Probabilistic Analysis Of A Parallel Algorithm For Finding Maximal Independent Sets

Neil Calkin
and
Alan Frieze
Department of Mathematics
Carnegie Mellon University
Pittsburgh, PA

Abstract

We consider a natural parallel version of the classical greedy algorithm for finding a maximal independent set in a graph. This version was studied in Coppersmith, Raghavan and Tompa [CRT] and they conjecture there that its expected running time on random graphs *of arbitrary edge density* is $O(\log n)$. We prove that conjecture.

1 Introduction

In this note we consider the problem of finding the lexicographically first maximal independent set (LFMIS) in a random graph. Coppersmith, Raghavan and Tompa [CRT] describe a parallel version of the standard greedy algorithm for this problem:
Suppose we are given a graph $G = (V, E)$, $V = [n] = \{1, 2, \dots, n\}$. For $Z \subseteq V$ we let

$$\Gamma^+(Z) = \{x \notin Z \ : \ xz \in E \text{ for some } z < x,\, z \in Z\},$$

and

$$\Gamma^-(Z) = \{x \notin Z \ : \ xz \in E \text{ for some } z > x,\, z \in Z\}.$$

Note that we have implicitly oriented the edges from low to high.

algorithm PARALLEL GREEDY (G);
    begin
        GIS $\leftarrow \emptyset$;
        until G has no vertices **do**
            begin
                let $A = \{a : \Gamma^-(a) = \emptyset\}$;

GIS $\leftarrow$ GIS$\cup A$;
remove $A \cup \Gamma(A)$ from G
end
output GIS
end

It is easy to see (Lemma 2.1 of [CRT]) that GIS is the LFMIS. Cook [CO] showed that the problem of computing the LFMIS of a graph is complete for P and so is not in NC unless NC=P. PARALLEL–GREEDY can be implemented on a CRCW PRAM in $O(1)$ time per iteration if one processor is allocate to each edge of G.

Coppersmith, Raghavan and Tompa showed that if $T(n,p)$ denotes the *expected* number of iterations $\tau = \tau(G)$ when $G = G_{n,p}$ then $T(n,p) = O(\frac{(\log n)^2}{\log\log n})$. ($G_{n,p}$ is the random graph with vertex set $[n]$ where each edge occurs independently with probability $p = p(n)$.).

They conjecture that $T(n,p) = O(\log n)$ and it is the aim of this paper to prove it. We also prove a lower bound $T(n,p) = \Omega(\frac{\log n}{\log\log n})$ for a range of values of p. More precisely

Theorem 1
(a) $\frac{\alpha \log n}{4 \log\log n} \leq T(n,p)$ *for* $\frac{1}{n} \leq p \leq \frac{1}{n^\alpha}$ *where* $0 < \alpha \leq 1$ *is constant*
(b) $T(n,p) = O(\log n)$.
The hidden constant in (b) is independent of p.

Note that our inequalities are only claimed for n large.

2 Lower Bound Proof

Let $m = \lfloor p^{-1} \rfloor$ and consider the subgraph H of G induced by $[m]$. If H contains a component which is a *directed* path of length $l-1$ then clearly $\tau \geq \frac{1}{2}l$. (The direction of an edge ij, $i<j$ is from i to j). Now let

$$l = \lfloor \frac{2\log m}{3\log\log m} \rfloor$$

and Z_l = the number of components of H which are directed paths of length l.

We show

$$\Pr(Z_l \neq 0) \geq \frac{1}{1+m^{-1/4}} \tag{1}$$

and the lower bound follows.

Now

$$E(Z_l) = \binom{m}{l} p^{l-1} (1-p)^{l(m-l)+\binom{l}{2}-l+1}$$

But

$$\binom{m}{l} = \frac{m^l e^{l-l^2/2m+O(l^3/m^2)}}{\sqrt{2\pi l} l^l}$$

and

$$(1-p)^{l(m-l)+\binom{l}{2}-l+1} = e^{-l+l^2/2m+O(l/m)}$$

and so

$$E(Z_l) \geq m^{1/3} \tag{2}$$

Furthermore

$$\begin{aligned} E(Z_l(Z_l-1)) &= \binom{m}{l}\binom{m-l}{l} p^{2l-2}(1-p)^{2l(m-2l)+\binom{2l}{2}-2l+2} \qquad (3) \\ &= \frac{\binom{m-l}{l}}{\binom{m}{l}}(1-p)^{-l^2}E(Z_l)^2 \\ &\leq (1-\frac{l}{m})^l(1-p)^{-l^2}E(Z_l)^2 \\ &\leq E(Z_l)^2. \end{aligned}$$

We can now use the useful inequality (Schwarz)

$$\begin{aligned} \Pr(Z_l \neq 0) &\geq \frac{E(Z_l)^2}{E(Z_l^2)} \\ &\geq \frac{1}{1+E(Z_l)^{-1}} \qquad \text{by (3)} \\ &\geq \frac{1}{1+m^{-1/3}} \qquad \text{by (2)} \end{aligned}$$

and (1) follows. □

3 Upper Bound Proof

We shall, as in [CRT] consider a modified algorithm which *always* takes at least as long as PARALLEL-GREEDY. It is similar to the one described there.
In what follows we may take $r = 5$.

algorithm MODIFIED PARALLEL-GREEDY (G);
    begin
        $X := V$; $Y :=$GIS$:= \emptyset$;
        while $|X| \geq m$ **do**
L1:   **begin**
            $Z := m$ lowest numbered elements of X; $X := X \setminus Z$
            run PARALLEL-GREEDY for r iterations on $Y \cup Z$
            and let A be the independent set constructed;
            $X := X - (A \cup \Gamma^+(A))$; $Y := (Y \cup Z) - (A \cup \Gamma^+(A))$;
            GIS:=GIS$\cup A$
L2:   **end**
L3:   run PARALLEL-GREEDY to completion on $X \cup Y$
    end

We claim that MODIFIED PARALLEL-GREEDY (MPG) constructs the LFMIS of G and always requires at least as many iterations as PARALLEL-GREEDY. This is because the difference between the modified algorithm and the original algorithm is that it does not necessarily add vertices to GIS the first time they become sinks. Instead we may have to wait until they become members of Z first. This does not change what goes into GIS, but instead delays the time when subsequent members of GIS become sinks.

We will show first under the assumption

$$p \leq (\log n)^{-7}$$

that with sufficiently high probability

(4) there are at most $\lceil \log_{10/9} n \rceil$ executions of the main loop L1–L2, and

(5) at most $\frac{3}{2}\lceil \log_{10/9} n \rceil$ iterations of PARALLEL-GREEDY are needed to execute statement L3.

Now let $X_t, Y_t, Z_t \subseteq X_t$ be the values of X, Y, Z at the start of the tth execution of the main loop A-B of MPG. Let A_t denote the set A constructed during the tth iteration.

The key to an analysis of MPG is an understanding of the distribution of the edges contained in $\bigcup_{i=1}^{k} Z_i$ for $k = 1, 2, \ldots$

Suppose we condition on the values of $Z_1, Z_2 \ldots, Z_k$. We claim that

(6a) the edges contained in each Z_i are unconditioned, i. e. are chosen independently with probability p.

(6b) a possible edge uv, $u \in Z_i$, $v \in Z_j$, $i < j$ will occur independently with probability 0 ($u \in A_i$) or p ($u \notin A_i$).

Knowing this we first prove

Lemma 1 *If m is large and $t \leq m^{1/4}$ then*

$$Pr(|Y_t| \geq 2\left(\frac{e^2}{2r}\right)^r m) \leq \frac{200 r^4 t^3}{m}.$$

Proof

Let $v \in Y_t$ and suppose $v \in Z_a, a \leq t-1$. We claim there exists $b \leq a$ and a directed path with head v, of length greater than $2r(t-b)$, contained in $\bigcup_{i=b}^{a} Z_i$. Now either Z_a contains a directed path of length greater than $2(t-a)r$ with head v, in which case we take $b = a$, or not. In the latter case there is a vertex $v' \in Z_c \cap Y_a$, where $c < a$, and a directed path of length greater than $2r(t-a)$ from v' to v which is contained in $\bigcup_{i=c}^{a} Z_i$. For otherwise, at the start of iteration a the longest directed path with head v contained in $Y_a \cup Z_a$ is of length at most $2r(t-a)$. But then v will be eliminated by the end of iteration $t-1$, contradiction. Inductively, for some $d \leq c$, v' is the head of a directed path of length greater than $2r(a-d)$ which is contained in $\bigcup_{i=d}^{c} Z_i$. The result now follows with $b = d$. (The base case, $t = 2$, for this induction is of course identical to the case where $b = a$ above.)

Let ξ_b denote the number of paths of length *exactly* $2r(t-b)$ contained in $\bigcup_{i=b}^{t-1} Z_i$. It follows that

$$|Y_t| \leq \sum_{b=1}^{t-1} \xi_b.$$

Now it follows from (6) that ξ_b is dominated stochastically by the number of *directed* paths η_b of length $2r(t-b)$ in the random graph $G_{m(t-b),p}$.

We prove in an appendix that

$$\Pr(\eta_b \geq \left(\frac{e^2}{2r}\right)^{r(t-b)} m) \leq \frac{200r^4t^2}{m}, \qquad 1 \leq b \leq t-1. \tag{7}$$

Hence

$$\Pr(|Y_t| \geq m \sum_{k=1}^{t-1} \left(\frac{e^2}{2r}\right)^{kr}) \leq \frac{200r^4t^3}{m}$$

and the lemma follows. □

If $|Y_t|$ is small compared with m then we can show that $|X_{t+1}| \leq \frac{9}{10}|X_t|$ with high probability.

Lemma 2 *If* $|Y_t| < 2\left(\frac{e^2}{2r}\right)^r m$ *and* $|X_t| \geq m$ *then*

$$Pr(|X_{t+1}| > \frac{9}{10}|X_t|) < e^{-\alpha m}$$

for some constant $\alpha > 0$.

Proof

We can proceed somewhat similarly to Lemma 2.2 of [CRT]. Let $\beta = 2\left(\frac{e^2}{2r}\right)^r$ and assume $|Y_t| \leq \beta m$. Let

$$Z_t' = \{z \in Z_t : \nexists y < z, y \in Y_t \cup Z_t \text{ and } yz \in E\}$$

and

$$X_t' = \{x \in X_t - Z_t : \exists z \in Z_t' \text{ such that } zx \in E\}.$$

Then clearly $X_{t+1} \subseteq X_t - (Z_t' \cup X_t')$. Now if $z \in Z_t$ we have

$$\Pr(z \in Z_t') \geq (1-p)^{(1+\beta)m} \geq .236\ldots$$

and since the events $\{z \in Z_t'\}$ for $z \in Z_t$ are independent we have

$$\Pr(|Z_t'| < \frac{1}{5}m) \leq e^{-\alpha' m} \qquad \text{for some } \alpha' > 0.$$

(Here we use the fact that the tails of the binomial distribution are exponentially small.)

On the other hand, for $x \in X_t - Z_t$ we have

$$\Pr(x \notin X_t' \mid |Z_t'| \geq \frac{1}{5}m) \leq (1-p)^{\frac{1}{5}m} \leq .82$$

and since, given Z_t', the events $\{x \in X_t'\}$ for $x \in X_t - Z_t$ are independent

$$\Pr(|X_t'| \leq \frac{1}{10}|X_t| \mid |Z_t'| \geq \frac{1}{5}m) \leq e^{-\alpha'' m} \qquad \text{for some } \alpha'' > 0.$$

The result follows. $\square$

Now let $l = \lceil \log_{10/9} n \rceil \leq m^{1/4}$ (since $p \leq (\log n)^{-7}$). For there to be more than l executions of the main loop of MPG we must have

$$|Y_t| \geq 2\left(\frac{e^2}{2r}\right)^r m \text{ for some } t \leq l$$

or

$$|Y_t| < 2\left(\frac{e^2}{2r}\right)^r m \text{ and } |X_{t+1}| > \frac{9}{10}|X_t| \text{ for some } t \leq l.$$

But, by Lemmas 3.1 and 3.2 the probability of one of these events is at most

$$\frac{200r^4l^4}{m} + le^{-\alpha m} = O((\log n)^{-3}). \tag{8}$$

Suppose now that $\tau \leq l$ iterations of the main loop of MPG suffice. Let $\lambda = \lambda(\tau) =$ the length of the longest directed path in $\bigcup_{i=1}^{\tau} Z_i$. Clearly the execution of Statement L3 of MPG requires at most $\lambda/2$ iterations of PARALLEL-GREEDY. But

$$\begin{aligned}
\Pr(\exists \tau \leq l : \lambda(\tau) \geq 3l) &\leq \sum_{t=1}^{l} \binom{mt}{3l} p^{3l-1} \\
&\leq m \sum_{t=1}^{l} \left(\frac{et}{3l}\right)^{3l} \\
&\leq mlm^{-2.5} \\
&\leq m^{-1}.
\end{aligned}$$

Together with (8) we have

Pr(MPG requires more than $(r+\frac{3}{2})\log_{10/9} n$ iterations of PARALLEL-GREEDY) $=O((\log n)^{-3})$.
Hence
(9)Pr(PARALLEL-GREEDY requires more than $(r+\frac{3}{2})\log_{10/9} n$ iterations) $=O((\log n)^{-3})$.

Now it follows easily from calculations done in [CRT] that
(10) Pr(PARALLEL-GREEDY requires more than $\frac{c(\log n)^2}{\log\log n}$ iterations)$=O(1/n)$, for some absolute constant $c > 0$.

It follows immediately from (9) and (10) that we can write

(11) $\tau(n,p) \leq c \log n$ for some absolute constant $c > 0$, provided $p \leq (\log n)^{-7}$.

We now consider $p \geq (\log n)^{-7}$. For this we introduce another modification of PARALLEL-GREEDY which we call MPG2. Let $m' = \lceil p^{-2} \rceil$.

Algorithm MPG2

begin $X := [n]$; GIS:= $\emptyset$;

while $X \neq \emptyset$ **do**

L1: **begin**

$Z :=$ min$\{m', |X|\}$ lowest numbered elements of X;
run PARALLEL-GREEDY to completion on Z and
let A be the independent set constructed;
$X := X - (A \cup \Gamma^+(A))$;
GIS:=GIS $\cup A$

L2: **end**

end

One can see as for MPG that MPG2 constructs the LFMIS of G and requires at least as many iterations as PARALLEL-GREEDY.

We will assume from now on that $p \leq (2 \log 1/p)^{-7}$. This is true for p sufficiently small, say $p \leq p_0$. For $p > p_0$ we know that with probability $1 - o(1/n)$ the size of GIS is less than say $2 \log_b n$ ($b = \frac{1}{1-p}$) and since PARALLEL-GREEDY finds at least one new member of GIS at each stage, it requires $O(\log n)$ iterations when $p > p_0$.

Let σ denote the number of iterations of the main loop L1–L2 of MPG2. Define X_t, Y_t, Z_t in analogy to MPG. Next let $\zeta_t, t = 1, 2, \ldots, \sigma$ denote the number of iterations of PARALLEL-GREEDY required for the tth execution of L1–L2, and let $\zeta_t = 0$ for $t \geq \sigma + 1$.

Then

$$T(n,p) \leq E\left(\sum_{t=1}^{n} \zeta_t\right). \tag{12}$$

Note by (11)

$$E(\zeta_t) \leq c \log 1/p, \quad t = 1, 2, \ldots n. \tag{13}$$

(The constant here is (roughly) twice that in (11). Also $E(\zeta_\sigma) = E_{m''}(T(m'', p))$ where $m'' \leq m'$ is a random variable. Clearly $T(m'', p) \leq T(m', p)$ always.)

Now let $\sigma_0 = \frac{3 \log n}{\log 1/p} + 1$; we will show that

$$\Pr(\sigma > \sigma_0) \leq \frac{1}{n}. \tag{14}$$

Hence

$$\begin{aligned} E(\sum_{t=1}^{n} \zeta_t) &= E(\sum_{t=1}^{\sigma_0} \zeta_t) + E(\sum_{t=\sigma_0+1}^{n} \zeta_t) \\ &\leq \sum_{t=1}^{\sigma_0} E(\zeta_t) + n\Pr(\sigma > \sigma_0) \\ &\leq \sigma_0 c \log 1/p \quad + 1 \\ &= O(\log n) \end{aligned}$$

We will therefore have finished the proof of our theorem once we have completed the
Proof of (14)

We will show in the appendix that if $t \le \sigma,\ b = \frac{1}{1-p}, \xi \ge 2,\ \alpha = \log_b(m'p+1) - 2\log_b\log_e(m'p+1)$, then

$$\Pr(|A_t| = \alpha - \log_b \xi) \le \begin{cases} e^{-\xi/p} & \xi \le \frac{m'p+1}{4(\log_e(m'p+1))^2} \\ e^{-1/4p^2} & \text{otherwise.} \end{cases} \tag{15}$$

Thus if $x \in X_t - Z_t$ and $\xi_0 = \frac{m'p+1}{4(\log_e(m'p+1))^2}$ then

$$\begin{aligned} \Pr(x \in X_{t+1}) &= E((1-p)^{|A_t|}) \\ &\le (1-p)^{\alpha - \log_b 2} + \sum_{k=\lceil \alpha - \log_b \xi_0 \rceil}^{\lfloor \alpha - \log_b 2 \rfloor} e^{-\frac{b^{\alpha-k}}{p}} + \alpha e^{-\frac{1}{4p^2}} \\ &\le \frac{1}{\xi_0} + \int_2^\infty e^{-\frac{\xi}{p}} \frac{\mathrm{d}\xi}{\xi \log_e b} + \frac{1}{p^2} e^{-\frac{1}{4p^2}} \\ &\le \frac{1}{\xi_0} + \int_2^\infty e^{-\frac{\xi}{2p}}\, \mathrm{d}\xi + \frac{1}{p^2} e^{-\frac{1}{4p^2}} \\ &= \frac{1}{\xi_0} + 2pe^{-\frac{1}{p}} + \frac{1}{p^2} e^{-\frac{1}{4p^2}} \\ &\le p^{2/3}. \end{aligned}$$

Hence if $|X_t| > m'$ then

$$E(|X_{t+1}| \mid |X_t|) \le p^{2/3}|X_t|. \tag{16}$$

So if we define $x_t = |X_t|$ for $t \le \sigma$ and $x_t = 0$ otherwise, then

$$E(x_t) \le np^{\frac{2(t-1)}{3}}. \tag{17}$$

This is true for $t = 1$ and assuming that it is true for t we have

$$\begin{aligned} E(x_{t+1}) &= \sum_{k=0}^{n} E(x_{t+1} \mid x_t = k)\Pr(x_t = k) \\ &\le \sum_{k=0}^{n} p^{2/3} k \Pr(x_t = k) \quad \text{by (16)} \\ &= p^{2/3} x_t \end{aligned}$$

and (17) follows. Note that the argument here works regardless of $t \le \sigma$ or $t > \sigma$.

It follows from (17) that

$$E(x_{\sigma_0}) \le \frac{1}{n}$$

and so

$$\Pr(x_{\sigma_0} \ge 1) \le \frac{1}{n}$$

and (14) follows.

We have thus verified the conjecture of [CRT]. It would be interesting to see if the same result can be proved for random regular graphs.

We are grateful to the refereees for their helpful comments, in particular for the simplification of the calculations in the proof of (1).

References

[CO] S. A. Cook, 'A taxonomy of problems with fast algorithms', Information and Control 64(1985) 2-22.

[CRT] D. Coppersmith, P. Raghavan and M. Tompa, 'Parallel graph algorithms that are efficient on average', Proceedings of 28'th Annual IEEE Symposium on Foundations of Computer Science (1987) 260-269.

APPENDIX

Proof of (7)

We first note that

$$E(\eta_b) = \binom{m(t-b)}{2r(t-b)+1} p^{2r(t-b)}$$

and so, since $t^2 = o(m)$ in our range of interest,

$$\frac{m}{4\sqrt{rt}} (\frac{1}{r})^{2r(t-b)} \leq E(\eta_b) \leq m(\frac{e}{2r})^{2r(t-b)}. \tag{1}$$

Furthermore, where $s_0 = 2r(t-b)$ and $m_0 = m(t-b)$, we have

$$E(\eta_b(\eta_b - 1)) \leq E(\eta_b) \left(E(\eta_b) + \sum_{s=1}^{s_0-1} \sum_{l=1}^{s} \binom{s_0}{2l-1} \binom{m_0}{s_0 - s - l} p^{s_0 - s} \right). \tag{2}$$

In the summation s denotes the number of edges that a general path of length s_0 has in common with a given fixed such path, l denotes the number of subpaths made up by these common edges.

Now let $u_l = \binom{s_0}{2l-1}\binom{m_0}{s_0-s-l}$. Then

$$\frac{u_{l+1}}{u_l} = \frac{(s_0-2l)(s_0-2l+1)(s_0-s-l)}{2l(2l+1)(m_0-s_0+s+l+1)} \leq \frac{s_0^3}{m_0}.$$

But since $t^3 = o(m)$ we have

$$\begin{aligned} E(\eta_b(\eta_b-1)) &\leq E(\eta_b) \left(E(\eta_b) + 2\sum_{s=1}^{s_0-1} s_0 \binom{m_0}{s_0-s} p^{s_0-s} \right) \\ &= E(\eta_b)^2 (1 + 2s_0 \sum_{s=1}^{s_0-1} v_s) \end{aligned}$$

where

$$v_s = \frac{\binom{m_0}{s_0-s}}{\binom{m_0}{s_0+1}} p^{-s}.$$

Observe now that

$$\frac{v_{s+1}}{v_s} = \frac{s_0 - s}{(m_0 - s_0 + s + 1)p} \leq \frac{s_0}{m_0 p} \leq 2r.$$

Also $v_1 \leq \frac{5r^2}{m}$. We deduce then that

$$E(\eta_b(\eta_b-1)) \leq E(\eta_b)^2 (1 + \frac{10r^2 s_0^2}{m}(2r)^{s_0})$$

and hence

$$\begin{aligned}
\frac{Var(\eta_b)}{E(\eta_b)^2} &\leq \frac{10r^2 s_0^2}{m}(2r)^{s_0} + \frac{1}{E(\eta_b)} \\
&\leq \frac{10r^2 s_0^2}{m}(2r)^{s_0} + \frac{4\sqrt{rt}}{m} r^{s_0} \\
&\leq \frac{11r^2 s_0^2}{m}(2r)^{s_0}
\end{aligned}$$

Now the Chebycheff inequality yields

$$\begin{aligned}
\Pr(\eta_b \geq \alpha m) &\leq \frac{Var(\eta_b)}{(\alpha m - E(\eta_b))^2} && \text{if } \alpha m \geq E(\eta_b) \\
&\leq \frac{44r^2 s_0^2}{\alpha^2 m}(\frac{e^2}{2r})^{s_0} && \text{if } \alpha \geq 2(\frac{e}{2r})^{s_0}.
\end{aligned}$$

Putting $\alpha = (\frac{e^2}{2r})^{s_0/2}$ yields (7). □

Proof of (15)

Now A_t is the set of vertices picked by the ordinary sequential greedy algorithm for finding a maximal independent set. Assume without loss of generality that $t = 1$ and $S = \{s_1 = 1, s_2, \ldots, s_k\} \subseteq [m']$. Then, where $s_{k+1} = m' + 1$,

$$\begin{aligned}
\Pr(A_t = S) &= (1-p)^{\binom{k}{2}} \prod_{i=1}^{k} (1 - (1-p)^i)^{s_{i+1} - s_i - 1} \\
&\leq (1-p)^{\binom{k}{2}} \prod_{i=1}^{k} (1 - (1-p)^k)^{s_{i+1} - s_i - 1} \\
&= (1-p)^{\binom{k}{2}} (1 - (1-p)^k)^{m'-k}.
\end{aligned}$$

Hence

$$\Pr(|A_t| = k) \leq \binom{m'-1}{k-1} (1-p)^{\binom{k}{2}} (1 - (1-p)^k)^{m'-k}.$$

Let now $\lambda = m'p + 1$ and $k = \log_b \lambda - 2\log_b \log_e \lambda - \log_b \xi, \xi \geq 2$. Assume first that $\xi \leq \frac{\lambda}{4(\log_e \lambda)^2}$. We then have

$$
\begin{aligned}
\Pr(|A_t| = k) &\leq \left(\frac{m'e}{\log_b \lambda}\right)^{\log_b \lambda} \left(\frac{(\log_e \lambda)^2 \xi}{\lambda}\right)^{\frac{k-1}{2}} \left(1 - \frac{(\log_e \lambda)^2 \xi}{\lambda}\right)^{m'-k} \\
&\leq \left(\frac{m'e}{\log_b \lambda}\right)^{\log_b \lambda} \left(1 - \frac{(\log_e \lambda)^2 \xi}{\lambda}\right)^{m'} \\
&\leq \left(\frac{3\lambda}{\log_e \lambda}\right)^{\log_b \lambda} exp\left\{-\xi m' \frac{(\log_e \lambda)^2}{\lambda}\right\} \\
&\leq \exp\left\{\frac{\log_e \lambda}{p}(\log_e \lambda + \log_e 3 - \log_e \log_e \lambda) - \xi m' \frac{(\log_e \lambda)^2}{\lambda}\right\} \\
&\leq \exp\left\{\frac{(\log_e \lambda)^2}{p}\left(1 - \frac{\xi}{2p}\right)\right\} \\
&\leq e^{-\frac{\xi}{p}} \qquad \text{as } p \text{ is small.}
\end{aligned}
$$

Finally, suppose $\xi \geq \frac{\lambda}{4(\log_e \lambda)^2}$, so that $k \leq \log_b 4$. Then

$$
\begin{aligned}
\Pr(|A_t| = k) &\leq \left(\frac{m'e}{\log_b \lambda}\right)^{\log_b \lambda} (\frac{3}{4})^{(m'-k)} \\
&\leq \left(\frac{6}{p \log_e \lambda}\right)^{\log_b \lambda} (\frac{3}{4})^{m'} \\
&\leq \exp\left\{\frac{1}{p}\log_e \lambda(\log_e 6 + \log_e \frac{1}{p}) - \frac{1}{p^2}\log_e \frac{4}{3}\right\} \\
&\leq e^{-\frac{1}{4p^2}}.
\end{aligned}
$$