LINEAR ALGEBRA AND BOOTSTRAP PERCOLATION

JÓZSEF BALOGH, BÉLA BOLLOBÁS, ROBERT MORRIS, AND OLIVER RIORDAN

Abstract. In $\mathcal{H}$-bootstrap percolation, a set $A \subset [n]$ of initially 'infected' vertices spreads by infecting vertices which are the only uninfected vertex in an edge of the hypergraph $\mathcal{H} \subset \mathcal{P}(n)$. A particular case of this is the H-bootstrap process, in which $\mathcal{H}$ encodes copies of H in a graph G. We find the minimum size of a set A that leads to complete infection when G is a power of a complete graph and H is a hypercube. The proof uses linear algebra, a technique that is new in bootstrap percolation, although standard in the study of weakly saturated graphs, which are equivalent to (edge) H-bootstrap percolation on a complete graph.

1. Introduction

Given a hypergraph $\mathcal{H} \subset \mathcal{P}(n)$, the *$\mathcal{H}$-bootstrap process* is defined as follows. Let $A \subset V(\mathcal{H})$ be a set of initially 'infected' vertices, and, at each time step, infect a vertex u if it lies in an edge of $\mathcal{H}$ in which all vertices other than u are already infected. To be precise, set $A_0 = A$, and, for each $t \geqslant 0$, set

$$A_{t+1} := A_t \cup \big\{ u \,:\, \exists\, S \in \mathcal{H} \text{ with } S \setminus A_t = \{u\} \big\}.$$

Let $[A]_{\mathcal{H}} = \bigcup_t A_t$, and say that A *percolates* (or $\mathcal{H}$-percolates) if $[A]_{\mathcal{H}} = V(\mathcal{H})$.

A large family of models of this type was introduced in [9]. Given graphs G and H, we obtain the *H-bootstrap process* on G by setting $\mathcal{H} = \{V(H') \,:\, H' \subset G \text{ and } H' \cong H\}$. The $\mathcal{H}$- and H-bootstrap processes can be seen as special cases of the 'cellular automata' introduced by von Neumann (see [16]) after a suggestion of Ulam [17], and generalize several previously studied models. For example, if G is a (finite) square grid and $H = C_4$, then we obtain the so-called 'Froböse process' (see [12] or [13]).

A fundamental question about bootstrap-type models is the following: given a hypergraph $\mathcal{H}$ (or a pair (G, H)), how large is the smallest percolating set in the $\mathcal{H}$-bootstrap process? We define

$$m(\mathcal{H}) := \min\big\{ |A| \,:\, A \subset V(\mathcal{H}),\ [A]_{\mathcal{H}} = V(\mathcal{H}) \big\}.$$

Fix $2 \leqslant r \leqslant d$, let K_n^d denote the graph with vertex set $[n]^d = \{0, \ldots, n-1\}^d$ in which uv is an edge if u and v differ in exactly one coordinate. Set

$$\mathcal{K}(n, d, t, r) := \big\{ S \subset [n]^d \,:\, K_n^d[S] = K_t^r \big\},$$

Key words and phrases. Bootstrap percolation, linear algebra, weak saturation.

Research supported in part by: (JB) NSF CAREER Grant DMS-0745185, UIUC Campus Research Board Grants 09072 and 11067, OTKA Grant K76099, and the TAMOP-4.2.1/B-09/1/KONV-2010-0005 project; (BB) NSF grants CNS-0721983, CCF-0728928 and DMS-0906634, and ARO grant W911NF-06-1-0076; (RM) CNPq bolsa de Produtividade em Pesquisa.

that is, the collection of induced copies of K_t^r in K_n^d. Note that $K_2^r = Q_r$, the r-dimensional hypercube.

Our main aim is to determine $m\big(\mathcal{K}(n,d,2,r)\big)$ precisely for every $n \in \mathbb{N}$ and every $d \geqslant r \geqslant 2$. We shall also consider the grid P_n^d with vertex set $[n]^d$, in which two vertices are adjacent if they differ by 1 in one coordinate, and agree in all others. (This graph is usually denoted $[n]^d$, but here this notation would cause confusion.) The corresponding hypergraph is

$$\mathcal{P}(n,d,t,r) := \big\{ S \subset [n]^d \,:\, P_n^d[S] = P_t^r \big\}.$$

Note that $\mathcal{P}(n,d,t,r) \subset \mathcal{K}(n,d,t,r)$. The following result is our main theorem.

Theorem 1. *For every $n \geqslant 2$ and $d \geqslant r \geqslant 2$, we have*

$$m\big(\mathcal{K}(n,d,2,r)\big) \;=\; m\big(\mathcal{P}(n,d,2,r)\big) \;=\; \sum_{t=0}^{r-1} \binom{d}{t} \big(n-1\big)^t.$$

We remark that the sum in the theorem is simply the number of vectors in $[n]^d$ having at most $r-1$ non-zero coordinates.

We shall also prove the following generalization of the case $r = d$ of Theorem 1.

Theorem 2. *For every $n \geqslant t \geqslant 2$ and $d \geqslant 2$ we have*

$$m\big(\mathcal{K}(n,d,t,d)\big) \;=\; m\big(\mathcal{P}(n,d,t,d)\big) \;=\; n^d - \big(n+1-t\big)^d.$$

Note that the formula above is simply the number of vectors in $[n]^d$ in which at least one coordinate takes one of the values $\{0,1,2,\ldots,t-2\}$.

The first extremal result related to bootstrap percolation was proved by Bollobás [8], and phrased in the language of 'weakly saturated graphs'. This is the natural edge version of the H-bootstrap percolation we have just defined (infect an edge if it is the last uninfected edge of a copy of H), with G complete. The main aim of [8] was to pose a conjecture concerning the extremal number when $H = K_k$ and $G = K_n$. This conjecture was proved by Alon [1], Frankl [11] and Kalai [15], using linear algebraic methods. We shall use the main lemma of [1] to prove Theorem 2.

The $\mathcal{H}$-bootstrap process is named after a closely related model, known as r-neighbour bootstrap percolation, which was introduced in 1979 by Chalupa, Leath and Reich [10] as a model of disordered magnetic systems. In this process, a vertex of a graph G becomes infected when it has at least r infected neighbours; we remark that this is similar to H-bootstrap percolation with H a star, except that a given copy of H can only be responsible for infecting its central vertex. The r-neighbour bootstrap process has been extensively studied by mathematicians and statistical physicists (see [2, 6, 14], for example, and the references therein). For further background see Bollobás [9].

In r-neighbour bootstrap percolation, one is mainly interested in estimating the *critical threshold* in the random setting: if the initially infected set A is formed by selecting vertices independently with probability p, for which p is it likely that eventually all vertices are infected? In the study of this probabilistic question, extremal results turn out to be important (see [4] or [13], for example). One of our main motivations in this work is to approach the following tantalizing open problem, which is our main stumbling block in attacking the probabilistic question on the hypercube. Let $m(G,r)$ denote the minimum size of a percolating set in r-neighbour bootstrap percolation on G. In [3], Balogh and Bollobás made the following conjecture.

Conjecture 1. *Let $r \geqslant 3$ be fixed. Then*

$$m(Q_d, r) \;=\; \left(\frac{1}{r} + o(1)\right)\binom{d}{r-1}$$

as $d \to \infty$.

The upper bound in Conjecture 1 follows by taking a Steiner system at level r, together with all of level $r-2$. Amazingly, we know of no super-linear lower bound. In the case $r = 2$ the situation is simpler, and $m(P_n^d, 2)$ is known exactly for all n and d (see [3] or [4]). At the other end of the range, Pete (see [7]) observed that $m(P_n^d, d) = n^{d-1}$. However, for fixed $2 < r < d$, $m(P_n^d, r)$ is known only up to a constant factor that depends on d.

Finally, we remark that the random questions are also interesting in the H-bootstrap model, and that some of the basic problems (in the 'edge version') are solved in [5] by the first three authors. As the reader might guess, however, there are still many more open problems than theorems.

The rest of this note is arranged as follows. In Section 2 we prove Theorem 1, and in Section 3 we prove Theorem 2.

2. Proof of Theorem 1

The proof of Theorem 1 is based on the following observation.

Lemma 3. *Let $\mathcal{H}$ be an arbitrary hypergraph. Suppose that we can find a vector space W spanned by vectors $\{f_v : v \in V(\mathcal{H})\}$ such that, for every edge $S \in \mathcal{H}$, the set $\{f_u : u \in S\}$ is linearly dependent. Then*

$$m(\mathcal{H}) \geqslant \dim W.$$

Proof. Once one thinks of the statement, the proof is essentially immediate. Indeed, suppose that $A \subset V(\mathcal{H})$ percolates in the $\mathcal{H}$-process. Then we can order the vertices $v_1, \dots, v_r$ in $V(\mathcal{H}) \setminus A$ so that each v_i is in an edge of $\mathcal{H}$ in $A_i = A \cup \{v_j : j \leqslant i\}$. Let W_i be the span of the vectors $\{f_v : v \in A_i\}$. The dependency condition implies that $v_i \in W_{i-1}$, so $W_i = W_{i-1}$ and hence $W_0 = W_r$. By assumption, $A_r = V(\mathcal{H})$, so $W_r = W$. Since W_0 is spanned by $|A|$ vectors, we have $|A| \geqslant \dim W$. $\square$

To prove Theorem 1, we must find the right vectors.

Proof of Theorem 1. Fix $n \in \mathbb{N}$ and $d \geqslant r \geqslant 2$, and set $\mathcal{K} = \mathcal{K}(n,d,2,r)$ and $\mathcal{P} = \mathcal{P}(n,d,2,r)$. Given $v \in [n]^d$, let $\sigma(v)$ denote the number of non-zero coordinates of v, and let $L_{\leqslant t} = \{v : \sigma(v) \leqslant t\}$ be the union of the first $t+1$ 'layers' of K_n^d. Note that Theorem 1 asserts exactly that $m(\mathcal{K}) = m(\mathcal{P}) = |L_{\leqslant r-1}|$.

Suppose that the set of initially infected vertices is exactly $L_{\leqslant r-1}$. Then every vertex v is eventually infected in the $\mathcal{P}$-process. Indeed, writing $|v|$ for the sum of the coordinates of v, for $v \notin L_{\leqslant r-1}$ we can use any r non-zero coordinates of v to construct a copy H of Q_r in P_n^d with v as the 'top' vertex, i.e., with $|u| < |v|$ for all other vertices u of H. It follows by induction on $|v|$ that all v are infected eventually. Hence, since $\mathcal{P} \subset \mathcal{K}$,

$$m(\mathcal{K}) \leqslant m(\mathcal{P}) \leqslant |L_{\leqslant r-1}|. \tag{1}$$

For the lower bound let W be a (real) vector space with basis $\{e_v : v \in L_{\leqslant r-1}\}$, so $\dim W = |L_{\leqslant r-1}|$. For $v \in [n]^d$ and $T \subset \{1, 2, \ldots, d\}$ with $|T| = r-1$, let v_T be the vector $(u_1, \ldots, u_d)$ with $u_i = v_i$ if $i \in T$ and $u_i = 0$ otherwise. Thus v_T is the 'downwards' projection of v onto a certain $(r-1)$-dimensional face of $[n]^d$, and $v_T \in L_{\leqslant r-1}$.

For each $v \in [n]^d$, define a vector f_v by

$$f_v := \sum_{|T|=r-1} e_{v_T} \in W, \tag{2}$$

where the sum is over all $T \subset \{1, 2, \ldots, d\}$ with $|T| = r-1$.

The vectors $\{f_v : v \in L_{\leqslant r-1}\}$ and the e_v are related in a 'triangular' way: if $\sigma(v) = t \leqslant r-1$, then the $\binom{d-t}{r-1-t} > 0$ terms in which T contains all of the non-zero coordinates of v each contribute e_v to the sum in (2), while all other terms are of the form e_w with $\sigma(w) < \sigma(v)$. Writing span$(\cdot)$ for the linear span of a set of vectors, it follows by induction on t that

$$\mathrm{span}(\{f_v : v \in L_{\leqslant t}\}) = \mathrm{span}(\{e_v : v \in L_{\leqslant t}\})$$

for $t = 0, 1, \ldots, r-1$. In particular, the vectors $\{f_v : v \in L_{\leqslant t-1}\}$ span W.

Let $d(u, v)$ denote the Hamming distance between u and v (i.e., the number of coordinates in which they differ). If $S \in \mathcal{K}$ and $u \in S$, then we claim that

$$\sum_{v \in S} (-1)^{d(u,v)} f_v = 0. \tag{3}$$

To see this, recall that $K_n^d[S]$ is a copy H of Q_r, and let $T \subset [d]$ with $|T| = r-1$. Then at least one edge of H involves changing a coordinate $i \notin T$. Group the vertices of H into pairs $\{v, v'\}$ differing only in the i^{th} coordinate. Then $d(v, v') = 1$ and $v_T = v'_T$, and it follows that

$$\sum_{v \in T} (-1)^{d(u,v)} e_{v_T} = 0.$$

Summing over T gives (3). Lemma 3 now implies that $m(\mathcal{K}) \geqslant \dim W = |L_{\leqslant r-1}|$. Together with (1) this completes the proof of the theorem. □

Let us remark that the special case $d = r = 2$ of Theorem 1 can be proved much more simply. Recall that $\mathcal{K}(n, 2, 2, 2)$ and $\mathcal{P}(n, 2, 2, 2)$ encode the induced copies of C_4 in K_n^d and P_n^d respectively.

Proposition 4. $m\big(\mathcal{K}(n,2,2,2)\big) = m\big(\mathcal{P}(n,2,2,2)\big) = 2n-1$.

The second equality, giving the minimum size of a percolating set in the Froböse model, is well known (see [13] for example).

Proof. Let $n \geqslant 2$, and set $\mathcal{K} = \mathcal{K}(n,2,2,2)$ and $\mathcal{P} = \mathcal{P}(n,2,2,2)$. The set $\{(x,y) : \min\{x,y\} = 0\}$ $\mathcal{P}$-percolates, and $\mathcal{P} \subset \mathcal{K}$. Thus $m(\mathcal{K}) \leqslant m(\mathcal{P}) \leqslant 2n-1$.

Suppose that $A \subset [n]^2$ $\mathcal{K}$-percolates. Associate to each vertex $(x,y) \in [n]^2$ an edge of the complete bipartite graph $K_{n,n}$ in the obvious way, and note that $\mathcal{K}$-percolation is equivalent to C_4-edge percolation in $K_{n,n}$. (This is because the line graph of C_4 is also C_4.) If the edges corresponding to A edge-percolate in $K_{n,n}$, then they must form a connected subgraph of $K_{n,n}$. So $|A| \geqslant 2n-1$, and hence $m(\mathcal{K}) \geqslant 2n-1$. $\square$

3. Proof of Theorem 2

In this section we shall prove Theorem 2 using of the following result of Alon [1], which was proved using methods from exterior algebra.

Theorem 5. *Let $d \in \mathbb{N}$, let $r_1, \ldots, r_d \in \mathbb{N}$ and $s_1, \ldots, s_d \in \mathbb{N}$, and let $X_1, \ldots, X_d$ be disjoint sets. Suppose there exist sets $A_{ij} \subset X_i$ and $B_{ij} \subset X_i$ for each $1 \leqslant i \leqslant d$ and $1 \leqslant j \leqslant h$, with $|A_{ij}| \leqslant r_i$ and $|B_{ij}| \leqslant s_i$, such that*

(a) $\left(\bigcup_i A_{ij}\right) \cap \left(\bigcup_i B_{ij}\right) = \emptyset \quad \text{for} \quad 1 \leqslant j \leqslant h.$

(b) $\left(\bigcup_i A_{ij}\right) \cap \left(\bigcup_i B_{i\ell}\right) \neq \emptyset \quad \text{for} \quad 1 \leqslant j < \ell \leqslant h.$

Then

$$h \;\leqslant\; \prod_{i=1}^{d} \binom{r_i + s_i}{r_i}.$$

Proof of Theorem 2. Fix $n \geqslant t \geqslant 2$ and $d \geqslant 2$, and set $\mathcal{K} = \mathcal{K}(n,d,t,d)$ and $\mathcal{P} = \mathcal{P}(n,d,t,d)$. For $v \in [n]^d$ and $t \in \mathbb{N}$, let $\sigma_{t-1}(v)$ denote the number of co-ordinates of v which are at least $t-1$.

For the upper bound, let $A = \big\{v \in [n]^d : \sigma_{t-1}(v) < d\big\}$, and note that $|A| = n^d - (n+1-t)^d$. Let $A \subset B \subsetneq [n]^d$, and let $v \in [n]^d \setminus B$ be chosen in order to minimize $|v|$. Then the (induced) copy H_v of P_t^d with co-ordinates $\{v_j - t + 1, \ldots, v_j\}$ in direction j is contained in $B \cup \{v\}$, and so v lies in the closure of B under the $\mathcal{P}$-process. It follows that A percolates under the $\mathcal{P}$-process, and so, since $\mathcal{P} \subset \mathcal{K}$,

$$m(\mathcal{K}) \;\leqslant\; m(\mathcal{P}) \;\leqslant\; n^d - (n+1-t)^d. \tag{4}$$

For the lower bound, let $A \subset [n]^d$, and suppose that A $\mathcal{K}$-percolates. Then there is at least one ordering $v(1), \ldots, v(h)$ of the $h = n^d - |A|$ vertices outside A such that, for every $1 \leqslant j \leqslant h$, there exists a copy H_j of K_t^d in $K_n^d[A \cup \{v(1), \ldots, v(j)\}]$ with $v(j) \in H_j$. (This is simply the order in which the vertices are infected, which is not uniquely defined.) Let $X_1, \ldots, X_d$ be disjoint sets of size n, and let $X_i = \{u_{i1}, \ldots, u_{in}\}$.

For $1 \leqslant i \leqslant d$ and $1 \leqslant j \leqslant h$, the i^{th} coordinates of the vertices of H_j take exactly t distinct values. Let A_{ij} be the set of all u_{ik} where k is *not* one of these values, so

$A_{ij} \subset X_i$ and $|A_{ij}| = n - t$. Let $B_{ij} = \{u_{ik}\}$ where k is the i^{th} coordinate of $v(j)$. Since $v(j)$ is a vertex of H_j, we see that $u_{ik} \notin A_{ij}$, so for all i and j the sets A_{ij} and B_{ij} are disjoint.

We claim that the conditions of Theorem 5 are satisfied by the A_{ij} and B_{ij}, with $r_i = n - t$ and $s_i = 1$ for all i. We have already verified all conditions apart from (b), so suppose that $1 \leqslant j < \ell \leqslant h$. Then by the definition of our order, $v(\ell) \notin H_j$. In other words, there is some coordinate i such that $v(\ell)_i$ is not one of the values of the i^{th} coordinate taken by H_j. Equivalently, there is some i such that $B_{i\ell} \subset A_{ij}$, giving $A_{ij} \cap B_{i\ell} \neq \emptyset$, as required.

Theorem 5 now gives

$$n^d \, - \, |A| \, = \, h \, \leqslant \, \big(n + 1 - t\big)^d,$$

so $|A| \geqslant n^d - (n+1-t)^d$, completing the proof. $\square$

References

[1] N. Alon, An extremal problem for sets with applications to graph theory, *J. Combin. Theory, Ser. A*, **40** (1985), 82–89.

[2] M. Aizenman and J.L. Lebowitz, Metastability effects in bootstrap percolation, *J. Phys. A.*, **21** (1988) 3801–3813.

[3] J. Balogh and B. Bollobás, Bootstrap percolation on the hypercube, *Prob. Theory Rel. Fields*, **134** (2006), 624–648.

[4] J. Balogh, B. Bollobás and R. Morris, Bootstrap percolation in high dimensions, *Combin. Prob. Computing*, **19** (2010), 643–692.

[5] J. Balogh, B. Bollobás and R. Morris, Graph bootstrap percolation, submitted.

[6] J. Balogh, B. Bollobás, H. Duminil-Copin and R. Morris, The sharp threshold for bootstrap percolation in all dimensions, to appear in *Trans. Amer. Math. Soc.*

[7] J. Balogh and G. Pete, Random disease on the square grid, *Random Structures Algorithms*, **13** (1998), 409–422.

[8] B. Bollobás, Weakly k-saturated graphs, in *Beiträge zur Graphentheorie* (H. Sachs, H.J. Voss, and H. Walther, Eds.), pp. 25–31, Leipzig, Barth, 1968.

[9] B. Bollobás, Graph bootstrap percolation, preprint (2011).

[10] J. Chalupa, P. L. Leath and G. R. Reich, Bootstrap percolation on a Bethe lattice, *J. Phys. C.*, **12** (1979), L31–L35.

[11] P. Frankl, An extremal problem for two families of sets, *Europ. J. Math.*, **3** (1982), 125–127.

[12] K. Froböse. Finite-size effects in a cellular automaton for diffusion, *J. Statist. Phys.*, **55** (1989) (5-6), 1285–1292.

[13] J. Gravner, A.E. Holroyd and R. Morris, A sharper threshold for bootstrap percolation in two dimensions, to appear in *Prob. Theory Rel. Fields.*

[14] A. Holroyd, Sharp Metastability Threshold for Two-Dimensional Bootstrap Percolation, *Prob. Theory Rel. Fields*, **125** (2003), 195–224.

[15] G. Kalai, Weakly saturated graphs are rigid, in Convexity and Graph Theory, M Rosenfeld and J. Zaks (eds), *Ann. Discrete Math.*, **20** (1984), 189–190.

[16] J. von Neumann, Theory of Self-Reproducing Automata. Univ. Illinois Press, Urbana, 1966.

[17] S. Ulam, Random processes and transformations, *Proc. Internat. Congr. Math.* (1950), 264–275.

Department of Mathematics, University of Illinois, 1409 W. Green Street, Urbana, IL 61801, and, Department of Mathematics, University of California, San Diego, La Jolla, CA 92093
E-mail address: jobal@math.uiuc.edu

Trinity College, Cambridge CB2 1TQ, England, and, Department of Mathematical Sciences, The University of Memphis, Memphis, TN 38152, USA
E-mail address: B.Bollobas@dpmms.cam.ac.uk

IMPA, Estrada Dona Castorina 110, Jardim Botânico, Rio de Janeiro, RJ, Brasil
E-mail address: rob@impa.br

Mathematical Institute, University of Oxford, 24–29 St Giles', Oxford OX1 3LB, UK
E-mail address: riordan@maths.ox.ac.uk