ISSN 1342-2812

Research Reports on Mathematical and Computing Sciences

Propagation Connectivity of Random Hypergraphs

Robert Berke and Mikael Onsjo

Oct. 2009, C–265

Department of
Mathematical and
Computing Sciences
Tokyo Institute of Technology

SERIES C: Computer Science

Propagation Connectivity of Random Hypergraphs

Robert Berke and Mikael Onsjö
Dept. of Mathematical and Computing Sciences
Tokyo Institute of Technology, Tokyo, Japan

Abstract

This report addresses the problem of identifying a threshold for propagation connectivity in random hypergraphs as specified in [BO09]. In that paper we gave upper and lower bounds for the threshold that left a gap of a factor $(\log n)(\log \log n)^2$. Unfortunately there is some uncertainty regarding a detail in the lemma that was used to provide the upper bound. Here we provide a simpler alternative lemma and a corresponding upper bound that is slightly less tight but still $\ll 1/n$ for the edge probability.

1 Introduction

We are concerned with 3-uniform random hypergraphs on n vertices, generated by the standard scheme $\mathcal{H}_{3,n,p}$. Each possible hyperedge on exactly three vertices is included in the graph independently with probability p. No other kinds of edges are allowed.

Propagation connectivity is defined in terms of a simple marking process: At some step, $i > 0$, if there is an edge $\{u, v, w\}$ such that u and v are marked (but not w) then we mark w. The process starts from some set of two vertices that are initially marked at step $i = 0$. A graph H is called *propagation connected* iff there is a pair of vertices from which the entire graph can be marked.

This process has been studied before, e.g. in [DN05]. Their result, however, is aimed at general hypergraphs and does not appear to include $\mathcal{H}_{3,n,p}$ for the range of p that we are interested in here.

In [BO09] we provided upper and lower bounds for p such that H is very likely to be, resp. not to be, propagation connected. Here we reiterate some of those results, however since there is some uncertainty regarding weather the result of a cited paper is actually applicable, we give a simpler version of the same lemma and a slightly less tight upper bound. The term "with high probability" (whp), denotes a probability that is $1 - o(1)$ with respect to n. We obtain evidence that if for some constant c, $p > \frac{1}{n(\log n)^{0.4}}$ then the graph is propagation connected whp. Conversely, if for some constant c, $p < \frac{c}{n(\log n)^2}$ then the graph is whp not propagation connected. The latter result is unchanged from [BO09].

2 Propagation Connectivity

For the standard connectivity of normal random graphs following $\mathcal{G}_{n,p}$, we know [ER61] that $\Theta(\log n/n)$ is a threshold for p such that a random graph following $\mathcal{G}_{n,p}$ is connected whp. We believe that a similar threshold exists for our generalized connectivity, which is of some interest by itself. Here we give supporting evidence by showing upper and lower bounds for p such that a random hypergraph following $\mathcal{H}_{n,p}$ is propagation connected whp.

A propagation connected component is a subgraph of H that is propagation connected and maximal in the sense that no other vertex can be included in the component without loosing the propagation connectivity. For our analysis, the following lemma plays an important role. This lemma states that either the marking process terminates earlier or it succeeds in marking the entire hypergraph.

Lemma 1. *Suppose that for some constant ϵ, $pn(\log n)^{1-\epsilon} \to \infty$ with n. Let k denote the size of a largest propagation connected component in H. Then whp, either $k = n$ or $k = o((\log n)^{2-\epsilon})$.*

Remark 1. *For future reference: The lemma holds even if ϵ is zero or negative.*

Proof. We estimate the probability, $\rho(k)$, that there is a propagation connected component on k vertices. By a union bound this is less than the number of choices of the vertices, multiplied by the probability that they are propagation connected and by the probability that no other vertex can be included. The first probability is bounded by one; the second probability is just the condition that there be no edge that has exactly two of its vertices among the k of the component.

$$\rho(k) \;\le\; \binom{n}{k} \cdot 1 \cdot (1-p)^{\binom{k}{2}(n-k)} \;\le\; \exp\left(2\min(k, n-k)\log n - pk^2(n-k)\right)$$

for sufficiently large n. If $n-1 > k > n/2$ then $\rho(k) = o(1)$ if e.g. $p > 12\log n/n^2$. If $k \le n/2$ we see that $\rho(k) = o(1)$ if e.g. $p > (6\log n)/(kn)$. It is apparent that this holds whenever $pn(\log n)^{1-\epsilon} \to \infty$ and $k = \Omega((\log n)^{2-\epsilon})$. The statement of the lemma follows. □

Now we prove our upper bound result. Before starting our analysis let us introduce some random variables and give some preliminary analysis. First note that our marking process can be split into stages. At each stage we only use marked vertices in the previous stage for propagation. For any $t \ge 1$, let L_t (resp., K_t) denote the number of newly marked (resp., all marked) vertices at stage t. Let $K_0 = L_0 = 2$ and $K_{-1} = 0$. K_t can be written as $K_t = \sum_{i=0}^{t} L_i$. Note that $L_{t+1} \sim \mathrm{Bin}((K_{t-1}L_t + \binom{L_t}{2})(n-K_t), p)$; but since we are considering relative small K_t, e.g., $O(\log n)$, we may well approximate it by $\mathrm{Bin}((K_{t-1}L_t + \binom{L_t}{2})n, p)$.

Theorem 1. *If we have $p > \frac{1}{n(\log n)^{0.4}}$, then $H \sim \mathcal{H}_{n,p}$ is whp propagation connected.*

Proof. Suppose that $pn = o(1)$, which is intuitively the most difficult case; the other is left to the reader. Consider the first step of the marking process and estimate L_1. Note that $L_1 \sim \mathrm{Bin}(n-2, p)$, that $\mu_1 = \mathrm{E}[L_1] \sim 1/(\log n)^{0.4}$, and $\mu_1 = o(1)$ by assumption. We show that the probability that L_1 is much larger than this expectation is not so small. For choosing two starting vertices, split the vertex set into subsets of some cardinality n' (that is to be defined later). Consider any one of these subsets; for any integer x in the appropriate range, let $\mathcal{E}_x$ be the event that there is no starting pair of vertices in the subset such that $L'_1 = x$, where L'_1 is the number of marked vertices in the subset from this starting pair. Let $p' = \Pr[L_1 = x]$, which is well approximated as $\binom{n}{x}p^x > \exp(x\log(pn/x))$ for sufficiently large n.

$$\Pr[\mathcal{E}_x] \;=\; (1-p')^{\binom{n'}{2}} \;\approx\; \exp\left(-\frac{n'^2}{2}\exp(x\log(pn'/x))\right) \;,$$

which is less than a given ϵ if $x\log(pn')^{-1} + x\log x < 2\log(n'/\log(\epsilon^{-1}))$. This condition is satisfied, e.g., by the following choice of parameters: $n' = n/(4\log n)$, $\epsilon = n^{-4}$, and $x = 0.01\log n/\log\log n$. Therefore, we conclude that whp there are at least $4\log n$ starting vertex pairs we can choose such that for each we may mark, in the first step of the marking process, a set of size $0.01\log n/\log\log n$ that is disjoint from any other.

Now consider L_2 but this time do not restrict the expansion to a particular subset. The expectation is $\mathrm{E}[L_2] = L_1^2 np > (\log n)^{1.5}$ and the probability of not reaching the expectation is obviously $\leq 1/2$. Thus, with prob. $> 1 - (1/2)^{4\log n} = 1 - n^{-4}$, there is certainly some choice of starting vertices such that L_2 is larger than $(\log n)^{1.5}$. But by applying Lemma 1 with e.g. $\epsilon = 0.6$, we have that if $(\log n)^{1.4}$ vertices can be marked, then so can the entire graph. Hence whp is H propagation connected. □

With the same line of reasoning but by applying upper estimates on L_i, we can show the asymptotic absence of propagation connectivity in graphs where p is somewhat smaller.

Theorem 2. *For any constant c, if we have $p < \frac{c}{n(\log n)^2}$, then $H \sim \mathcal{H}_{n,p}$ is whp* not *propagation connected.*

Proof. Suppose (for induction) that for a given i, it holds that $K_i = O(\log n)$ and $L_i = o(\log n)$. Then since $L_{i+1} \sim \mathrm{Bin}(K_i L_i n, p)$, we have $\mu = \mathrm{E}[L_{i+1}] = K_i L_i np = o(1)$. For some $\delta = \delta(n)$ that tends to infinity, from the standard Chernoff bound, we have $\Pr[\, L_{i+1} \geq \mu\delta \,] \leq \exp(-\mu\delta\log\delta)$. This is less than n^{-10}, e.g., if δ is chosen so that $\delta\mu = (11\log n)/\log\log n$. Hence if $K_i = O(\log n)$ for some i, then we may assume that indeed all L_i, $i \leq i+1$ are $o(\log n)$. Note that $K_{i+1} = \sum_{j=0}^{i+1} L_j$ and that $\Pr[L_j = x_j] \leq (o(1)\mathrm{e}/x_j)^{x_j} < x_j^{-x_j}$ for any x_j (where $0^0 = 1$ in this notation). Using these, we estimate $\Pr[K_{i+1} = s]$ for some number s. For this, consider the ways of picking $i+1$ non-negative integers $x_1, \ldots, x_{i+1}$ such that their sum becomes s; there are $\binom{s+i}{i} < \exp(i\log(\exp(s+i)/i))$ ways of doing so. In each case, we have

$$\prod_{j=1}^{i+1} \Pr[L_j = x_j] \;\leq\; \exp\left(-\sum_{j=0}^{i+1} x_j \log x_j\right) \;\leq\; \exp\left(-s\log\frac{s}{i+1}\right) \; .$$

Hence by a union bound, $\Pr[K_{i+1} = s] \;\leq\; \exp\left(i\log\frac{e(s+i)}{i} - s\log\frac{s}{i+1}\right)$ and again, this is less than n^{-10}, if, e.g., $i \leq 10\log n$ and $s = 40\log n$. But then by recursive application of the union bound, we have with prob. $> 1 - n^{-9}$ that $K_i = O(\log n)$ and $L_i = o(\log n)$ for all $i \leq 10\log n$, which was the assumption made at the beginning of this proof. Hence with prob. $> 1 - n^9$ we have a situation such that in each of $10\log n$ steps the expectation of the increment $\mathrm{E}[L_i]$ is $o(1)$. Therefore with prob. $> 1 - n^{-9}(o(1))^{10\log n} > 1 - n^{-8}$, one of the increments is zero, which means that the propagation stopped. □

Acknowledgement

This research was supported in part by the JSPS Global COE program "Computationism as a Foundation for the Sciences".

References

[BO09] Berke, R., Onsjö, M.: Propagation Connectivity of Random Hypergraphs. In: Proc SAGA'09, LNCS vol. 5792, (2009)

[DN05] Darling, R. W. R., Norris J. R.: Structure of Large Random Hypergraphs. In: Ann. of Applied Prob 15, no. 1A pp. 125–152 (2005)

[ER61] Erdős, P., Rényi, A.: On the evolution of random graphs. In: Bull. Inst. Internat. Statist., vol. 38, pp. 343–347 (1961)