

# PRIME CHAINS AND PRATT TREES

KEVIN FORD, SERGEI V. KONYAGIN AND FLORIAN LUCA

ABSTRACT. We study the distribution of prime chains, which are sequences  $p_1, \dots, p_k$  of primes for which  $p_{j+1} \equiv 1 \pmod{p_j}$  for each  $j$ . We first give conditional upper bounds on the length of Cunningham chains, chains with  $p_{j+1} = 2p_j + 1$  for each  $j$ . We give estimates for  $P(x)$ , the number of chains with  $p_k \leq x$  ( $k$  variable), and  $P(x; p)$ , the number of chains with  $p_1 = p$  and  $p_k \leq px$ . The majority of the paper concerns the distribution of  $H(p)$ , the length of the longest chain with  $p_k = p$ , which is also the height of the Pratt tree for  $p$ . We show  $H(p) \geq c \log \log p$  and  $H(p) \leq (\log p)^{1-c'}$  for almost all  $p$ , with  $c, c'$  explicit positive constants. We can take, for any  $\varepsilon > 0$ ,  $c = e - \varepsilon$  assuming the Elliott-Halberstam conjecture. A stochastic model of the Pratt tree is introduced and analyzed. The model suggests that for most  $p \leq x$ ,  $H(p)$  stays very close to  $e \log \log x$ .

## 1. INTRODUCTION

Impose on the set of positive integers a relation as follows:  $a \prec b$  if there is a positive integer  $m$  with  $b = am + 1$ . We are interested in properties of the chains with respect to this partial ordering, especially the chains consisting only of primes, the *prime chains*. A simple example is 3, 7, 29, 59. In a chain  $n_1 \prec n_2 \prec \dots \prec n_k$  of length  $k$ , there are positive integers  $m_1, \dots, m_{k-1}$  with  $n_{i+1} = m_i n_i + 1$  (which we refer to as *links*) for  $1 \leq i \leq k-1$ .

**1.1. Cunningham Chains.** When  $k \geq 2$  is fixed, the links  $m_1, \dots, m_{k-1}$  are fixed and there are no obvious congruential obstructions, it is unknown whether there are infinitely many such prime chains, an affirmative answer being implied by Dickson's prime  $k$ -tuples conjecture. In the special case where  $m_1 = \dots = m_{k-1} = 2$  there are no congruential obstructions and the chains in question are known as Cunningham Chains (these are sometimes called Cunningham Chains of the first kind, whereas a sequence  $p_1, p_2 = 2p_1 - 1, \dots, p_k = 2p_{k-1} - 1$  is called a Cunningham Chain of the second kind). A basic example is 2, 5, 11, 23, 47, while the longest one known is the chain with  $p_1 = 810433818265726529159$  and  $k = 16$ , discovered by Carmody and Jobling in 2002

---

*Date:* August 3, 2009.

*2000 Mathematics Subject Classification.* Primary 11N05, 11N36; Secondary 60J80.

The research of K. F. was supported in part by National Science Foundation grants DMS-0555367 and DMS-0901339, that of S. K. was supported in part by Grants 08-01-00208 from the Russian Foundation for Basic Research and NSH-3233.2008.1 from the Program Supporting Leading Scientific Schools, and that of F. L. was supported in part by projects PAPIIT 100508 and SEP-CONACyT 79685.

(see [4]). In such a chain  $p_1, \dots, p_k$  we have

$$(1.1) \quad p_j = 2^{j-1}p_1 + 2^{j-1} - 1 = 2^{j-1}(p_1 + 1) - 1.$$

Let  $k(p)$  be the length of the longest Cunningham Chain starting from  $p$ . By a standard upper bound for Sophie Germain primes coming from sieve theory ([30], Theorem 2.4),  $k(p) = 1$  for all primes  $p \leq x$  except for  $O(x/\log^2 x)$  of them. How large can  $k(p)$  be in terms of  $p$ ? By the first equality in (1.1) and Fermat's little theorem, we have  $k(p) \leq \text{ord}_p 2$  and in particular  $k(p) \leq p - 1$  for all  $p$ . The distribution of  $\text{ord}_p 2$  is largely unknown, and it is conjectured (Artin's Conjecture) that  $\text{ord}_p 2 = p - 1$  for infinitely many  $p$ ; that is, 2 is a primitive root of infinitely many primes. We quote a well-known result of Hooley [33].

**Theorem H.** (Hooley) *Assume the Riemann Hypothesis for the Dedekind zeta functions  $\zeta_{K_r}(s)$  attached to the number fields  $K_r = \mathbb{Q}(2^{1/r}, e^{2\pi i/r})$ , where  $r$  runs over the primes. Then  $A(x)$ , the number of primes  $p \leq x$  for which 2 is a primitive root, satisfies*

$$A(x) \sim C\pi(x), \quad C = \prod_r \left(1 - \frac{1}{r(r-1)}\right) = 0.3739\dots$$

where  $\pi(x)$  is the number of primes  $\leq x$ .

Using the second equality in (1.1) together with Hooley's result, we greatly improve (conditionally) the bound on  $k(p)$ .

**Theorem 1.** *Assume the conclusion of Theorem H, that is,  $A(x) \sim C\pi(x)$ . Then*

$$\limsup_{p \rightarrow \infty} \frac{k(p)}{\log p} \leq \frac{1}{C}.$$

We next show, unconditionally, that substantially larger values of  $k(p)$  are very rare. To state the results, let  $\theta^+$  be the supremum of real numbers  $\theta$  so that there are  $\gg x^{1+o(1)}$  primes  $p \leq x$  such that  $p - 1$  has a prime factor  $> x^\theta$ . The best known result is due to Baker and Harman, who showed in [6] that  $\theta^+ \geq 0.677$ . It is conjectured that  $\theta^+ = 1$ .

**Theorem 2.** *For each fixed  $A > 1/\theta^+$ , there are  $O(x^{1-\frac{\theta^+}{2} + \frac{1}{2A} + o(1)})$  primes  $p \leq x$  with  $k(p) > (\log p)^A$ .*

**1.2. General prime chains.** We now turn to problems about counting prime chains with variable links, which will be the main focus of this paper. Dirichlet's theorem on primes in arithmetic progressions implies that there are infinitely long prime chains. We consider problems where the links are allowed to vary, but additional constraints are put on the chains, e.g. given beginning  $p_1$ , given end  $p_k$ ,  $p_k/p_1 \leq x$ , etc. The study of such restricted types of prime chains has arisen, for example, in investigations of iterates of Euler's totient function  $\phi(n)$  and Carmichael's function  $\lambda(n)$  (e.g. [7], [8], [9], [22], [38], [39], [40]), the value distribution of  $\lambda(n)$  [25], common values of  $\phi(n)$  and the sum-of-divisors function  $\sigma(n)$  [26], and the complexity of the Pratt primality certificate ([10], [44]).

Some basic counting functions of general prime chains are  $P_k(x)$ , the number of prime chains with fixed length  $k$  and  $p_k \leq x$ , and  $P(x)$ , the total number of prime chains with  $p_k \leq x$ , so that  $P(x) = \sum_{k \geq 1} P_k(x)$ . For fixed  $k$  we have

$$(1.2) \quad P_k(x) \sim \frac{x(\log_2 x)^{k-1}}{(k-1)! \log x}.$$

When  $k = 1$ , this is a restatement of the Prime Number Theorem. The asymptotic (1.2) for  $k = 2$  is implicit in the work of Erdős [21], for  $k = 3$  it follows from Erdős and Pomerance [24], and the general case was shown by Bassily, Kátai and Wijsmuller [9], where (1.2) is proved uniformly for  $k \ll \frac{\log_3 x}{\log_4 x}$ . Here and throughout,  $\log_k x$  is the  $k$ -th iterate of the natural logarithm. The idea behind (1.2) is that for most primes  $p$ ,  $p - 1$  has about  $\log_2 p$  prime factors, uniformly distributed on a log log-scale. A related result is proved in [16]. It is straightforward to prove by induction the uniform upper bound

$$P_k(x) \ll \frac{x(\log_2 x)^{k-1}}{\log x} \quad (x \geq 3, k \geq 1)$$

using the methods of [22] (Theorem 3.5 therein, for example), but this is only non-trivial for  $k \ll \frac{\log_2 x}{\log_3 x}$  in light of Theorem 3 below.

If the asymptotic (1.2) is true uniformly for  $k \leq c \log_2 x$  with  $c > 1$ , this suggests that most prime chains with  $p_k \leq x$  have length about  $\log_2 x$ , and that  $P(x) \approx x$ . An upper bound of this order is easy to prove, but lower bounds are more difficult. The asymptotic (1.2) provides a lower bound for  $P(x)$ , but we can do better by considering all  $k$  at once.

**Theorem 3.** *We have  $x(\log x)^{-0.36} \ll P(x) \leq (\frac{2}{\log 2} + o(1))x$  as  $x \rightarrow \infty$ .*

Note that  $p \prec q$  means  $q \equiv 1 \pmod{p}$ . Thus, our study of prime chains is intimately connected with the distribution of primes in arithmetic progressions, fundamental results about which can be found in the monograph of Davenport [18]. Let  $\pi(x; q, a)$  be the number of primes  $p \leq x$  with  $p \equiv a \pmod{q}$ , and let  $\text{li}(x) = \int_2^x dt / \log t$ . For the lower bound in Theorem 3, we require that  $\pi(x; p, 1) \sim \frac{\text{li}(x)}{p-1}$  for most  $p \leq x^\beta$ , where  $\beta > 0$  is fixed. Consider the statement

$$(1.3) \quad \sum_{m \leq Q} \max_{y \leq x} \left| \pi(y; m, 1) - \frac{\text{li}(y)}{\phi(m)} \right| \ll \frac{x}{R}.$$

The Bombieri-Vinogradov Theorem ([18], Ch. 28) implies that for every  $A > 0$  there is a  $B > 0$  so that (1.3) holds with  $Q = x^{1/2}(\log x)^{-B}$  and  $R = (\log x)^A$ . If (1.3) holds with  $Q = x^\theta$  for some  $\theta > \frac{1}{2}$  and suitable  $R$ , then we can prove a stronger lower bound for  $P(x)$ . Assuming the Elliott-Halberstam Conjecture, for any fixed  $\varepsilon > 0$ , (1.3) holds with  $Q = x^{1-\varepsilon}$  and  $R = (\log x)^2$ , and we can deduce  $P(x) \gg x/(\log x)^{o(1)}$  (see the proof of Theorem 3 in Section 3). With a little stronger hypothesis (but still one which is widely believed), we can prove an asymptotic for  $P(x)$ . Moreover, we deduce the normal order

of the function  $f(p)$ , the number of prime chains with a given end  $p_k = p$  (again, with  $k$  variable).

**Theorem 4.** *Assume that (1.3) holds with  $Q = x^{1-(\log_2 x)^{-1-\delta}}$ ,  $R = (\log x)^2$  for some fixed  $\delta > 0$ . Then  $P(x) \sim cx$  for some constant  $c > 0$ . Moreover,  $f(p)$  has normal order  $c \log p$ .*

We do not, however, have any guess as to the size of  $c$ , other than the upper bound provided by Theorem 3. Theorems 3 and 4 will be proved in section 3.

**1.3. Chains with a given starting prime.** Our next objective is to count prime chains with a given *starting* prime  $p_1$ . As the number of such chains is infinite, we consider  $P(x; p)$ , the number of prime chains with  $p_1 = p$  and  $p_k/p \leq x$ . When  $p$  is very small compared with  $x$ , almost all primes  $q \leq x$  “lie above”  $p$  in some chain. The proof of Theorem 4.5 of [22] implies that if  $p \leq (\log x)^c$  for some small absolute constant  $c > 0$ , then  $P(x; p) \gg \pi(px) \sim px/\log x$ . Also, by the argument leading to Theorem 4, if  $p$  is fixed, then we expect  $P(x; p) \asymp x$ . For substantially larger  $p$  we expect that fewer primes  $q \leq px$  will lie above  $p$ , especially if  $p$  is a fixed power of  $x$ .

**Theorem 5.** *For  $2 \leq p \leq x$ , we have the effective estimate*

$$P(x; p) \ll x \exp \left\{ \frac{\log x (\log_3 x + O(1))}{\log_2 x} \right\} \quad (x \rightarrow \infty).$$

*In particular, for every  $\varepsilon > 0$  there is an effective constant  $C(\varepsilon)$  so that*

$$P(x; p) \leq C(\varepsilon)x^{1+\varepsilon}.$$

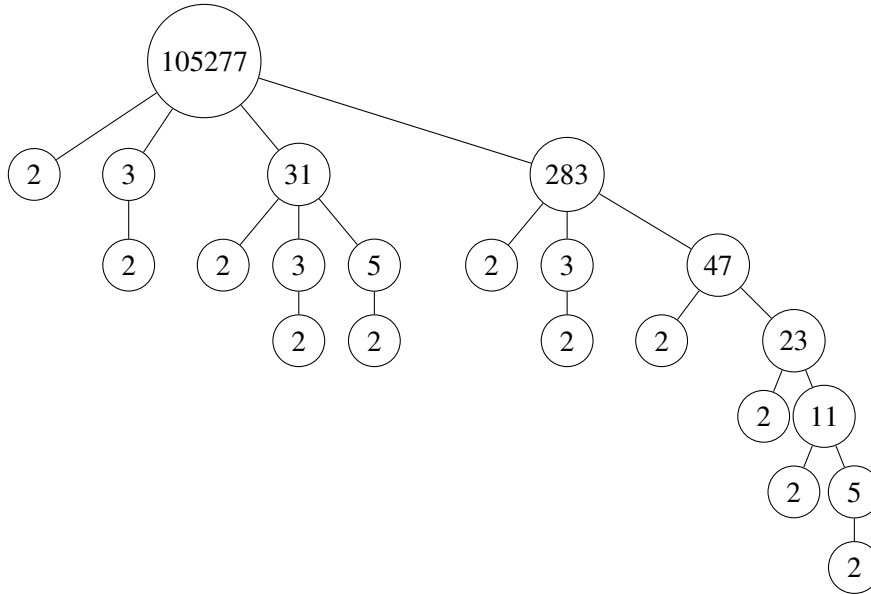
Theorem 5 will be proved in Section 4 using a novel sieve method based on matrices of Dirichlet series. It is an important tool in the recent proof by Ford, Luca and Pomerance [26] that the equation  $\phi(a) = \sigma(b)$  has infinitely many solutions, settling a well-known 50-year old problem of Erdős. In [25], Theorem 5 is applied to the problem of determining if for every positive integer  $n$ , there is another positive integer  $m$  with  $\lambda(n) = \lambda(m)$ .

**Remarks.** Theorem 5 is only nontrivial for large  $p$ , namely for

$$p \geq \exp \left\{ \frac{\log x (\log_3 x + O(1))}{\log_2 x} \right\}.$$

On the other hand, considering only chains of length 2 gives the lower bound  $P(x; p) \geq \pi(xp, p, 1)$  and we therefore expect that  $P(x; p) \gg x/\log x$  if  $x \geq p^\delta$  for fixed  $\delta > 0$ . We conjecture an even stronger upper bound.

**Conjecture 1.** *We have  $P(x; p) \ll x$ .*

FIGURE 1. Pratt tree for  $p = 105277$ 

1.4. **Pratt trees.** The *Pratt tree* for a prime  $p$  is the tree with root node  $p$ , and below  $p$  are links to the Pratt trees for primes  $q$  which divide  $p - 1$ . This tree was introduced by V. Pratt in 1975 [44], who showed how to use it in conjunction with Lucas' primality test ([17], §4.1) to create an efficient certificate of primality for a given odd prime  $p$ . The certificate consists of a primitive root  $g$  of  $p$ , which is proved to be such by verifying  $g^{(p-1)/q} \not\equiv 1 \pmod{p}$  for each prime  $q|(p-1)$ . To prove each of these  $q$  is prime, one iterates this procedure. Pratt showed that this primality certificate provides a *proof* that  $p$  is prime in polynomial time (the number of "bit operations" is  $O((\log p)^C)$  for some constant  $C$ ). For a general number  $n$ , this method is not a practical for *determining* if  $n$  is prime, since it requires factoring all the shifted primes in the tree. Today, there are polynomial time algorithms for determining whether or not a given integer is prime [2]. Pomerance [43] gave another method for producing primality certificates with a smaller upper bound on the complexity. It is likely that the Pratt certificate is more complex for some primes, but it is an open problem whether the Pratt certificate has longer complexity for most primes (see §1 of [43]).

All the leaves of the Pratt tree will be labeled with the prime 2. Figure 1 shows one example.

One measure of the complexity of the Pratt tree is the total number of nodes, which is the quantity  $f(p)$  introduced in §1.2. Another important statistic is the height of the tree, denoted  $H(p)$ . We can also interpret  $H(p)$  as the length of the longest prime chain with  $p_k = p$ . For example,  $H(105277) = 7$ . In the next graphs, we show histograms of  $H(p)$  for (i) all primes  $p \leq 10^9$  and (ii) for 1000 randomly chosen primes near  $10^{40}$ .

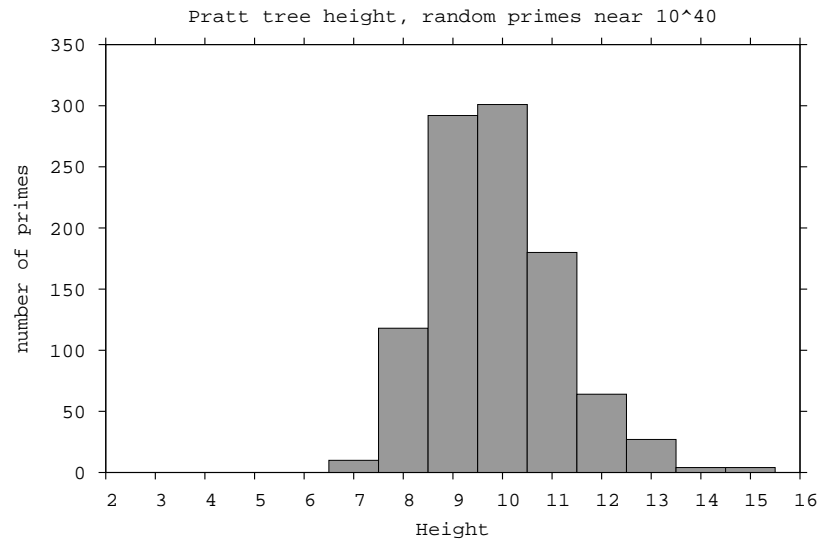
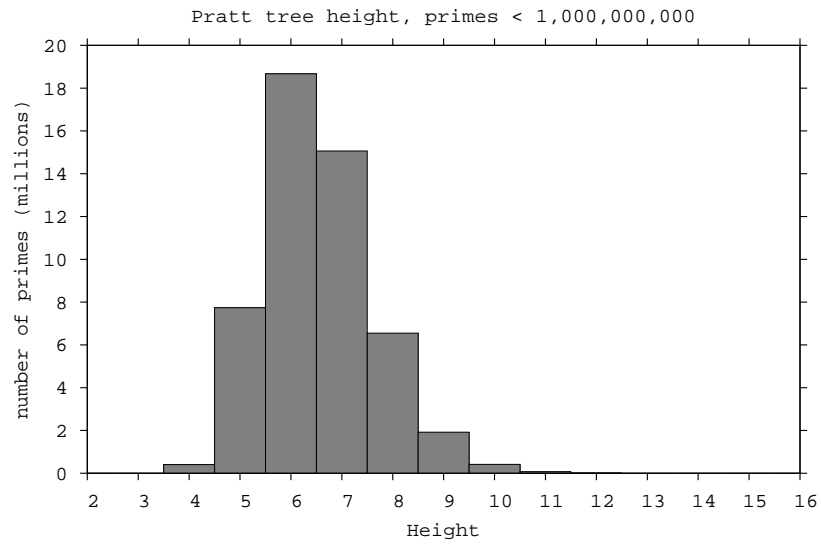


FIGURE 2. Pratt tree height

Understanding the extreme values of  $H(p)$  is notoriously difficult. At one extreme, a prime with  $H(p) = 2$  is a Fermat prime, that is,  $p = 2^{2^m} + 1$  for some integer  $m$ . Only 5 such primes are known, corresponding to  $m \in \{0, 1, 2, 3, 4\}$ , and it is widely believed that there are only finitely many (see [17], §1.3.2). On the other hand, the following conjecture seems plausible.

**Conjecture 2.** *For every  $k \geq 3$ , there are infinitely many primes with  $H(p) = k$ .*

If, for example, there are infinitely many primes of the form  $2^m 3^n 5^r 17^s 257^t 65537^u + 1$ , then Conjecture 2 holds for  $k = 3$ .

At the other extreme, we have the trivial bound  $H(p) \leq \lfloor \frac{\log p}{\log 2} \rfloor + 1$ , and pose the following question.

**Problem 1.** *Is it true that  $H(p) \gg \log p$  for infinitely many  $p$ ?*

The answer is yes if (2.2) holds, but (2.2) seems very difficult to settle. It is plausible that there is some  $C > 0$  so that for every prime  $p$ , the least prime  $q \equiv 1 \pmod{p}$  satisfies  $q \ll p(\log p)^C$ . If so, then the special chains  $p_1 = 2 \prec 3 \prec \cdots \prec p_k$  where for each  $j$ ,  $p_{j+1}$  is the least prime  $\equiv 1 \pmod{p_j}$ , have the property that  $H(p_k) \gg \frac{\log p_k}{\log_2 p_k}$ .

Primes with  $H(p)$  very small or very large should be rare. One main goal in this paper is to understand the behavior of  $H(p)$  for a typical prime  $p$ . A seemingly natural candidate for the longest chain in the Pratt tree is the special chain  $p = p_0 \succ p_1 \succ \cdots \succ p_k = 2$ , with  $p_j$  the largest prime factor of  $p_{j-1} - 1$  for each  $j$ . Let  $L(p)$  denote the length of this chain. Recently, Banks and Shparlinski [8] proved that for almost all  $p$ ,

$$(1.4) \quad L(p) \geq (1 - o(1)) \frac{\log_2 p}{\log_3 p} \quad (p \rightarrow \infty)$$

We can do better, at least heuristically. For a randomly chosen integer  $n$ , the largest prime factor of  $n$  is  $\leq n^{1/u}$  with probability  $\rho(u)$ , where  $\rho$  is the *Dickman function*, the unique continuous solution of the differential-delay equations (see Section 1 of [32])

$$\rho(u) = 1 \quad (0 \leq u \leq 1), \quad u\rho'(u) = -\rho(u-1) \quad (u > 1).$$

It is natural to conjecture that the largest prime factor of the shifted primes  $p-1$  has the same distribution. We can thus model the special chain by assuming that  $\log_2 p_j - \log_2 p_{j+1}$  are independent random variables with distribution function  $1 - \rho(e^x)$  for  $x \geq 0$ . These random variables have mean

$$\mu = \int_0^\infty x \frac{d}{dx} (1 - \rho(e^x)) dx = \int_1^\infty \frac{\rho(u)}{u} du.$$

By the law of large numbers, if  $k$  is large then we expect  $\log_2 p - \log_2 p_k \approx k\mu$  most of the time. Hence, we should have  $L(p) \sim (1/\mu) \log_2 p$  for most  $p$ . Numerically,  $1/\mu = 1.916045\dots$

Better unconditional lower bounds can be obtained for  $H(p)$  than that provided by (1.4). In [35], Kátai showed that for some small unspecified positive constant  $c$ ,  $H(p) \geq c \log_2 p$

for almost all primes  $p$ . In fact, his bound holds for a more general type of Pratt-like tree (see the end of this introduction for more). We show a quantitatively stronger bound, using (1.3). First, however, we show that, over primes  $p \leq x$ , the distribution of  $H(p)$  is *tight* on the left of its median. That is, the “width” of the distribution, at least to the left of the median, remains bounded as  $x \rightarrow \infty$ , in contrast to what might be expected. We conjecture that the distribution is also tight to the right of the mean (see Conjecture 4 below).

**Theorem 6.** *Suppose  $g$  and  $h$  are increasing functions, satisfying  $0 \leq g(x) \leq h(x)$ ,  $h(x^2) - h(x) \leq K$  and  $g(x^2) - g(x) \leq K$  for  $x \geq 1$ . Suppose, for large  $x$ , that  $H(p) \geq h(p)$  for at least  $c\pi(x)$  primes  $p \leq x$ . Then*

$$(1.5) \quad H(p) \geq h(p) - g(p)$$

for all primes  $p \leq x$  with at most  $O(\pi(x) \exp\{-\frac{c \log^2 g(x)}{K}\})$  exceptions. Consequently, if  $g(x) \rightarrow \infty$  as  $x \rightarrow \infty$ , then (1.5) holds for almost all  $p$ .

**Theorem 7.** *Suppose that (1.3) holds with  $Q = x^\theta$  and  $R = (\log x)^2$ . For any  $c < \frac{1}{e^{-1} + \log(1/\theta)}$ , there is an  $\varepsilon > 0$  so that  $H(p) > c \log_2 p$  for all but  $O(x/(\log x)^{1+\varepsilon})$  primes  $p \leq x$ .*

**Corollary 1.** *Assume the Elliott-Halberstam Conjecture. Then, for every  $c < e$ , almost all primes satisfy  $H(p) > c \log_2 p$ .*

In light of the heuristic for  $L(p)$ , we see that  $L(p)$  should be much smaller than  $H(p)$  for most  $p$ .

We believe that the conclusion of Corollary 1 is best possible; see Conjecture 3 below. If  $p_1 \prec \cdots \prec p_k = p$  with  $k = H(p)$ , we have

$$\log p = \log p_1 \prod_{j=1}^{k-1} \frac{\log p_{j+1}}{\log p_j} \gg \left( \min_{1 \leq j \leq k-1} \frac{\log p_{j+1}}{\log p_j} \right)^{k-1},$$

hence

$$(1.6) \quad H(p) \leq \frac{\log_2 p}{\log \min_{1 \leq j \leq k-1} \frac{\log p_{j+1}}{\log p_j}} + O(1).$$

It is not known that there is an infinite set of primes  $p$  with  $H(p)/\log_2 p \rightarrow \infty$ . If this is the case, then (1.6) implies that for all  $\varepsilon > 0$ , there are an infinite number of primes  $p$  such that  $p - 1$  has a prime factor  $> p^{1-\varepsilon}$  (cf. section 1.1). In the opposite direction, we can show, for infinitely many  $p$ , a lower bound for  $H(p)$  which is quantitatively better than the bound in Theorem 7. In particular, if the Elliott-Halberstam conjecture is true, then  $\limsup H(p)/\log_2 p = \infty$ .

**Theorem 8.** *Assume that  $0 < \theta < 1$  and for every  $A > 0$ , (1.3) holds with  $Q = x^\theta$  and  $R = (\log x)^A$ . Then, for every  $c < \frac{1}{\log(1/\theta)}$ , there is a constant  $K$  so that for large  $x$ , there are  $\gg x/(\log x)^K$  primes  $p \leq x$  satisfying  $H(p) > c \log_2 p$ .*

Theorems 6, 7 and 8 will be proven in Section 5.

What can be said about upper bounds on  $H(p)$ , valid for almost all  $p$ ? Before our work it was not known that  $H(p) = o(\log p)$  for infinitely many primes  $p$ . A seemingly easy method for finding  $p$  with small  $H(p)$  is to take  $p$  with  $p - 1$  having only small prime factors, and use

$$(1.7) \quad H(p) \leq 1 + \max_{q|(p-1)} \frac{\log q}{\log 2}.$$

However, the best we know is that there are infinitely many  $p$  such that all the prime factors of  $p - 1$  are  $< p^{0.2931}$  [6]. In order to show  $H(p) = o(\log p)$  using (1.7), we would need a prime  $p$  so that all prime factors of  $p - 1$  are  $\leq p^{o(1)}$ . Using a different method, we prove that  $H(p)$  is substantially smaller than  $\log p$  for most primes  $p$ .

**Theorem 9.** *For  $c = 0.0378$  and some  $\delta > 0$ , we have  $H(p) \ll (\log p)^{1-c}$  for all but  $O(x \exp\{-(\log x)^\delta\})$  primes  $p \leq x$ .*

The proof of Theorem 9 is given in Section 6. Later, in Section 7, we apply a different method to deduce weaker upper estimates, but ones which hold with a much smaller exceptional set of primes (Theorem 11 there).

As  $\delta \rightarrow 0^+$ , the probability that the largest prime factor of  $n$  is between  $n^\delta$  and  $n^{1-\delta}$  tends to 1. The same is expected for the largest prime factor of shifted primes  $p - 1$ . It is reasonable to conjecture that  $H(p) \sim c \log_2 p$  for some constant  $c$ . In section 8 we give a heuristic argument, based on a stochastic model of the Pratt trees, for the assertion that this is true with  $c = e$ . More crudely, by the argument leading to (1.2), we expect that the number of primes at level  $k$  of a Pratt tree for  $p \leq x$  has normal order  $(\log_2 x)^k / k!$ . By Stirling's formula, this quantity drops below 1 when  $k \approx e \log_2 x$ .

**Conjecture 3.**  *$H(p)$  has normal order  $e \log_2 p$ . That is, for every  $\varepsilon > 0$ , we have*

$$|H(p) - e \log_2 p| < \varepsilon \log_2 p$$

*for almost all primes  $p$ .*

Recall Corollary 1, which says that the Elliott-Halberstam conjecture implies the lower bound implicit in Conjecture 3.

For the primes considered in Figure 1, we have  $e \log_2 10^9 \approx 8.2$  and  $e \log_2 10^{40} \approx 12.3$ . The peak of the distribution in these two cases is a bit smaller than  $e \log_2 p$ , and the argument in Section 8 explains this phenomenon. Moreover, our model predicts the tightness of the distribution of  $H(p)$  (cf Theorem 6) and also a large asymmetry in the distribution with respect to its median. The following conjecture makes this more precise.

**Conjecture 4.**  *$H(p) = e \log_2 p - \frac{3}{2} \log_3 p + E(p)$ , where  $E(p)$  is an exponentially tight sequence. More precisely, for some fixed  $c > 0$  we have, for any  $z \geq 0$ ,  $E(p) \leq z$  for all but  $O(e^{-cz} \pi(x))$  primes  $\leq x$ , and  $E(p) \geq -z$  for all but  $O(\exp\{-e^{cz}\} \pi(x))$  primes  $p \leq x$ .*

The asymmetry in the distribution is already evident in the graphs in Figure 1. Numerically,  $e \log_2 10^9 - \frac{3}{2} \log_3 10^9 \approx 6.57$  and  $e \log_2 10^{40} - \frac{3}{2} \log_3 10^{40} \approx 10.03$ .

Due to the fact that  $H(p)$  is integer valued, it is unlikely that  $E(p)$  behaves like a particular random variable. In particular, depending on the size of  $x$ , the second most popular value of  $H(p)$  for  $p \leq x$  may be one less or one greater than the most popular value (already evident in Figure 2). It is possible that  $E(p)$  behaves like a discrete approximation to a random variable.

The model in Section 8 can also be used to guess the distribution of other statistics of Pratt trees. For example, the quantity  $f(p)$  defined earlier is the total number of primes in the tree, and our model predicts that  $f(p) \asymp \log p$  for most primes  $p$  (see Theorem 4). Other interesting statistics are the width (maximum number of primes at some level of the tree) and mass (product of all primes in the tree; see [10]). We leave these topics to the interested reader.

**1.5. Further problems.** One can define a more general type of prime chain, by fixing a non-zero integer  $a$  and putting  $p \prec q$  if  $p|(q+a)$ . Everything we have shown in the case  $a = -1$  holds in the general case as well, with one small caveat. One must avoid the special prime  $q = -a$ , if  $-a$  is a prime, and avoid primes occurring in chain cycles when  $a > 0$ . For example, when  $a = 6$  we have the cycle  $5 \prec 19 \prec 13 \prec 7 \prec 29 \prec 23 \prec 17 \prec 11 \prec 5$ . Problems concerning the structure of the cycles for various  $a$  are interesting in their own right, and are investigated in [35], [36] and [42].

We conclude this section with a conjecture about prime chains, which follows from the prime  $k$ -tuples conjecture but should be “easier”. It can be considered a multiplicative analog of the statement that the primes contain arbitrarily long arithmetic progressions, recently proved by Green and Tao [29].

**Conjecture 5.** *For each  $k \geq 3$ , there are infinitely many prime  $k$ -tuples  $(p_1, \dots, p_k)$  where, for some  $m$ ,  $p_{j+1} = mp_j + 1$  for  $1 \leq j \leq k - 1$ .*

Even the case  $k = 3$  is not known.

**1.6. Notation.** Some notational conventions have already been mentioned, e.g.  $\pi(x)$  for the number of primes  $\leq x$  and  $\log_k x$  for the  $k$ -th iterate of the logarithm of  $x$ . Other number theoretic functions we need are the Möbius function  $\mu(n)$ , Euler’s totient function  $\phi(n)$ , the number of distinct prime divisors  $\omega(n)$  of  $n$ , and  $\Omega(n)$ , the number of prime power divisors of  $n$ . The letter  $p$ , with or without subscripts, always denotes a prime. Constants implied by the  $O$ - and  $\ll$ -symbols do not depend on any parameter unless indicated. In Section 8, we use  $\mathbf{P}$  for probability and  $\mathbf{E}$  for probabilistic expectation.

## 2. CUNNINGHAM CHAINS

*Proof of Theorem 1.* Let  $q \leq p$  be a prime such that  $\text{ord}_q 2 = q - 1$  and  $q \nmid (p + 1)$ . The powers of 2 generate all reduced residues modulo  $q$ , so there is some  $j \leq q - 1$  for which

$k$	$p_1$	$k/\log p_1$	discoverer
16	810433818265726529159	0.3323	Carmody and Jobling (2002)
15	$\leq 113220800675069784839$	$\geq 0.3248$	Carmody (2003)
14	95405042230542329	0.3580	Jobling (1999)
13	4090932431513069	0.3616	Brennen (1998)
12	554688278429	0.4437	Löh (1989)
11	665043081119	0.4040	Löh (1989)
10	26089808579	0.4169	Löh (1989)
9	85864769	0.4926	Nelson (1980); Sumiyama (1983)
8	19099919	0.4771	Nelson (1980)
7	1122659	0.5024	Lehmer (1965)
6	89	1.3367	Cunningham (1907)

TABLE 1. Smallest prime  $p$  with  $k(p) = k$ 

$2^{j-1}(p+1) \equiv 1 \pmod{q}$ . Hence,  $k(p) \leq q-1$ . More generally, let  $\Gamma_q$  be the orbit of 2 modulo  $q$ , i.e.,  $\Gamma_q = \{2^j \pmod{q} : j \in \mathbb{Z}\}$ . Then

$$(2.1) \quad k(p) \leq \min\{\text{ord}_q 2 : q \leq p, p+1 \in \Gamma_q\}.$$

Let  $0 < \varepsilon \leq \frac{1}{2}$  and put  $x = (\frac{1+\varepsilon}{C}) \log p$ . By Lemma 1.1 and partial summation, if  $p$  is large enough then

$$\sum_{\substack{q \leq x \\ \text{ord}_q 2 = q-1}} \log q \geq C \left(1 - \frac{\varepsilon}{2}\right) x > \log(p+1).$$

Thus, there is some  $q \leq x$  with  $\text{ord}_q 2 = q-1$  and  $q \nmid (p+1)$ . By (2.1),  $k(p) \leq q-1 \leq x$ .  $\square$

In light of Theorem 1, it is natural to ask if

$$(2.2) \quad \limsup_{p \rightarrow \infty} \frac{k(p)}{\log p} > 0.$$

Table 1 gives values of  $k(p)/\log p$  for the smallest  $p$  for which  $k(p) = k$  (the smallest  $p$  for  $k = 15$  apparently isn't known). These were taken from a web site maintained by Dirk Augustin (third column values truncated).

The proof of Theorem 2 uses Montgomery's large sieve estimate ([15], Théorème 6):

**Lemma 2.1.** *Suppose a set of positive integers  $\leq x$  avoids  $\xi(p)$  residue classes modulo  $p$  for each prime  $p$ . Then the cardinality of the set is*

$$\leq \frac{2x}{G}, \quad G = \sum_{n \leq x^{1/2}} \mu^2(n) \prod_{p|n} \frac{\xi(p)}{p - \xi(p)}.$$

*Proof of Theorem 2.* Let  $\varepsilon > 0$  be very small. Suppose that  $x \leq p \leq 2x$  and  $k(p) > (\log x)^A =: Q$ . If  $x$  is large enough, there are  $\gg Q^{1-\varepsilon/2}$  primes  $q \leq Q$  with  $q-1$  having a prime factor  $> Q^{\theta^+-\varepsilon}$ . Let  $\mathcal{Q}$  be the set of primes  $q \leq Q$  such that  $|\Gamma_q| > Q^{\theta^+-\varepsilon}$ . For any prime  $q$  with  $q-1$  having a prime factor  $> Q^{\theta^+-\varepsilon}$ , we have either  $|\Gamma_q| > Q^{\theta^+-\varepsilon}$  or  $|\Gamma_q| < Q^{1-\theta^++\varepsilon} =: Q_0$ . Any prime in the latter category divides  $M = \prod_{j \leq Q_0} (2^j - 1)$ . As  $M \leq 2^{Q_0^2}$  and  $\theta^+ > \frac{1}{2}$ , the number of such  $q$  is  $\leq Q_0^2 \ll Q^{1-\varepsilon}$  if  $\varepsilon$  is small enough. Therefore,  $|\mathcal{Q}| > Q^{1-\varepsilon}$  for all sufficiently large  $x$ . Let  $\omega := \lfloor (\log x)/(2 \log Q) \rfloor$  and let  $\mathcal{N}$  be the set of squarefree positive integers  $n$  having precisely  $\omega$  prime factors all from  $\mathcal{Q}$ . Clearly,  $n \leq x^{1/2}$  for all  $n \in \mathcal{N}$ . Further,

$$|\mathcal{N}| \geq \binom{|\mathcal{Q}|}{\omega} \geq \left( \frac{|\mathcal{Q}| - \omega}{\omega} \right)^\omega \gg x^{\frac{1}{2} - \frac{1}{2A} - \frac{\varepsilon}{2}}.$$

Finally, if  $n \in \mathcal{N}$ , then taking  $\xi(q) = |\Gamma_q|$ , we have

$$\prod_{q|n} \frac{\xi(q)}{q - \xi(q)} \geq \left( Q^{\theta^+-\varepsilon-1} \right)^\omega = Q^{-\omega(1-\theta^++\varepsilon)} \gg x^{\frac{\theta^+-1-2\varepsilon}{2}}.$$

Thus,

$$G \gg x^{\frac{\theta^+-1}{2} + \frac{1}{2} - \frac{1}{2A} - 2\varepsilon},$$

and the lemma follows.  $\square$

**Remarks.** The link 2 is special. If we consider generalized Cunningham Chains  $p_1, p_2 = mp_2 + 1, \dots, p_k = mp_{k-1} + 1$ , where  $m$  is an even integer, it is trivial that  $k$  is bounded in terms of  $m$  for  $m > 2$ . Indeed, if  $q$  is a prime factor of  $m-1$ , then  $p_j \equiv p_{j-1} + 1 \pmod{q}$ , and thus  $k \leq q-1$  if  $p_1 > q$  and  $k \leq q$  if  $p_1 < q$ .

**Problem 2.** *In light of the fact that  $k(p) = 1$  for most  $p$ , can it be proved unconditionally that  $\sum_{p \leq x} k(p) \sim \pi(x)$ ?*

### 3. ESTIMATES OF $P(x)$

Throughout this section, variables  $q, q_1$  and  $q_2$  denote primes.

*Proof of Theorem 3.* Since  $f(2) = 1$  and

$$(3.1) \quad f(p) = 1 + \sum_{q|(p-1)} f(q),$$

an easy induction shows that

$$(3.2) \quad f(p) \leq \frac{2 \log p}{\log 2} - 1,$$

and hence

$$P(x) = \sum_{p \leq x} f(p) \leq \frac{2}{\log 2} \sum_{p \leq x} \log p \leq \left( \frac{2}{\log 2} + o(1) \right) x \quad (x \rightarrow \infty)$$

by the Prime Number Theorem.

We next show that if  $\theta < 1$ , (1.3) holds with  $Q = x^\theta$  and  $R = (\log x)^2$ , and if  $\alpha$  satisfies  $0 < \alpha < 1$  and  $\theta^{1-\alpha} > 1 - \alpha$ , then

$$(3.3) \quad P(x) \gg x(\log x)^{-\alpha}.$$

The lower bound in Theorem 3 follows by taking  $\theta = 0.499$  (admissible by the Bombieri-Vinogradov theorem) and  $\alpha = 0.36$ . Define the numbers  $a_j$  by

$$a_j = \inf_{2 \leq y \leq \exp\{(1/\theta)^j\}} \frac{P(y)(\log y)^\alpha}{y}.$$

We have  $a_j > 0$  and  $a_{j+1} \leq a_j$  for all  $j$ . To show that  $\lim a_j > 0$ , suppose that  $j$  is large and  $\exp\{(1/\theta)^j\} < z \leq \exp\{(1/\theta)^{j+1}\}$ . By (1.3) and (3.1),

$$\begin{aligned} P(z) &= \sum_{p \leq z} \left( 1 + \sum_{q|(p-1)} f(q) \right) \\ &\geq \sum_{q \leq z^\theta} f(q) \pi(z; q, 1) \\ &= \text{li}(z) \sum_{q \leq z^\theta} \frac{f(q)}{q-1} + O\left(\frac{z}{\log z}\right). \end{aligned}$$

Since  $f(q) \geq 1$  for all  $q$  and  $\sum_q \frac{1}{q-1}$  diverges, for some  $q_0$  we have

$$P(z) \geq \text{li}(z) \sum_{q_0 < q \leq z^\theta} \frac{f(q)}{q-1}.$$

By partial summation and the definition of  $a_j$ ,

$$\begin{aligned} P(z) &\geq \text{li}(z) \int_{q_0}^{z^\theta} \frac{P(t)}{t^2} dt \\ &\geq a_j \frac{z}{\log z} \int_{q_0}^{z^\theta} \frac{dt}{t(\log t)^\alpha} \\ &= a_j \left[ \frac{\theta^{1-\alpha}}{1-\alpha} \frac{z}{(\log z)^\alpha} + O\left(\frac{z}{\log z}\right) \right]. \end{aligned}$$

Since  $\theta^{1-\alpha} > 1 - \alpha$ , the right side is  $> a_j z (\log z)^{-\alpha}$  for large  $j$ . Hence, the sequence  $\{a_j\}$  is eventually constant.  $\square$

*Proof of Theorem 4.* By the proof of Theorem 2.1 in [22] (more precisely, combine Lemma 2.4, Corollary 2.5, (2.8) and Theorem 2.1), we obtain

$$(3.4) \quad P(x) = cx + O(x/(\log_2 x)^\delta)$$

for some  $c \geq 0$ . Section 2 of [22] is devoted to the study of the normal behavior of the function  $\ell(n)$ , the number of times one must iterate the Euler function, starting from  $n$ , in order to reach 1. For example,  $\ell(9) = 4$  since  $\phi(9) = 8$ ,  $\phi(8) = 4$ ,  $\phi(4) = 2$  and  $\phi(2) = 1$ . It turns out that  $F(n) = \ell(n) - \begin{cases} 1 & n \text{ odd} \\ 0 & n \text{ even} \end{cases}$  is completely additive, and for odd primes  $p$ ,

$$F(p) = F(p-1) = \sum_{q^a \parallel p-1} aF(q).$$

This is very similar to relation (3.1), the only difference being the behavior at proper prime powers, which play an insignificant role in the arguments in [22].

In [22],  $F(p) \gg \log p$ , however we do not have  $f(p) \gg \log p$  from (3.1), exactly because of the influence of prime powers; from (3.2),  $f(p) \ll \log(\prod_{q|(p-1)} q)$ . Hence, we cannot immediately deduce that  $c > 0$  in (3.4). To show this, we use the method of the proof of Theorem 3. For  $j \geq 1$  let

$$a_j = \inf_{2 \leq y \leq e^j} \frac{P(y) \log_3 y}{y}.$$

For  $e^j < z \leq e^{j+1}$ , let  $z_0 = z^{1-(\log_2 z)^{-1-\delta}}$ . For large enough  $z$ ,  $z_0 \leq z/e$ . For such  $z$ , we obtain for some fixed  $q_0 \geq 100$

$$\begin{aligned} P(z) &\geq \text{li}(z) \sum_{q \leq z_0} \frac{f(q)}{q} + O\left(\frac{z}{\log z}\right) \\ &\geq \text{li}(z) \sum_{q_0 < q \leq z_0} \frac{f(q)}{q} \\ &\geq \frac{z}{\log z} \int_{q_0}^{z_0} \frac{P(t)}{t^2} dt \\ &\geq \frac{a_j z}{\log z} \int_{q_0}^{z_0} \frac{dt}{t \log_3 t}. \end{aligned}$$

By integration by parts,

$$\int \frac{dt}{t \log_3 t} = \frac{\log t}{\log_3 t} + \int \frac{dt}{t \log_2 t (\log_3 t)^2};$$

hence,

$$\int_{q_0}^{z_0} \frac{dt}{t \log_3 t} \geq \frac{\log z_0}{\log_3 z_0} \left(1 + \frac{1}{\log_2 z_0 \log_3 z_0}\right) - O(1) \geq \frac{\log z}{\log_3 z}$$

for large  $z$ . Hence,  $P(z) \geq a_j z / \log_3 z$  and thus the sequence  $\{a_j\}$  is eventually constant. Thus,  $P(z) \gg z / \log_3 z$  and therefore  $c > 0$  by (3.4).

For the second part of the theorem, start with

$$\begin{aligned} \sum_{p \leq x} (f(p) - c \log x)^2 &= \sum_{p \leq x} f(p)^2 - 2c \log x \left( cx + O\left(\frac{x}{(\log_2 x)^\delta}\right) \right) + c^2 \pi(x) \log^2 x \\ &= \sum_{p \leq x} f(p)^2 - c^2 x \log x + O\left(\frac{x \log x}{(\log_2 x)^\delta}\right). \end{aligned}$$

The sum on the right side is

$$\sum_{q \leq x} f(q)^2 \pi(x; q, 1) + 2 \sum_{q_1 < q_2 \leq x} f(q_1) f(q_2) \pi(x; q_1 q_2, 1).$$

Put  $x_0 = x^{1 - (\log_2 x)^{-1 - \delta}}$ . We use (1.3) to handle the terms with  $q \leq x_0$  and  $q_1 q_2 \leq x_0$ . The terms with  $q > x_0$  are dealt with using  $f(q) \ll \log q$  and sieve methods. By Theorem 2.4 of [30], we have

$$\begin{aligned} \sum_{x_0 < q \leq x} \pi(x; q, 1) &\leq \sum_{k \leq x/x_0} \#\{q \leq x/k : q, qk + 1 \text{ both prime}\} \\ &\ll \sum_{k \leq x/x_0} \frac{x}{\phi(k) \log^2 x} \ll \frac{x}{(\log x)(\log_2 x)^{1 + \delta}}, \end{aligned}$$

and

$$\begin{aligned} \sum_{\substack{x_0 < q_1 q_2 \leq x \\ q_1 < q_2}} \pi(x; q_1 q_2, 1) &\leq \sum_{q_1 \leq \sqrt{x}} \sum_{k \leq x/x_0} \#\{q_2 \leq x/(kq_1) : q_2, kq_1 q_2 + 1 \text{ both prime}\} \\ &\ll \sum_{q_1 \leq \sqrt{x}} \sum_{k \leq x/x_0} \frac{x}{\phi(kq_1) \log^2 x} \\ &\ll \frac{x}{(\log x)(\log_2 x)^\delta}. \end{aligned}$$

Hence,

$$\sum_{p \leq x} f(p)^2 = \frac{x}{\log x} \sum_{q \leq x_0} \frac{f(q)^2}{q} + \frac{x}{\log x} \sum_{\substack{q_1 < q_2 \leq x \\ q_1 q_2 \leq x_0}} \frac{2f(q_1) f(q_2)}{q_1 q_2} + O\left(\frac{x \log x}{(\log_2 x)^\delta}\right).$$

By the proof of Proposition 2.6 of [22], the two sums above are each

$$\frac{1}{2} c^2 \log^2 x + O((\log x)^2 (\log_2 x)^{-\delta}),$$

and we conclude that

$$\sum_{p \leq x} (f(p) - c \log x)^2 \ll \left(\frac{x \log x}{(\log_2 x)^\delta}\right).$$

The second statement in the theorem follows.  $\square$

#### 4. SIFTED CHAINS

The proof of Theorem 5 uses a novel method based on matrices of Dirichlet series. The underlying idea is a simple sieve; relax the condition that the numbers in the chain are prime, and instead only require that they do not have small prime factors. Let  $y \geq 2$  and let  $q$  be the product of the primes  $\leq y$ . For  $(a, q) = 1$ , let  $N_a(x; y)$  be the number of chains with  $n_1 = a$ , with  $n_k/n_1 \leq x$  and consisting of numbers coprime to  $q$ . Also let

$$N(x, y) = \max_{(a, q)=1} N_a(x, y).$$

If  $p > y$ , we have

$$P(x; p) \leq N_p(x, y) \leq N(x, y).$$

For example, when  $y = q = 2$ , the links in the chains counted by  $N(x; 2)$  are even and have product  $\leq x$ . A variant of Kalmár's argument from [34] shows that  $N(x; 2) \ll x^{\rho'}$ , where  $\rho'$  is the unique positive solution of  $2^{-s}\zeta(s) = 1$ . For  $y \geq 3$ , however, the restrictions on the links  $m_i$  are more complex to deal with.

For positive integers  $a, b, q$  and real  $s > 1$ , let

$$S(a, b) = S(a, b; q, s) = \sum_{\substack{m \geq 1 \\ am+1 \equiv b \pmod{q}}} m^{-s}.$$

This Dirichlet series encodes the possible links  $m$  from a number  $n_i \equiv a \pmod{q}$  to a number  $n_{i+1} \equiv b \pmod{q}$ .

Let  $U_q = (\mathbb{Z}/q\mathbb{Z})^*$  be the multiplicative group of integers coprime to  $q$ . Let  $A_k(a_1, a_k)$  be the sum of  $(m_1 \cdots m_{k-1})^{-s}$  over all  $(k-1)$ -tuples  $(m_1, \dots, m_{k-1})$  which could serve as links in a chain starting from a number  $n_1 \equiv a_1 \pmod{q}$ , ending with a number  $n_k \equiv a_k \pmod{q}$  and with all numbers in the chain coprime to  $q$  (we suppress the dependence on  $q$  and  $s$  in this and future definitions). Then  $A_2(a_1, a_2) = S(a_1, a_2)$  and for  $k \geq 3$ ,

$$A_k(a_1, a_k) = \sum_{a_2, \dots, a_{k-1} \in U_q} S(a_1, a_2) S(a_2, a_3) \cdots S(a_{k-1}, a_k).$$

Let  $V_k(a_1)$  be the column vector  $(A_k(a_1, a_k) : a_k \in U_q)$ . For consistency, let  $V_1(a_1)$  be a vector with all zero entries except for an entry of 1 in the  $a_1$  position. Since

$$A_{k+1}(a_1, a_{k+1}) = \sum_{a_k \in U_q} A_k(a_1, a_k) S(a_k, a_{k+1}),$$

we obtain

$$V_{k+1}(a_1) = M V_k(a_1),$$

where  $M = M(q, s)$  is the transition matrix

$$M = (S(a, b))_{b, a \in U_q}.$$

The rows of  $M$  are indexed by  $b$  and the columns are indexed by  $a$ . Finally, let  $F_k(a_1) = \sum_{a_k} A_k(a_1, a_k)$ , so that

$$(4.1) \quad F_k(a_1) = (1, \dots, 1)V_k(a_1) = (1, \dots, 1)M^{k-1}V_1(a_1);$$

i.e.,  $F_k(a_1)$  is the sum of the entries of column  $a_1$  in  $M^{k-1}$ .

**Lemma 4.1.** *We have*

$$N_a(x; y) \leq \inf_{s > 1} \left( x^s \sum_{1 \leq k \leq \frac{\log x}{\log 2} + 1} F_k(a) \right).$$

*Proof.* If  $n_k/n_1 \leq x$ , then  $m_1 \cdots m_{k-1} \leq x$  and hence  $(x/m_1 \cdots m_{k-1})^s \geq 1$ .  $\square$

Observe that the sum on  $k$  in Lemma 4.1, if extended to  $k = \infty$ , is convergent if and only if  $M$  is a contracting matrix, i.e., all the eigenvalues of  $M$  have modulus  $< 1$ . Since  $M$  has positive real entries, the Perron-Frobenius Theorem implies that the eigenvalue with largest modulus is positive, real and simple. Call this eigenvalue  $\lambda(s; y)$ .

**Problem 3.** *Obtain good estimates for  $\lambda(s; y)$ . In particular, determine the smallest  $s$ , as a function of  $y$ , for which  $\lambda(s; y) < 1$ .*

We show below that if  $y$  is large and  $s \geq 1 + \frac{\log_2 y}{\log y}$ , then  $\lambda(s; y) < 1$ . Accurate estimation of  $\lambda(s; y)$  is difficult for large  $y$ , but the largest row sum of  $M$  serves as an upper bound. For a generic matrix  $A$ , let  $R_b(A)$  be the sum of the entries in the row indexed by  $b$ , and let  $R(A)$  be the maximum row sum of  $A$ .

**Lemma 4.2.** *With  $M$  defined above, we have  $R_b(M) = \alpha(q)\beta((b-1, q))$ , where*

$$\alpha(q) = \prod_{p > y} (1 - p^{-s})^{-1}, \quad \beta(d) = \prod_{p|d} \frac{p-1}{p^s-1}.$$

*Proof.* Write  $d = (b-1, q)$  and  $b' = \frac{b-1}{d}$ . Then

$$R_b(M) = \sum_{a \in U_q} \sum_{am \equiv b-1 \pmod{q}} m^{-s} = d^{-s} \sum_{(k, q/d)=1} k^{-s} \#\{a \in U_q : ak \equiv b' \pmod{q/d}\}.$$

The congruence  $ak \equiv b' \pmod{q/d}$  has a unique solution modulo  $q/d$ , and hence has  $\phi(d)$  solutions  $a \in U_q$ . Thus,

$$R_b(M) = \frac{\phi(d)}{d^s} \sum_{(k, q/d)=1} k^{-s} = \frac{\phi(d)}{d^s} \prod_{p|(q/d)} (1 - p^{-s})^{-1} = \alpha(q)\beta(d).$$

$\square$

By Lemma 4.2,

$$(4.2) \quad R(M) = \alpha(q)\beta(2) = \frac{1}{2^s - 1} \prod_{p>y} (1 - p^{-s})^{-1}.$$

Since  $R(AB) \leq R(A)R(B)$ ,  $R(M^{k-1}) \leq R(M)^{k-1}$ . To bound  $N(x, y)$ , however, we require a bound on the largest *column* sum of  $M^{k-1}$ . Lacking a better method, we'll just use the trivial bound  $\phi(q)R(M)^{k-1}$ . We then have

$$(4.3) \quad F_k(a) \leq \sum_{a \in U_q} F_k(a) \leq \phi(q)R(M^{k-1}) \leq \phi(q)R(M)^{k-1},$$

and, applying Lemma 4.1,

$$(4.4) \quad N(x, y) \leq \sum_{a \in U_q} N_a(x, y) \leq \phi(q) \inf_{s: R(M) < 1} \frac{x^s}{1 - R(M)}.$$

By standard prime number estimates, if  $1 < s \leq 2$ , then

$$\begin{aligned} - \sum_{p>y} \log(1 - p^{-s}) &= O(1/y^{2s-1}) + \sum_{p>y} p^{-s} \\ &\ll \int_y^\infty \frac{dt}{t^s \log t} \ll \frac{e^{-(s-1) \log y}}{(s-1) \log y}. \end{aligned}$$

Also,  $2^s - 1 = 1 + (2 \log 2)(s - 1) + O((s - 1)^2)$  for  $1 \leq s \leq 2$ . By (4.2), for  $y \rightarrow \infty$ ,  $s - 1 \geq \log_2 y / \log y$  and  $s = 1 + o(1)$ , we have

$$(4.5) \quad 1 - R(M) \sim (2 \log 2)(s - 1).$$

Take  $y = (\log x) / \log_2 x$  and  $s = 1 + \frac{\log_2 y}{\log y}$ . Using (4.4) and the Prime Number Theorem bound

$$\phi(q) \leq q = e^{(1+o(1))y},$$

we obtain the following.

**Theorem 10.** *If  $y = \frac{\log x}{\log_2 x}$ , then*

$$N(x; y) \leq \sum_{a \in U_q} N_a(x, y) \ll x \exp \left\{ \frac{\log x (\log_3 x + O(1))}{\log_2 x} \right\}.$$

## 5. PRATT TREE HEIGHT: LOWER BOUNDS

*Proof of Theorem 6.* The conclusion is trivial if  $g(x^{1/2}) \leq 3K$ , so we will assume that  $g(x^{1/2}) > 3K$ . Let

$$m = \left\lfloor \frac{g(x^{1/2})}{K} \right\rfloor$$

so that  $m \geq 3$ . Put  $Q = x^{2^{-m}}$  and let  $T$  be the set of primes  $x^{1/2} < p \leq x$  such that there is a prime  $q|(p-1)$  with  $Q < q \leq x^{1/4}$  and  $H(q) \geq h(q)$ . For  $p \in T$ ,

$$H(p) \geq 1 + h(q) \geq h(Q) \geq h(x) - mK \geq h(p) - g(p).$$

By sieve methods (Theorem 4.2 of [30]), for large  $x$

$$\begin{aligned} |\{x^{1/2} < p \leq x : p \notin T\}| &\ll \frac{x}{\log x} \prod_{\substack{Q < q \leq x^{1/4} \\ H(q) \geq h(q)}} \left(1 - \frac{1}{q}\right) \\ &\ll \frac{x}{\log x} 2^{-mc} \end{aligned}$$

and the theorem follows.  $\square$

In the proof of Theorem 7, we need some further estimates from sieve theory. The first is the Brun-Titchmarsh inequality

$$(5.1) \quad \pi(x, q, 1) \ll \frac{x}{\phi(q) \log(x/q)},$$

and the second is a way to handle primes in progressions to large moduli.

**Lemma 5.1.** *Uniformly for  $z \geq 2$  and  $4 \leq A \leq x^{1/2}$ , we have*

$$\sum_{\substack{p_1, p_2 \geq z \\ p_1 p_2 \geq x/A}} \pi(x, p_1 p_2, 1) \ll \frac{x(\log^2 x) \log A}{\log^4 z}.$$

*Proof.* Without loss of generality, assume that  $z \leq x^{1/4}$ . If  $p \leq x$  and  $p \equiv 1 \pmod{p_1 p_2}$  with  $p_1 p_2 \geq x/A$ , then  $p = 1 + k p_1 p_2$ , with  $1 \leq k \leq A$ . For each  $p$ , there are  $O((\frac{\log x}{\log z})^2)$  choices for the pair  $p_1, p_2$ . Hence, by sieve methods (e.g. Theorem 2.2 of [30]),

$$\begin{aligned} \sum_{\substack{p_1, p_2 \geq z \\ p_1 p_2 \leq x/A}} \pi(x, p_1 p_2, 1) &\ll \left(\frac{\log x}{\log z}\right)^2 \sum_{1 \leq k \leq A} |\{n \leq x : n \equiv 1 \pmod{k}, \\ &\quad \text{all prime factors of } n(\frac{n-1}{k}) \text{ are } > z\}| \\ &\ll \left(\frac{\log x}{\log z}\right)^2 \sum_{1 \leq k \leq A} \frac{x}{\phi(k) \log^2 z} \\ &\ll \frac{x(\log^2 x) \log A}{\log^4 z}. \end{aligned}$$

$\square$

*Proof of Theorem 7.* Suppose that

$$c < h < c' < \frac{1}{e^{-1} + \log(1/\theta)}.$$

For some constant  $c''$ , described below, let

$$\mathcal{P} = \{p : H(p) \geq c' \log_2 p - c''\}.$$

We will show, for some  $\delta > 0$ , that

$$(5.2) \quad P(x) := |\{p \leq x : p \in \mathcal{P}\}| \geq \delta \frac{x}{\log x}.$$

Consequently, a positive proportion of primes  $p$  satisfy  $H(p) > h \log_2 p$ , and Theorem 7 follows from Theorem 6.

Since  $\lim_{k \rightarrow \infty} \frac{1}{k} (k!)^{1/k} = e^{-1}$ , there is an integer  $k \geq 2$  such that

$$\frac{1}{c'} > \frac{(k!)^{1/k}}{k} + \log(1/\theta).$$

Let  $\alpha, \beta$  satisfy

$$e^{-k/c'} < \beta < \theta^k \exp(-(k!)^{1/k})$$

and

$$\beta \exp((k!)^{1/k}) < \alpha < \theta^k.$$

Suppose that  $\delta$  is sufficiently small, depending only on the choice of  $c', \theta, k, \alpha, \beta$ . Let  $x_0$  be sufficiently large, depending on  $c', \theta, k, \alpha, \beta, \delta$ , and put  $c'' = c' \log_2(x_0)$ . Observe that (5.2) holds trivially for  $2 \leq x \leq x_0$ , provided  $\delta$  is small enough. Throughout this proof, constants implied by the  $O$ - and  $\ll$ -symbols may depend on  $c', \theta, k, \alpha, \beta$ , but not on  $\delta$ .

Next, suppose that  $Y \geq x_0$  and that inequality (5.2) holds for  $2 \leq x \leq Y$ . Let  $S$  be a subset of the primes in  $\mathcal{P} \cap [Y^\beta, Y^{\theta^k}]$ . Let  $N(S)$  be the number of primes  $p_0 \in (Y, 2Y]$  so that there is a prime chain

$$(5.3) \quad p_k \prec p_{k-1} \prec \cdots \prec p_0$$

with  $p_k \in S$ . For such  $p_0$ , we have

$$\begin{aligned} H(p_0) &\geq k + H(p_k) \\ &\geq k + c' \log_2 p_k - c'' \\ &= c' \log_2(2Y) - c'' + c' \log \beta + k + O\left(\frac{1}{\log Y}\right) \\ &\geq c' \log_2 p_0 - c'' \end{aligned}$$

if  $x_0$  is large enough. We will show, for appropriate  $S$ , that

$$(5.4) \quad N(S) \geq \delta \frac{Y}{\log Y},$$

which implies

$$P(2Y) \geq P(Y) + \delta \frac{Y}{\log Y} \geq \delta \frac{2Y}{\log 2Y}.$$

Therefore, by induction over dyadic intervals, (5.4) implies (5.2), and hence the theorem.

To prove (5.4), we will consider chains (5.3) satisfying not only  $p_k \in S$ , but also

$$(5.5) \quad p_{j+1} \leq p_j^\theta \quad (0 \leq j \leq k-1), \quad p_1 \leq Y^\theta.$$

With (5.5), we can use (1.3) to accurately count such chains. We have

$$N(S) \geq N_1(S) - N_2(S),$$

where  $N_1(S)$  is the number of chains (5.3) satisfying  $p_k \in S$  and (5.5), and  $N_2(S)$  is the number of pairs of distinct chains satisfying these conditions with the same  $p_0$ . For brevity, write

$$E(x, p) = \pi(x, p, 1) - \frac{\text{li}(x)}{p-1}.$$

We first have

$$N_1(S) = \sum_{p_k \in S} \sum_{p_{k-1}} \cdots \sum_{p_1} \left( \pi(2Y, p_1, 1) - \pi(Y, p_1, 1) \right),$$

where we assume (5.3) and (5.5) in the summations. By induction on  $1 \leq j \leq k$ , we shall now prove

$$(5.6) \quad N_1(S) = \frac{Y}{\log Y} \sum_{p_k} \sum_{p_{k-1}} \cdots \sum_{p_j} \frac{(\log_2 Y^{\theta^j} - \log_2 p_j)^{j-1}}{p_j(j-1)!} + O\left(\frac{Y}{\log^2 Y}\right).$$

First,

$$\pi(2Y, p_1, 1) - \pi(Y, p_1, 1) = \frac{\text{li}(2Y) - \text{li}(Y)}{p_1} + E(2Y, p_1) - E(Y, p_1) + O\left(\frac{Y/\log Y}{p_1^2}\right).$$

For a given  $p_1$ , there are  $O(1)$  chains  $p_k \prec \cdots \prec p_1$  with  $p_k \geq Y^\beta$ . Also,  $\sum 1/p_1 = O(1)$ . By (1.3), the totality of the error terms is  $O(Y/\log^2 Y)$ , and we obtain (5.6) for  $j = 1$ . Now suppose (5.6) is true for some particular  $j \geq k-1$ . By partial summation, for  $Y^\beta \leq p_{j+1} \leq Y^{\theta^{j+1}}$ ,

$$\begin{aligned} \sum_{\substack{p_j \equiv 1 \pmod{p_{j+1}} \\ p_{j+1}^{1/\theta} \leq p_j \leq Y^{\theta^j}}} \frac{(\log_2 Y^{\theta^j} - \log_2 p_j)^{j-1}}{p_j(j-1)!} &= \int_{p_{j+1}^{1/\theta}}^{Y^{\theta^j}} \frac{(\log_2 Y^{\theta^j} - \log_2 t)^{j-1}}{(p_{j+1}-1)(j-1)!t \log t} dt \\ &+ O\left(\frac{|E(Y^{\theta^j}, p_{j+1})|}{Y^{\theta^j}} + \frac{|E(p_{j+1}^{1/\theta}, p_{j+1})|}{p_{j+1}^{1/\theta}} + \int_{p_{j+1}^{1/\theta}}^{Y^{\theta^j}} \frac{|E(t, p_{j+1})|}{t^2} dt\right). \end{aligned}$$

The first integral is

$$\frac{(\log_2 Y^{\theta^{j+1}} - \log_2 p_{j+1})^j}{j! p_{j+1}} + O\left(\frac{1}{p_{j+1}^2}\right).$$

To estimate the aggregate of the error terms, note that for each  $p_{j+1}$ , there are at most  $O(1)$  chains  $p_k \prec \cdots \prec p_{j+1}$  with  $p_k \geq Y^\beta$ . For fixed  $t \in [Y^\beta, Y^{\theta j}]$ , (1.3) gives

$$\sum_{p_{j+1} \leq t^\theta} |E(t, p_{j+1})| = O\left(\frac{t}{\log^2 t}\right).$$

By another application of (1.3),

$$\begin{aligned} \sum_{Y^\beta \leq p_{j+1} \leq Y^{\theta j+1}} \frac{|E(p_{j+1}^{1/\theta}, p_{j+1})|}{p_{j+1}^{1/\theta}} &\ll (\log Y) \max_{Y^\beta \leq P \leq Y^{\theta j+1}} \sum_{P \leq p_{j+1} \leq 2P} \frac{|E(p_{j+1}^{1/\theta}, p_{j+1})|}{p_{j+1}^{1/\theta}} \\ &\ll \max_{Y^\beta \leq P \leq Y^{\theta j+1}} \frac{\log Y}{P^{1/\theta}} \sum_{p \leq 2P} \max_{y \leq (2P)^{1/\theta}} |E(y, p)| \ll \frac{1}{\log Y}. \end{aligned}$$

This completes the proof of (5.6) with  $j$  replaced by  $j + 1$ . By (5.6) with  $j = k$  and partial summation,

$$\begin{aligned} N_1(\mathcal{P} \cap [Y^\beta, Y^\alpha]) &\geq \frac{\delta Y}{\log Y} \int_{Y^\beta}^{Y^\alpha} \frac{(\log_2 Y^{\theta k} - \log_2 t)^{k-1}}{(k-1)! t \log t} dt + O\left(\frac{Y}{\log^2 Y}\right) \\ &> \frac{\delta Y}{\log Y} \int_{Y^\beta}^{Y^\alpha} \frac{(\log_2 Y^\alpha - \log_2 t)^{k-1}}{(k-1)! t \log t} dt + O\left(\frac{Y}{\log^2 Y}\right) \\ &= \frac{\delta Y}{\log Y} \frac{(\log(\alpha/\beta))^k}{k!} + O\left(\frac{Y}{\log^2 Y}\right). \end{aligned}$$

By hypothesis,  $\log(\alpha/\beta) > (k!)^{1/k}$ . Also, note that the summands in (5.6) (with  $j = k$ ) are  $\asymp 1/p_k$  for  $Y^\beta \leq p_k \leq Y^\alpha$ . Hence, if  $\delta$  is small enough, there is a set  $S \subseteq \mathcal{P} \cap [Y^\beta, Y^\alpha]$  such that

$$(5.7) \quad N_1(S) \geq (\delta + \delta^{3/2}) \frac{Y}{\log Y}$$

and

$$(5.8) \quad \sum_{p_k \in S} \frac{1}{p_k} \ll \delta.$$

We have

$$N_2 = \sum_{j=0}^{k-1} N_{2,j},$$

where  $N_{2,j}$  counts pairs of connected chains

$$\left. \begin{array}{l} p_k \prec \cdots \prec p_{j+1} \\ p'_k \prec \cdots \prec p'_{j+1} \end{array} \right\} p_j \prec \cdots \prec p_0$$

with each of the two chains satisfying (5.5),  $p_{j+1} \neq p'_{j+1}$  and  $p_k, p'_k \in S$ . We further write  $N_{2,j} = N'_{2,j} + N''_{2,j}$ , where  $N'_{2,j}$  counts pairs of such chains with  $p_j \leq p_{j+1}p'_{j+1}Y^{\delta^2}$ . As before, for each pair  $(p_{j+1}, p'_{j+1})$ , there are  $O(1)$  choices for  $p_k, p'_k, \dots, p_{j+2}, p'_{j+2}$ . By Lemma 5.1,

$$N'_{2,0} \ll \sum_{\substack{p_1, p'_1 \geq Y^\beta \\ p_1 p'_1 \geq Y^{1-\delta^2}}} \pi(2Y, p_1 p'_1, 1) \ll \delta^2 \frac{Y}{\log Y}.$$

When  $j \geq 1$ , an argument similar to that leading to (5.6) yields

$$N'_{2,j} \ll \frac{Y}{\log Y} \sum_{p_{j+1}, p'_{j+1}} \sum_{p_j} \frac{1}{p_j}.$$

Using (5.1) and Lemma 5.1, we get (writing  $q = p_{j+1}p'_{j+1}$ )

$$\begin{aligned} N'_{2,j} &\ll \frac{Y}{\log Y} \sum_{p_{j+1}, p'_{j+1}} \left[ \frac{\pi(qY^{\delta^2}, q, 1)}{q} Y^{\delta^2} + \int_{2q}^{qY^{\delta^2}} \frac{\pi(t, q, 1)}{t^2} dt \right] \\ &\ll \delta^2 \frac{Y}{\log Y} + \frac{Y}{\delta^2 \log^2 Y} \ll \delta^2 \frac{Y}{\log Y} \end{aligned}$$

provided that  $\log x_0 \geq \delta^{-4}$ . Hence,

$$(5.9) \quad \sum_j N'_{2,j} \ll \delta^2 \frac{Y}{\log Y}.$$

For chains counted by  $N''_{2,j}$ , the Brun-Titchmarsh inequality suffices for the estimations. When  $j \geq 1$  and  $p_j$  is given, as before we have

$$\sum_{p_{j-1}} \cdots \sum_{p_1} \pi(2Y, p_1, 1) \ll \frac{Y}{p_j \log Y}.$$

By partial summation and (5.1), given  $p_{j+1}$  and  $p'_{j+1}$ ,

$$\sum_{p_j} \frac{1}{p_j} \ll \frac{1}{p_{j+1}p'_{j+1}\delta^2 \log Y} + \frac{\log(1/\delta)}{p_{j+1}p'_{j+1}} \ll \frac{\log(1/\delta)}{p_{j+1}p'_{j+1}}.$$

For  $j+1 \leq r \leq k-1$ ,

$$(5.10) \quad \sum_{p_r} \frac{1}{p_r} \ll \frac{1}{p_{r+1}}, \quad \sum_{p'_r} \frac{1}{p'_r} \ll \frac{1}{p'_{r+1}}.$$

Finally, by (5.8), we arrive at

$$(5.11) \quad N''_{2,j} \ll \delta^2 \log(1/\delta) \frac{Y}{\log Y}.$$

In a similar way, when  $j = 0$  we have by (5.1) and partial summation,

$$\sum_{p_1 p'_1 \leq 2Y^{1-\delta^2}} \pi(2Y, p_1 p'_1, 1) \ll \frac{\log(1/\delta)}{p_2 p'_2}.$$

A second application of (5.10) then gives (5.11) in this case.

Finally, combining (5.9) and (5.11), we obtain

$$N_2(S) \ll \delta^2 \log(1/\delta) \frac{Y}{\log Y}.$$

Together with (5.7), if  $\delta$  is small enough then (5.4) holds, and this completes the proof.  $\square$

*Proof of Theorem 8.* We proceed by induction as in the proof of Theorem 7. However, the proof is much simpler. Let  $c'$  and  $\theta'$  satisfy  $\theta' > 1/3$  and

$$c < c' < \frac{1}{\log(1/\theta')} < \frac{1}{\log(1/\theta)},$$

and define  $K$  by

$$\theta^K = \frac{\theta - \theta'}{8}.$$

Let  $x_0$  be large, depending on  $K, c, c', \theta, \theta'$  and put  $c'' = c' \log_2(x_0^3)$ . Let

$$\mathcal{P} = \{p : H(p) \geq c' \log_2 p - c''\}.$$

In particular,  $\mathcal{P}$  contains all primes  $\leq x_0^3$ . We shall prove that

$$(5.12) \quad Q(x) := |\mathcal{P} \cap (x/2, x]| \geq \frac{x}{(\log x)^K}$$

for  $x \geq x_0$ , which implies the lemma (since  $c' > c$ ). By the Prime Number Theorem and the fact that  $K > 1$ , if  $x_0$  is large enough then (5.12) holds for  $x_0 \leq x \leq x_0^3$ . Suppose  $y \geq x_0^3$  and (5.12) holds for  $x_0 \leq x \leq y$ . Assume  $y < x \leq 2y$  and put  $I = \mathcal{P} \cap (x^{\theta'}, x^\theta]$ . Suppose that  $x/2 < p \leq x$ , and that  $q|p-1$ , where  $q \in I$ . Then

$$\begin{aligned} H(p) &\geq 1 + H(q) \geq 1 + c' \log_2 q - c'' \\ &\geq 1 + c' \log_2 x + c' \log \theta' - c'' > c' \log_2 p - c'', \end{aligned}$$

so that  $p \in \mathcal{P}$ . For  $x/2 < p \leq x$ ,  $p-1$  is divisible by at most two primes from  $I$ , and hence by our hypothesis (1.3),

$$\begin{aligned} Q(x) &\geq \frac{1}{2} \sum_{q \in I} \left( \pi(x; q, 1) - \pi(x/2; q, 1) \right) \\ &\geq \frac{\text{li}(x) - \text{li}(x/2)}{2} \sum_{q \in I} \frac{1}{q-1} + O\left(\frac{x}{(\log x)^{K+1}}\right) \\ &\geq \frac{x}{4 \log x} \sum_{q \in I} \frac{1}{q} + O\left(\frac{x}{(\log x)^{K+1}}\right). \end{aligned}$$

Since (5.12) holds for  $x^{\theta'} < y \leq x^\theta$ , we have

$$\begin{aligned} \sum_{q \in I} \frac{1}{q} &\geq \sum_{\substack{j \geq 1 \\ 2^j \leq x^{\theta-\theta'}}} \frac{Q(2^j x^{\theta'})}{2^j x^{\theta'}} \\ &\geq \left( \frac{(\theta - \theta') \log x}{\log 2} - 1 \right) \frac{1}{(\log x^\theta)^K} \\ &\geq \frac{\theta - \theta'}{\theta^K (\log x)^{K-1}} \\ &= \frac{8}{(\log x)^{K-1}}. \end{aligned}$$

Therefore,

$$Q(x) \geq \frac{2x}{(\log x)^K} + O\left(\frac{x}{(\log x)^{K+1}}\right) \geq \frac{x}{(\log x)^K}$$

if  $x_0$  is large enough. By induction on dyadic intervals, (5.12) holds for all  $x \geq x_0$ .  $\square$

**Remark.** If (1.3) holds with  $Q = x^{1-\varepsilon(x)}$  for a function  $\varepsilon(x) \rightarrow 0$  as  $x \rightarrow \infty$ , and where  $R(x) = (\log x)^{f(x)}$ , where  $f(x) \rightarrow \infty$  as  $x \rightarrow \infty$  (sufficiently fast depending on the decay of  $\varepsilon(x)$ ), we can deduce that  $H(p) \geq g(p) \log_2 p$  for infinitely many  $p$ , where  $g(p) \rightarrow \infty$  as  $p \rightarrow \infty$  (depending on the decay of  $\varepsilon(x)$ ).

## 6. PRATT TREE HEIGHT: UPPER BOUNDS, I

The proof of Theorem 9 is more complex than the proofs of the lower bounds in Theorems 7 and 8. Rather than trying to construct friable shifted primes, we utilize sieve methods, which tell us that the largest prime factor of  $p-1$  cannot be too large too often. At the core is a sieve upper bound for  $k$ -tuples of primes which is uniform in  $k$ . These bounds involve a factor, the singular series, for which average estimates are required. There is another factor,  $(c_1 k)^k$ , in the sieve estimate, which has the potential to derail any attempt to bound  $H(p)$  non-trivially. We get around this problem by observing that if  $H(p)$  is

large, then there must be a prime chain in the Pratt tree for  $p$  which is very condensed in a multiplicative sense.

**Lemma 6.1.** *There is a  $\delta > 0$  so that the following holds. Let  $a_1, \dots, a_k$  be positive integers, let  $b_1, \dots, b_k$  be integers and let  $\xi(p)$  be the number of incongruent solutions of  $\prod_{i=1}^k (a_i n + b_i) \equiv 0 \pmod{p}$ . If  $x \geq 10$ ,  $1 \leq k \leq \delta \frac{\log x}{\log_2 x}$  and*

$$B := \sum_p \frac{k - \xi(p)}{p} \log p \leq \delta \log x,$$

then the number of integers  $n \leq x$  for which  $a_1 n + b_1, \dots, a_k n + b_k$  are all prime and  $> k$  is

$$\ll \frac{2^k k!}{(\log x)^k} x \mathfrak{S} \cdot \exp\left(O\left(\frac{kB + k^2 \log_2 x}{\log x}\right)\right), \quad \mathfrak{S} = \prod_p \left(1 - \frac{\xi(p)}{p}\right) \left(1 - \frac{1}{p}\right)^{-k}.$$

**Remarks.** When  $a_j = 1$  for all  $j$  and  $k$  is of order  $\log x$ , better upper bounds are possible using methods of Elsholtz [20].

*Proof.* Since  $\xi(p) = k$  for large  $p$ ,  $\mathfrak{S}$  converges to a nonzero number if and only if  $\xi(p) < p$  for all  $p$ . Also,  $\xi(p) \leq k$  for all  $p$ . Hence, if  $\mathfrak{S} = 0$ , the number of  $n$  in question is zero and there is nothing to prove. Now assume  $\mathfrak{S} > 0$ . By Lemma 2.1, the number of  $n$  in question is  $\ll x/G(\sqrt{x})$ , where

$$G(z) = \sum_{n \leq z} g(n), \quad g(n) = \mu^2(n) \prod_{p|n} \frac{\xi(p)}{p - \xi(p)}.$$

When  $k$  is fixed, the argument in §5.3 of [30] implies that  $G(z) \sim (\log z)^k / (k! \mathfrak{S})$ . We modify the argument to make the estimate uniform in  $k$  in the stated range.

By the argument on p. 147–148 of [30], we have

$$\begin{aligned} \sum_{d \leq z} g(d) \log d &= \sum_{d \leq z} g(d) \sum_{p \leq z/d} \frac{\xi(p) \log p}{p} + \sum_{p \leq z} \frac{g(p) \xi(p)}{p} \log p \sum_{\substack{z/p^2 < d \leq z/p \\ p|d}} g(d) \\ &= \sum_{d \leq z} g(d) \sum_{p \leq z/d} \frac{\xi(p) \log p}{p} + \sum_{h \leq z} g(h) \sum_{\substack{p|h \\ p > z/h}} \frac{\xi(p) \log p}{p}. \end{aligned}$$

On the right side, we have

$$\sum_{p \leq z/d} \frac{\xi(p) \log p}{p} = \sum_{p \leq z/d} \frac{k \log p}{p} + O(B) = k \log(z/d) + O(B + k)$$

and

$$\sum_{\substack{p|h \\ p > z/h}} \frac{\xi(p) \log p}{p} \leq k \sum_{p|h} \frac{\log p}{p} \leq k \sum_{p \leq \log z} \frac{\log p}{p} + k \frac{\log_2 z}{\log z} \sum_{p|h} 1 \ll k \log_2 z.$$

Thus,

$$\sum_{d \leq z} g(d) \log d = k \sum_{d \leq z} g(d) \log \frac{z}{d} + O(G(z)(B + k \log_2 z)).$$

Adding the sum on the right side to both sides yields

$$(6.1) \quad \begin{aligned} G(z) \log z &= (k+1) \sum_{d \leq z} g(d) \log \frac{z}{d} + r(z)G(z) \log z \\ &= (k+1) \int_1^z \frac{G(t)}{t} dt + r(z)G(z) \log z, \end{aligned}$$

where  $r(z) \ll \frac{B+k \log_2 z}{\log z}$ . By assumption, if  $\delta$  is small enough and  $z \geq \sqrt{x}$ , then  $|r(z)| \leq \frac{1}{2}$ . As on p. 150 of [30], if we define

$$E(y) = \log \left( \frac{k+1}{\log^{k+1} y} \int_1^y \frac{G(t)}{t} dt \right),$$

then we obtain from (6.1) the estimate

$$E'(y) = \frac{k+1}{y \log y} \frac{r(y)}{1-r(y)} \ll \frac{k(B+k \log_2 y)}{y \log^2 y}.$$

It follows that the integral

$$\int_{\sqrt{x}}^{\infty} E'(y) dy$$

is absolutely convergent. Hence, for some constant  $D$  and for  $z \geq \sqrt{x}$ ,

$$\begin{aligned} (1-r(z)) \frac{G(z)}{\log^k z} &= \frac{k+1}{\log^{k+1} z} \int_1^z \frac{G(t)}{t} dt \\ &= e^{E(z)} = D \exp \left\{ - \int_z^{\infty} E'(y) dy \right\} \\ &= D \exp \left\{ O \left( \frac{kB + k^2 \log_2 z}{\log z} \right) \right\}. \end{aligned}$$

By the argument on p. 151–152 of [30],  $D^{-1} = k! \mathfrak{S}$ . Therefore,

$$G(\sqrt{x}) \geq \frac{\log^k \sqrt{x}}{k! \mathfrak{S}} \exp \left\{ O \left( \frac{kB + k^2 \log_2 x}{\log x} \right) \right\},$$

and the proof is complete.  $\square$

For given positive integers  $m_1, \dots, m_{k-1}$ , we will apply Lemma 6.1 with the forms

$$f_1(n) = n, \quad f_{j+1}(n) = m_j f_j(n) + 1 \quad (1 \leq j \leq k-1).$$

Write  $f_j(n) = a_j n + b_j$ , so that

$$(6.2) \quad a_j = m_1 \cdots m_{j-1} \quad (j \geq 1), \quad b_1 = 0, \quad b_j = 1 + \sum_{i=2}^{j-1} m_i \cdots m_{j-1} \quad (j \geq 2).$$

Let  $\mathfrak{S}(\mathbf{m}) = \mathfrak{S}$  be the associated singular series and let  $\xi(p, \mathbf{m}) = \xi(p)$ . We first give a global upper bound on  $\mathfrak{S}(\mathbf{m})$ , and then show that  $\mathfrak{S}(\mathbf{m})$  is bounded in a suitable average sense. Our result can be compared to the result of Gallagher [28], where an average of the singular series for prime  $k$ -tuples with fixed  $k$  is estimated.

**Lemma 6.2.** *There is a constant  $c_2 > 0$  so that*

$$(a) \quad \mathfrak{S}(\mathbf{m}) \ll (c_2 \log_2(4m_1 \cdots m_{k-1}))^{k-1}.$$

Furthermore,

$$(b) \quad \sum_p \frac{k - \xi(p, \mathbf{m})}{p} \log p \leq k (\log_2(4m_1 \cdots m_{k-1}) + O(1)).$$

*Proof.* Let  $x = m_1 \cdots m_{k-1}$ . Assume that  $\mathfrak{S}(\mathbf{m}) \neq 0$ , so that all  $m_i$  are even and  $2^{k-1} \leq x$ . Then  $\xi(p, \mathbf{m}) = k$  if  $p \nmid N$ , where

$$N = m_1 \cdots m_{k-1} \prod_{i < j} |a_i b_j - a_j b_i|.$$

By (6.2),  $a_j \leq x$  and  $b_j \leq 1 + \sum_{j=1}^{k-2} x/2^j \leq x$  for each  $j$ . Thus,  $N \leq x^{k(k-1)+1} \leq \exp\{O(\log^3 x)\}$ . Then, since  $1 - k/p \leq (1 - 1/p)^k$  for  $p > k$ ,

$$\mathfrak{S}(\mathbf{m}) \leq \prod_{p|N} \left(1 - \frac{1}{p}\right)^{1-k} = \left(\frac{N}{\phi(N)}\right)^{k-1}.$$

Part (a) follows from the elementary bound  $N/\phi(N) \ll \log_2(N+2)$ . For part (b), the Mertens' estimates give

$$\begin{aligned} \sum_p \frac{k - \xi(p, \mathbf{m})}{p} \log p &\leq k \sum_{p|N} \frac{\log p}{p} \leq k \sum_{p \leq \log N} \frac{\log p}{p} + \frac{k \log_2 N}{\log N} \sum_{p|N, p > \log N} 1 \\ &\leq k (\log_2 N + O(1)). \end{aligned}$$

□

**Lemma 6.3.** *If  $n$  and  $k$  are positive integers, then there exists  $d|n$  with  $d \leq n^{1/k}$  and  $\omega(n) \leq k(\omega(d) + 1)$ .*

*Proof.* If  $\omega(n) \leq k$ , then take  $d = 1$ . Otherwise, write the prime factorization of  $n$  as  $n = p_1^{e_1} \cdots p_r^{e_r}$ , where  $r > k$  and  $p_1^{e_1} < \cdots < p_r^{e_r}$ . Define  $j$  by  $p_1^{e_1} \cdots p_j^{e_j} \leq n^{1/k} < p_1^{e_1} \cdots p_{j+1}^{e_{j+1}}$  and put  $d = p_1^{e_1} \cdots p_j^{e_j}$ . Then  $r \leq k(j+1)$ , for otherwise

$$n > \prod_{i=1}^{k(j+1)} p_i^{e_i} = \prod_{l=0}^{k-1} \prod_{i=l(j+1)+1}^{(l+1)(j+1)} p_i^{e_i} \geq (p_1^{e_1} \cdots p_{j+1}^{e_{j+1}})^k > n.$$

□

**Lemma 6.4.** *Let  $k \geq 2$  and  $I \subseteq \{1, \dots, k-1\}$ . If the variables  $m_i$  are fixed ( $1 \leq i \leq k-1, i \notin I$ ), then for any prime  $p$ ,*

$$\sum_{0 \leq m_i < p \ (i \in I)} \xi(p, (m_1, \dots, m_{k-1})) \geq p^{|I|+1} - (p-1)^{|I|+1}.$$

*Proof.* Put  $\mathbf{m}_l = (m_1, \dots, m_l)$  for  $0 \leq l \leq k-1$  and let

$$(6.3) \quad N_l(p) := \sum_{0 \leq m_i < p \ (i \in I, i \leq l)} \xi(p, \mathbf{m}_l).$$

We have  $N_0(p) = 1$ . If  $a+1 \notin I$ , then  $\xi(p, \mathbf{m}_{a+1}) \geq \xi(p, \mathbf{m}_a)$  and hence  $N_{a+1}(p) \geq N_a(p)$ . If  $a+1 \in I$ , we have

$$\begin{aligned} N_{a+1}(p) &= \sum_{0 \leq m_i < p \ (i \in I, i \leq a+1)} |\{n \bmod p : p | f_1(n) \cdots f_a(n)(m_a f_a(n) + 1)\}| \\ &= \sum_{0 \leq m_i < p \ (i \in I, i \leq a)} \left[ p \cdot |\{n \bmod p : p | f_1(n) \cdots f_a(n)\}| \right. \\ &\quad \left. + |\{n \bmod p : p \nmid f_1(n) \cdots f_a(n)\}| \right] \\ &= p^{1+g(a)} + (p-1)N_a(p), \end{aligned}$$

where  $g(a) = |\{i \in I : i \leq a\}|$ . It follows, by induction, that

$$N_a(p) \geq p^{1+g(a)} - (p-1)^{1+g(a)} \quad (0 \leq a \leq k-1),$$

and the lemma follows from the case  $a = k-1$ . □

**Lemma 6.5.** *Let  $k \geq 4$ , and suppose that  $M_i, N_i$  are integers satisfying  $M_i \geq 2$  and  $2 \leq N_i \leq 2kM_i$  for  $1 \leq i \leq k-1$ . We have*

$$\sum_{\substack{N_i < m_i \leq N_i + M_i \\ (1 \leq i \leq k-1)}} \mathfrak{S}(\mathbf{m}) \ll M_1 \cdots M_{k-1} (3e^\gamma \log k + O(1))^b \exp \left\{ O \left( \frac{k \log_2 k}{\log k} \right) \right\},$$

where  $b$  is the number of variables  $M_i$  which are  $\leq 2k^3$ .

*Proof.* Let  $L = \lfloor \log k \rfloor + 1$  and  $r = k^3 L + 1$ . We will perform a precise averaging of the factors in  $\mathfrak{S}(\mathbf{m})$  for primes  $p \leq r$ , and we will use crude estimates for larger  $p$ .

If  $p > r$  and  $p | m_1 \cdots m_{k-1}$ , we'll use the trivial bound  $\xi(p, \mathbf{m}) \geq 1$ . For other primes  $p$ , each congruence  $f_j(n) \equiv 0 \pmod{p}$  has exactly one solution. For  $h > j$ ,  $f_j(n) \equiv 0 \pmod{p}$  and  $f_h(n) \equiv 0 \pmod{n}$  have a common solution if and only if  $p | (a_j b_h - a_h b_j)$ . Write

$$a_j b_h - a_h b_j = m_1 \cdots m_{j-1} g_{j,h}(\mathbf{m}),$$

where

$$g_{j,j+1}(\mathbf{m}) = 1, \quad g_{j,h}(\mathbf{m}) = 1 + \sum_{i=j+1}^{h-1} m_i \cdots m_{h-1} \quad (h \geq j+2).$$

It follows that

$$\xi(p, \mathbf{m}) \geq k - w(p, \mathbf{m}),$$

where

$$w(p, \mathbf{m}) = |\{3 \leq h \leq k : \exists j \leq h-2 \text{ with } p | g_{j,h}(\mathbf{m})\}|.$$

Hence,

$$(6.4) \quad \left(1 - \frac{\xi(p, \mathbf{m})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k} \leq \left(1 - \frac{1}{p}\right)^{\xi(p, \mathbf{m}) - k} \leq \left(1 - \frac{1}{p}\right)^{-w(p, \mathbf{m})}.$$

Let

$$\psi_r(n) = \prod_{p|n, p>r} \frac{p}{p-1},$$

and

$$G_{j,h}(\mathbf{m}) = \prod_{\substack{p>r \\ p|g_{j,h}(\mathbf{m}) \\ p \nmid g_{i,h}(\mathbf{m}) \ (j+1 \leq i \leq h-2) \\ p \nmid m_1 \cdots m_{k-1}}} p.$$

We then have

$$(6.5) \quad \prod_{\substack{p \nmid m_1 \cdots m_{k-1} \\ p>r}} \left(1 - \frac{1}{p}\right)^{-w(p, \mathbf{m})} = \prod_{\substack{1 \leq j < h \leq k-1 \\ h \geq j+2}} \psi_r(G_{j,h}(\mathbf{m})).$$

The number of pairs  $(j, h)$  in the product on the right is  $J = \frac{(k-2)(k-3)}{2}$ . By (6.4), (6.5) and Hölder's inequality,

(6.6)

$$\begin{aligned} \sum_{\mathbf{m}} \mathfrak{G}(\mathbf{m}) &\leq \sum_{\mathbf{m}} \prod_{p \leq r} \left(1 - \frac{\xi(p, \mathbf{m})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k} \prod_{i=1}^{k-1} \psi_r(m_i)^{k-1} \prod_{j,h} \psi_r(G_{j,h}(\mathbf{m})) \\ &\leq \left( \sum_{\mathbf{m}} \left[ \prod_{p \leq r} \left(1 - \frac{\xi(p, \mathbf{m})}{p}\right) \left(1 - \frac{1}{p}\right)^{-k} \right]^{\frac{L}{L-1}} \right)^{1 - \frac{1}{L}} \\ &\quad \times \prod_{i=1}^{k-1} \left( \sum_{\mathbf{m}} \psi_r(m_i)^{2L(k-1)^2} \right)^{\frac{1}{2L(k-1)}} \prod_{j,h} \left( \sum_{\mathbf{m}} \psi_r(G_{j,h}(\mathbf{m}))^{2JL} \right)^{\frac{1}{2JL}}. \end{aligned}$$

If we write

$$\psi_r(m)^s = \sum_{d|m} \beta_s(d),$$

then  $\beta_s$  is multiplicative and supported on square-free integers composed of primes  $> r$ . Furthermore, if  $p > r \geq s + 1$ , then

$$\beta_s(p) = \left( \frac{p}{p-1} \right)^s - 1 \leq e^{s/(p-1)} - 1 \leq \frac{4s}{p}.$$

We then have for each  $i$

$$\begin{aligned} \sum_{N_i < m_i \leq N_i + M_i} \psi_r(m_i)^{2L(k-1)^2} &\leq \sum_{d \leq N_i + M_i} \beta_{2L(k-1)^2}(d) \frac{N_i + M_i}{d} \\ &\leq (2k+1)M_i \prod_{p>r} \left(1 + \frac{\beta_{2L(k-1)^2}(p)}{p}\right) \\ &\leq (2k+1)M_i \prod_{p>r} \left(1 + \frac{8Lk^2}{p^2}\right) \ll kM_i. \end{aligned}$$

Hence,

$$\begin{aligned} (6.7) \quad \prod_{i=1}^{k-1} \left( \sum_{\mathbf{m}} \psi_r(m_i)^{2L(k-1)^2} \right)^{\frac{1}{2L(k-1)}} &\ll (kM_1 \cdots M_{k-1})^{\frac{1}{2L}} \\ &\ll (M_1 \cdots M_{k-1})^{\frac{1}{2L}}. \end{aligned}$$

For fixed  $(j, h)$ , let  $M_l = \max(M_{j+1}, \dots, M_{h-1})$  and write

$$(6.8) \quad g_{j,h}(\mathbf{m}) = m_l(m_{l+1} \cdots m_{h-1})g_{j,l}(\mathbf{m}) + g_{l,h}(\mathbf{m}).$$

We'll use

$$G_{j,h}(\mathbf{m})|G'_{j,h}(\mathbf{m}) := \prod_{\substack{p>r \\ p|g_{j,h}(\mathbf{m}) \\ p \nmid m_1 \cdots m_{k-1} g_{l,h}(\mathbf{m})}} p,$$

and note that crudely  $G'_{j,h}(\mathbf{m}) \leq g_{j,h}(\mathbf{m}) \leq k(M_l + N_l)^k \leq (6kM_l)^k$ . Fix all  $m_i$  satisfying  $j+1 \leq i \leq h-1$  and  $i \neq l$ . If  $M_l \geq 2^{k^3}$ , then we have, by (6.8),

$$\begin{aligned} \sum_{N_l < m_l \leq N_l + M_l} \psi_r(G'_{j,h}(\mathbf{m}))^{2JL} &= \sum_{\substack{d \leq (6kM_l)^k \\ (d, \prod_{i \neq l} m_i) = 1}} \beta_{2JL}(d) |\{N_l < m_l \leq N_l + M_l : d|G'_{j,h}(\mathbf{m})\}| \\ &\leq \sum_{d \leq (6kM_l)^k} \left(\frac{M_l}{d} + 1\right) \beta_{2JL}(d) \\ &\leq M_l \prod_{p>r} \left(1 + \frac{8JL}{p^2}\right) + \prod_{r < p \leq (6kM_l)^k} \left(1 + \frac{8JL}{p}\right) \\ &\ll M_l + \exp[8JL(\log_2(6kM_l) + \log k - \log_2 r + O(1))] \\ &\ll M_l. \end{aligned}$$

If  $M_l < 2^{k^3}$ , then Lemma 6.3 gives

$$\begin{aligned} \sum_{N_l < m_l \leq N_l + M_l} \psi_r(G'_{j,h}(\mathbf{m}))^{2JL} &\leq \sum_{m_l \leq (2k+1)M_l} \exp\left(\frac{2JL}{r-1} \omega(G'_{j,h}(\mathbf{m}))\right) \\ &\leq \sum_{m_l \leq 3kM_l} \sum_{\substack{d|G'_{j,h}(\mathbf{m}) \\ d \leq 6kM_l}} \exp\left(\frac{k^3 L}{r-1} (\omega(d) + 1)\right) \\ &\leq e \sum_{d \leq 6kM_l} e^{\omega(d)} |\{m_l \leq 3kM_l : d|G'_{j,h}(\mathbf{m})\}| \\ &\ll kM_l \sum_{d \leq 6kM_l} \frac{e^{\omega(d)}}{d} \\ &\ll kM_l \prod_{p \leq 6kM_l} \left(1 + \frac{e}{p-1}\right) \\ &\ll kM_l (\log(6kM_l))^e \ll M_l k^{3e+1}. \end{aligned}$$

We conclude that

$$(6.9) \quad \prod_{j,h} \left( \sum_{\mathbf{m}} \psi_r(G_{j,h}(\mathbf{m}))^{2jL} \right)^{\frac{1}{2jL}} \ll (k^{3e+1} M_1 \cdots M_{k-1})^{\frac{1}{2L}} \\ \ll (M_1 \cdots M_{k-1})^{\frac{1}{2L}}.$$

Note that trivially

$$\left[ \prod_{p \leq r} \left( 1 - \frac{\xi(p, \mathbf{m})}{p} \right) \left( 1 - \frac{1}{p} \right)^{-k} \right]^{\frac{L}{L-1}-1} \leq \prod_{p \leq r} \left( 1 - \frac{1}{p} \right)^{-\frac{k}{L-1}} \\ = \exp \left\{ O \left( \frac{k \log_2 k}{\log k} \right) \right\}.$$

Put  $r' = \frac{1}{4}k^3$ . We have similarly

$$\prod_{r' < p \leq r} \left( 1 - \frac{\xi(p, \mathbf{m})}{p} \right) \left( 1 - \frac{1}{p} \right)^{-k} = \exp \left\{ O \left( \frac{k \log_2 k}{\log k} \right) \right\}.$$

Therefore, by (6.6), (6.7) and (6.9),

$$\sum_{\mathbf{m}} \mathfrak{S}(\mathbf{m}) \ll (M_1 \cdots M_{k-1})^{\frac{1}{L}} S^{1-\frac{1}{L}} \exp \left\{ O \left( \frac{k \log_2 k}{\log k} \right) \right\},$$

where

$$S = \sum_{\mathbf{m}} \prod_{p \leq r'} \left( 1 - \frac{\xi(p, \mathbf{m})}{p} \right) \left( 1 - \frac{1}{p} \right)^{-k}.$$

Let  $I = \{i : M_i \geq 2^{k^3}\}$  and  $M' = \prod_{p \leq r'} p$ . By an elementary estimate,  $M' \leq e^{2r'}$ . Thus, for  $i \in I$ ,  $M_i \geq kM'$ . Hence, the number of  $N_i \leq m_i < N_i + M_i$  in a given residue class modulo  $M'$  is  $\leq M_i/M' + 1 \leq (1 + O(1/k))M_i/M'$ . Thus, by Lemma 6.4 and the Chinese Remainder Theorem,

$$S \leq \sum_{\substack{N_i < m_i \leq N_i + M_i \\ (i \notin I)}} \prod_{i \in I} \frac{M_i}{M'} \left( 1 + O \left( \frac{1}{k} \right) \right) \sum_{\substack{m_i \bmod M' \\ (i \in I)}} \prod_{p \leq r'} \left( 1 - \frac{\xi(p, \mathbf{m})}{p} \right) \left( 1 - \frac{1}{p} \right)^{-k} \\ \ll \prod_{i \in I} M_i \sum_{\substack{N_i < m_i \leq N_i + M_i \\ (i \notin I)}} \prod_{p \leq r'} \frac{1}{p^{|I|}} \left( \frac{p}{p-1} \right)^k \left[ p^{|I|} - \frac{1}{p} \sum_{\substack{0 \leq m_i < p \\ (i \in I)}} \xi(p, \mathbf{m}) \right] \\ \ll M_1 \cdots M_{k-1} \prod_{p \leq r'} \left( \frac{p}{p-1} \right)^{k-1-|I|}.$$

Noting that  $b = k - 1 - |I|$ , the lemma follows from Mertens' estimate.  $\square$

*Proof of Theorem 9.* Suppose that  $x$  is large, and let  $r \geq 1$ ,  $l \geq 2$  be integers to be chosen later, and depending on  $x$ . Put

$$(6.10) \quad k_j = jl \quad (0 \leq j \leq r)$$

and, for  $0 < \eta \leq \frac{1}{2}$  fixed, let

$$(6.11) \quad x_j = x^{(j+1)^{-\eta}} \quad (0 \leq j \leq r).$$

If  $p \leq x$  and  $H(p) \geq \frac{\log x_r}{\log 2} + k_r$ , then there are even integers  $h_1, \dots, h_{k_r}$  so that

$$(6.12) \quad p = p_0 = h_1 p_1 + 1, p_1 = h_2 p_2 + 1, \dots, p_{k_r-1} = h_{k_r} p_{k_r} + 1,$$

with  $p_0, \dots, p_{k_r}$  prime and  $p_{k_r} \geq x_r$ . We further suppose that

$$(6.13) \quad k_r \leq \frac{\log x_r}{(\log_2 x)^2},$$

The vector  $(h_1, \dots, h_{k_r})$  may not be unique, but we associate to each such  $p$  a single such vector. For  $1 \leq j \leq r$ , let  $Q_j$  be the set of primes  $p$  so that

$$(6.14) \quad p_{k_i} < x_i \quad (i < j), \quad p_{k_j} \geq x_j.$$

Then,

$$(6.15) \quad \left| \left\{ p < x : H(p) \geq \frac{\log x_r}{\log 2} + k_r \right\} \right| \leq \sum_{j=1}^r |Q_j|.$$

Fix  $1 \leq j \leq r$  and even integers  $h_1, \dots, h_{k_j}$  satisfying  $h_1 \cdots h_{k_j} \leq x/x_j$ . By Lemma 6.2 (b) and (6.13),

$$\sum_p \frac{k_j - \xi(p; (h_{k_j}, \dots, h_1))}{p} \log p \ll k_j \log_2 x \ll \frac{\log x_r}{\log_2 x} \leq \frac{\log x_j}{\log_2 x}.$$

Fix  $c_1 > 2e^{-1}$ . By Lemma 6.1, (6.13) and Stirling's formula, the number of  $p = p_0 \leq x$  satisfying (6.12) is

$$(6.16) \quad \ll \frac{x}{h_1 \cdots h_{k_j}} \frac{(c_1 k_j)^{k_j+1} \mathfrak{S}(h_{k_j}, \dots, h_1)}{(\log x_j)^{k_j+1}}$$

if  $x$  is large enough. Now we estimate  $|Q_j|$ . Let  $1 \leq b_j \leq k_j$ , a parameter to be chosen later. Let  $Q_{j,1}$  be the set of  $p \in Q_j$  for which at least  $b_j$  of the variables  $h_1, \dots, h_{k_j}$  are  $\leq 2^{2k_j^3}$ , and  $Q_{j,2} = Q_j \setminus Q_{j,1}$ .

To estimate  $|Q_{j,1}|$ , fix a set  $\mathcal{B} \subseteq \{1, \dots, k_j\}$  of size  $b_j$  so that  $h_i \leq 2^{2k_j^3}$  for  $i \in \mathcal{B}$ . Let  $I = \{1 \leq i \leq k_j : i \notin \mathcal{B}\}$  and put

$$a_i = |\mathcal{B} \cap \{k_i + 1, \dots, k_{i+1}\}| \quad (0 \leq i \leq j-1),$$

$$I_i = I \cap \{k_i + 1, \dots, k_{i+1}\} \quad (0 \leq i \leq j-1).$$

By (6.14),

$$(6.17) \quad \prod_{g \in I_i \cup \dots \cup I_{j-1}} h_g \leq h_{k_{i+1}} \cdots h_{k_j} \leq \frac{x_i}{x_j} \quad (0 \leq i \leq j-1).$$

Since  $h_i \geq 2$  for all  $i$ ,

$$(6.18) \quad \prod_{g \in \mathcal{B}} \sum_{2 \leq h_g \leq 2^{2k_j^3}} \frac{1}{h_g} \leq (2k_j^3)^{b_j}.$$

Let  $\alpha = \frac{l}{\log x_j}$ . By (6.17) and the elementary estimate  $\sum_{2 \leq h \leq y} h^{-1-s} \leq 1/s$ ,

$$\begin{aligned} \sum_{h_g (g \in I)} \frac{1}{\prod_{g \in I} h_g} &\leq \sum_{h_g \geq 2 (g \in I)} \frac{1}{\prod_{g \in I} h_g} \prod_{i=0}^{j-1} \left( \frac{x_i}{x_j} \right)^\alpha \frac{1}{\prod_{g \in I_i \cup \dots \cup I_{j-1}} h_g^\alpha} \\ &= \prod_{i=0}^{j-1} \left( \frac{x_i}{x_j} \right)^\alpha \sum_{h_g \geq 2 (g \in I_i)} \frac{1}{\prod_{g \in I_i} h_g^{1+(i+1)\alpha}} \\ &\leq \prod_{i=0}^{j-1} \left( \frac{x_i}{x_j} \right)^\alpha \left( \frac{1}{(i+1)\alpha} \right)^{k_{i+1}-k_i-\alpha_i} \\ &= \left( \frac{1}{\alpha} \right)^{k_j-b_j} \frac{1^{a_0} 2^{a_1} \cdots j^{a_{j-1}}}{(j!)^l} \exp \left\{ l \sum_{i=0}^j \left[ \left( \frac{j+1}{i+1} \right)^\eta - 1 \right] \right\}. \end{aligned}$$

Since

$$\sum_{i=0}^j (i+1)^{-\eta} \leq 1 + \int_1^{j+1} t^{-\eta} dt \leq \frac{(j+1)^{1-\eta}}{1-\eta},$$

we have

$$\sum_{i=0}^j \left[ \left( \frac{j+1}{i+1} \right)^\eta - 1 \right] \leq \frac{\eta}{1-\eta} (j+1) \leq (j+1) \leq 2j.$$

Also,  $1^{a_0} 2^{a_1} \cdots j^{a_{j-1}} \leq j^{b_j}$  and  $j! \geq (j/e)^j$ . Hence,

$$(6.19) \quad \sum_{h_g (g \in I)} \frac{1}{\prod_{g \in I} h_g} \leq e^{3k_j} \left( \frac{\log x_j}{k_j} \right)^{k_j-b_j}.$$

The number of choices for  $\mathcal{B}$  is

$$\binom{k_j}{b_j} \leq \frac{k_j^{b_j}}{b_j!} \leq \left( \frac{ek_j}{b_j} \right)^{b_j}.$$

By (6.16), (6.18), (6.19), and Lemma 6.2, we obtain

$$\begin{aligned}
(6.20) \quad |Q_{j,1}| &\ll \frac{x(c_1 c_2 k_j \log_2 x)^{k_j+1}}{(\log x_j)^{k_j+1}} \left(\frac{2ek_j^4}{b_j}\right)^{b_j} e^{3k_j} \left(\frac{\log x_j}{k_j}\right)^{k_j-b_j} \\
&= \frac{x}{\log x_j} k_j (c_1 c_2 e^3 \log_2 x)^{k_j+1} \left(\frac{2ek_j^4(k_j/b_j)(j+1)^\eta}{\log x}\right)^{b_j} \\
&\ll x \exp\left\{O(k_j \log_3 x) + b_j \left[(4+\eta) \log k_j + \log\left(\frac{ek_j}{b_j}\right) - \log_2 x\right]\right\}.
\end{aligned}$$

We next estimate  $|Q_{j,2}|$ . Place each variable  $h_i$  into an interval  $J_i$ . If  $h_i \leq 2^{2k_j^3}$ , then take  $J_i = (2^{l_i-1}, 2^{l_i}]$  for an integer  $l_i \geq 1$ , and if  $h_i > 2^{2k_j^3}$ , then take

$$J_i = \left( \lfloor 2^{2k_j^3} (1 + 1/k_j)^{l_i-1} \rfloor, \lfloor 2^{2k_j^3} (1 + 1/k_j)^{l_i} \rfloor \right]$$

for some integer  $l_i \geq 1$ . For brevity, write  $J_i = (H_i, K_i]$  for each  $i$ . Since  $K_i - H_i \geq H_i/(2k)$ , there are at most  $b_j$  values of  $i$  with  $K_i - H_i \leq 2^{k_j^3}$ . Lemma 6.5 then gives

$$\begin{aligned}
\sum_{h_1 \in J_1} \cdots \sum_{h_{k_j} \in J_{k_j}} \frac{\mathfrak{S}(h_{k_j}, \dots, k_1)}{h_1 \cdots h_{k_j}} &\leq \frac{1}{H_1 \cdots H_{k_j}} \sum_{h_1 \in J_1} \cdots \sum_{h_{k_j} \in J_{k_j}} \mathfrak{S}(h_{k_j}, \dots, h_1) \\
&\ll (3e^\gamma \log k_j + O(1))^{b_j} \exp\left\{O\left(\frac{k_j \log_2 k_j}{\log k_j}\right)\right\} \prod_{i=1}^{k_j} \frac{K_i - H_i}{H_i}.
\end{aligned}$$

By our definition of the intervals  $J_i$ ,

$$\begin{aligned}
\prod_{i=1}^{k_j} \frac{K_i - H_i}{H_i} &\leq \prod_{\substack{1 \leq i \leq k_j \\ K_i - H_i < 2^{k_j^3}}} 2 \sum_{H_i < h_i \leq K_i} \frac{1}{h_i} \prod_{\substack{1 \leq i \leq k_j \\ K_i - H_i \geq 2^{k_j^3}}} \left(1 + O\left(\frac{1}{k_j}\right)\right) \sum_{H_i < h_i \leq K_i} \frac{1}{h_i} \\
&\ll 2^{b_j} \sum_{h_1 \in J_1} \cdots \sum_{h_{k_j} \in J_{k_j}} \frac{1}{h_1 \cdots h_{k_j}}.
\end{aligned}$$

Thus, after summing over all possibilities for  $J_1, \dots, J_{k_j}$ , we obtain by (6.16)

$$|Q_{j,2}| \ll \frac{x(c_1 k_j)^{k_j+1}}{(\log x_j)^{k_j+1}} \exp\left\{O\left(\frac{k_j \log_2 k_j}{\log k_j} + b_j \log_2 k_j\right)\right\} \sum_{h'_1, \dots, h'_{k_j}} \frac{1}{h'_1 \cdots h'_{k_j}},$$

where, by (6.14),  $h'_1, \dots, h'_{k_j}$  satisfy

$$(6.21) \quad h'_{k_i+1} \cdots h'_{k_j} \leq 2^{k_j-k_i} h_{k_i+1} \cdots h_{k_j} \leq 2^{k_j} \frac{x_i}{x_j} \quad (0 \leq i \leq j-1).$$

To handle the sum on  $h'_1, \dots, h'_{k_j}$ , let  $\alpha_0, \dots, \alpha_{j-1}$  be positive real numbers. By (6.21),

$$\begin{aligned} \sum_{h'_1, \dots, h'_{k_j}} \frac{1}{h'_1 \cdots h'_{k_j}} &\leq \prod_{i=0}^{j-1} \left[ \left( 2^{k_j} \frac{x_i}{x_j} \right)^{\alpha_i} \sum_{h'_{k_{i+1}}, \dots, h'_{k_{i+1}}=2}^{\infty} \frac{1}{(h'_{k_{i+1}} \cdots h'_{k_{i+1}})^{1+\alpha_0+\dots+\alpha_i}} \right] \\ &\leq \prod_{i=0}^{j-1} \left( 2^{k_j} \frac{x_i}{x_j} \right)^{\alpha_i} \left( \frac{1}{\alpha_0 + \dots + \alpha_i} \right)^l. \end{aligned}$$

If we ignore the factors  $2^{k_j \alpha_i}$ , the optimal choice of parameters is

$$\alpha_i = \frac{l}{(j+1)^\eta \log x_j} \left[ \frac{(i+2)^\eta (i+1)^\eta}{(i+2)^\eta - (i+1)^\eta} - \frac{(i+1)^\eta i^\eta}{(i+1)^\eta - i^\eta} \right].$$

Since  $(i+2)^\eta - (i+1)^\eta \geq \eta(i+2)^{\eta-1}$ , we have

$$\alpha_0 + \dots + \alpha_i = \frac{l}{(j+1)^\eta \log x_j} \frac{(i+2)^\eta (i+1)^\eta}{(i+2)^\eta - (i+1)^\eta} \leq \frac{l(i+2)(i+1)^\eta}{\eta(j+1)^\eta \log x_j}.$$

In the opposite direction, we have

$$\frac{1}{(i+1)^\eta} - \frac{1}{(i+2)^\eta} = \frac{1}{(i+2)^\eta} \left[ \left( 1 + \frac{1}{i+1} \right)^\eta - 1 \right] \leq \frac{\eta}{(i+1)(i+2)^\eta}.$$

Thus,

$$\frac{1}{\alpha_0 + \dots + \alpha_i} \leq \frac{\eta(j+1)^\eta \log x_j}{l(i+1)(i+2)^\eta}.$$

We therefore obtain (recall that  $k_j = jl$ )

$$\begin{aligned} \sum_{h'_1, \dots, h'_{k_j}} \frac{1}{h'_1 \cdots h'_{k_j}} &\leq 2^{k_j(\alpha_0 + \dots + \alpha_{j-1})} \exp \left\{ \sum_{i=0}^{j-1} \alpha_i \left[ \left( \frac{j+1}{i+1} \right)^\eta - 1 \right] \log x_j \right\} \\ &\quad \times \left( \frac{\eta(j+1)^\eta \log x_j}{l} \right)^{k_j} \frac{1}{(j!)^l ((j+1)!)^{\eta l}}. \end{aligned}$$

By the definition of  $\alpha_0, \dots, \alpha_{j-1}$ , we have

$$\sum_{i=0}^{j-1} \alpha_i \left[ \left( \frac{j+1}{i+1} \right)^\eta - 1 \right] \log x_j = lj = k_j.$$

Again,  $j! \geq (j/e)^j$  and also

$$\frac{(j+1)^{\eta k_j}}{((j+1)!)^{\eta l}} \leq \frac{e^{\eta l(j+1)}}{(j+1)^{\eta l}} \leq e^{\eta k_j} \quad (j \geq 2).$$

Hence,

$$\sum_{h'_1, \dots, h'_{k_j}} \frac{1}{h'_1 \cdots h'_{k_j}} \leq 2^{2k_j^2/(\eta \log x_j)} \left( \frac{\eta e^{2+\eta} \log x_j}{k_j} \right)^{k_j}.$$

By (6.13),  $\log x_j \geq \log x_r \gg k_j (\log k_j)^2$ , and thus

$$(6.22) \quad |Q_{j,2}| \ll \frac{k_j x}{\log x_j} (c_1 \eta e^{2+\eta})^{k_j} \exp \left\{ O \left( \frac{k_j \log_2 k_j}{\log k_j} + b_j \log_2 k_j \right) \right\}.$$

Take  $\eta$  so that  $c_1 \eta e^{2+\eta} < 1$ . If we take  $c_1$  arbitrarily close to  $2e^{-1}$ , we may take  $\eta = 0.15718$ . Put also

$$l = \lfloor (\log x)^\varepsilon \rfloor, \quad r = \lfloor (\log x)^\beta \rfloor,$$

where  $\varepsilon$  and  $\beta$  are fixed positive real numbers satisfying

$$0 < \varepsilon + \beta < \frac{1}{4 + \eta}.$$

Then  $\log x_r \asymp (\log x)^{1-\eta\beta}$ . Also put

$$b_j = \left\lfloor \frac{k_j}{(\log_2 k_j)^2} \right\rfloor.$$

Since  $\eta \leq \frac{1}{2}$  and  $\beta + \varepsilon < \frac{1}{4}$ , (6.13) holds. By (6.20) and (6.22), we conclude that

$$\sum_{j=1}^{r-1} |Q_j| \ll x \exp\{-(\log x)^\delta\},$$

where  $\delta > 0$  depends on  $\eta$ ,  $\varepsilon$ , and  $\beta$ . Take  $\varepsilon$  very close to 0 and  $\beta$  very close to  $\frac{1}{4+\eta} - \varepsilon$ . Then, for  $p$  not in some  $Q_j$ ,

$$H(p) \leq \frac{\log x_r}{\log 2} + k_r \ll (\log x)^{1-0.0378}.$$

□

## 7. PRATT TREE HEIGHT: UPPER BOUNDS, II

Here, we take a different approach to the problem of bounding the height of Pratt trees, focusing on counting prime chains via the method from Section 4. Recall from Section 4, the definition of the matrices  $M(q)$  and dominant eigenvalue  $\lambda(s; y)$ . Define

$$B = \sup_{s>1, y \geq 2} \frac{1 - \lambda(s; y)}{s - 1}.$$

**Theorem 11.** *For every  $c > \frac{1}{B}$ , there is a positive constant  $\delta = \delta(c)$  so that the number of primes  $p \leq x$  with  $H(p) > c \log x$  is  $O_c(x^{1-\delta})$ .*

*Proof.* The number of primes  $p \leq x$  with  $H(p) \geq K$  is at most the number of prime chains with  $p_1 = 2$ ,  $p_k \leq x$ , and  $k \geq K - 1$ . To deal with prime entries which are  $\leq y$ , we'll allow any even link  $m_i$  for  $1 \leq i \leq l := \left\lfloor \frac{\log y}{\log 2} \right\rfloor$ . Let  $q$  be the product of the primes  $\leq y$ . In the same manner that Lemma 4.1 was proved, we deduce

$$P_k(x) \leq q \inf_{s>1} x^s (2^{-s} \zeta(s))^l R(M(q)^{k-1-l}).$$

Let  $0 < \varepsilon < B - \frac{1}{c}$ , and fix  $s$  and  $y$  such that  $\frac{1-\lambda(s;y)}{s-1} \geq B - \varepsilon$ . Then,  $P_k(x) \ll_{\varepsilon} x^s \lambda(s, y)^{k-1}$  for  $k \geq l$ . Summing over  $k > c \log x$  yields a total of

$$\ll_{\varepsilon} x^{1+(s-1)(1-c(B-\varepsilon))}$$

chains. Clearly this is also an upper bound for the number of primes  $p \leq x$  with  $H(p) > c \log x$ .  $\square$

**Remark.** Inequality (4.2) implies that  $B \geq 2 \log 2$ .

**Conjecture 6.**  $B = \infty$ .

In some sense, the entries of  $M(q)$  are “randomly” distributed. From the theory of random matrices with i.i.d. entries (e.g. [27]), one expects the Perron eigenvalue to be close to the average row sum of the matrix. A simple calculation reveals that the average row sum of  $M(q)$  is

$$(7.1) \quad \alpha(q)\beta(2) \prod_{p|\frac{q}{2}} \left( 1 - \left( \frac{1}{p-1} - \frac{1}{p^s-1} \right) \right).$$

If  $\lambda(s; y)$  is close to this value, standard prime number estimates then imply Conjecture 6.

Here, we give an argument giving a better bound for  $\lambda(s, y)$  than that given by  $\lambda(s, y) \leq R(M)$ . There is another approach, using  $\lambda(s; y) \leq R(M^k)^{1/k}$  and obtaining estimates for  $R(M^k)$ , which, together with some computer calculation, has yielded  $B \geq 3.2$ .

**Lemma 7.1.**  $B \geq 2 \log 2 + \frac{3}{5} \log 3 = 2.045 \dots$

*Proof.* Write  $q = \prod_{p \leq y} p$  and  $M = M(q, s)$ . We begin with the familiar fact that  $M$  and  $P^{-1}MP$  have the same eigenvalues for an invertible matrix  $P$ . In particular, this holds if  $\mathbf{z} = (z_i : i \in U_q)$  and  $P$  is the diagonal matrix with diagonal entries  $z_i$ . In particular,  $\lambda(s; y) \leq R(P^{-1}MP) = \max_j z_j^{-1} (M\mathbf{z})_j$ . If  $\mathbf{z}$  is an eigenvector corresponding to  $\lambda(s; y)$ , then  $\lambda(s; y) = R(P^{-1}MP)$ . Thus, our goal is to use a vector  $\mathbf{z}$  which is “close” to an eigenvector, while at the same time simple enough that we can estimate  $R(P^{-1}MP)$ .

One method of specializing  $\mathbf{z}$  is as follows. Let  $Q|q$ , and let  $\{x_j : j \in U_Q\}$  be positive real numbers. Put  $z_b = x_f$  if  $(b, q) = 1$  and  $b \equiv f \pmod{Q}$ . For any  $b$ , write  $d = (b-1, q)$

and  $b' = \frac{b-1}{d}$ . We have

$$\begin{aligned} \frac{(M\mathbf{z})_b}{z_b} &= z_b^{-1} \sum_{a \in U_q} S(a, b) z_a \\ &= z_b^{-1} d^{-s} \sum_{(k, q/d)=1} k^{-s} \sum_{\substack{a \in U_q \\ b' \equiv ak \pmod{q/d}}} z_a \\ &= z_b^{-1} d^{-s} \sum_{(k, q/d)=1} k^{-s} \sum_{g \in U_Q} x_g |\{a \in U_q : a \equiv g \pmod{Q}, b' \equiv ak \pmod{q/d}\}|. \end{aligned}$$

The system of congruences  $a \equiv g \pmod{Q}$ ,  $b' \equiv ak \pmod{q/d}$  has a solution if and only if  $b' \equiv gk \pmod{q'}$ , where  $q' = (q/d, Q) = Q/(d, Q)$ . In the case where solutions exists, there are precisely  $\phi(q/[Q, q/d]) = \phi(d/(d, Q))$  such solutions  $a \in U_q$ . Writing  $b \equiv f \pmod{Q}$ , we have

$$(7.2) \quad \frac{(M\mathbf{z})_b}{z_b} = \frac{\phi(d)d^{-s}}{x_f \phi((d, Q))} \sum_{g \in U_Q} x_g \sum_{\substack{(k, q/d)=1 \\ kg \equiv b' \pmod{q'}}} k^{-s}.$$

We specialize to the case  $Q = 3$ ,  $x_1 = x$ ,  $x_2 = 1$ . Suppose  $b \in U_q$  and  $b \equiv f \pmod{Q}$ . Write  $d = (b-1, q)$  and  $q' = Q/(Q, d)$ . If  $f = 1$ , then  $3|d$  and  $q' = 1$ . By (7.2),

$$\frac{(M\mathbf{z})_b}{z_b} = \frac{\phi(d)d^{-s}}{2x} (1+x) \sum_{(k, q/d)=1} k^{-s} = \alpha(q)\beta(d) \frac{1+x}{2x} \leq \alpha(q)\beta(2) \left( \frac{1+x}{2x} \beta(3) \right).$$

If  $f = 2$ , then  $3 \nmid d$  and  $q' = 3$ . We also have  $b' \equiv d \pmod{3}$ . Hence, by (7.2),

$$\begin{aligned} \frac{(M\mathbf{z})_b}{z_b} &= \phi(d)d^{-s} \left[ x \sum_{\substack{(k, q/d)=1 \\ k \equiv d \pmod{3}}} k^{-s} + \sum_{\substack{(k, q/d)=1 \\ k \equiv 2d \pmod{3}}} k^{-s} \right] \\ &= \phi(d)d^{-s} \left[ \sum_{(k, q/d)=1} k^{-s} - (1-x) \sum_{\substack{(k, q/d)=1 \\ k \equiv d \pmod{3}}} k^{-s} \right]. \end{aligned}$$

In the second sum, let  $k = 2^t k_1$ , where  $(k_1, 2q/d) = 1$ . Given  $k_1$ , we sum  $2^{-ts}$  over  $t$  such that  $2^t \equiv dk_1^{-1} \pmod{3}$ , i.e., either over odd  $t$  or over even  $t$ . The sum over odd  $t$  of  $2^{-ts}$  is smaller than the corresponding sum over even  $t$ , hence

$$\sum_{\substack{(k, q/d)=1 \\ k \equiv d \pmod{3}}} k^{-s} \geq \sum_{(k_1, 2q/d)=1} k_1^{-s} \frac{2^{-s}}{1-2^{-2s}} = \frac{2^{-s}}{1+2^{-s}} \sum_{(k, q/d)=1} k^{-s}.$$

Moreover, we have not lost much in general with this seemingly crude inequality, since it is nearly equality when  $d = 2p$ , where  $p$  is a large prime such that  $p \equiv 1 \pmod{3}$ . In this

special case, either  $k_1 = 1$  and we sum over all odd  $t$ , or  $k_1 \geq p$ . We obtain

$$\begin{aligned} \frac{(M\mathbf{z})_b}{z_b} &\leq \phi(d)d^{-s} \left( \sum_{(k,q/d)=1} k^{-s} \right) \left( 1 - \frac{(1-x)2^{-s}}{2^{-s}+1} \right) \\ &= \alpha(q)\beta(d) \left( 1 - \frac{1-x}{2^s+1} \right) \leq \alpha(q)\beta(2) \left( 1 - \frac{1-x}{2^s+1} \right). \end{aligned}$$

Again, we lose little in the last inequality by considering the case when  $d = 2p$  and  $p$  is a large prime.

Considering both cases,  $f = 1$  and  $f = 2$ , we conclude that

$$\lambda(s, y) \leq \alpha(q)\beta(2) \max \left[ \frac{1+x}{2x} \beta(3), 1 - \frac{1-x}{2^s+1} \right].$$

If  $s = 1 + \eta$  for small  $\eta > 0$ , then the optimal value of  $x$  satisfies

$$x = 1 - \left( \frac{9}{5} \log 3 \right) \eta + O(\eta^2).$$

Note that  $\alpha(q)\beta(2) = R(M)$ . Hence

$$\lambda(s, y) \leq R(M) - \left( \frac{3}{5} \log 3 \right) \eta + O(\eta^2).$$

Combined with (4.2), this concludes the proof.  $\square$

## 8. STOCHASTIC MODEL OF PRATT TREES

In this section, we develop a model of the Pratt trees which explains Conjectures 3 and 4. Our heuristic argument is based on the distribution of the large prime factors of a random integer  $n \leq x$ . Factor  $n$  as  $n = \prod_{j=1}^{\Omega(n)} p_j(n)$ , with  $p_1(n) \geq p_2(n) \geq \dots$ . Put  $p_j(n) = 1$  for  $j > \Omega(n)$  and consider the infinite dimensional vector

$$S(n) = \left( \frac{\log p_1(n)}{\log n}, \frac{\log p_2(n)}{\log n}, \dots \right).$$

Note that the sum of the components of  $S(n)$  is 1. The distribution of the first component of  $S(n)$  has been greatly studied, the results having wide application in the theory of numbers; see for instance the comprehensive survey article [32]. We have

$$\mathbf{P} \left( \log p_1(n) \leq \frac{1}{u} \log n \right) = \rho(u),$$

where  $\rho$  is the *Dickman function*, the unique continuous solution of the differential-delay equations (see Section 1 of [32])

$$(8.1) \quad \rho(u) = 1 \quad (0 \leq u \leq 1), \quad u\rho'(u) = -\rho(u-1) \quad (u > 1).$$

More precisely, if  $\mathcal{A}$  is a set of positive integers, by  $\mathbf{P}(n \in \mathcal{A})$  we mean  $\lim_{x \rightarrow \infty} \frac{1}{x} |\{n \leq x : n \in \mathcal{A}\}|$ . It follows that

$$(8.2) \quad \rho(u) = 1 - \log u \quad (1 \leq u \leq 2).$$

and, with a little bit more work (see [32], Cor. 2.3), that

$$(8.3) \quad \rho(u) = e^{-(1+o(1))u \log u} \quad (u \rightarrow \infty).$$

The complete distribution of  $S(n)$ , found by Billingsly in 1972 [14], corresponds to the Poisson-Dirichlet distribution with parameter 1,  $PD(1)$  for short (more precisely, for each  $k$ , the first  $k$  components of  $S(n)$  are distributed as the first  $k$  components of the  $PD(1)$  distribution). Thanks to Tenenbaum [45], we now know strong uniform versions, where  $k \rightarrow \infty$  as  $n \rightarrow \infty$ . The distribution of all the components in the  $PD(1)$  distribution can easily be expressed in terms of  $\rho$ ; namely, the first  $k$  components  $(y_1, \dots, y_k)$  of  $S(n)$  have joint density function

$$(8.4) \quad \frac{1}{y_1 \cdots y_k} \rho \left( \frac{1 - y_1 - \cdots - y_k}{y_k} \right)$$

There is a simpler, and very useful, characterization of the distribution, found by Donnelly and Grimmett [19]. Let  $U_1, U_2, \dots$  be independent random variables with uniform distribution on  $[0, 1]$ . Let  $\mathbf{x} = (x_1, x_2, \dots)$  be the infinite dimensional vector formed from the decreasing rearrangement of the numbers

$$(8.5) \quad y_1 = U_1, y_2 = (1 - U_1)U_2, y_3 = (1 - U_1)(1 - U_2)U_3, \dots$$

Then  $\mathbf{x}$  has the  $PD(1)$  distribution. The paper [19] gives a simple, transparent proof that  $(x_1, \dots, x_k)$  and the first  $k$  components of  $S(n)$  have the same distribution. For a discussion of other realizations of the  $PD(1)$  distribution, see Section 1 of [45].

Notice that  $\sum x_i = 1$  with probability 1. Thus, we can interpret the  $PD(1)$  distribution geometrically as a random partition of the unit interval  $[0, 1]$  into an infinite number of parts achieved by cutting  $[0, 1]$  at a random place (with uniform distribution), then cutting the right sub-interval at a random place, and so on.

**Conjecture 7.** *As  $p$  runs over the set of primes,  $S(p-1)$  has  $PD(1)$  distribution.*

Conjecture 7 is widely believed, and is a simple consequence of the Elliott-Halberstam conjecture. Unconditionally, we know little about primes in progressions to very large moduli; the distribution of small prime factors of  $p-1$  is handled by the Bombieri-Vinogradov theorem. Assuming Conjecture 7, we can construct a reasonable model of the distribution of the primes appearing at a given level of the Pratt tree for  $p$ . That is, assume that  $S(p-1)$  has  $PD(1)$  distribution, for each prime  $q|(p-1)$ ,  $S(q-1)$  has  $PD(1)$  distribution, the vectors  $S(q-1)$  for  $q|(p-1)$  are independent, and so forth. In simple language, the primes in the first level of the tree, on a logarithmic scale, correspond to a random partition of  $[0, 1]$ . The primes in the second level correspond to randomly partitioning each of the parts of the original partition, and so on. The entire procedure corresponds to what is

known as a discrete-time *random fragmentation process*, where one may also think of a process whereby objects are repeatedly broken up. Random fragmentation processes are a type of *branching process* in probability theory, and have been used to model a variety of physical phenomena (e.g., genetic mutations, planet formation) and the growth of certain data structures in computer science. The monograph [11] details the theory of continuous-time fragmentation processes. Discrete time fragmentation processes may be recast in the language of *branching random walks*, which we describe below.

Let  $M_n$  be the size of the largest object at time  $n$ . Then  $M_n$  is a model of  $Q_n := \frac{\log q_n}{\log p}$ , where  $q_n$  is the largest prime at level  $k$  of the tree. The event  $\{M_n < \frac{\log 2}{\log p}\}$  is a model of the statement “all the primes at level  $n$  of the Pratt tree for  $p$  are  $< 2$ ”, that is,  $H(p) < n$ . Thus,  $H(p)$  is modeled by the random variable  $T(\frac{\log 2}{\log p})$ , where

$$(8.6) \quad T(\varepsilon) = \min\{n : M_n \leq \varepsilon\}.$$

Under the assumption of the Elliott-Halberstam conjecture, Lamzouri [38] showed that  $Q_n$  has the same distribution as  $M_n$  (Lamzouri studies the distribution of  $P^+(\phi_n(m))$ , where  $\phi_n$  is the  $n$ -th iterate of  $\phi$ ; the same proofs give the distribution of  $P^+(\phi_n(p-1))$ ). Further, he gives explicit expressions for the distribution of  $M_n$  in terms of differential-delay equations: Let  $\rho_1(u) = \rho(u)$  from (8.1), and for  $n \geq 2$  let  $\rho_n(u) = 1$  ( $0 \leq u \leq 1$ ) and

$$(8.7) \quad u\rho_n(u) = \int_0^\infty \rho_n(u-t)\rho_{n-1}(t) dt \quad (u > 1).$$

On the Elliott-Halberstam conjecture, Lamzouri shows that

$$(8.8) \quad \mathbf{P}\{Q_n \leq 1/u\} = \mathbf{P}\{M_n \leq 1/u\} = \rho_n(u).$$

Using (8.7), he also established the right-tail estimate

$$(8.9) \quad \rho_n(u) = \left( \frac{1 + o(1)}{\log_{n-1}(u) \log_n(u)} \right)^u \quad (u \rightarrow \infty),$$

valid for each *fixed*  $n \geq 1$  (here  $\log_0(u) = u$ ). Our goal in this paper is to understand the distribution of  $M_n$  as  $n \rightarrow \infty$ . Our methods do not use (8.7); it is interesting to see if (8.7) can be used to obtain similar (or stronger) results to those below.

Create a tree structure from the random fragmentation process as follows: label the root node with zero, beneath the root node put an infinite number of child nodes, each corresponding to one of the fragments of the initial segment  $[0, 1]$ . Each of these nodes has an infinite number of child nodes, corresponding to the fragments in the second step of the process, and so on. Each node is labeled with the number  $-\log x$ , where  $x$  is the fragment size. This randomly labeled tree corresponds to a *branching random walk* (BRW); alternatively, if we labeled each node with  $x$  instead, the resulting tree is known as a *multiplicative cascade*. More generally, an initial ancestor is at the origin, and who forms the zeroth generation. This parent then produces children, the first generation, which are randomly displaced from the parent according to some law. Each of these children behaves

like an independent copy of the parent, their children randomly displaced from their parent according to the same law, and forming the second generation, and so on. In our case, each parent produces an infinite number of offspring, the displacements from their parent given by  $V = \{-\log y : y \in Z\}$ , where  $Z$  is a point set with  $PD(1)$  distribution. We'll say that  $V$  has  $LPD$  (logarithmic Poisson-Dirichlet) distribution from now on.

Let  $B_n$  be the minimum label (position) of an individual at time  $n$ , so that  $B_n = -\log M_n$ . The distribution of the analog of  $B_n$  for a general BRW has received much attention during the last 35 years. The first order behavior of  $B_n$  (law of large numbers) was determined in the 1970s by Biggins, Hammersley and Kingman ([12], [31], [37]). Under very general conditions on the law governing the walk,  $B_n \sim \gamma n$  almost surely (i.e., with probability 1), for some real number  $\gamma$ . Let  $\mathbf{z}$  denote the random set of displacements of children from their parent. In the notation of Biggins [12], if

$$m(\theta) = \mathbf{E} \sum_{z \in \mathbf{z}} e^{-\theta z}, \quad \mu(a) = \inf\{e^{\theta a} m(\theta) : \theta > 0\},$$

then  $\gamma = \inf\{a : \mu(a) > 1\}$ . In our case, we have  $m(\theta) = 1/\theta$  (see (8.13) below) and hence  $\mu(a) = ae$  and  $\gamma = 1/e$ . Hence, almost surely

$$(8.10) \quad B_n \sim \frac{n}{e}.$$

This implies immediately that almost surely  $T(\varepsilon) \sim -e \log \varepsilon$  as  $\varepsilon \rightarrow 0$ . In particular,

$$T\left(\frac{\log 2}{\log p}\right) \sim e \log_2 p,$$

which justifies Conjecture 3.

Let  $b_n = \text{median}(B_n)$ . The study of  $B_n$  naturally breaks into two parts: (i) global behavior: asymptotics for  $b_n$ , and (ii) local behavior: the distribution of  $B_n - b_n$ .

We first address the global behavior of  $B_n$  and discuss more precise asymptotics for  $b_n$ . The sharpest known estimates which apply to our situation are due to McDiarmid [41]; there is a mistake in the proof of the lower bound for  $b_n$  in ([41], Theorem 2(a)), pointed out to us by Louigi Addario-Berry, but we need only the upper bound (Theorem 2(b)).

**Theorem 12.** *We have*

$$b_n = \frac{n}{e} + O(\log n).$$

Applying Theorem 12 and the definition (8.6), we arrive at

**Corollary 2.** *We have*

$$\text{median}(T(\varepsilon)) = e \log(1/\varepsilon) + O(\log_2(1/\varepsilon)).$$

The distribution of  $T(\varepsilon)$  for computer simulations of the random fragmentation process are given in Figure 2. The first is for  $\varepsilon = 0.0075 \approx \frac{\log 2}{\log 10^{40}}$ , which corresponds to the distribution of  $H(p)$  for primes around  $10^{40}$ . The second graph shows histograms with  $\varepsilon = 10^{-m}$  for  $m = 1, 2, 3, 4, 5$ .

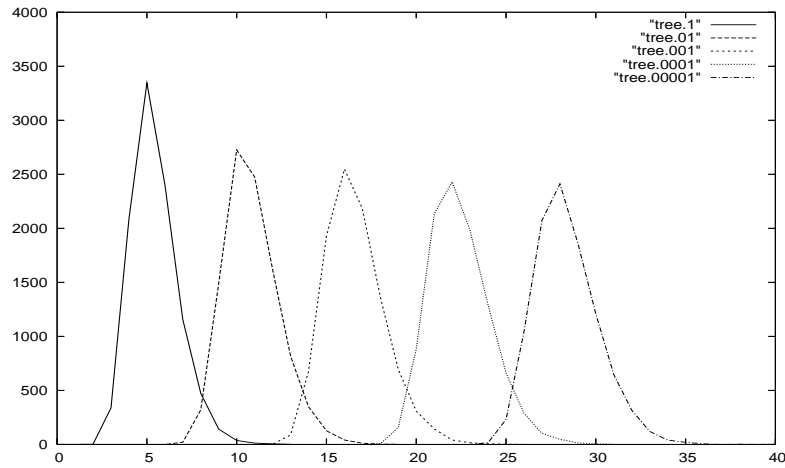
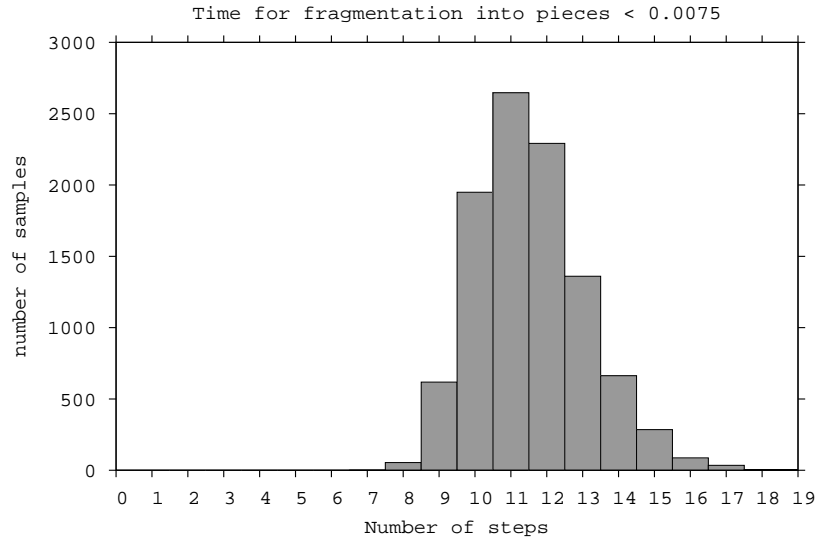


FIGURE 3.  $T(\varepsilon)$  for random fragmentation simulations. (i)  $\varepsilon = 0.0075$   
(ii)  $\varepsilon = 10^{-m}$ ,  $1 \leq m \leq 5$

Let  $Z_n(t)$  be the number of generation  $n$  individuals with position  $\leq t$ , and let  $\mathbf{z}^{(n)}$  be the set of positions of generation  $n$  individuals. For every  $n$ , with probability 1

$$(8.11) \quad \sum_{z \in \mathbf{z}^{(n)}} e^{-z} = 1.$$

**Lemma 8.1.** *For real  $s > 0$  and  $n \geq 1$ ,*

$$\mathbf{E} \sum_{z \in \mathbf{z}^{(n)}} e^{-sz} = \frac{1}{s^n},$$

Also,

$$(8.12) \quad \mathbf{E} Z_n(t) = \frac{t^n}{n!} \quad (n \geq 1, t \geq 0).$$

*Proof.* If  $\mathbf{v} = (v_1, v_2, \dots)$  has  $PD(1)$  distribution, then, by (8.5),

$$(8.13) \quad \begin{aligned} \mathbf{E} \sum_{j=1}^{\infty} v_j^s &= \mathbf{E} \sum_{k=1}^{\infty} ((1 - U_1) \cdots (1 - U_{k-1}) U_k)^s \\ &= \sum_{k=1}^{\infty} \int_0^1 \cdots \int_0^1 ((1 - u_1) \cdots (1 - u_{k-1}) u_k)^s du_1 \cdots du_k \\ &= \sum_{k=1}^{\infty} \frac{1}{(1+s)^k} = \frac{1}{s}. \end{aligned}$$

By the branching property,

$$\begin{aligned} \mathbf{E} \sum_{z_n \in \mathbf{z}^{(n)}} e^{-sz_n} &= \mathbf{E} \sum_{z_{n-1} \in \mathbf{z}^{(n-1)}} e^{-sz_{n-1}} \sum_{z \in \mathbf{z}} e^{-sz} \\ &= \frac{1}{s} \mathbf{E} \sum_{z_{n-1} \in \mathbf{z}^{(n-1)}} e^{-sz_{n-1}}. \end{aligned}$$

The first claim follows by induction on  $n$ . We have

$$\begin{aligned} \int_0^{\infty} e^{-st} d(\mathbf{E} Z_n(t)) &= \mathbf{E} \int_0^{\infty} e^{-st} dZ_n(t) \\ &= \mathbf{E} \sum_{z \in \mathbf{z}^{(n)}} e^{-sz} = \frac{1}{s^n}. \end{aligned}$$

Identity (8.12) follows from the uniqueness of the Laplace transform and the fact that  $\mathbf{E} Z_n(t)$  is increasing as a function of  $t$ . Incidentally, when  $s$  is fixed, the sequence of random variables

$$X_n = s^n \sum_{z_n \in \mathbf{z}^{(n)}} e^{-sz_n}$$

forms a martingale which is fundamentally important in the study of branching random walks.  $\square$

**Lemma 8.2.** *For real  $t \geq 1$  and integer  $k \geq 1$ , we have*

$$\mathbf{P}\{Z_1(t) \geq k\} \leq \left(\frac{et}{k}\right)^{k-1}.$$

*Proof.* The conclusion is trivial if  $k \leq et$ , so we suppose  $k > et$ . By (8.5),  $Z_1(t) \geq k$  implies  $(1 - U_1) \cdots (1 - U_{k-1}) \geq e^{-t}$ . Thus, for  $s \geq 0$ ,

$$\begin{aligned} \mathbf{P}\{Z_1(t) \geq k\} &\leq \mathbf{P}\{(1 - U_1) \cdots (1 - U_{k-1}) \geq e^{-t}\} \\ &\leq e^{st} \int_0^1 \cdots \int_0^1 [(1 - u_1) \cdots (1 - u_{k-1})]^s du_1 \cdots du_{k-1} \\ &= \frac{e^{st}}{(1 + s)^{k-1}}. \end{aligned}$$

Taking  $s = \frac{k}{t} - 1$  completes the proof.  $\square$

*Proof of Theorem 12.* By (8.12) and Stirling's formula,

$$\mathbf{P}\{B_n \leq \lambda\} = \mathbf{P}\{Z_n(\lambda) \geq 1\} \leq \mathbf{E}Z_n(\lambda) = \frac{\lambda^n}{n!} \asymp n^{-1/2} \left(\frac{e\lambda}{n}\right)^n.$$

Taking a large constant  $C_3$  and  $\lambda = \frac{n}{e} + \frac{\log n}{2e} - C_3$ , we find that  $\mathbf{P}(B_n \leq \lambda) \ll e^{-eC_3}$ . This gives

$$(8.14) \quad b_n \geq \frac{n}{e} + \frac{\log n}{2e} - O(1).$$

We next apply Theorem 2 (b) of [41]. In the notation of that paper, we have  $m = \infty$ ,  $F(t) = t$ ,  $\alpha = 0$ ,  $\phi(s) = 1/s$ ,  $\gamma = 1/e$  and  $\tau = e$ . If  $t \leq 1$ , the  $Z_1(t) \leq 2$ . By Lemma 8.2, if  $t > 1$  then

$$\mathbf{E}\{Z_1(t)^2\} \leq 9t^2 + \sum_{b=3}^{\infty} (b+1)^2 t^2 \mathbf{P}\{bt < Z_1(t) \leq (b+1)t\} \ll t^2,$$

so that condition (i) in Theorem 2 (b) of [41] is satisfied. For some constants  $C_4 > 0$  and  $\delta > 0$ , we conclude that

$$\mathbf{P}\left\{B_n \geq \frac{n}{e} + C_4 \log n\right\} \ll n^{-\delta}$$

and it follows that  $b_n \leq \frac{n}{e} + C_4 \log n$  for large  $n$ .  $\square$

Sharper estimates for the analog of  $b_n$  exist in the literature for special types of branching random walks. Our walk is sufficiently "nice" that we conjecture that similar behavior holds in our case. This leads to the

**Conjecture 8.** *We have*

$$b_n = \frac{n}{e} + \frac{3}{2e} \log n + C + o(1) \quad (n \rightarrow \infty)$$

for some constant  $C$ .

The nature of the second term in the conjectured asymptotic needs some explanation. Recall (8.12). The quantity  $t^n/n!$  is near 1 when  $t = \frac{n}{e} + \frac{1}{2e} \log n + O(1)$ , and a naive guess would be  $b_n = \frac{n}{e} + \frac{1}{2e} \log n + O(1)$ . However, for reasons clearly explained in [1], the leftmost point in the  $n$ -th generation of a branching random walk usually has an atypical ancestry. If the locations of points in the ancestral line of this leftmost point are  $0, w_1, w_2, \dots, w_n = B_n$  with  $B_n$  close to  $b_n$ , then we expect that

$$(8.15) \quad w_j \geq \frac{j}{n} w_n - O(1) \quad (1 \leq j \leq n/2).$$

We'll see later in Theorem 13 that the sequence  $B_n - b_n$  is *tight* (a sequence of random variables  $X_1, X_2, \dots$  is tight if for every  $\varepsilon > 0$  there is an  $K$  with  $\mathbf{P}\{|X_j| > K\} \leq \varepsilon$  for every  $j$ ). Thus, if  $w_j$  is much smaller than  $\frac{j}{n} w_n \approx \frac{j}{n} b_n$ , then with high probability there is an  $(n - j)$ -st generation descendant of  $w_j$  which has position  $\leq w_j + b_{n-j} + O(1)$ . Assuming that  $b_n - b_{n-j} = \frac{j}{n} b_n + O(1)$ , we have

$$w_j + b_{n-j} + O(1) = b_n + O(1) - \left( \frac{j}{n} b_n - w_j \right).$$

That is, failure of (8.15) implies that  $w_n$  is very likely much less than  $b_n$ .

Finally, we address how likely (8.15) is. Fix numbers  $g_j$  ( $1 \leq j \leq n/2$ ) and suppose  $w_1$  is the  $g_1$ -st child (measured from left to right) of  $w_1$ ,  $w_2$  is the  $g_2$ -st child of  $w_1$ , and so forth. We note that the sequence of displacements  $w_1, w_{j+1} - w_j$  ( $1 \leq j \leq n/2$ ) are independent random variables. They have different distributions (depending on the numbers  $g_j$ ), but for the sake of simplicity consider a sequence of independent, identically distributed continuous random variables  $X_1, \dots, X_k$  with given sum  $S$ . By considering all cyclic permutations of  $X_1, \dots, X_k$ , and using the cycle lemma from combinatorial theory, we see that the probability that  $X_1 + \dots + X_j \geq \frac{j}{k} S$  for every  $j$  is  $1/k$ . So, we expect that (8.15) occurs with probability of order  $1/n$ . If we denote by  $Z_n^*(t)$  the number of individuals  $w_n$  in generation  $n$  satisfying (8.15), we anticipate that

$$\mathbf{E}Z_n^*(t) \approx \frac{\mathbf{E}Z_n(t)}{n} = \frac{t^n}{n \cdot n!},$$

and the right side is  $\geq 1$  when  $t \geq \frac{n}{e} + \frac{3}{2e} \log n + O(1)$ .

It is possible that by adapting the methods in [1], one can prove the weaker bound

$$(8.16) \quad b_n = \frac{n}{e} + \frac{3}{2e} \log n + O(1).$$

We next discuss the local behavior of  $B_n$ . Under very general conditions on the BRW, it is known that  $B_n - b_n$  is a tight sequence. The basic idea is that a single individual will, with high probability, produce many offspring a few generations later which are close by. Previously, we have seen (Theorem 6) that  $H(p)$  does have a tight distribution with respect to its median, at least on one side. It is also known that under certain conditions

on the displacement law of  $\mathbf{z}$  (e.g., [5]), the analog of  $B_n - b_n$  converges in probability to a random variable as  $n \rightarrow \infty$ , and this random variable is a mixture of extreme value (Gumbel) distributions. In particular, the random variable has exponential decay on one side and double-exponential decay on the other. The existence of a limiting distribution is not known in our case, but we conjecture that it exists.

**Conjecture 9.**  $B_n - b_n \rightarrow X$  for a random variable  $X$  with continuous distribution. That is, there is a continuous distribution function  $F$  such that for each real number  $x$ ,  $\mathbf{P}\{B_n - b_n \leq x\} \rightarrow F(x)$  as  $n \rightarrow \infty$ .

If  $X$  exists, and the medians satisfy  $b_{n+1} - b_n \rightarrow e^{-1}$  (plausible in light of (8.10)), it is easy to see that  $X$  satisfies

$$(8.17) \quad X \stackrel{d}{=} -\frac{1}{e} + \min_i (z_i + X_i),$$

where  $(z_1, z_2, \dots)$  has *LPD* distribution,  $X_1, X_2, \dots$  are independent copies of  $X$ , and  $\stackrel{d}{=}$  means “has the same distribution as”. The relation (8.17) follows by conditioning on the positions of the first generation individuals (the points  $z_i$ ); that is, using

$$B_n \stackrel{d}{=} \min_i \left( z_i + B_{n-1}^{(i)} \right),$$

where  $B_{n-1}^{(i)}$  are independent copies of  $B_{n-1}$ . For a survey of results about equations like (8.17), we refer the reader to [3], especially Section 5 therein.

**Problem 4.** Solve “explicitly” the equation (8.17). Find asymptotics for the tails of the distribution.

We next prove tail probabilities of  $B_n - b_n$ . Unconditionally (whether  $X$  exists or not) we prove that the tails are exponentially decreasing. Consequently, if Conjecture 9 holds, then all moments of  $X$  exist.

**Theorem 13.** (a) For any  $c_1 < 1$ , we have

$$\mathbf{P}\{B_n - b_n \leq -x\} \ll e^{-c_1 x} \quad (n \geq 1, x \geq 0).$$

(b) For any  $c_2 < \frac{1}{8e}$ , we have

$$\mathbf{P}\{B_n - b_n \geq x\} \ll e^{-c_2 x} \quad (n \geq 1, x \geq 0).$$

An immediate consequence of Theorem 13 is that the collection of variables ( $T(\varepsilon)$  – median  $T(\varepsilon)$ ), for  $0 < \varepsilon \leq 1$ , is also exponentially tight.

The proof of Theorem 13 is based on the following lemmas. The first two lemmas hold for very general branching random walks. A notable feature is that they are *local* results, and tightness of  $B_n - b_n$  can be proved without knowing anything about the growth of  $b_n$ . We will use some growth information for  $b_n$  to prove the stronger “exponential tightness” in Theorem 13.

**Lemma 8.3.** *For positive integers  $m, n$  and positive real numbers  $M, N$ ,*

$$\mathbf{P}\{B_{m+n} \geq M + N\} \leq \mathbf{E}[(\mathbf{P}\{B_n \geq N\})^{Z_m(M)}].$$

*Proof.* Suppose  $B_{m+n} \geq M + N$  and  $Z_m(M) = k$ . For each of these  $k$  individuals, all of their descendants in generation  $m + n$  are offset from their generation  $m$  ancestor by at least  $N$ .  $\square$

**Lemma 8.4.** *Let  $m, n$  be positive integers and let  $M > 0$  be real. If  $\mathbf{E}\{(1 - \varepsilon)^{Z_m(M)}\} \leq \frac{1}{2}$ , where  $\varepsilon > 0$ , then*

$$\mathbf{P}\{B_n \leq b_{n+m} - M\} \leq \varepsilon.$$

*In particular, if  $\mathbf{P}\{Z_m(M) < 1/\varepsilon\} \leq \frac{1}{5}$ , then*

$$\mathbf{P}\{B_n \leq b_{n+m} - M\} \leq \varepsilon.$$

*Proof.* Let  $q$  be the  $\varepsilon$ -quantile of  $B_n$ , that is,  $\mathbf{P}\{B_n \leq q\} = \varepsilon$ . By Lemma 8.3,

$$\mathbf{P}\{B_{m+n} \geq M + q\} \leq \mathbf{E}\left[(\mathbf{P}\{B_n \geq q\})^{Z_m(M)}\right] \leq \frac{1}{2}.$$

Therefore,  $M + q \geq b_{m+n}$ , and thus  $\mathbf{P}\{B_n \leq b_{m+n} - M\} \leq \mathbf{P}\{B_n \leq q\} = \varepsilon$ . To prove the second part, observe that

$$\begin{aligned} \mathbf{E}\{(1 - \varepsilon)^{Z_m(M)}\} &\leq \mathbf{P}\{Z_m(m) < \frac{1}{\varepsilon}\} + (1 - \mathbf{P}\{Z_m(M) < \frac{1}{\varepsilon}\})(1 - \varepsilon)^{1/\varepsilon} \\ &\leq \frac{1}{5} + \frac{4}{5e} < \frac{1}{2}. \end{aligned}$$

$\square$

The next lemma concerns the growth of our branching random walk to the right of the minimal position, and is due to Biggins [13, Theorem 2]. It is complementary to (8.12).

**Lemma 8.5.** *For any  $a > 1/e$  and  $\eta > 0$ , we have*

$$\mathbf{P}\{Z_n(an) < (ae - \eta)^n\} \rightarrow 0 \quad (n \rightarrow \infty).$$

*Proof of Theorem 13.* Let  $a > 1/e$  and  $\eta > 0$ . By Lemma 8.5, if  $k$  is large enough,

$$\mathbf{P}\{Z_k(ak) \leq (ae - \eta)^k\} \leq \frac{1}{5}.$$

Apply Lemma 8.4 with  $M = ak$ ,  $m = k$ ,  $\varepsilon = (ae - \eta)^{-k}$ . We have, for large integers  $k$ ,

$$\mathbf{P}\{B_n \leq b_n - ak\} \leq \mathbf{P}\{B_n \leq b_{n+k} - ak\} \leq (ae - \eta)^{-k}.$$

The first estimate follows with  $c_1 = \frac{\log(ae - \eta)}{a}$ . We optimize by taking  $a = 1$  and  $\eta$  sufficiently small.

For the second part, we prove a version of a special case of Lemma 8.5, but with an explicit probability estimate. We claim

$$(8.18) \quad \mathbf{P}\{Z_m(4m) \leq e^{m/2e}\} \ll e^{-c_1 m/2e}.$$

From (8.18), we take  $m = \lfloor x/4 \rfloor + 1$ . By Lemma 8.3,

$$\mathbf{P}\{B_n \geq b_n + x\} \leq \mathbf{P}\{Z_m(4m) \leq e^{m/2e}\} + 2^{-\exp\{m/2e\}} \ll e^{-c_1 m/2e} \ll e^{-c_1 x/8e}.$$

To prove (8.18), suppose that  $Z_m(4m) \leq e^{m/2e}$ . Since  $c_1 < 1$ , the probability that  $B_m \leq \frac{m}{2e} + 1$  is  $O(e^{-c_1 m/2e})$  by (8.14) and part (a) of Theorem 13. Now suppose  $B_m > \frac{m}{2e} + 1$ . By (8.11),

$$1 = \sum_{z \in \mathbf{z}^{(m)}} e^{-z} \leq e^{m/2e} e^{-m/2e-1} + \sum_{\substack{z \in \mathbf{z}^{(m)} \\ z \geq 4m}} e^{-z},$$

hence the sum on the right side is  $\geq \frac{1}{2}$ . For some  $k \geq 4m$ , there are at least  $\frac{1}{10} \left(\frac{e}{2}\right)^{4m} 2^k$  points of  $\mathbf{z}^{(m)}$  lying in  $[k, k+1)$ , since otherwise

$$\sum_{\substack{z \in \mathbf{z}^{(m)} \\ z \geq 4m}} e^{-z} \leq \frac{1}{10} \sum_{k \geq 4m} \left(\frac{e}{2}\right)^{4m-k} < \frac{1}{2}.$$

By (8.12) and Stirling's formula, for each  $k$ ,

$$\begin{aligned} \mathbf{P}\left\{Z_m(k+1) \geq \frac{1}{10} \left(\frac{e}{2}\right)^{4m} 2^k\right\} &\leq \frac{\mathbf{E}Z_m(k+1)}{\frac{1}{10} \left(\frac{e}{2}\right)^{4m} 2^k} \\ &= \frac{10(k+1)^m}{m!(e/2)^{4m} 2^k} \ll k \left(\frac{k}{m}\right)^m 2^{-k} \ll 2^{-k/2}. \end{aligned}$$

Summing on  $k$  gives  $\mathbf{P}\{B_m > m/2e + 1, Z_m(4m) \leq e^{m/2e}\} \ll 2^{-2m}$  and completes the proof of (8.18).  $\square$

The local behavior of the sequence  $b_1, b_2, \dots$  is also very important.

**Lemma 8.6.** *We have  $b_{n+1} - b_n \leq \log 3$  for all  $n \geq 1$ . Further,*

$$b_{m+n} \leq b_m + b_n + O(1),$$

*uniformly in  $m \geq 1, n \geq 1$ .*

*Proof.* By (8.1) and (8.4),  $\mathbf{P}\{Z_1(M) = 0\} = \rho(e^M)$  and

$$\mathbf{P}\{Z_1(M) = 1\} = \int_{e^{-M}}^1 \frac{\rho(e^M(1-y))}{y} dy.$$

Taking  $M = \log 3$  and using numerical integration, we have  $\mathbf{P}\{Z_1(M) = 0\} = \rho(3) \leq 0.0487$  and  $\mathbf{P}\{Z_1(M) = 1\} \leq 0.8042$ . Thus,

$$\mathbf{E}\{2^{-Z_1(M)}\} \leq \rho(e^M) + \frac{\mathbf{P}\{Z_1(M) = 1\}}{2} + \frac{1 - \rho(e^M) - \mathbf{P}\{Z_1(M) = 1\}}{4} \leq \frac{1}{2}.$$

By Lemma 8.3 with  $N = b_n$ ,  $\mathbf{P}(B_{n+1} \geq b_n + M) \leq \frac{1}{2}$ , hence  $b_{n+1} - b_n \leq \log 3$ .

Now suppose  $m \geq 1$  and  $n \geq 1$ . By Theorem 13, there is a constant  $C$  so that for all  $k$ ,  $\mathbf{P}\{B_k \leq b_k - C\} \leq \frac{1}{4}$  and  $\mathbf{P}\{B_k \leq b_k + C\} \geq \frac{3}{4}$ . Applying Lemma 8.3,

$$\begin{aligned} \mathbf{P}\{B_{m+n} \geq b_m + b_n + C\} &\leq \mathbf{P}\{Z_m(b_m + C) = 0\} + \mathbf{P}\{B_n \geq b_n\} \\ &= \mathbf{P}\{B_m > b_m + C\} + \frac{1}{2} \leq \frac{3}{4}. \end{aligned}$$

Thus,  $b_{m+n} - C \leq b_m + b_n + C$ . □

**Remark.** Likely  $b_{n+1} - b_n \rightarrow e^{-1}$  as  $n \rightarrow \infty$  (this follows from Conjecture 8).

We are unable to prove that  $b_{n+1} - b_n$  is bounded away from zero. Such an estimate has profound implications for the tails of  $B_n - b_n$ . Actually, a weaker estimate suffices.

**Conjecture 10.** For some  $r \geq 1$  and  $\delta > 0$ ,  $b_{n+r} - b_n \geq \delta$  for all  $n \geq 1$ .

**Theorem 14.** Assume Conjecture 10. Then, (a) for any  $c_1 < \frac{1+\log a_0}{a_0-\delta/r}$ , where  $a_0 > 1/e$  satisfies  $a_0 \log a_0 > -\delta/r$ , the conclusion of Theorem 13 (a) holds; (b) for  $c'_1 = \frac{2r}{\delta} \log \frac{2r}{\delta}$ , we have

$$\mathbf{P}\{B_n \leq b_n - x\} \gg e^{-c'_1 x} \quad \left( n \geq 1, 0 \leq x \leq \frac{\delta n}{2r} \right);$$

and (c) for any  $c_3 < 1$  there is a constant  $c_4$ , depending on  $r, \delta$  and  $c_3$ , so that

$$\mathbf{P}\{B_n \geq b_n + x\} \leq \exp(-e^{c_3(x-c_4)}) \quad (n \geq 1, x \geq 0).$$

For part (c), we need an additional estimate about the growth of  $Z_n(t)$ .

**Lemma 8.7.** For all  $r \geq 1, \theta > 1$  and  $\varepsilon > 0$ , if  $x$  is large enough, then

$$\mathbf{P}\{Z_r(x) \geq \theta^x\} \leq \exp\{-(\theta - \varepsilon)^x\}.$$

*Proof.* When  $r = 1$ , this follows from Lemma 8.2. Assume true for some  $r \geq 1$ , let  $\theta, \varepsilon$  be given, and assume without loss of generality that  $\theta - \varepsilon > 1$ . The probability that  $Z_r(x) \geq (\theta - \varepsilon/3)^x$  is  $\leq \exp\{-(\theta - \varepsilon/2)^x\}$  for large  $x$ . Now suppose  $Z_r(x) = j < (\theta - \varepsilon/3)^x$  and  $Z_{r+1}(x) \geq \theta^x$ . Let  $m_i$  be the number of children of the  $i$ -th largest point in  $\mathbf{z}^{(r)}$  which are offset at most  $x$  from their parent. Let  $\mathcal{I}$  be the set of indices with  $m_i \geq 100x$ . Since  $m_1 + \dots + m_j \geq \theta^x$ , we have

$$\sum_{i \in \mathcal{I}} m_i \geq \theta^x - 100xj \geq \frac{\theta^x}{2}.$$

With  $j, m_1, \dots, m_j$  fixed, by Lemma 8.2, the probability that  $Z_{r+1}(x) \geq \theta^x$  is at most

$$\prod_{i=1}^j \mathbf{P}\{Z_1(x) \geq m_i\} \leq \prod_{i \in \mathcal{I}} e^{-2m_i} \leq \exp\{-\theta^x\}.$$

Since each  $m_i \leq e^x$ , the number of choices for  $j, m_1, \dots, m_j$  is at most

$$\sum_{j < (\theta - \varepsilon/3)^x} e^{xj} < \exp\{(\theta - \varepsilon/4)^x\}.$$

Therefore,

$$\mathbf{P}\{Z_r(x) \geq \theta^x\} \leq \exp\{-(\theta - \varepsilon/2)^x\} + \exp\{(\theta - \varepsilon/4)^x - \theta^x\} \leq \exp\{-(\theta - \varepsilon)^x\}$$

for large enough  $x$ .  $\square$

*Proof of Theorem 14.* Recall from the proof of Theorem 13 (a), that for any  $a > 1/e$  and any  $\eta > 0$ , if  $k$  is large enough, then

$$\mathbf{P}(B_n \leq b_{n+k} - ak) \leq (ae - \eta)^{-k}.$$

By assumption,  $b_{n+k} \geq b_n + (k/r - 1)\delta$ , so the conclusion of Theorem 13 (a) holds with  $c_1 = \frac{\log(ae - \eta)}{a - \delta/r}$ . By Theorem 12, we have  $\delta/r \leq 1/e$ . Take  $a_0 > 1/e$  so that  $a_0 \log a_0$  is close to  $-\delta/r$  and  $\eta$  sufficiently small to conclude part (a).

By (8.2), for  $0 \leq \varepsilon \leq \log 2$ ,

$$\mathbf{P}(Z_1(\varepsilon) \geq 1) = 1 - \rho(e^\varepsilon) = \varepsilon.$$

Considering the “leftmost child of the leftmost child of the . . . of the initial ancestor” in the branching random walk (corresponding to the special prime chain with length  $L(p)$  in the Pratt tree for  $p$ ), we have

$$\mathbf{P}(B_{kr} \leq \delta k/2) \geq \mathbf{P}(Z_1(\delta/2r) \geq 1)^{kr} = (\delta/2r)^{kr}$$

for every  $k \geq 1$ . Hence,

$$\mathbf{P}\{B_n \leq b_{n-kr} + \delta k/2\} \geq \mathbf{P}\{B_{n-kr} \leq b_{n-kr}\} \mathbf{P}\{B_{kr} \leq \delta k/2\} \geq \frac{1}{2} \left(\frac{\delta}{2r}\right)^{kr}.$$

By assumption,  $b_{n-kr} + \delta k/2 \leq b_n - \delta k/2$ . Hence, for  $0 \leq k \leq n/r$  we have

$$\mathbf{P}(B_n \leq b_n - \delta k/2) \geq \frac{1}{2} \left(\frac{\delta}{2r}\right)^{kr}.$$

This implies the lower bound in part (b) when  $0 \leq x \leq \frac{\delta n}{2r}$ , with  $c'_1 = \frac{2r}{\delta} \log \frac{2r}{\delta}$ .

To show part (c), we use induction on  $n$  to show that

$$(8.19) \quad \mathbf{P}(B_n \geq b_n + x) \leq 2^{-\exp\{c_3(x - c_5)\}}$$

for  $n \geq 1$  and  $x \geq 0$ , where  $c_5$  is sufficiently large. Theorem 13 (c) then follows with  $c_4 = c_5 - \frac{\log \log 2}{c_3}$ . Since (8.19) is trivial for  $0 \leq x \leq c_5$ , it suffices to consider the case  $x > c_5$ . Let  $A$  be a large integer, so that if  $R = Ar$  and  $\Delta = A\delta$ , then

$$(8.20) \quad e^{2-\Delta} \leq \frac{1 - c_3}{2}.$$

Also, without loss of generality, suppose  $c_3 \geq \frac{1}{2}$ . To get things started, observe that by (8.9), when  $1 \leq n \leq R$ ,

$$\mathbf{P}\{B_n \geq b_n + x\} \leq \mathbf{P}\{B_n \geq x\} = \rho_n(e^x) \leq \exp\{-e^x\}$$

if  $c_5$  is large enough. Suppose now that (8.19) has been proved for  $1 \leq n \leq m-1$ , where  $m-1 \geq R$ . Define

$$\lambda_j = \Delta + \frac{\log j - 1}{c_3} \quad (j \geq 1).$$

Let  $j_0$  be the largest index  $j$  with  $\lambda_j \leq x + \Delta$ . Let  $z_1 \leq z_2 \leq \dots$  be the points in  $\mathbf{z}^{(R)}$ . For  $1 \leq j \leq j_0$ , let  $P_j$  be the event  $\{z_i > \lambda_i \ (i < j), z_j \leq \lambda_j\}$ , and the  $Q$  be the event  $\{z_i > \lambda_i \ (1 \leq i \leq j_0)\}$ . If  $P_j$ , then the generation  $m$  points descending from each of the  $j$  points  $z_1, \dots, z_j$  are offset from their generation  $R$  ancestor by at least  $b_m + x - \lambda_j$ . By the induction hypothesis,

$$\begin{aligned} \mathbf{P}\{B_m \geq b_m + x\} &\leq \sum_{j=1}^{j_0} \mathbf{P}[P_j] \mathbf{P}\{B_{m-R} \geq b_m + x - \lambda_j\}^j + \mathbf{P}[Q] \\ &\leq \sum_{j=1}^{j_0} \mathbf{P}[P_j] \mathbf{P}\{B_{m-R} \geq b_{m-R} + x + \Delta - \lambda_j\}^j + \mathbf{P}[Q] \\ &\leq \sum_{j=1}^{j_0} \mathbf{P}[P_j] 2^{-j \exp\{c_3(x-c_5+\Delta-\lambda_j)\}} + \mathbf{P}[Q] \\ &\leq \sum_{j=1}^{j_0} \mathbf{P}[P_j] 2^{-\exp\{c_3(x-c_5)+1\}} + \mathbf{P}[Q] \\ &\leq 2^{-1-\exp\{c_3(x-c_5)\}} + \mathbf{P}[Q]. \end{aligned}$$

Now suppose  $Q$  holds. Then, by (8.20),

$$\sum_{j \leq j_0} e^{-z_j} \leq \sum_{j=1}^{j_0} e^{-\lambda_j} \leq e^{-\Delta+1/c_3} \sum_{j=1}^{\infty} j^{-1/c_3} \leq \frac{1}{2}.$$

As  $\lambda_{j_0} \geq x + \Delta - (\lambda_{j_0+1} - \lambda_{j_0}) \geq x$ , (8.11) implies

$$\sum_{\substack{z \in \mathbf{z}^{(R)} \\ z \geq x}} e^{-z} = 1 - \sum_{\substack{z \in \mathbf{z}^{(R)} \\ z < x}} e^{-z} \geq \frac{1}{2}.$$

Let  $\varepsilon = \frac{1}{3}(e - e^{c_3})$  and  $\theta = e - \varepsilon$ , so that  $e^{c_3} < \theta - \varepsilon < \theta < e$ . For some integer  $k \geq x$ , there are  $\geq \theta^k$  points of  $\mathbf{z}^{(R)}$  in  $[k-1, k)$ , for otherwise

$$\sum_{\substack{z \in \mathbf{z}^{(R)} \\ z \geq x}} e^{-z} \leq \sum_{k \geq x} e \left(\frac{\theta}{e}\right)^k < \frac{1}{2}.$$

By Lemma 8.7,  $\mathbf{P}\{Z_R(k) \geq \theta^k\} \leq \exp\{-(\theta - \varepsilon)^k\}$ . Summing over  $k$  gives

$$\mathbf{P}[Q] \leq 2 \exp\{-(\theta - \varepsilon)^x\} \leq 2^{-1-\exp\{c_3 x\}}.$$

This completes the proof of (c). □

**Remarks.** 1. Assuming Conjecture 10, the distribution of  $B_n - b_n$  is highly asymmetric, resembling a Gumbel-type extreme value distribution. Further, Theorem 14 (c) is best possible in the sense that the conclusion is false if  $c_3 > 1$ . This follows from (8.9):

$$\mathbf{P}\{B_n \geq b_n + x\} = \rho_n(e^{b_n+x}) = \exp\{-(1 + o(1))e^{x+b_n} \log_{n-1}(x + b_n)\}.$$

2. If (8.16) holds, then Conjecture 10 holds for all  $r$  sufficiently large and with  $\delta = r/e - O(1)$ . Consequently, we may take any  $c_1 < e$ , and any  $c'_1 > 2e \log(2e)$ . We conjecture that any  $c'_1 > e$  is admissible.

**Acknowledgments.** The authors wish to thank Louigi Addario-Berry, Jean Bertoin, Yakov Sinai, and Renming Song for helpful discussions about branching random walks. The authors thank Christian Elsholtz and Carl Pomerance for discussions about prime  $k$ -tuples, and Adolf Hildebrand for discussions about the distribution of prime factors integers. The authors also thank Imre Kátaı for bringing papers [35] and [36] to our attention.

Some of this work was accomplished while the third author visited the University of Illinois at Urbana-Champaign in February 2007, supported by NSF grant DMS-0555367. The first and second authors enjoyed the hospitality of the Institute for Advanced Study, where some of this work was done in November 2007. The first author was a participant in the NSF supported Workshop in Linear Analysis and Probability, Texas A&M University, August 2008.

## REFERENCES

- [1] L. Addario-Berry and B. Reed, *Minima in branching random walks*, Ann. Prob. (to appear).
- [2] M. Agrawal, N. Kayal and N. Saxena, *PRIMES is in P*, Ann. Math. **160** (2004), 781–793.
- [3] D. J. Aldous and A. Bandyopadhyay, *A survey of max-type recursive distributional equations*, Ann. Appl. Prob. **15** (2005), 1047–1110.
- [4] <http://hjem.get2net.dk/jka/math/Cunningham.Chain.records.htm>, web site maintained by Dirk Augustin.
- [5] M. Bachman, *Limit theorems for the minimal position in a branching random walk with independent logconcave displacements*, Adv. Appl. Prob. **32** (2000), 159–176.
- [6] R. C. Baker and G. Harman, *Shifted primes without large prime factors*, Acta Arith. **83** (1998), 331–361.
- [7] W. D. Banks, J. Friedlander, F. Luca, F. Pappalardi and I. E. Shparlinski, *Coincidences in the values of the Euler and Carmichael functions*, Acta Arith. **122** (2006), 207–234.
- [8] W. D. Banks and I. E. Shparlinski, *On values taken by the largest prime factor of shifted primes*, J. Aust. Math. Soc. **82** (2007), no. 1, 133–147.
- [9] N. L. Bassily, I. Kátaı and M. Wijsmuller, *Number of prime divisors of  $\phi_k(n)$ , where  $\phi_k$  is the  $k$ -fold iterate of  $\phi$* , J. Number Theory **65** (1997), 226–239.
- [10] J. Bayless, *The Lucas-Pratt primality tree*, Math. Comp. **77** (2008), 495–502.
- [11] J. Bertoin, *Random fragmentation and coagulation processes*, Cambridge Studies in Advanced Mathematics, 102. Cambridge University Press, Cambridge, 2006. viii+280 pp.
- [12] J. D. Biggins, *The first- and last-birth problems for a multitype age-dependent branching process*, Adv. Appl. Prob. **31** (1976), 446–459.
- [13] J. D. Biggins, *Chernoff’s theorem in the branching random walk*, J. Appl. Prob. **14** (1977), 630–636.

- [14] P. Billingsly, *On the distribution of large prime divisors*, Collection of articles dedicated to the memory of Alfréd Rényi, I. Period. Math. Hungar. **2** (1972), 283–289.
- [15] E. Bombieri, *Le grand crible dans la théorie analytique des nombres*, 2nd ed. (French. English summary) Astérisque No. 18. Société Mathématique de France, Paris, 1987.
- [16] M. A. Cherepnev, *Some properties of large prime divisors of numbers of the form  $p - 1$* , Mat. Zametki **80** (2006), 920–925. (Russian). English translation in Math. Notes **80** (2006), 863–867.
- [17] R. Crandall and C. Pomerance, *Prime numbers: a computational perspective*, 2nd ed., Springer-Verlag, 2005.
- [18] H. Davenport, *Multiplicative number theory*, 3rd ed., Graduate Texts in Mathematics vol. 74, Springer-Verlag, New York, 2000.
- [19] P. Donnelly and G. Grimmett, *On the asymptotic distribution of large prime factors*, J. London Math. Soc. (2) **47** (1993), 395–404.
- [20] C. Elsholtz, *Upper bounds for prime  $k$ -tuples of size  $\log N$  and oscillations*, Arch. Math. (Basel) **82** (2004), no. 1, 33–39.
- [21] P. Erdős, *On the normal number of prime factors of  $p - 1$  and some related problems concerning Euler's  $\phi$ -function*, Quart. J. Math. Oxford **6** (1935), 205–213.
- [22] P. Erdős, A. Granville, C. Pomerance, and C. Spiro, *On the normal behavior of the iterates of some arithmetic functions*, in Analytic Number Theory, Proceedings of a conference in honor of Paul T. Bateman, Birkhäuser, Boston, 1990, 165–204.
- [23] P. Erdős and M. Ram Murty, *On the order of  $a \pmod{p}$* . Number theory (Ottawa, ON, 1996), 87–97, CRM Proc. Lecture Notes, **19**, Amer. Math. Soc., Providence, RI, 1999.
- [24] P. Erdős and C. Pomerance, *On the normal number of prime factors of  $\phi(n)$* , Number theory (Winnipeg, Man., 1983). Rocky Mountain J. Math. **15** (1985), no. 2, 343–352.
- [25] K. Ford and F. Luca, *The number of solutions of  $\lambda(x) = n$* , preprint (2009).
- [26] K. Ford, F. Luca and C. Pomerance, *Common values of the arithmetic functions  $\phi$  and  $\sigma$* , submitted. arXiv:0906.3380
- [27] Z. Füredi and J. Komlós, *The eigenvalues of random symmetric matrices*, Combinatorica **1** (3) (1981), 233–241.
- [28] P. X. Gallagher, *On the distribution of primes in short intervals*, Mathematika **23** (1976), no. 1, 4–9.
- [29] B. Green and T. Tao, *The primes contain arbitrarily long arithmetic progressions*, Ann. Math. **167** (2008), 481–547.
- [30] H. Halberstam and H.-E. Richert, *Sieve methods*, Academic Press, London, 1974.
- [31] J. M. Hammersley, *Postulates for subadditive processes*, Ann. Prob. **2** (4) (1974), 652–680.
- [32] A. Hildebrand and G. Tenenbaum, *Integers without large prime factors*, J. de Théorie des Nombres de Bordeaux, **5** (1993), 411–484.
- [33] C. Hooley, *On Artin's conjecture*, J. Reine Angew. Math. **225** (1967), 209–220.
- [34] L. Kalmár, *Über die mittlere Anzahl Produktdarstellungen der Zahlen*, Acta Sci. Math. (Szeged) **5** (1930–32), 95–107.
- [35] I. Kátai, *Some problems on the iteration of multiplicative number-theoretical functions*, Acta Math. Acad. Scient. Hung. **19** (1968), 441–450.
- [36] I. Kátai, *On the iteration of multiplicative functions*, Publ. Math. Debrecen, **36** (1989), 129–134.
- [37] J. F. C. Kingman, *The first birth problem for an age-dependent branching process*, Ann. Prob. **3** (5) (1975), 790–801.
- [38] Y. Lamzouri, *Smooth values of iterates of the Euler phi-function*, Canad. J. Math. **59** (2007), 127–147.
- [39] F. Luca and C. Pomerance, *Irreducible radical extensions and Euler-function chains*, in “Combinatorial number theory”, 351–361, de Gruyter, Berlin (2007).

- [40] G. Martin and C. Pomerance, *The iterated Carmichael  $\lambda$ -function and the number of cycles of the power generator*, Acta Arith. **188** (2005), 305–335.
- [41] C. McDiarmid, *Minimal positions in a branching random walk*, Ann. Appl. Prob. **5** (1995), no. 1, 128–139.
- [42] R. M. Pollack, H. N. Shapiro, and G. N. Sparer, *On the graphs of I. Kátai*, Commun. Pure and Appl. Math., **27** (1974) 669–713.
- [43] C. Pomerance, *Very short primality proofs*, Math. Comp. **48** (1987), 315–322.
- [44] V. Pratt, *Every prime has a succinct certificate*, SIAM J. Comput. **4** (1975), no. 3, 214–220.
- [45] G. Tenenbaum, *A rate estimate in Billingsley's theorem for the size distribution of large prime factors*, Quart. J. Math. Oxford **51** (2000), no. 3, 385–403.

KF: DEPARTMENT OF MATHEMATICS, UNIVERSITY OF ILLINOIS AT URBANA-CHAMPAIGN URBANA,  
1409 WEST GREEN ST., URBANA, IL 61801, USA

*E-mail address:* ford@math.uiuc.edu

SVK: DEPARTMENT OF MECHANICS AND MATHEMATICS, MOSCOW STATE UNIVERSITY, MOSCOW,  
119992, RUSSIA

*E-mail address:* konyagin@ok.ru

FL : INSTITUTO DE MATEMÁTICAS, UNIVERSIDAD NACIONAL AUTONOMA DE MÉXICO, C.P. 58180,  
MORELIA, MICHOACÁN, MÉXICO

*E-mail address:* fluca@matmor.unam.mx