange the sequence of moves to make it be of the form  $[2]^*[1]^*[3]^*$  (or in fact  $[2]^*[3]^*$ ) without changing percolation time.

Assume that Move 1 occurs only once in  $(m_1, m_2, \dots, m_r)$ , say that  $m_k = 1$  and  $m_j \in \{2, 3\}$  for  $j \in [r] \setminus \{k\}$ . Assume first that  $P_{k-1} \in \text{Rec}(s_{k-1}, t_{k-1})$  with  $s_{k-1} > t_{k-1}$ , so that the score of Move 1 at step  $k$  equals  $s_{k-1} - t_{k-1}$ . For a contradiction, let  $m_j = 3$  for some  $1 \leq j < k \leq r$ . Consider a new sequence of moves  $(m'_1, m'_2, \dots, m'_r)$ , obtained from  $(m_i)_{i=1}^r$  by moving  $m_j$  to position  $k$  and shifting  $m_{j+1}, \dots, m_k$  to positions  $j, \dots, k - 1$  respectively: more formally let  $m'_i = m_i$  if  $i < j$  or  $i > k$ ,  $m'_i = m_{i+1}$  for  $j \leq i \leq k - 1$  and  $m'_k = m_j = 3$ .

Let  $P'_0 \subset P'_1 \subset \dots \subset P'_r = R$  be a sequence of rectangles obtained using the sequence of moves  $(m'_i)_{i=1}^r$  (since for all  $i \geq k$  the dimensions of rectangles  $P'_i$  equal the dimensions of rectangles  $P_i$  we indeed have  $P'_r = R$ ). Then the only Move 1 in this new sequence is applied to the rectangle  $P'_{k-2} \in \text{Rec}(s_{k-1}, t_{k-1} - 2)$  (this is because there is one less Move 3 among  $(m'_1, m'_2, \dots, m'_{k-2})$  as compared to  $(m_1, m_2, \dots, m_{k-1})$ ) and so this Move 1 has score  $s_{k-1} - t_{k-1} + 2$ . Note that the scores of other moves do not change, as they are still equal 3. Thus the cumulative score of the sequence  $(m'_i)_{i=1}^r$  is greater than the one of the sequence  $(m_i)_{i=1}^r$  and consequently percolation time of  $A$  is not maximum, contradicting the fact that  $A$  is  $(k, \ell)$ -good. Thus for all  $1 \leq i \leq k - 1$  we must have  $m_i = 2$ . In an analogous way we prove that  $m_i = 3$  for all  $k + 1 \leq i \leq r$ . So in this case  $(m_i)_{i=1}^r$  must be of the form  $[2]^*[1]^*[3]^*$ .

If Move 1 occurs only once in  $(m_1, m_2, \dots, m_r)$ , say that again  $m_k = 1$  and  $m_j \in \{2, 3\}$  for  $j \in [r] \setminus \{k\}$ , and additionally we have  $P_{k-1} \in \text{Rec}(s_{k-1}, t_{k-1})$  with  $s_{k-1} < t_{k-1}$ , then in an analogous way we show that, this time,  $(m_i)_{i=1}^r$  must be of the form  $[3]^*[1]^*[2]^*$ .

In the remaining case where  $P_{k-1}$  is a square, i.e.,  $s_{k-1} = t_{k-1}$ , also in a completely

analogous way, we can show that we must have  $r = 1$  and  $m_1 = 1$ . If that was not the case, i.e., if we had  $r \geq 2$  and there was some  $m_j \in \{2, 3\}$  then moving  $m_j$  to the opposite side of the only occurrence of Move 1 in  $(m_i)_{i=1}^r$  would increase cumulative score (as Move 1 would no longer have score 0) contradicting the  $(k, \ell)$ -goodness of  $A$ . Thus in this case  $(m_i)_{i=1}^r = (1)$ , which is at the same time of the form  $[2]^*[1]^*[3]^*$  and of the form  $[3]^*[1]^*[2]^*$ .

Thus assume that Move 1 occurs more than once in  $(m_1, m_2, \dots, m_r)$ . If all occurrences of it constitute a subsequence of consecutive  $m_i$ 's then we deal with this case exactly like we did with the one where Move 1 occurred only once. This is straightforward because, as we already noticed, using Move 1 does not change the difference between the length of the longer side and the length of the shorter side of the rectangle it is applied to.

Thus assume that there is some  $1 \leq j < t < k \leq r$  such that  $m_j = m_k = 1$  and  $m_t \in \{2, 3\}$ . For  $P_j \in \text{Rec}(s_j, t_j)$  and  $P_k \in \text{Rec}(s_k, t_k)$  assume that  $|s_j - t_j| \geq |s_k - t_k|$ . Consider a new sequence of moves  $(m'_1, m'_2, \dots, m'_r)$  obtained from  $(m_i)_{i=1}^r$  by moving  $m_k$  to position  $j + 1$ , and shifting  $m_{j+1}, \dots, m_{k-1}$  to positions  $j + 2, \dots, k$  respectively, that is, let  $m'_i = m_i$  if  $i \leq j$  or  $i > k$ ,  $m'_{j+1} = m_k = 1$  and for  $j + 2 \leq i \leq k$  let  $m'_i = m_{i-1}$ .

Let  $P'_0 \subset P'_1 \subset \dots \subset P'_r = R$  be a sequence of rectangles obtained using the sequence  $(m'_i)_{i=1}^r$  (note that as previously  $P'_r = P_r = R$ ). Then using Move  $m'_{j+1} = 1$  at step  $j + 1$  we finish the infection of  $P'_{j+1} \in \text{Rec}(s_j + 1, t_j + 1)$  and so this move has score  $|s_j - t_j|$  which is at least as big as the score of the move  $m_k$  at time  $k$ . Note that if  $i \leq j$  or  $i > k$  then the score of the move  $m'_i$  at time  $i$  equals the score of the move  $m_i$  at time  $i$ . Finally if  $j + 2 \leq i \leq k$  then the score of the move  $m'_i$  at time  $i$  equals the score of the move  $m_{i-1}$  at time  $i - 1$ . Thus the cumulative score of the sequence  $(m'_i)_{i=1}^r$  is at least as big as the one of the sequence  $(m_i)_{i=1}^r$ . Thus applying this modification (which does not decrease the score) of the sequence  $(m_i)_{i=1}^r$  we could obtain a sequence describing another  $(k, \ell)$ -good set in which all moves 1 occur in consecutive positions of the sequence. However, we already know that such sequence must be of the form  $[2]^*[1]^*[3]^*$  or of the form  $[3]^*[1]^*[2]^*$ . When  $|s_j - t_j| < |s_k - t_k|$  we proceed analogously, moving  $m_j = 1$  to position  $k - 1$ . This completes the proof of the claim.  $\square$

By Lemma 10 there exists an  $(k, \ell)$ -good set  $A$  for which  $P_0 \in \text{Rec}(s, 2) \cup \text{Rec}(2, s)$ , with  $s \geq 4$ . The construction we give in Lemma 8 shows that in this case  $P_0$  can be obtained from some  $P' \in \text{Rec}(2, 2)$  either by, if  $P_0 \in \text{Rec}(s, 2)$ , applying  $(s - 2)/2$  times Move 2, or by applying Move 3 if  $P_0 \in \text{Rec}(2, s)$ . Note that indeed for all these occurrences of move 2 or 3 we infect one new site at exactly three time steps.

**Observation 14.** *The proof of Claim 13 actually tells us that, for a brief moment slightly abusing the notation (relaxing condition (1) in Definition 1) and for  $i \geq 1$  allowing  $P'_i \in \text{Rec}(s'_i, 2) \cup \text{Rec}(2, s'_i)$  for  $s'_i \geq 4$  and even, there exists a  $(k, \ell)$ -good set  $A$  and a good sequence of rectangles  $P'_0 \subset P'_1 \subset \dots \subset P'_r \in \text{Rec}(k, \ell)$  associated with it, with  $P'_0 \in \text{Rec}(2, 2)$ , such that the sequence of moves  $(m'_1, m'_2, \dots, m'_r)$  used to fully infect  $P'_r$  from  $P'_0$  is of the form  $[2]^*[1]^*[3]^*$  or of the form  $[3]^*[1]^*[2]^*$ . Since in Claim 13 we have  $P_0 \in \text{Rec}(s, 2) \cup \text{Rec}(2, s)$ , with  $s \geq 4$ , we see that if  $(m'_i)_{i=1}^r$  is of the form  $[2]^*[1]^*[3]^*$  then the subsequence of moves 2 is nonempty. Analogously, if  $(m'_i)_{i=1}^r$  is of the form  $[3]^*[1]^*[2]^*$  then the subsequence of moves 3 is nonempty. Applying a nonempty sequence of moves*

2 to  $P'_0 \in \text{Rec}(2, 2)$  fully infects a rectangle  $P''_0 \in \text{Rec}(s'', 2)$ , with  $s'' \geq 4$  and even. Analogously, applying a nonempty sequence of moves 3 to  $P'_0 \in \text{Rec}(2, 2)$  fully infects a rectangle  $P''_0 \in \text{Rec}(2, s'')$ , with  $s'' \geq 4$  and even.

Thus by Observation 14 we obtain the following lemma which, for any  $k$  and  $\ell$  such that  $k \geq 4, \ell \geq 2$  and  $k + \ell$  is even, fully characterizes a good sequence of rectangles associated with at least one  $(k, \ell)$ -good set of initially infected sites.

**Lemma 15.** *Let  $k$  and  $\ell$  be such that  $k \geq 4, \ell \geq 2$  and  $k + \ell$  is even. Then there exists a  $(k, \ell)$ -good set  $A$  and a good sequence of rectangles  $P_0 \subset P_1 \subset \dots \subset P_r \in \text{Rec}(k, \ell)$  associated with it, with  $P_0 \in \text{Rec}(s, 2) \cup \text{Rec}(2, s)$  for some  $s \geq 4$ , such that the sequence of moves  $(m_1, m_2, \dots, m_r)$  used to fully infect  $P_r$  from  $P_0$  is either of the form  $[1]^*[3]^*$  if  $P_0 \in \text{Rec}(s, 2)$ , or of the form  $[1]^*[2]^*$  if  $P_0 \in \text{Rec}(2, s)$ .*

**Corollary 16.** *For  $n \geq 4$ , there is a  $(n, n)$ -good set  $A$ , whose good sequence of rectangles  $P_0 \subset P_1 \subset \dots \subset P_r \in \text{Rec}(n, n)$  is such that  $P_0 \in \text{Rec}(s, 2)$  and that the sequence of moves used to build it is of the form  $[1]^*[3]^*$ . Furthermore, if the number of times we use Move 1 equals  $m$  then  $m = n - s$  and we must use Move 3 exactly  $\frac{n-2-m}{2}$  times.*

*Proof.* Apply Lemma 15 with  $k = \ell = n$ . By symmetry, we can assume that  $P_0 \in \text{Rec}(s, 2)$  and the sequence of moves obtained is of type  $[1]^*[3]^*$ . It is trivial to check that, in order to obtain  $P_r \in \text{Rec}(n, n)$ , we must have  $m = n - s$  and we must use Move 3 exactly  $\frac{n-2-m}{2}$  times.  $\square$

We are now ready to prove the exact formula for  $M(n)$  for  $n \geq 4$ .

*Proof of Theorem 1.* Given  $m \geq 0$ , let  $A_m^n$  be (if it exists) the  $(n, n)$ -good set described in Corollary 16 for which during the infection process Move 1 is used exactly  $m$  times (note that when  $n$  and  $m$  have different parities then  $A_m^n$  definitely does not exist). For example, Figure 6 shows the set  $A_4^{12}$ . Now, we notice that for every  $n \geq 4$  and  $0 \leq m \leq n - 4$  (with

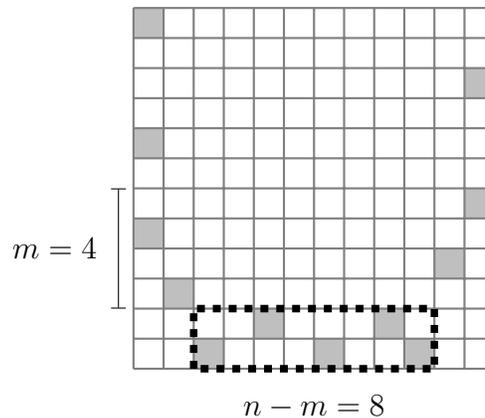


Figure 6: Set  $A_4^{12}$ .

$n$  and  $m$  having the same parity) the percolation time of  $A_m^n$  (if it exists) can be given explicitly as follows:

1. Infection of the rectangle  $P_0 \in \text{Rec}(n - m, 2)$  takes time

$$M(n - m, 2) = \frac{3(n - m) - 4}{2} = \frac{3(n - m)}{2} - 2;$$

2. Applying  $m$  times Move 1 takes time

$$\sum_{i=0}^{m-1} (n - m + i) = mn - m^2 + \frac{m(m - 1)}{2} = mn - \frac{m(m + 1)}{2};$$

3. Finishing the infection with  $\frac{n-m-2}{2}$  applications of Move 3 takes time

$$\frac{n - m - 2}{2}(n + 1) = \frac{n^2 - n - mn - m - 2}{2}.$$

Letting  $f(n, m)$  denote the percolation time of  $A_m^n$ , by the above calculations we have

$$f(n, m) = \frac{n^2 + n(m + 2) - (m^2 + 5m + 6)}{2}.$$

For a given  $n$ , the function  $f_n(m) = f(n, m)$  is a quadratic function of  $m$  with maximum value at  $m = \frac{n-5}{2}$ . As our  $(n, n)$ -good set maximizes  $f_n(m)$  subject to  $m \in \mathbb{N}$  and  $m$  having the same parity as  $n$ , maximal percolation time is obtained for

$$m = m_0 = \left\lfloor \frac{n}{2} - \frac{5}{2} \right\rfloor + \mathbb{1}_{\{4|n-1\}} + \mathbb{1}_{\{4|n\}},$$

where in the above statement we use  $\mathbb{1}_{\{\phi\}}$  to denote the indicator function,

$$\mathbb{1}_{\{\phi\}} = \begin{cases} 1, & \text{if the sentence } \phi \text{ is true,} \\ 0, & \text{otherwise.} \end{cases}$$

Now, by considering the possible values of  $n \pmod{4}$  we see that for all  $n \geq 4$  we have  $f(n, m_0) = \lfloor \frac{5n^2 - 2n}{8} \rfloor$ . This completes our proof.  $\square$

Using Lemma 15, given  $\alpha \in (0, 1)$  and  $n$  large, assuming  $(1 + \alpha)n$  is an even natural number we can determine an asymptotic value of  $M(n, \alpha n)$ . All we need to do is, for both  $P_0 \in \text{Rec}(s, 2)$  and  $P_0 \in \text{Rec}(2, s)$ , to optimize  $s$  to maximize the cumulative score of our sequence of moves knowing that the number of times we use Move 1 is fully determined by  $s$  and the horizontal or vertical alignment of  $P_0$  in  $[n] \times [\alpha n]$ , and that the score of all occurrences of Move 1 equals  $s - 2$ .

**Corollary 17.** *We have:*

1. *If  $\frac{1}{2} \leq \alpha < 1$  then*

$$M(n, \alpha n) = \left( \frac{\alpha}{2} + \frac{1}{8} \right) n^2 + O(n).$$

*To maximize percolation time we should first infect a roughly  $\frac{n}{2} \times 2$  rectangle in time  $O(n)$ , then using Move 1  $(\frac{n}{2} + O(1))$  times extend it to a roughly  $n \times \frac{n}{2}$  one in time  $\frac{3n^2}{8} + O(n)$ , and then finish the infection in additional  $(\frac{\alpha}{2} - \frac{1}{4}) n^2 + O(n)$  time steps using Move 3  $((\frac{\alpha}{2} - \frac{1}{4}) n + O(1))$  times.*

2. *If  $0 < \alpha < \frac{1}{2}$  then*

$$M(n, \alpha n) = \left( \alpha - \frac{\alpha^2}{2} \right) n^2 + O(n).$$

*To maximize percolation time we should first infect a roughly  $(1 - \alpha)n \times 2$  rectangle in time  $O(n)$ , and then finish the infection in additional  $(\alpha - \frac{\alpha^2}{2}) n^2 + O(n)$  time steps using Move 1  $(\alpha n + O(1))$  times.*

## References

- [1] M. Aizenman and J. Lebowitz. Metastability effects in bootstrap percolation. *Journal of Physics A*, 21:3801–3813, 1988.
- [2] J. Balogh, B. Bollobás, H. Duminil-Copin, and R. Morris. The sharp threshold for bootstrap percolation in all dimensions. *Transactions of the American Mathematical Society*, 364:2667–2701, 2012.
- [3] J. Balogh and G. Pete. Random disease on the square grid. *Random Structures and Algorithms*, 13:409–422, 1998.
- [4] F. S. Benevides and M. Przykucki. Maximum percolation time in two-dimensional bootstrap percolation. Preprint, 2013.
- [5] B. Bollobás, C. Holmgren, P. Smith, and A.J. Uzzell. The time of bootstrap percolation for dense initial sets. To appear in *Annals of Probability*, <http://arxiv.org/abs/1205.3922>.
- [6] B. Bollobás, P. Smith, and A.J. Uzzell. The time of bootstrap percolation with dense initial sets for all thresholds. Preprint, <http://arxiv.org/abs/1209.4339>.
- [7] J. Chalupa, P.L. Leath, and G.R. Reich. Bootstrap percolation on a Bethe lattice. *Journal of Physics C*, 12:L31–L35, 1979.
- [8] A. E. Holroyd. Sharp metastability threshold for two-dimensional bootstrap percolation. *Probability Theory and Related Fields*, 125:195–224, 2003.
- [9] R. Morris. Minimal percolating sets in bootstrap percolation. *The Electronic Journal of Combinatorics*, 16(1):1–20, 2009.

- [10] J. von Neumann. *Theory of Self-Reproducing Automata*. University of Illinois Press, 1966.
- [11] M. Przykucki. Maximal percolation time in hypercubes under 2-bootstrap percolation. *The Electronic Journal of Combinatorics*, 19(2):1–13, 2012.
- [12] E. Riedl. Largest minimal percolating sets in hypercubes under 2-bootstrap percolation. *The Electronic Journal of Combinatorics*, 17(1):1–13, 2010.
- [13] L. Shapiro and A.B. Stephens. Bootstrap percolation, the Schröder numbers, and the  $N$ -kings problem. *SIAM Journal on Discrete Mathematics*, 4:275–280, 1991.