Analysis of Parallel Algorithms for Finding A Maximal Independent Set in A Random Hypergraph

H. CHEN*   A.M.FRIEZE†

February 20, 1996

Abstract

It is well known [9] that finding a maximal independent set in a graph is in class $\mathcal{NC}$, and [10] that finding a maximal independent set in a hypergraph with fixed dimension is in $\mathcal{RNC}$. It is not known whether this latter problem remains in $\mathcal{NC}$ when the dimension is part of the input. We will study the problem when the problem instances are randomly chosen.

It was shown in [6] that the expected running time of a simple parallel algorithm for finding the lexicographically first maximal independent set (lfmis) in a random simple graph is logarithmic in the input size. In this paper, we will prove a generalization of this result. We show that if a random k-uniform hypergraph has vertex set $\{1, 2, \ldots, n\}$ and its edges are chosen independently with probability p from the set of $\binom{n}{k}$ possible edges, then our algorithm finds the lfmis

*Department of Mathematics, Carnegie Mellon University.Pittsburgh, PA 15213

†Department of Mathematics, Carnegie Mellon University. Pittsburgh, PA 15213. Supported in part by NSF Grant CCR-9024935.

in $O(\frac{\log^3 n}{\log\log n})$ expected time. The hidden constant is independent of k, p.

1 Introduction

A *hypergraph* $H(V,E)$ will have $V = V(H)$ as a vertex set, edge $e \in E(H)$ is simply a subset of V. The *dimension* of the hypergraph H is the cardinality of its largest edge. If all the edges in the hypergraph have the same size k then we say that the hypergraph is k-uniform. Thus a 2-uniform hypergraph is simply a graph.

An independent set is a subset S of V which contains no edge. An independent set is maximal if it is not a proper subset of another independent set. We discuss the problem of finding a maximal independent set in parallel. Karp and Wigderson [9] were the first to design a deterministic polylog parallel algorithm to find a maximal independent set in a graph. Beame and Luby [1] also provided a randomized parallel algorithm to find a maximal independent set in a bounded dimension hypergraph. However their analysis only worked for the case of a hypergraph of dimension three. Kelsen [10] later showed that their algorithm is actually an $\mathcal{RNC}$ algorithm for hypergraphs with fixed dimension. Unfortunately the derandomization of their algorithm [10] only gives a sublinear time bound for the deterministic version. Dalhaus, Karpinski and Kelsen [7] extended the idea of Goldberg and Spencer [8] to an $\mathcal{NC}$ algorithm for finding a maximal independent set in hypergraphs of dimension three.

In this paper, we are interested in finding the *lexicographically first* maximal independent set of a hypergraph. It is known that this problem is $\mathcal{P}$-complete even for graphs. Coppersmith, Raghavan and Tompa [6] show that for a random graph the lfmis can be found in $o(\log^2 n)$ expected time. This was later improved to $\Theta(\log n)$ by Calkin, Frieze and Kučera [4] [5]. These algorithms requires only linearly many processors.

In our model $H_{n,k,p}$ of random hypergraphs, all the hypergraphs are k-uniform. The vertex set V is $\{1, 2, \ldots, n\}$. Each of the $\binom{n}{k}$ edges has the same independent probability $p = p(n)$ of appearing. Thus when $k = 2$, we have the commonly used random graph model $G_{n,p}$. The main contribution of this paper is to show that the lfmis of a random k-uniform hypergraph can be found in $o(\log^3 n)$ expected time where there are no hidden factors which depend on p or k. Our model of computation is a Concurrent Read Concurrent Write (CRCW) PRAM with $|V| + |E|$ processors.

We will assume $k \geq 3$ from now on as the case $k = 2$ has been dealt with previously.

We describe an algorithm PHMIS and prove

Theorem 1 *On input $H_{n,k,p}$,* PHMIS *produces its* lfmis *in $O((\log n)^3/\log\log n)$ expected time, where the upper bound is uniform in p and k.*

□

It is conceivable that an improved analysis such as that carried out in [4] for graphs, could lead to a reduction in the upper bound on the expected running time.

2 Algorithm PHMIS and Its Modification

In this section, we give an algorithm to find the lfmis of a hypergraph, and also give a modification of the algorithm needed for the analysis.

2.1 The Parallel Greedy Algorithm

Our algorithm is a natural generalization of the algorithm of [6]. To describe the algorithm, we need the following definitions:

Definition For $x \in V$ and $S \subset V$, we let

$$\Gamma(x, S) = \{e \backslash S : e \in E, x = \max e\}.$$

Definition For $M \subset V$, we let

$$N(M) = N_H(M) = \{x : \exists B \subseteq M : \ x > \max B \ \text{ and } \ B \cup \{x\} \in E\}.$$

Below we give the pseudo code of the algorithm.

Explanation: each execution of the loop, Steps 3 to 8, contributes a set A of vertices which do not form an edge with the set of previously chosen vertices M.

The procedure DELETE removes A and any vertices which form an edge with M - line 12. The edges which contain vertices from $N(M)$ are removed, as they can never be subsets of the finally chosen set M - line 13. Vertices of A are removed from the remaining edges, so that we have a hypergraph with vertices X and edges Y.

Thus at the end of procedure DELETE Y contains all sets $e \backslash M (e \in M)$ which contains no vertex in $N(M)$.

Algorithm PHMIS

0 Input $H = H(V, E)$.
1 **begin**
2 $\quad M \leftarrow \emptyset; X \leftarrow V; Y \leftarrow E.$
3 $\quad$ **while** $X \neq \emptyset$ **do**
4 $\qquad$ **begin**
5 $\qquad\quad A \leftarrow \{a \mid \Gamma(a, M) \cap Y = \emptyset\};$
6 $\qquad\quad M \leftarrow M \cup A;$
7 $\qquad\quad$ DELETE;
8 $\qquad$ **end**
9 $\quad$ Output M;
10 **end**

DELETE

11 **begin**
12 $\quad X \leftarrow X \backslash (A \cup N(M));$

13 $\qquad Y \leftarrow Y\backslash\{e \in Y : e \cap N(M) \neq \emptyset\}$;
14 $\qquad Y \leftarrow \{e\backslash A : e \in Y\}$.
15 **end**

Lemma 1 PHMIS *produces the lexicographically first maximal independent set.*

Proof The algorithm must terminate because the lowest remaining vertex is always added to M at line 6.

M is independent: Suppose M contains the edge $e = \{v_1 < v_2 < \ldots < v_k\}$. Consider the iteration when v_k is added to M. At the start of this iteration $\Gamma(v_k, M) \cap Y = \emptyset$ and so $e\backslash M \notin Y$. Thus during some previous iteration $e \cap N(M) \neq \emptyset$. This implies $e \not\subseteq M$, contradiction.

M is maximal: Consider $v \notin M$. At some stage we must find $v \in N(M)$ which means that $B \cup \{v\}$ is an edge of H for some $B \subseteq M$.

M is lexicographically first: let L be the lexicographically first maximal independent set. If $M \neq L$, then by maximality, $L\backslash M \neq \emptyset$. Let $v = \min L\backslash M$. Since $v \notin M, v \in N(M)$ at some iteration, and so at this time there exists $e = B \cup \{v\} \in E$ with $B \subseteq M$ and $v > \max B$. Furthermore $B \not\subseteq L$ since L is independent. But this implies M precedes L in the lexicographic order. $\Box$

It is easily seen that even if the original graph only has edges of size k, that after the first iteration of the algorithm, edges of size i, $2 \leq i \leq k-1$ may appear. It turns out that the key to the analysis of the algorithm will be knowledge of the distribution of edges of different sizes at each stage of the algorithm.

2.2 Modified-PHMIS(H)

As in the case of graphs, it is difficult to analyze the running time of PHMIS directly. After the first iteration of the algorithm, the remaining graph is conditioned and it becomes difficult to handle this. To circumvent this, we use the same trick as [6]. Thus at each time, intead of running PHMIS on the whole graph, we pick a block of m least numbered elements and run PHMIS until completion on this block. Then after a clean up step, we proceed by picking the next m least numbered elements and so on. This procedure mimics the behavior of the sequential algorithm which does not condition the remaining graph so severely. The number of vertices m is carefully chosen so as to achieve an efficient time bound.

It will be shown that the running time of this modified algorithm dominates the running time of PHMIS, and that it also produces the lfmis.

We give the pseudo code for the modified version of PHMIS below. Let $\alpha = \frac{1}{2e}$.

The quantity p_i will later be seen to be the probability that a given edge of size i exists in the graph that remains. Recall that k is the dimension of the hypergraph H.

We define $H[Z]$, the subgraph induced by Z, as $(Z, \{e \in E : e \subset Z\})$.

Algorithm Modified-PHMIS(H)

begin
    $M \leftarrow \emptyset$; $X \leftarrow V; Y \leftarrow E; m \leftarrow \lfloor k\alpha p^{-\frac{1}{k-1}} \rfloor$.
        If $|X| \geq m$ **do begin**
            $Z \leftarrow$ the m lowest elements of X;
            Run PHMIS to completion on $H[Z]$;
            Let the independent set found be A;
            $M \leftarrow M \cup A$;
            DELETE;
            $p_i = 1 - (1-p)^{\binom{|M|}{k-i}}, 2 \leq i \leq k$.
            $m \leftarrow \lfloor \min\{2\alpha p_2^{-1}, 3\alpha p_3^{-\frac{1}{2}}, \ldots, k\alpha p_k^{-\frac{1}{k-1}}\} \rfloor$;
            end
        Run PHMIS until completion on $H[X]$;
        Let the independent set found be A;
        $M \leftarrow M \cup A$;
    Output M;

16 **end**

Remark 1 *Note that when line 9 is first executed,* $|M| \geq k-1$.

Remark 2 *The value of* i *for which* $i\alpha p_i^{-1/(i-1)}$ *is minimized in the definition of* m *in line 10, is called the* dominant edge size ι.

Remark 3 *In both algorithms,* $v \in X$ *implies there exists an edge* $e = \{v_1, v_2, \ldots, v_{k-1}, v\} \in E$ *where each* v_i *is either in* X *or in the current independent set, at least one being in* X.

Remark 4 *We let an execution of* DELETE *denote the end of an* iteration *in both algorithms. Thus the execution of Step 5 of* Modified-PHMIS *may involve several iterations. An execution of Step 8 in* Modified-PHMIS *signifies the end of a* round *of this algorithm.*

Lemma 2 *The number of iterations executed by* Modified-PHMIS *is at least the number of iterations of* PHMIS.

Proof Let the subscripts p and mp refer to corresponding sets in algorithms PHMIS and Modified-PHMIS. Let the superscript (i) denote the set at the end of iteration i. Thus the sets X and M are initially $X_p^{(0)}$ and $M_p^{(0)}$ for PHMIS, and $X_{mp}^{(0)}$ and $M_{mp}^{(0)}$ for Modified-PHMIS.

It suffices to prove the following statements for each $i \geq 0 : X_p^{(i)} \subseteq X_{mp}^{(i)}, M_{mp}^{(i)} \subseteq M_p^{(i)}$ and $M_p^{(i)} - M_{mp}^{(i)} \subseteq X_{mp}^{(i)}$. These statements clearly hold for $i = 0$. Assuming they hold for i let us establish them for $i+1$.

First we show that $M_{mp}^{(i+1)} \subseteq M_p^{(i+1)}$. Let $v \in X_{mp}^{(i)} - M_p^{(i+1)}$ be added to $M_{mp}^{(i+1)}$ at iteration $i+1$. If $v \in X_p^{(i)}$, then v was not added at this step to $M_p^{(i+1)}$ because v is maximal in an edge $e = (v_1 < v_2 < \ldots < v_{k-1} < v)$ all of those vertices belong to $X_p^{(i)} \cup M_p^{(i)}$. By the inductive claim $e \subseteq X_{mp}^{(i)} \cup M_{mp}^{(i)}$. But then v will not be added to $M_{mp}^{(i+1)}$ because of e, a contradiction.

Now consider the case $v \notin X_p^{(i)}$. That means v was deleted from $X_p^{(j)}$ for some $j < i$. Thus there exists an edge $e = (v_1 < v_2 < \ldots < v_{k-1} < v)$ in E such that all the v_i's belong to $M_p^{(i)}$. By the inductive claim they also belong to $X_{mp}^{(i)} \cup M_{mp}^{(i)}$. But that implies that v will not be put into $M_{mp}^{(i+1)}$ because of e, again a contradiction. We conclude that $M_{mp}^{(i+1)} \subseteq M_p^{(i+1)}$.

Next we need to show that $X_p^{(i)} \subseteq X_{mp}^{(i)}$ and $M_p^{(i)} - M_{mp}^{(i)} \subseteq X_{mp}^{(i)}$. Suppose that v is deleted from $X_{mp}^{(i)}$ in iteration $i+1$. This is because $v \in N(M_{mp}^{(i+1)})$. But this implies $v \in N(M_p^{(i+1)})$, since we just showed $M_{mp}^{(i+1)} \subseteq M_p^{(i+1)}$. Hence v is deleted by PHMIS at iteration $i+1$ or at an earlier iteration. We conclude $X_p^{(i)} \subseteq X_{mp}^{(i)}$ and $M_p^{(i)} - M_{mp}^{(i)} \subseteq X_{mp}^{(i)}$. This completes the proof of the lemma. □

Corollary 1 *The independent set produced by* PHMIS *is the same as the independent set produced by* Modified-PHMIS.

Proof Let M denote the independent set produced by PHMIS and $\tilde{M}$ denote the independent set produced by Modified-PHMIS. By **Statement 1**, $\tilde{M} \subseteq M$. Let $v = \min M \backslash \tilde{M}$. Since v is not in $\tilde{M}$, it has to be deleted by Modified-PHMIS as a neighbor of $k-1$ members of $\tilde{M}$. But then v cannot be placed in M either since $\tilde{M} \subseteq M$ - contradiction. □

3 Analysis of Modified-PHMIS(H)

The analysis will be in two parts. In Section 3.1, we prove that **whp**[1] there will be no long chain in each block Z, so that each round of Modified-PHMIS will finish in $O(\log n / \log\log n)$ steps.

In Section 3.2, we study the change of ι, and show that **whp** after $O(\log^2 n)$ rounds, either the algorithm terminates; or the ι will be 2. The rapid convergence of the algorithm then follows from a similar analysis to that of [6]. Combining these two parts gives a polylogarithmic time bound.

[1]By **whp**, we mean that an event occurs with probability $1 - o(1/n)$

Outline of the proof The second part of the proof of our result is rather technical and it will be helpful to give an outline. All statements in this section will later be shown to hold **whp** and we will drop this qualification for the outline.

We first show (Lemma 4) that if at any stage $m \leq (\log n)^2/2\log\log n$ then the algorithm will finish in $O((\log n)^3/\log\log n)$ more iterations. So assume for the remainder of this section that $m > (\log n)^2/2\log\log n$.

We show next (Lemma 5) that either the independent set M increases by at least m/e during any round or failing this, the algorithm terminates in $O(\log n)$ more iterations.

We show (Lemma 7) that

$$|\mathsf{lfmis}| \leq e^{-1}(k + 2\log n)\theta\mu.$$

where $\mu = p^{-1/(k-1)}$ and $1 \leq \theta \leq e$ is the solution of the transcendental equation (10).

We then show (Lemma 8) that if

$$|M| \;=\; e^{-1}(k - \Delta)\theta\mu \tag{1}$$

where $k > \Delta > 4(\log n)^2$, then it takes $O(\log n)$ rounds to increase $|M|$ to size at least $e^{-1}(k - \Delta/\log n)\theta\mu$.

Thus after $O((\log n)^2/\log\log n)$ rounds, the value of Δ in (1) is at most $4\log n$.

We then show (Lemma 10) that after a further $O(\log^2 n)$ rounds $\iota = 2$ and remains so for the rest of the algorithm.

Once $\iota = 2$ we can apply essentially the same argument as in [6] to show that a final $O(\log n)$ rounds suffice.

3.1 Distribution of Edge Sizes

Next let us consider the distribution of edge sizes in the hypergraph $K = (X, Y)$ remaining at each round. If $f \subseteq X$ with $|f| = i$ then f appears as an edge in K if and only if there is some $g \subseteq M$, $|g| = k - i$ such that $e = f \cup g \in E$. Since distinct g yield distinct e, we have, by definition of $H_{n,k,p}$

$$\begin{aligned}\mathbf{Pr}[\ f \subseteq X \text{is not an edge of } (X,Y)\ ] &= (1-p)^{\binom{|M|}{k-i}} \\ &= 1 - p_i\end{aligned}$$

Furthermore, edges of different size can only arise from the contraction of different original edges of H. So the distribution of edges at any stage can be summarized: for $2 \leq i \leq k$, each of the $\binom{|X|}{i}$ possible remaining edges appears independently with probability p_i.

3.2 Long Forcing Sequences in an m-Block

In simple graphs, as studied in [6], for PHMIS to require l rounds to deal with an m-block Z (chosen in line 4 of Modified-PHMIS), there has to be a path of length $2l$. We define a structure with the same effect in the case of a k-uniform hypergraph.

We show that if at least l iterations in one round of Modified-PHMIS are needed then there is a *forcing sequence* $v_1 < v_2 < \ldots < v_l$ where v_i is deleted in iteration i and certain other properties hold.

Let $v = v_l$ be added to M in the lth iteration. Let Y_{l-1} denote the set at the start of iteration $l - 1$. Since v survives the $(l-1)$st iteration, the following is true:

v is the highest vertex of a t-edge $e = \{w_1 < w_2 < \ldots < w_t = v\} \in Y_{l-1}$ for some t, $2 \leq t \leq k$. At least one of $w_1, w_2, \ldots, w_{t-1}$ is deleted as a member of $N(M)$ in iteration $l - 1$. Otherwise v won't be placed in M in iteration l.

Let $\gamma_l = w_\tau$ be such a deleted vertex. There has to be another t'-edge $e' = \{u_1 < u_2 < \ldots < u_{t'} = w_\tau\} \in Y_{l-1}$ for some t', $2 \leq t' \leq k$, such that $u_1, u_2, \ldots, u_{t'-1}$ are placed in the M in the $(l-1)$st iteration. Let $v_{l-1} = u_1$.

Clearly $v_l > \gamma_l > v_{l-1}$. This argument is repeated to define $v_{l-2}, v_{l-3}, \ldots, v_1$.

Let $\mathcal{E}(x, y)$ denote the event that vertices x and y are related in the way v_{l-1}

and v_l are as described above. Then we claim that for all x, y,

$$\mathbf{Pr}[\overline{\mathcal{E}}(x,y)] \;\geq\; \prod_{t=2}^{k}\prod_{t'=2}^{k}(1-p_{t'}p_t)^{(t-1)\binom{m-1}{t-1}\binom{m-t}{t'-2}}$$

This is because $\overline{\mathcal{E}}(x,y)$ occurs if there is no triple (e, e', w) where e is a t-edge containing y, w and e' is a t'-edge containing x and w.

Lemma 3 **whp** *a given m-block Z contains no forcing sequence of length $l \geq \lceil 4\log n/\log\log n\rceil$.*

Proof Clearly the probability of the existence of a forcing sequence of length l is bounded by

$$\begin{aligned}
\sum_{v_1<v_2<\ldots<v_l}\mathbf{Pr}\Big(\bigcap_{i=1}^{l-1}\mathcal{E}(v_i,v_{i+1})\Big) &\leq \binom{m}{l}\left[1-\prod_{t=2}^{k}\prod_{t'=2}^{k}(1-p_tp_{t'})^{(t-1)\binom{m-1}{t-1}\binom{m-t}{t'-2}}\right]^{l-1} && (2)\\
&\leq \frac{m^le^l}{l^l}\left[\sum_{t=2}^{k}\sum_{t'=2}^{k}(t-1)\binom{m-1}{t-1}\binom{m-t}{t'-2}p_tp_{t'}\right]^{l-1}\\
&\leq \frac{m^le^l}{l^l}\left[\sum_{t=2}^{k}\sum_{t'=2}^{k}m^{-1}\frac{(\alpha t)^{t-1}(\alpha t')^{t'-1}}{(t-2)!(t'-2)}\right]^{l-1} && (3)\\
&\leq \frac{m^le^l}{l^l}\left[\frac{1}{m}\left(\sum_{t=2}^{\infty}\frac{(\alpha t)^{t-1}}{(t-2)!}\right)^2\right]^{l-1} && (4)\\
&\leq \frac{m^le^l}{l^l}\left(\frac{c_1}{m}\right)^{l-1} && \text{for some constant } c_1>0\\
&= m\left(\frac{c_1e}{l}\right)^l\\
&= o(n^{-2}).
\end{aligned}$$

For (3) use

$$\binom{m-1}{t-1}p_t \;\leq\; \frac{(m-1)^{t-1}}{(t-1)!}p_t$$

and

$$m \;\leq\; t\alpha p_t^{-\frac{1}{t-1}}, \qquad \text{for } 2\leq t\leq k. \tag{5}$$

A justification is needed for inequality (2). We claim that the edge sets involved in different events $\mathcal{E}(v_i, v_{i+1})$ are disjoint. This is because edges causing $\mathcal{E}(v_i, v_{i+1})$ will either have v_{i+1} or γ_{i+1} as their highest element. $\square$

We show next that **whp** a substantial portion of m will be placed in the lfmis in each round. This is stated as Lemma 5. Before proving the lemma, we show that we can assume that at any stage of the algorithm $m \geq \frac{1}{2}\frac{\log^2 n}{\log\log n}$.

Lemma 4 *If at some stage $m \leq \frac{1}{2}\frac{\log^2 n}{\log\log n}$, then* **whp**, *the number of vertices added in later rounds to the independent set is $O(\frac{\log^3 n}{\log\log n})$.*

Proof Suppose that at some point $m = \lfloor \frac{l}{2e} p_l^{-\frac{1}{l-1}} \rfloor \leq \frac{1}{2}\frac{\log^2 n}{\log\log n}$. This implies that $l = o(\log^2 n)$. Once the size of the independent set subsequently reaches $a = \lceil \frac{10\log^3 n}{\log\log n} \rceil$, then **whp** no more vertices can be added. This follows as the expected number of possible additions n_0 after this time satisfies,

$$
\begin{aligned}
n_0 &\leq n(1-p_l)^{\binom{a}{l-1}} \\
&\leq n\exp\left\{-\binom{a}{l-1}p_l\right\} \\
&\leq n\exp\left\{-\frac{(a-l+2)^{l-1}}{(l-1)!}p_l\right\} \\
&\leq n\exp\left\{-\frac{1}{\sqrt{2\pi(l-1)}}\left[\frac{(a-l+2)e}{l-1}p_l^{\frac{1}{l-1}}\right]^{l-1}\right\} \\
&\leq n\exp\left\{-\frac{1}{\sqrt{2\pi(l-1)}}(9\log n)^{l-1}\right\} \quad (\text{ Since } p_l \text{ only increases}) \\
&\leq n\exp\{-3\log n\} \\
&= n^{-2}.
\end{aligned}
$$

Each iteration of $\mathsf{Modified\text{-}PHMIS}$ adds at least one vertex to the independent set. So if $m \leq \frac{1}{2}\frac{\log^2 n}{\log\log n}$ then **whp** there are fewer than $\lceil 10\log^3 n/\log\log n \rceil$ more iterations until completion. $\square$

So from now on we assume that throughout the algorithm

$$m \geq \frac{\log^2 n}{2 \log\log n}. \tag{6}$$

Lemma 5 *Let Z be the set of elements we run* Modified-PHMIS *on in some round, and let $A \subseteq Z$ be the independent set that is placed in* lfmis *in this round. Then* **whp** *either*
(i) the algorithm terminates in at most $25 \log n$ more iterations; or
(ii) $|A| \geq m/e$.

Proof The probability of an element being placed in A is bounded below by the probability it is placed in $\tilde{A} \subseteq A$ where $\tilde{A}$ is the set of elements in Z with $\Gamma(z, M) \cap Y = \emptyset$. But

$$\begin{aligned}
\mathbf{Pr}[z \in \tilde{A}] &\geq \prod_{i=2}^{m} (1-p_i)^{\binom{m}{i-1}} \\
&\geq \exp\left\{-\sum_{i=2}^{m} \binom{m}{i-1} \frac{p_i}{1-p_i}\right\} \\
&\geq \exp\left\{-\sum_{i=2}^{m} \frac{m^{i-1}}{(i-1)!} \frac{p_i}{1-p_i}\right\} \\
&\geq \exp\left\{-\sum_{i=2}^{m} \frac{(\alpha i)^{i-1}}{(i-1)!} \frac{1}{1-p_i}\right\}.
\end{aligned} \tag{7}$$

Here we use the inequality

$$1 - x \geq e^{-x/(1-x)}, \quad \text{for } 0 \leq x \leq 1. \tag{8}$$

Case 1: $\exists\, 2 \leq i \leq \log n$ such that $p_i \geq \frac{1}{10}$. In this case at most $25 \log n$ more elements will be added to the independent set. Indeed for $t = \lceil 25 \log n \rceil$, t elements can be added only if they contain no edge of size i. Hence

$$\begin{aligned}
\mathbf{Pr}[\ t \text{ more elements can be added }] &\leq \binom{n}{t} (1-p_i)^{\binom{t}{i}} \\
&\leq n^t \exp\left\{-\binom{t}{i}/10\right\} \\
&\leq n^{-3}.
\end{aligned}$$

Case 2: $p_i \leq 1/10$ for $2 \leq i \leq \log n$. Suppose next that $p_i \geq 1/10$ for some $i > \log n$. Then

$$\begin{aligned} m &= \lfloor \min\{2\alpha p_2^{-1}, 3\alpha p_3^{-\frac{1}{2}}, \ldots, k\alpha p^{-\frac{1}{k-1}}\} \rfloor \\ &\leq \frac{i}{e}, \qquad\qquad \text{since } i \geq \log n. \end{aligned}$$

This is a contradiction if $i \leq m$ and thus shows that $p_i \leq \frac{1}{10}$ for $i \leq m$.

So, from (7),

$$\begin{aligned} \mathbf{Pr}[z \in \tilde{A}] &\geq \exp\left\{-\frac{10}{9}\sum_{i=2}^{m}\frac{(\frac{i}{2e})^{i-1}}{(i-1)!}\right\} \\ &\geq e^{-0.9}. \end{aligned} \tag{9}$$

Now the events $z \in \tilde{A}$ are mutually independent since they depend on disjoint sets of edges of $H_{n,k,p}$. Using (6) and standard estimates for the tails of the binomial distribution and the fact that $m = \Omega(\log^2 n/\log\log n)$, it is easy to see that **whp**, $|\tilde{A}| \geq m/e$. □

3.3 Size of lfmis

In this section, we will give an upper bound on the size of the lfmis of a k-uniform random hypergraph with edge density p.

Definition. For $0 < x < 1$, let

$$\begin{aligned} \Psi(x) &= x^{-1}(x + (1-x)\log(1-x)) \\ &= \frac{x}{1\cdot 2} + \frac{x^2}{2\cdot 3} + \cdots + \frac{x^i}{i(i+1)} + \cdots \\ \mu &= p^{-\frac{1}{k-1}} \end{aligned}$$

and θ be the unique positive solution of the equation,

$$\Phi(x) = \frac{1}{x}\exp\left\{\Psi\left(\frac{e}{\mu x}\right)\right\} = 1. \tag{10}$$

θ exists as the LHS of (10) decreases from $\mu > 1$ to 0 as x increases from e/μ to ∞.

Lemma 6 $1 \leq \theta \leq e$.

Proof If $e \geq \mu$ then $\theta \geq e/\mu \geq 1$. If $e < \mu$, $\theta \geq 1$ follows from $\Phi(1) = \exp\{\Psi(e/\mu)\} \geq 1$.

For the upper bound, notice that if $\theta > e$ then $\mu > 1$ implies,

$$\frac{1}{\theta}\exp\left\{\Psi\left(\frac{e}{\theta}\right)\right\} > 1.$$

A contradiction. $\square$

In the following proof and some later ones, C represents a generic (absolute) large constant and c a small constant. Their values may change from time to time but the reader can easily substitute fixed values to make the arguments correct.

Lemma 7 **whp**, *the size of the* lfmis *of* $H_{n,k,p}$ *is bounded above by* $e^{-1}(k + 2\log n)\theta\mu$.

Proof Let $a = |M|$ denote the size of the independent set at the start of some round and $b = a/\mu$. The expected number of nodes n_0 that now can be added satisfies

$$\begin{aligned}
n_0 &\leq n(1-p)^{\binom{a}{k-1}} \\
&\leq n\exp\left\{-\binom{a}{k-1}p\right\} \\
&\leq n\exp\left\{-\frac{C}{\sqrt{k-1}}\frac{a^a}{(k-1)^{k-1}(a-k+1)^{a-k+1}}p\right\} &(11)\\
&= n\exp\left\{-\frac{C}{\sqrt{k-1}}\left[\frac{b}{k-1}\left(1-\frac{k-1}{a}\right)^{-\left(\frac{a}{k-1}-1\right)}\right]^{k-1}\right\} &(12)\\
&= n\exp\left\{-\frac{C}{\sqrt{k-1}}\left[\frac{be}{k-1}\exp\left\{-\Psi\left(\frac{k-1}{a}\right)\right\}\right]^{k-1}\right\}.
\end{aligned}$$

[(11) follows from substituting Stirling's formula for factorials.]

Let $a = e^{-1}(k + 2\log n)\theta\mu$. Since

$$\begin{aligned} \exp\left\{-\Psi\left(\frac{k-1}{a}\right)\right\} &= \exp\left\{-\Psi\left(\frac{(k-1)e}{(k+2\log n)\theta\mu}\right)\right\} \\ &\geq \exp\left\{-\Psi\left(\frac{e}{\theta\mu}\right)\right\}, \end{aligned}$$

we have

$$\theta \exp\left\{-\Psi\left(\frac{k-1}{a}\right)\right\} \geq 1.$$

This implies

$$\begin{aligned} n_0 &\leq n\exp\left\{-\frac{C}{\sqrt{k-1}}\left[\frac{k+2\log n}{k-1}\theta\exp\left\{-\Psi\left(\frac{k-1}{a}\right)\right\}\right]^{k-1}\right\} \\ &\leq n\exp\left\{-\frac{C}{\sqrt{k-1}}\left(1+\frac{2\log n}{k-1}\right)^{k-1}\right\} \\ &= o(n^{-1}). \end{aligned}$$

So, **whp**, $|\mathsf{lfmis}| \leq e^{-1}(k+2\log n)\theta\mu$. □

3.4 Size of Dominant Edges

In this section we will study the change of edge densities as the algorithm proceeds. The main technical lemma we are going to prove is the following:

Lemma 8 *Let the independent set M be of size $a = e^{-1}(k-\Delta)\theta\mu$ at the start of some round, where $k > \Delta > 4\log n$. Let ι be as in Remark 2 Thereafter* **whp**,

$$\Delta/5 \leq \quad \iota \quad \leq 5e^2\Delta \tag{13}$$

and

$$\alpha\Delta\mu/5 \leq \quad \alpha\iota p_\iota^{-\frac{1}{\iota-1}} \quad \leq (1+o(1))\alpha\Delta\mu \tag{14}$$

except for possibly $O(\log n)$ rounds.

Proof We first observe that if $0 < x' = 1-(1-x)^m < \frac{1}{2}$, $m \geq 1$ is integer then

$$x' \leq \quad mx \quad \leq x' + x'^2 \tag{15}$$

The LHS of (15) follows from $(1-x)^m \geq 1-mx$. For the RHS we use $x' \geq 1-e^{-mx}$ which implies

$$\begin{aligned} mx &\leq -\log(1-x') \\ &\leq x' + x'^2. \end{aligned}$$

We consider two cases:

Case 1: $\mu \geq 4e^2$. (Recall $\mu = p_k^{-\frac{1}{k-1}}$)

$m = \lfloor \alpha k\mu \rfloor$ in the first round, and **whp** (Lemma 5) , at least m/e elements will be placed in M. The assumption guarantees that **whp**, the independent set after round 1 will be of size at least $2k$.

Now at any stage $p = p_k \leq p_{k-1} \leq \cdots \leq p_2$. So if any p_i is greater than $\sqrt{1/\log n}$ then $p_2 \geq \sqrt{1/\log n}$.

In this case the current independent set is maximal **whp**, since the expected number of possible additional members is at most

$$n(1-p_2)^{|M|} \quad \leq \quad n \exp\left\{-\frac{(\log n)^{3/2}}{\log\log n}\right\} \tag{16}$$

using (6) and Lemma 5.

It follows from the LHS of (15) and $p_i \leq (\log n)^{-1/2}$ that, if $a = |M| = b\mu$ and $k > i$ then

$$p_i \quad \leq \quad \binom{a}{k-i} p$$

$$
\begin{aligned}
&\leq \frac{2}{\sqrt{2\pi(k-i)}}\left(\frac{b}{k-i}\right)^{k-i}\left(1-\frac{k-i}{a}\right)^{-a+k-i} p^{\frac{i-1}{k-1}} \\
&= \frac{2}{\sqrt{2\pi(k-i)}}\left[\frac{b}{k-i}\exp\left\{1-\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{k-i} p^{\frac{i-1}{k-1}}.
\end{aligned}
$$

We apply the upper bound of (15), with $x = p$, $m = \binom{a}{k-i}$ and $x' + x'^2 = (1+o(1))x'$, since $x' = p_i = o(1)$. Then a similar calculation shows that

$$
p_i \geq \frac{1}{2\sqrt{2\pi(k-i)}}\left[\frac{b}{k-i}\exp\left\{1-\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{k-i} p^{\frac{i-1}{k-1}}.
$$

Thus we have

$$
p_i^{-\frac{1}{i-1}} = (1+\epsilon_i)\left[\frac{k-i}{be}\exp\left\{\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{\frac{k-i}{i-1}} \mu \tag{17}
$$

where $1/2 \leq (1+\epsilon_i)^{i-1}[2\pi(k-i)]^{-1/2} \leq 2$.

Equation (17) is clearly true for $k = i$ as $\epsilon_i = 0$ in this case.

Let $i_o = \lceil \Delta \rceil$. Then

Case 1a: $i \geq 4\Delta$.

$$
\begin{aligned}
\alpha i p_i^{-\frac{1}{i-1}} &= \alpha i(1+\epsilon_i)\left[\frac{k-i}{be}\exp\left\{\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{\frac{k-i}{i-1}} \mu \\
&= \alpha i(1+\epsilon_i)\left[\frac{k-i_o}{be}\exp\left\{\Psi\left(\frac{k-i_o}{a}\right)\right\}\right. \\
&\qquad \times \left.\frac{k-i}{k-i_o}\exp\left\{\Psi\left(\frac{k-i}{a}\right)-\Psi\left(\frac{k-i_o}{a}\right)\right\}\right]^{\frac{k-i}{i-1}} \mu \\
&= \alpha i(1+\epsilon_i+o(1))\left[\frac{k-i}{k-i_o}\exp\left\{\Psi\left(\frac{k-i}{a}\right)-\Psi\left(\frac{k-i_o}{a}\right)\right\}\right]^{\frac{k-i}{i-1}} \mu.
\end{aligned} \tag{18}
$$

Note that we have used (10) to obtain the expression (18) from the previous line.

Let

$$M_1 = \left[\frac{k-i}{k-i_o}\right]^{\frac{k-i}{i-1}} \quad \text{and} \quad M_2 = \exp\left\{\Psi\left(\frac{k-i}{a}\right) - \Psi\left(\frac{k-i_o}{a}\right)\right\}^{\frac{k-i}{i-1}}$$

We show that M_1 and M_2 are bounded below by constants.

$$\begin{aligned}
M_1 &= \left(\frac{k-i}{k-i_o}\right)^{\frac{k-i}{i-1}} \\
&= \left(1 - \frac{i-i_o}{k-i_o}\right)^{\frac{k-i}{i-1}} \\
&\geq \exp\left\{-\frac{i-i_o}{k-i_o}\frac{k-i}{i-1}\frac{k-i_o}{k-i}\right\} \qquad \text{from } 8 && (19) \\
&= \exp\left\{-\frac{i-i_o}{i-1}\right\} \\
&\geq e^{-1}. && (20)
\end{aligned}$$

Now let $\tau = e/\mu\theta \leq e/\mu \leq 1/4e$.

Then

$$\begin{aligned}
M_2 &= \exp\left\{\Psi\left(\frac{k-i}{a}\right) - \Psi\left(\frac{k-i_o}{a}\right)\right\}^{\frac{k-i}{i-1}} \\
&= \exp\left\{(i_o - i)\sum_{t=1}^{\infty}\frac{1}{at(t+1)}\sum_{s=1}^{t}\left(\frac{k-i}{a}\right)^{t-s}\left(\frac{k-i_o}{a}\right)^{s-1}\right\}^{\frac{k-i}{i-1}} \\
&\geq \exp\left\{(i_o - i)\sum_{t=1}^{\infty}\frac{\tau^{t-1}}{at(t+1)}\sum_{s=1}^{t}\left(\frac{k-i}{k-\Delta}\right)^{t-s}\right\}^{\frac{k-i}{i-1}} \\
&\geq \exp\left\{\frac{i_o - i}{a}\sum_{t=1}^{\infty}\frac{\tau^{t-1}}{t+1}\right\}^{\frac{k-i}{i-1}} \\
&\geq \exp\left\{\frac{i_o - i}{a}\right\}^{\frac{k-i}{i-1}} \\
&\geq 9/10, \qquad \text{since } a \geq k-1. && (21)
\end{aligned}$$

It follows from (17) and (18) that

$$\alpha i_o p_{i_o}^{-\frac{1}{i_o-1}} \leq (1+o(1))\alpha\Delta\mu \tag{22}$$

and that for $i \geq 4\Delta$

$$\alpha i p_i^{-\frac{1}{i-1}} \geq \frac{36}{10e}\alpha\Delta\mu \tag{23}$$

which confirms the RHS of (13).

Case 1b: $i \leq \Delta/5$.

Applying (17) we obtain

$$\begin{aligned} \alpha i p_i^{-\frac{1}{i-1}} &= \alpha(1+\epsilon_i)i\left[\frac{k-i}{be}\exp\left\{\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{\frac{k-i}{i-1}}\mu \\ &\geq \alpha i\left(\frac{k-i}{k-\Delta}\right)^{\frac{k-i}{i-1}}\mu \\ &= \alpha i\left(1+\frac{\Delta-i}{k-\Delta}\right)^{\frac{k-i}{i-1}}\mu \end{aligned} \tag{24}$$

Consider two subcases here.

If $\frac{\Delta-i}{k-\Delta} \geq 1$, then

$$\begin{aligned} \alpha i p_i^{-\frac{1}{i-1}} &\geq \alpha i 2^{\frac{k-i}{i-1}}\mu \\ &\geq 16\alpha\Delta\mu/5. \end{aligned} \tag{25}$$

If $\frac{\Delta-i}{k-\Delta} < 1$, then from (24) and $1+x \geq e^{x/2}, 0 \leq x \leq 1$,

$$\begin{aligned} \alpha i p_i^{-\frac{1}{i-1}} &\geq \alpha i \exp\left\{\frac{1}{2}\frac{\Delta-i}{i-1}\frac{k-i}{k-\Delta}\right\}\mu \\ &\geq \alpha i \exp\left\{\frac{\Delta-i}{2i}\right\}\mu \\ &\geq e^2\alpha\Delta\mu/5 \end{aligned} \tag{26}$$

which follows from the fact that $i\exp\left\{\frac{1}{2}\frac{\Delta-i}{i}\right\}$ is monotone increasing for $i \leq \Delta/5$ as i decreases. Since $\frac{16}{5} > \frac{e^2}{5} > 1$, this confirms the LHS of (13) in Case 1.

Case 2: $1 \leq \mu \leq 4e^2$. (Recall that $k \geq 4\log n$).

Case 2a: $\iota > k/\log n$. Now **whp** each round adds at least

$$\begin{aligned} m/e &\geq \iota\alpha/e \\ &= \iota/(2e^2) \end{aligned}$$

elements into our independent set.

Furthermore, Lemma 7 and $\mu \leq 4e^2$ imply that **whp** the size of the lfmis is at most $6e^2k$. So this case can occur at most $12e^4\log n$ times.

Case 2b: $\iota \leq k/\log n$.

$$\begin{aligned} \binom{a}{k-i}p &\leq \frac{C}{\sqrt{k-i}}\left[\frac{b}{k-i}\exp\left\{1-\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{k-i} p^{\frac{i-1}{k-1}} \\ &\leq \frac{C}{\sqrt{(k-i)}}p^{\frac{i-1}{k-1}} \\ &= o(1), \end{aligned}$$

provided $k-i \to \infty$. In particular this is true for $i \leq \Delta/3$. In which case we can use the analysis of Case 1b. The only problem is that when $i \geq 4\Delta$, $\binom{a}{k-i}p$ might be large and so cannot used as an approximation for p_i.

Let now $i_o = \lfloor \Delta - \sqrt{\Delta} \rfloor$. Then

$$\alpha i_o p_{i_o}^{-\frac{1}{i_o-1}} = (1+o(1))\alpha i_o \mu \tag{27}$$

So when $i \geq 5e^2 i_o$,

$$\begin{aligned} \alpha i p_i^{-\frac{1}{i-1}} &\geq \alpha i \\ &\geq 5e^2\alpha i_o \\ &> 5\alpha i_o\mu/4, \end{aligned}$$

since $\mu \leq 4e^2$. This completes the proof of the first part of Lemma 8.

For the second part, it follows from (22) and (27) that **whp**,
$\alpha\iota p_\iota^{-\frac{1}{\iota-1}} \leq (1+o(1))\alpha\Delta\mu$, which confirms the RHS of (14).

For the LHS of (14), consider the following two cases:

(a) $\iota \geq \Delta$. Since $M_1 \geq e^{-1}, M_2 \geq 0.9$, we obtain from (17) that

$$\alpha \iota p_\iota^{-\frac{1}{\iota-1}} \geq M_1 M_2 \alpha \iota \mu \geq \frac{1}{5}\alpha \Delta \mu.$$

(b) $\iota \leq \Delta$. From (25) (26) (which are valid for all i) and the fact that $i \exp\left\{\frac{1}{2}\frac{\Delta - i}{i}\right\}$ reaches its minimum for $i \leq \Delta$ at $i = \Delta/2$,

$$\alpha \iota p_\iota^{-\frac{1}{\iota-1}} \geq \frac{\sqrt{e}}{2}\alpha \Delta \mu.$$

This completes the proof of the LHS of (14). □

With this lemma, we can bound the change of ι.

Lemma 9 **Whp** *it takes $O(\log n)$ rounds for the size of the independent set to increase from $e^{-1}(k - \Delta)\theta\mu$ to at least $e^{-1}(k - \frac{\Delta}{\log n})\theta\mu$, assuming that $\Delta \geq 4\log^2 n$. Equivalently, it takes $O(\log n)$ rounds to decrease ι from $\Theta(\Delta)$ to $\Theta(\frac{\Delta}{\log n})$.*

Proof We estimate the number of rounds required to increase $|M|$ from $e^{-1}(k - \Delta)\theta\mu$ to at least $e^{-1}(k - \Delta')\theta\mu$, where $k > \Delta = \Delta' \log n > 4\log^2 n$.

By the previous Lemma and Lemma 5, in each round while $|M| \leq e^{-1}(k - \Delta')\theta\mu$ it increases in size by at least $\alpha\Delta'\mu/10e$. So the number of rounds required is at most $10e(\Delta - \Delta')\alpha\theta/\Delta'$. Consequently, only $O(\log n)$ rounds transpire before $|M|$ increases from $e^{-1}(k - \Delta)\theta\mu$ to at least $e^{-1}(k - \Delta')\theta\mu$. □

Now we will show that **whp** in a further $O(\log^2 n)$ rounds, ι will be reduced to two. Then the convergence analysis of the algorithm will be similar to that in [6].

Lemma 10 *If Δ of Lemma 8 satisfies $\Delta \leq 4\log n$, then* **whp** *after another $O(\log^2 n)$ rounds (or $O(\log^3 n/\log\log n)$ iterations), either the algorithm halts or ι drops to two and remains at two for the rest of the algorithm.*

Remark 5 *Notice that since $\Delta < k$, this Lemma also takes care of the case when the dimension k of the hypergraph is smaller than $4\log n$.*

Proof Since $\Delta \leq 4\log n$, the size of M will be at least $e^{-1}(k-4\log n)\theta\mu$. Since w.h.p (Lemma 7) the size of the lfmis is at most $e^{-1}(k+2\log n)\theta\mu$, the difference is $O(\mu \log n)$ **whp**. We can assume that $\mu \geq \frac{\log^2 n}{\log\log n}$, for otherwise the number of elements remaining to be added is $O(\frac{\log^3 n}{\log\log n})$, and they will be found in $O(\frac{\log^3 n}{\log\log n})$ further iterations.

We will now show that if ι exceeds two, then **whp**, in each round we can add at least $\lceil \mu/\log n \rceil$ elements into our independent set. This means the procedure will end in $O(\log^2 n)$ rounds.

Once again let a be the current size of M and $b = a/\mu$. We can assume that p_2 is smaller than $\sqrt{1/\log n}$ as otherwise, **whp** the current independent set is maximal, see (16).

We will break the analysis into two parts. Let

$$\lambda \;=\; \frac{1}{\sqrt{2\pi(k-2)}}\left[\frac{b}{k-2}\exp\left\{1-\Psi\left(\frac{k-2}{a}\right)\right\}\right]^{k-2}. \qquad (28)$$

Now

$$\begin{aligned}
p_i &\leq \binom{a}{k-i}p \\
&\leq \frac{C}{\sqrt{k-i}}\left[\frac{b}{k-i}\exp\left\{1-\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{k-i} p^{\frac{i-1}{k-1}} \\
&= \frac{C}{\sqrt{k-i}}\left[\frac{k-2}{k-i}\exp\left\{\Psi\left(\frac{k-2}{a}\right)-\Psi\left(\frac{k-i}{a}\right)\right\}\right]^{k-i}\left[\lambda\sqrt{k-2}\right]^{\frac{k-i}{k-2}} p^{\frac{i-1}{k-1}}.
\end{aligned}$$

So for $i \geq 2$,

$$\alpha i p_i^{-\frac{1}{i-1}} \;\geq\; \frac{i\mu}{2e} M_1 M_2 M_3 \lambda^{-\frac{k-i}{(k-2)(i-1)}},$$

where

$$M_1 \;=\; \left(\frac{k-i}{k-2}\right)^{\frac{k-i}{i-1}}$$

$$
\begin{aligned}
M_2 &= \exp\left\{\Psi\left(\frac{k-i}{a}\right) - \Psi\left(\frac{k-2}{a}\right)\right\}^{\frac{k-i}{i-1}} \\
M_3 &= \left[\frac{\sqrt{2\pi(k-i)}}{(\sqrt{2\pi(k-2)})^{\frac{k-i}{k-2}}}\right]^{\frac{1}{i-1}}
\end{aligned}
$$

We show next that M_1, M_2 and M_3 are bounded below by a constant. Now

$$
\begin{aligned}
M_3 &= \left[\frac{\sqrt{2\pi(k-i)}}{\left(\sqrt{2\pi(k-2)}\right)^{\frac{k-i}{k-2}}}\right]^{\frac{1}{i-1}} \\
&\geq \left[\frac{\sqrt{2\pi(k-i)}}{\sqrt{2\pi(k-2)}}\right]^{\frac{1}{i-1}} \\
&\geq M_1.
\end{aligned}
$$

So we only have to show that M_1 and M_2 are not too small.

But $M_1 \geq e^{-1}$ and $M_2 \geq 9/10$ can be proven as in (20), (21) by replacing i_o by 2.

So

$$
\alpha i p_i^{-\frac{1}{i-1}} \geq C\alpha i \mu \lambda^{-\frac{k-i}{(k-2)(i-1)}} \tag{29}
$$

Case 1: $\lambda \leq \log n$.

It follows from (29) that for $i \geq 2$,

$$
\alpha i p_i^{-\frac{1}{i-1}} \geq C\mu/\log n.
$$

Thus $m \geq C\mu/\log n$ and then **whp** at least $C\mu/(e\log n)$ elements are added to our independent set. By Lemma 7, we are only going to add $O(\mu \log n)$ more elements. Hence this can only happen for $O(\log^2 n)$ rounds.

Case 2. $\lambda > \log n$.

Using (15) and $p_2 = o(1)$ we see that

$$
p_2 \geq C\lambda p^{\frac{1}{k-1}}.
$$

and so

$$e^{-1}p_2^{-1} \leq C\lambda^{-1}\mu.$$

on the other hand, (29) implies that

$$\frac{i}{2e}p_i^{-\frac{1}{i-1}} \geq C\lambda^{-\frac{k-i}{(k-2)(i-1)}}\mu$$

Since $\lambda = \Omega(\log n)$, $\lambda^{-1} \ll \lambda^{-\frac{k-i}{(k-2)(i-1)}}$, for $i \geq 3$ and so

$$e^{-1}p_2^{-1} \leq \frac{i}{2e}p_i^{-\frac{1}{i-1}} \quad (\text{for} \;\; i \geq 3).$$

Hence $\iota = 2$. □

When ι drops to two, we adapt the analysis in [6]. It can be stated as the following lemma,

Lemma 11 *If $\iota = 2$, then* **whp** *the algorithm will stop in $O(\log n)$ more rounds.*

Proof Let m be the number of elements we pick. From Lemma 5 **whp** the algorithm finishes in $\log n$ more rounds or the number of elements placed in M is at least m/e. So if $x \in X$ then

$$\begin{aligned} Pr[x \notin N(M)] &\leq (1-p_2)^{\binom{m/e}{1}} \ldots (1-p_k)^{\binom{m/e}{k-1}} \\ &\leq \exp\left\{-p_2\binom{m/e}{1} \ldots - p_k\binom{m/e}{k-1}\right\} \\ &\leq \exp\left\{-p_2\binom{m/e}{1}\right\} \\ &= \exp\{-p_2\lfloor p_2^{-1}\rfloor/e\} \\ &\leq e^{-1/2e}. \end{aligned}$$

Thus the expected size of the remaining vertices X shrinks by a constant factor in each round. So the algorithm finishes **whp** in $O(\log n)$ rounds. □

This completes the proof of Theorem 1.

Acknowledgement: We thank Marek Karpinski for bringing this problem to our attention. We also thank the referees for their detailed reports. We especially thank the referee who gave corrections to two places of the proof.

References

[1] P.Beame and M.Luby,Parallel Search for Maximal Independence Given Minimal Independence, *Proceedings of the First Annual ACM/SIAM Symposium on Discrete Algorithms*, (1990) 212-218.

[2] B.Berger and J.Rompel. Simulating $(\log^c n)$-wise Independence in NC, *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, (1989) 2-7.

[3] B.Berger, J.Rompel and P.Shor, Efficient NC Algorithms for Set Cover with Applications to Learning and Geometry, *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, (1989) 54-59.

[4] Neil Calkin and Alan Frieze , Probabilistic Analysis of a Parallel Algorithm for Finding Maximal Independent Sets, *Random Stuctures and Algorithms 1*, (1990) 39-50.

[5] N.Calkin, A. Frieze and L.Kučera , On the Expected Performance of a Parallel Algorithm for Finding Maximal Independent Subset of Random Graph, *Random Structures and Algorithms 3*, (1992) 215-221.

[6] D. Coppersmith,P. Raghavan and M. Tompa, Parallel Graph Algorithms that are Efficient on Average, *Information and Computation 81*, (1989) 318-333.

[7] E.Dalhaus, M.Karpinski, P.Kelsen, An Efficient Parallel Algorithm for Computing a Maximal Independent Set in a Hypergraph of Dimension 3, *Information Processing Letters* 42 (1992), pp.309-313

[8] M.Goldberg and T.Spencer, A New Parallel Algorithm for the Maximal Independent Set Problem, *SIAM Journal on Computing 18*, (1989) 419-427.

[9] R. Karp and Avi Wigderson, A fast parallel algorithm for the maximal independent set problem, *Journal of the Association for Computing* 32:762-773, 1985

[10] P.Kelsen , On the Parallel Complexity of Computing a Maximal Independent Set in a Hypergraph, *Proceedings of the 24th Annual ACM Symposium on Theory of Computing*, (1992) 339-350.

[11] R.Motwani, J.Naor and M.Naor, The Probabilistic Method Yields Deterministic Parallel Algorithms, *Proceedings of the 30th Annual IEEE Symposium on Foundations of Computer Science*, (1989) 8-13.

[12] R.Motwani and P.Raghavan, Randomized Algorithms, Cambridge University Press, 1995