Deterministic bootstrap percolation in high dimensional grids

Hao Huang∗ Choongbum Lee †

Abstract

In this paper, we study the $k$-neighbor bootstrap percolation process on the $d$-dimensional grid $[n]^d$, and show that the minimum number of initial vertices that percolate is $(1 - \frac{d}{k})n^d + O(n^{d-1})$ when $d \leq k \leq 2d$. This confirms a conjecture of Pete.

AMS Subject classification: 05C35, 82B43

1 Introduction

Consider the following process on a graph $G$. We start with an initial set of infected vertices $S \subseteq V(G)$. At each step, a vertex becomes infected if at least $k$ of its neighbors are already infected, and infected vertices remain infected forever. The process terminates when no further vertex can be infected. This process is known as $k$-neighbor bootstrap percolation on $G$. If the entire vertex set becomes infected in the end, then we say that $S$ percolates (or $k$-percolates) on $G$. Bootstrap percolation was first introduced and studied in the statistical physics literature by Chalupa et al. [10]. One can also view it as an example of a cellular automaton, a concept introduced by von Neumann [14] after a suggestion of Ulam [19]. See [1] for more on motivations and applications of bootstrap percolation.

The main question of interest in bootstrap percolation is to determine the initial sets that percolate for a given graph. When the initial set is chosen randomly, the goal is to determine the critical probability at which percolation happens with high probability. The random model has been extensively studied in various works [2, 3, 4, 5, 6, 7, 8, 9, 11, 12]. In contrast, relatively fewer results are known around extremal problems for bootstrap percolation, i.e., when the initial set is a deterministic set. Morris [13] and Riedl [18] studied the maximal size of a minimal percolating set, and Przykucki [17] and some later results studied the maximal percolation time.

In this paper, we study the minimum size of the initial set that percolates. The following folklore result is perhaps the best starting point of this subject and it motivates our work.

Proposition 1.1. In an $n \times n$ all-white chessboard, start with a set $S$ of squares colored black. At each step we color a white square black if it is adjacent to at least two black squares. If the whole chessboard becomes black in the end, then $|S| \geq n$.

∗Institute for Advanced Study, Princeton, NJ 08540 and DIMACS at Rutgers University. Email: huanghao@math.ias.edu. Research supported in part by NSF grant DMS-1128155.

†Department of Mathematics, Massachusetts Institute of Technology, Cambridge 02139. Email: cb_lee@math.mit.edu.
The key idea of the proof of Proposition 1.1 is to consider the total perimeter of the black regions on the chessboard. At each step, a white square becomes black only if it is adjacent to at least two black squares, therefore the perimeter increases by at most \((4 - 2) - 2 = 0\), i.e., is non-increasing. Also, if the whole chessboard becomes black, then the perimeter is 4n. Since in the beginning the total perimeter is at most \(4|S|\), we see that \(4|S|\) must be at least \(4n\) and thus \(|S| \geq n\). Further note that Proposition 1.1 is best possible, since there exists a set \(S\) of size \(n\) for which the whole chessboard becomes black: let \(S\) consist of the squares on the diagonal, and observe that in each round the white squares on a new subdiagonal become black.

Now a natural question arises: can this result be generalized to chessboards of higher dimensions? A \(d\)-dimensional grid is the graph whose vertex set is \([n]^d\), and two vertices are adjacent if they differ by 1 in exactly one coordinate and are identical everywhere else. Denote by \(f_{d,k}(n)\) the minimum number of initial vertices that \(k\)-percolates on the \(d\)-dimensional lattice \([n]^d\) (note that the definition only makes sense when \(1 \leq k \leq 2d\)). Proposition 1.1 says that \(f_{2,2}(n) = n\). For the range \(d \leq k \leq 2d\), Pete [15] proved that \(f_{d,k}(n) \geq (1 - \frac{d}{k})n^d + \frac{d}{k}n^{d-1}\). He also proved that this bound is tight for \(d = k\), and asymptotically tight for \((d,k) = (2,3), (2,4), (3,4), (3,6)\) and when \(k = 2d\). He then conjectured that it is asymptotically tight for all \(d + 1 \leq k \leq 2d\). In this paper, we verify this conjecture.

**Theorem 1.2.** For \(d \leq k \leq 2d\), we have

\[
f_{d,k}(n) = \left(1 - \frac{d}{k}\right)n^d + O(n^{d-1}).
\]

In the rest of this paper we prove Theorem 1.2. We also discuss some potential further generalizations in the concluding remarks.

## 2 A short proof of the lower bound

In this section, we give a different proof of the lower bound in Theorem 1.2. We first define a polynomial that is a natural generalization of the perimeter used in the proof of Proposition 1.1.

**Lemma 2.1.** For \(d \leq k \leq 2d\),

\[
f_{d,k}(n) \geq \left(1 - \frac{d}{k}\right)n^d + \frac{d}{k}n^{d-1}.
\]

**Proof.** Let \(G\) be the \(d\)-dimensional grid graph on the vertex set \([n]^d\), and consider the \(k\)-neighbor bootstrap percolation on \(G\). The number of vertices of \(G\) is \(v(G) = n^d\), and the number of edges of \(G\) is \(e(G) = d(n-1)n^{d-1}\), since in each of the \(d\) directions there are precisely \((n-1)n^{d-1}\) edges.

Consider the following polynomial \(p : \{0,1\}^{V(G)} \to \mathbb{R}\):

\[
p(\{x_v\}_{v \in V(G)}) = \sum_{v \in V(G)} x_v - \frac{1}{k} \sum_{uv \in E(G)} x_u x_v.
\]
When \( \{x_v\}_{v \in V(G)} \) is the indicating vector of a subset \( S \) of vertices of \([n]^d\), this function in some sense characterizes the “perimeter” of \( S \). Suppose that the infection happens one vertex at a time, and consider the moment at which a vertex \( v \) becomes infected. Let \( S \) be the set of infected vertices at this time, and let \( \vec{S} \) be the indicating vector of \( S \). Let \( \vec{e}_v \) be the indicating vector of \( \{v\} \). Then \( p \) changes by

\[
 p(\vec{S} + \vec{e}_v) - p(\vec{S}) = \frac{\partial p}{\partial x_v} \bigg|_{\vec{s}} = 1 - \frac{1}{k} \sum_{u: v \sim u} S_u.
\]

Note that since \( v \) becomes infected, it has at least \( k \) neighbors in \( S \). Hence \( \sum_{u: v \sim u} S_u \geq k \), and \( p(\vec{S} + \vec{e}_v) - p(\vec{S}) \leq 0 \), thus \( p \) does not increase during the entire process. As a result, if the initial set \( T \) percolates, then

\[
 |T| \geq p(\vec{T}) \geq p(\vec{1}) = v(G) - \frac{e(G)}{k} = n^d - \frac{d(n-1)n^{d-1}}{k} = \left( 1 - \frac{d}{k} \right) n^d + \frac{d}{k} n^{d-1},
\]

and this proves the lemma. \( \square \)

3 Construction for the upper bound

In this section, we will give an explicit construction of an initial set that \( k \)-percolates on the \( d \)-dimensional grid, whose size attains the lower bound proven earlier in Section 2.

**Theorem 3.1.** For every positive integers \( d, k \) with \( d \leq k \leq 2d \), there exists an initial set of \((1 - \frac{d}{k})n^d + O(n^{d-1})\) vertices that \( k \)-percolates on the \( d \)-dimensional grid graph over the vertex set \([n]^d\).

Before going into details, we begin with some simple arguments to show that Lemma 3.1 holds for two special cases: \( k = d \) and \( k = 2d \). The construction for these cases will provide some insights into finding the general construction. For convenience, we call the vertices in \([n]^d\) with some of its coordinates equal to 1 or \( n \) the boundary vertices, and the rest interior vertices.

For \( k = d \) take the initial set \( S \) to be the boundary vertices. It is easy to check that \( |S| = O(n^{d-1}) \). Note that interior vertices have \( 2d \leq x_1 + \cdots + x_d \leq (n-1)d \), and all vertices having \( x_1 + \cdots + x_d \leq 2d - 1 \) are infected in the beginning. Observe that every interior vertex \( x = (x_1, \cdots, x_d) \) satisfying \( x_1 + \cdots + x_d = 2d \) are adjacent to at least \( d \) vertices in \( S \) and gets infected in the first step, since by decreasing any of the \( d \) coordinates by 1 we obtain a vertex still in \([n]^d\) that satisfies \( x_1 + \cdots + x_d = 2d - 1 \). Similarly, in the next step the vertices satisfying \( x_1 + \cdots + x_d = 2d + 1 \) will be infected, and this process continues until every vertex of the graph becomes infected.

When \( k = 2d \), let \( S \) consist of vertices which are either on the boundary of \([n]^d\) or those whose coordinates sum up to be an even number. Note that every interior vertex not in \( S \) is adjacent to exactly \( 2d \) vertices from \( S \), and thus \( S \) percolates. The number of vertices lying on the boundary is \( O(n^{d-1}) \), and the number of vertices whose coordinates sum up to an even number is at most \( \lceil \frac{1}{2} n^d \rceil \). Therefore we have \( |S| = \frac{1}{2} n^d + O(n^{d-1}) \). Now we prove the lemma for all cases.
Proof of Theorem 3.1. As in the $k = d$ case, we first let $S_1$ be the set of boundary vertices. We consider two subcases according to the parity of $k$:

(i) $k = 2t + 1$ and thus $t \leq d - 1$. Define two functions $X$ and $Y$ on the vertex set as follows. For a vertex $x = (x_1, \ldots, x_d) \in [n]^d$, let

\[
X(x) = x_1 + 2x_2 + \cdots + tx_t \mod (2t + 1) \\
Y(x) = x_1 + x_2 + \cdots + x_d \mod 2.
\]

Let $I = [1, k - d]$, and $A = I \cup (I + t)$ be the union of two intervals. From $t \leq d - 1$, it follows that $k - d = 2t + 1 - d \leq t$ and $k - d + t = 3t + 1 - d \leq 2t$. Hence the two intervals do not overlap, and

\[
|A| = 2|I| = 2(k - d).
\]

Define

\[
S_2 = \{x : X(x) \in A \text{ and } Y(x) = 0\}.
\]

Note that if $X(x) \in A$ and $Y(x) = 0$, then for $e_d = (0, \ldots, 0, 1)$, we have $X(x + e_d) = X(x) \in A$ and $Y(x + e_d) = 1$. Therefore,

\[
|S_2| = \frac{1}{2} \cdot \left| \{x : X(x) \in A\} \right| + O(n^{d-1}) = \frac{1}{2} \cdot \frac{|A|}{2t + 1} \cdot n^d + O(n^{d-1}) = \left(1 - \frac{d}{k}\right) n^d + O(n^{d-1}).
\]

Since $|S_1| = O(n^{d-1})$, the set $S = S_1 \cup S_2$ has the claimed asymptotic density. It remains to check that $S$ percolates. We partition the initially non-infected vertex set into $S_3 \cup S_4$ as follows:

\[
S_3 = \{x : X(x) \in A \text{ and } Y(x) = 1\} \quad \text{and} \quad S_4 = \{x : X(x) \notin A\}.
\]

The vertices in $S_3$ gets infected as follows. For an arbitrary interior vertex $x \in S_3$, if we change any of its $d$ coordinates by 1 or $-1$, the value of $Y$ becomes 0. Also, for $1 \leq i \leq t$, if we change the $i$-th coordinates by 1 (or $-1$), then $X$ changes by $i$ or $-i$. Since

\[
\bigcup_{i=1}^{t} \{i, -i\} = \{1, 2, \ldots, 2t\} \mod (2t + 1)
\]

and $X(x) \in A$, among these $2d$ neighbors there are exactly

\[
|A^c| = (2t + 1) - |A| = (2t + 1) - (4t + 2 - 2d) = 2d - 2t - 1
\]

neighbors having $X \notin A$. Hence each vertex in $S_3$ is adjacent to at least $2d - (2d - 2t - 1) = 2t + 1$ initial vertices, and gets infected (the reason we have at least instead of exactly is because some neighbors may have $X \notin A$ but be a boundary vertex).

Now we want to show that the vertices in $S_4$ gets infected. Note that all interior vertices have $2d \leq x_1 + \cdots + x_n \leq (n - 1)d$. Consider an interior vertex $x$ having $x_1 + x_2 + \cdots + x_d = 2d$. Call a neighbor of $x$ negative if the value of some coordinate decreases, and positive otherwise. Note that all negative neighbors of $x$ are infected (since they have $x_1 + \cdots x_d = 2d - 1$). Thus it suffices
to show that \( x \) has \( 2t + 1 - d \) infected positive neighbors. Note that the set \( A \) has the property that for all \( a \not\in A \), the intersection \( (a + [t]) \cap A \) is either \( I \) or \( I + t \). Since \( X(x) \not\in A \) and for each \( i = 1, 2, \cdots, t \), there exists a positive neighbor of \( x \) whose value of \( X \) increases by \( i \), we see that there are precisely \( |I| \) positive neighbors that are in \( S_2 \cup S_3 \). Therefore, \( x \) gets infected. The same argument shows that the interior vertices having \( x_1 + \cdots + x_d = 2d + 1 \) get infected in the next step, and this continues as in the \( k = d \) case until all the vertices are infected.

(ii) \( k = 2t \) and thus \( t \leq d \). Let

\[
X = x_1 + 2x_2 + \cdots + (t - 1)x_{t-1} \mod 2t \\
Y = x_1 + x_2 + \cdots + x_d \mod 2.
\]

Let \( I = [1, k - d] \), and \( A = I \cup (I + t) \) be the union of two intervals. From \( t \leq d \), it follows that \( k - d = 2t - d \leq t \) and \( k - d + t = 3t - d \leq 2t \). Hence the two intervals do not overlap, and

\[
|A| = 2|I| = 2(k - d).
\]

Again let

\[ S_2 = \{x : X(x) \in A \text{ and } Y(x) = 0\} \]

Then

\[
|S_2| = \left(1 - \frac{d}{k}\right)n^d + O(n^{d-1}).
\]

As in the previous case, in order to prove that \( S_1 \cup S_2 \) percolates on the grid, it suffices to check that for all modular classes \( a \not\in A \), the intersection \( (a + [t - 1]) \cap A \) is either \( I \) or \( I + t \). This can be easily verified. \( \square \)

4 Concluding remarks

- In this paper we proved that the minimum fraction of initial vertices needed to \( k \)-percolate the \( d \)-dimensional grid \([n]^d\) is asymptotically \( 1 - \frac{d}{k} \). The upper bound is given by an explicit construction and the lower bound is proved using a monovariant polynomial. It would be interesting to obtain an exact formula for \( f_{d,k}(n) \), maybe with the extra assumption that \( k \mid n \). Alternatively one can consider the same problem for a slightly denser and more symmetric graph \( \mathbb{Z}_n^d \) instead of \([n]^d\).

- Pete [15] proved that when \( k \leq d \), \( f_{d,k}(n) = \Theta(n^{k-1}) \) by projecting the infected set to subgrids. For this range the polynomial we used does not give useful lower bound since \( 1 - \frac{d}{k} < 0 \). The following question is still open: what is the exact constant \( C_{d,k} \) such that \( f_{d,k}(n) = (C_{d,k} + o(1))n^{k-1} \)? It seems that the only nontrivial result known is \( C_{d,2} = d/2 \) and \( C_{d,d} = 1 \) from [15].

- In general we could ask for which graphs, the polynomial we used in Section 2 can provide the correct asymptotic density of the minimum initial set. It is also worth mentioning that the polynomial we used only has degree 2. Is it possible to use polynomials of higher degrees to solve other deterministic percolation problems, say percolation in other vertex-transitive graphs?
Acknowledgment: We would like to thank Jie Ma and Benny Sudakov for very helpful initial discussions.

References


