On the Tails of the Limiting QuickSort Density

James Allen Fill
Department of Applied Mathematics and Statistics, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218-2682, USA
jimfill@jhu.edu

Wei-Chun Hung
Department of Applied Mathematics and Statistics, The Johns Hopkins University, 3400 N. Charles Street, Baltimore, MD 21218-2682, USA
whung6@jhu.edu

Abstract

We give upper and lower asymptotic bounds for the left tail and for the right tail of the continuous limiting `QuickSort` density f that are nearly matching in each tail. The bounds strengthen results from a paper of Svante Janson (2015) concerning the corresponding distribution function F. Furthermore, we obtain similar upper bounds on absolute values of derivatives of f of each order.

2012 ACM Subject Classification Theory of computation → Sorting and searching, Mathematics of computing → Distribution functions

Keywords and phrases Quicksort, density tails, asymptotic bounds, Landau–Kolmogorov inequality

Digital Object Identifier 10.4230/LIPIcs.AofA.2018.21

Funding Partial support for both authors was provided by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.

Acknowledgements We are grateful to an anonymous referee for alerting us to the Landau–Kolmogorov inequality, which for upper bounds greatly simplified our proofs and improved our results somewhat. We also thank Svante Janson and another anonymous referee for helpful comments.

1 Introduction

Let X_n denote the (random) number of comparisons when sorting n distinct numbers using the algorithm `QuickSort`. Clearly $X_0 = 0$, and for $n \geq 1$ we have the recurrence relation

$$X_n \stackrel{\mathcal{L}}{=} X_{U_n - 1} + X^*_{n-U_n} + n - 1,$$

where $\stackrel{\mathcal{L}}{=}$ denotes equality in law (i.e., in distribution); $X_k \stackrel{\mathcal{L}}{=} X^*_k$; the random variable U_n is uniformly distributed on $\{1, \ldots, n\}$; and $U_n, X_0, \ldots, X_{n-1}, X^*_0, \ldots, X^*_{n-1}$ are all independent. It is well known that

$$\mathbb{E}X_n = 2\,(n+1)\,H_n - 4n,$$

where H_n is the nth harmonic number $H_n := \sum_{k=1}^{n} k^{-1}$ and (from a simple exact expression) that $\operatorname{Var} X_n = (1 + o(1))(7 - \frac{2\pi^2}{3})n^2$. To study distributional asymptotics, we first center and scale X_n as follows:

$$Z_n = \frac{X_n - \mathbb{E}X_n}{n}.$$

29th International Conference on Probabilistic, Combinatorial and Asymptotic Methods for the Analysis of Algorithms (AofA 2018).
Editors: James Allen Fill and Mark Daniel Ward; Article No. 21; pp. 21:1–21:10
Leibniz International Proceedings in Informatics
LIPICS Schloss Dagstuhl – Leibniz-Zentrum für Informatik, Dagstuhl Publishing, Germany

Using the Wasserstein d_2-metric, Rösler [8] proved that Z_n converges to Z weakly as $n \to \infty$. Using a martingale argument, Régnier [7] proved that the slightly renormalized $\frac{n}{n+1} Z_n$ converges to Z in L^p for every finite p, and thus in distribution; equivalently, the same conclusions hold for Z_n. The random variable Z has everywhere finite moment generating function with $\mathbb{E}Z = 0$ and $\operatorname{Var} Z = 7 - \left(2\pi^2/3\right)$. Moreover, Z satisfies the distributional identity

$$Z \stackrel{\mathcal{L}}{=} UZ + (1-U)Z^* + g(U).$$

On the right, $Z^* \stackrel{\mathcal{L}}{=} Z$; U is uniformly distributed on $(0,1)$; U, Z, Z^* are independent; and

$$g(u) := 2u \ln u + 2(1-u)\ln(1-u) + 1.$$

Further, the distributional identity together with the condition that $\mathbb{E}Z$ (exists and) vanishes characterizes the limiting `Quicksort` distribution; this was first shown by Rösler [8] under the additional condition that $\operatorname{Var} Z < \infty$, and later in full by Fill and Janson [1].

Fill and Janson [2] derived basic properties of the limiting `QuickSort` distribution $\mathcal{L}(Z)$. In particular, they proved that $\mathcal{L}(Z)$ has a (unique) continuous density f which is everywhere positive and infinitely differentiable, and for every $k \geq 0$ that $f^{(k)}$ is bounded and enjoys superpolynomial decay in both tails, that is, for each $p \geq 0$ and $k \geq 0$ there exists a finite constant $C_{p,k}$ such that $\left|f^{(k)}(x)\right| \leq C_{p,k}|x|^{-p}$ for all $x \in \mathbb{R}$.

In this paper, we study asymptotics of $f(-x)$ and $f(x)$ as $x \to \infty$. Janson [3] concerned himself with the corresponding asymptotics for the distribution function F and wrote this: "Using non-rigorous methods from applied mathematics (assuming an as yet unverified regularity hypothesis), Knessl and Szpankowski [4] found very precise asymptotics of both the left tail and the right tail." Janson specifies these Knessl–Szpankowski asymptotics for F in his equations (1.6)–(1.7). But Knessl and Szpankowski actually did more, producing asymptotics for f, which were integrated by Janson to get corresponding asymptotics for F. We utilize the same abbreviation $\gamma := (2 - \frac{1}{\ln 2})^{-1}$ as Janson [3]. With the *same* constant c_3 as in (1.6) of [3], the density analogues of (1.6) (omitting the middle expression) and (1.7) of [3] are that, as $x \to \infty$, Knessl and Szpankowski [4] find

$$f(-x) = \exp\left[-e^{\gamma x + c_3 + o(1)}\right] \tag{1}$$

for the left tail and

$$f(x) = \exp[-x \ln x - x \ln \ln x + (1 + \ln 2)x + o(x)] \tag{2}$$

for the right tail.

We will come as close to these non-rigorous results for the density as Janson [3] does for the distribution function and we also obtain corresponding asymptotic upper bounds for absolute values of derivatives of the density. Although our asymptotics for f imply the asymptotics for F in Janson's Theorem 1.1, it is important to note that in the case of upper bounds (but not lower bounds) we use his results in the proofs of ours.

The next two theorems are our main results.

▶ **Theorem 1.1.** *Let $\gamma := (2 - \frac{1}{\ln 2})^{-1}$. As $x \to \infty$, the limiting* `QuickSort` *density function f satisfies*

$$\exp\left[-e^{\gamma x + \ln \ln x + O(1)}\right] \leq f(-x) \leq \exp\left[-e^{\gamma x + O(1)}\right], \tag{3}$$

$$\exp[-x \ln x - x \ln \ln x + O(x)] \leq f(x) \leq \exp[-x \ln x + O(x)]. \tag{4}$$

▶ **Theorem 1.2.** *Given an integer $k \geq 0$, as $x \to \infty$ the k^{th} derivative of the limiting* `QuickSort` *density function f satisfies*

$$\left|f^{(k)}(-x)\right| \leq \exp\left[-e^{\gamma x + O(1)}\right], \tag{5}$$

$$\left|f^{(k)}(x)\right| \leq \exp[-x \ln x + O(x)]. \tag{6}$$

▶ Remark. The non-rigorous arguments of Knessl and Szpankowski [4] suggest that the following asymptotics as $x \to \infty$ obtained by repeated formal differentiation of (1)–(2) are correct for every $k \geq 0$:

$$f^{(k)}(-x) = \exp\left[-e^{\gamma x + c_3 + o(1)}\right], \tag{7}$$

$$f^{(k)}(x) = (-1)^k \exp[-x \ln x - x \ln \ln x + (1 + \ln 2)x + o(x)]. \tag{8}$$

But these remain conjectures for now. Unfortunately, for $k \geq 1$ we don't even know how to identify rigorously the asymptotic signs of $f^{(k)}(\mp x)$! Concerning $k = 1$, it has long been conjectured that f is unimodal. This would of course imply that $f'(-x) > 0$ and $f'(x) < 0$ for sufficiently large x.

As already mentioned, Fill and Janson [2] proved that or each $p \geq 0$ and $k \geq 0$ there exists a finite constant $C_{p,k}$ such that $\left|f^{(k)}(x)\right| \leq C_{p,k}|x|^{-p}$ for all $x \in \mathbb{R}$. Our technique for proving the upper bounds in Theorems 1.1 and 1.2 is to use explicit bounds on the constants $C_k := C_{0,k}$ together with the Landau–Kolmogorov inequality (see, for example, [9]).

Our extended abstract is organized as follows. In Section 2 we deal with preliminaries: We restate (to render this extended abstract self-contained) the asymptotic results of Janson [3, Theorem 1.1], bound C_k explicitly in terms of k, review the Landau–Kolmogorov inequality, and recall an integral equation for f that is the starting point for our lower-bound results. In Section 3 we establish the left-tail upper bounds on $|f^{(k)}|$ for $k \geq 0$ claimed in (3) and (5). In Section 4, we establish the right-tail upper bounds on $|f^{(k)}|$ for $k \geq 0$ claimed in (4) and (6). Sections 5 and 6 derive the stated lower bounds on the left and right tails, respectively, of f using an iterative approach similar to that of Janson [3] for the distribution function.

2 Preliminaries

2.1 Janson's asymptotic bounds on F

The upper bounds in the following main Theorem 1.1 of Janson [3] are used in our proof of the upper bounds in our Theorems 1.1 and 1.2.

▶ **Proposition 2.1.** *Let $\gamma := (2 - \frac{1}{\ln 2})^{-1}$. As $x \to \infty$, the limiting* `QuickSort` *distribution function F satisfies*

$$\exp\left[-e^{\gamma x + \ln \ln x + O(1)}\right] \leq F(-x) \leq \exp\left[-e^{\gamma x + O(1)}\right], \tag{9}$$

$$\exp[-x \ln x - x \ln \ln x + O(x)] \leq 1 - F(x) \leq \exp[-x \ln x + O(x)]. \tag{10}$$

2.2 Explicit constant bounds for absolute derivatives

We also make use of the following two results extracted from [2, Theorem 2.1 and (3.3)].

▶ **Lemma 2.2.** *Let ϕ denote the characteristic function corresponding to f. Then for every real $p \geq 0$ we have*

$$|\phi(t)| \leq 2^{p^2+6p}|t|^{-p} \quad \text{for all } t \in \mathbb{R}.$$

▶ **Lemma 2.3.** *For every integer $k \geq 0$ we have*

$$\sup_{x \in \mathbb{R}} |f^{(k)}(x)| \leq \frac{1}{2\pi} \int_{t=-\infty}^{\infty} |t|^k \, |\phi(t)| \, dt.$$

Using these two results, it is now easy to bound $f^{(k)}$.

▶ **Proposition 2.4.** *For every integer $k \geq 0$ we have*

$$\sup_{x \in \mathbb{R}} |f^{(k)}(x)| \leq 2^{k^2+10k+17}.$$

Proof. For every integer $k \geq 0$ we have

$$\begin{aligned}
\sup_{x \in \mathbb{R}} |f^{(k)}(x)| &\leq \frac{1}{2\pi} \int_{t=-\infty}^{\infty} |t|^k \, |\phi(t)| \, dt \\
&\leq \frac{1}{2\pi} \left[\int_{|t|>1} |t|^k \, |\phi(t)| \, dt + \int_{|t|\leq 1} |t|^k \, |\phi(t)| \, dt \right] \\
&\leq \frac{1}{2\pi} \left[\int_{|t|>1} 2^{(k+2)^2+6(k+2)} t^{-2} \, dt + \int_{|t|\leq 1} |t|^k \, dt \right] \\
&\leq \frac{1}{\pi} \left[2^{k^2+10k+16} + \frac{1}{k+1} \right] \leq 2^{k^2+10k+17},
\end{aligned}$$

as desired. ◀

2.3 Landau–Kolmogorov inequality

For an overview of the Landau–Kolmogorov inequality, see [6, Chapter 1]. Here we state a version of the inequality well-suited to our purposes; see [5] and [9, display (21) and the display following (17)].

▶ **Lemma 2.5.** *Let $n \geq 2$, and suppose $h : (0, \infty) \to \mathbb{R}$ has n derivatives. If h and $h^{(n)}$ are both bounded, then for $1 \leq k < n$ so is $h^{(k)}$. Moreover, there exist constants $c_{n,k}$ (not depending on h) such that the supremum norm $\| \cdot \|$ satisfies*

$$\|h^{(k)}\| \leq c_{n,k} \, \|h\|^{1-(k/n)} \, \|h^{(n)}\|^{k/n}, \quad 1 \leq k < n.$$

Further, for $1 \leq k \leq n/2$ the best constants $c_{n,k}$ satisfy

$$c_{n,k} \leq n^{(1/2)[1-(k/n)]} (n-k)^{-1/2} \left(\frac{e^2 n}{4k} \right)^k \leq \left(\frac{e^2 n}{4k} \right)^k.$$

2.4 An integral equation for f

Fill and Janson [2, Theorem 4.1 and (4.2)] produced an integral equation satisfied by f, namely,

$$f(x) = \int_{u=0}^{1} \int_{z \in \mathbb{R}} f(z) \, f\left(\frac{x - g(u) - (1-u)z}{u} \right) \frac{1}{u} \, dz \, du. \tag{11}$$

3 Left Tail Upper Bound for Absolute Derivatives

The left-tail upper bound (5) in Theorem 1.2 can be written in the equivalent form that, for each fixed integer $k \geq 0$, we have

$$\limsup_{x\to\infty} \left(\gamma x - \ln\left[-\ln\left|f^{(k)}(-x)\right|\right]\right) < \infty, \tag{12}$$

just as Janson's left-tail upper-bound on F in (9) can be written

$$\limsup_{x\to\infty} \left(\gamma x - \ln\left[-\ln F(-x)\right]\right) < \infty. \tag{13}$$

In this section we prove (5) $\equiv$ (12) in the strengthened form LHS(3.1) $\leq$ LHS(3.2), for which the following proposition is clearly sufficient.

▶ **Proposition 3.1.** *For each fixed $k \geq 0$ we have*

$$\limsup_{x\to\infty} \left(-\ln\left[-\ln\left|f^{(k)}(-x)\right|\right] + \ln[-\ln F(-x)]\right) \leq 0. \tag{14}$$

Proof. Choosing any x and applying the Landau–Kolmogorov inequality Lemma 2.5 to the function h defined for $t \geq 0$ by $h(t) := F(-x-t)$, we find for $0 \leq k \leq (n/2) - 1$ that

$$\begin{aligned}\left|f^{(k)}(-x)\right| &\leq \sup_{t\geq x}\left|f^{(k)}(-t)\right| \\ &\leq \left[\frac{e^2 n}{4(k+1)}\right]^{k+1} [F(-x)]^{1-[(k+1)/n]} \left[\sup_{t\geq x}\left|f^{(n-1)}(-t)\right|\right]^{(k+1)/n}.\end{aligned}$$

For $n \geq 2$ we can bound the last supremum using Proposition 2.4 simply by

$$2^{(n-1)^2+10(n-1)+17} = 2^{n^2+8n+8} \leq 2^{7n^2}. \tag{15}$$

Thus the argument of the lim sup in (14) can be bounded above by

$$-\ln\left[1 - \frac{k+1}{n} - \frac{\ln a_k + (k+1)(7n\ln 2 + \ln n)}{-\ln F(-x)}\right],$$

with $a_k := [e^2/(4(k+1))]^{k+1}$. Letting $n \equiv n(x) \to \infty$ with $n(x) = o(e^{\gamma x})$ and again using the upper bound from (9), the claimed inequality follows. ◀

4 Right Tail Upper Bound for Absolute Derivatives

In this section we establish the next proposition, a right-tail analogue of Proposition 3.1, which [by Janson's right-tail upper bound on F in (10)] implies the following strengthened form of (6):

$$\limsup_{x\to\infty} x^{-1}\left(x\ln x + \ln\left|f^{(k)}(x)\right|\right) \leq \limsup_{x\to\infty} x^{-1}\left(x\ln x + \ln[1-F(x)]\right) < \infty.$$

▶ **Proposition 4.1.** *For each fixed $k \geq 0$ we have*

$$\limsup_{x\to\infty} x^{-1}\left(\ln\left|f^{(k)}(x)\right| - \ln[1-F(x)]\right) \leq 0. \tag{16}$$

Proof. Proceeding as in the proof of Proposition 3.1, for any x and any $0 \le k \le (n/2) - 1$ we have

$$\left|f^{(k)}(x)\right| \le \left[\frac{e^2 n}{4(k+1)}\right]^{k+1} [1 - F(x)]^{1-[(k+1)/n]} \left[\sup_{t \ge x} \left|f^{(n-1)}(t)\right|\right]^{(k+1)/n};$$

we again bound the third factor by (15).

Thus the argument of the lim sup in (16) can be bounded above by

$$x^{-1}[\tfrac{k+1}{n}(-\ln[1 - F(x)]) + \ln a_k + (k+1)(7n \ln 2 + \ln n)],$$

again with $a_k := [e^2/(4(k+1))]^{k+1}$. Letting $n \equiv n(x)$ satisfy $n(x) = \omega(\log x)$ and $n(x) = o(x)$, and now using the right-tail *lower* bound on F from (10), the claimed inequality follows. ◀

5 Left Tail Lower Bound on f

Our iterative approach to finding the left tail lower bound on f is similar to the method used by Janson [3] for F. The following lemma gives us an inequality that is essential in this section; as we shall see, it is established from a recurrence inequality. For $z \ge 0$ define

$$m_z := \left(\min_{x \in [-z,0]} f(x)\right) \wedge 1.$$

▶ **Lemma 5.1.** *Given $\epsilon \in (0, 1/10)$, let $a \equiv a(\epsilon) := -g\left(\frac{1}{2} - \epsilon\right) > 0$. Then for any integer $k \ge 2$ we have*

$$m_{ka} \ge \left(2\epsilon^3 m_{2a}\right)^{2^{k-2}}.$$

We delay the proof of Lemma 5.1 in order to show next how the lemma leads us to the desired lower bound in (3) on the left tail of f by using the same technique as in [3] for F.

▶ **Proposition 5.2.** *As $x \to \infty$ we have*

$$\ln f(-x) \ge -e^{\gamma x + \ln \ln x + O(1)}.$$

Proof. By Lemma 5.1, for $x > a$ we have

$$f(-x) \ge m_x \ge m\left(\left\lceil \frac{x}{a} \right\rceil a\right) \ge \left(2\epsilon^3 m_{2a}\right)^{2^{\lceil x/a \rceil - 2}} \ge \left(2\epsilon^3 m_{2a}\right)^{2^{x/a}},$$

provided ϵ is sufficiently small that $2\epsilon^3 m_{2a} < 1$. The same as Janson [3], we pick $\epsilon = x^{-1/2}$ and, setting $\gamma = (2 - \frac{1}{\ln 2})^{-1}$, get $\frac{1}{a} = \frac{\gamma}{\ln 2} + O(x^{-1})$ and

$$\begin{aligned}
\ln f(-x) &\ge 2^{\frac{\gamma}{\ln 2} x + O(1)} \cdot \ln\left(2\epsilon^3 m_{2a}\right) \\
&= e^{\gamma x + O(1)} \cdot \left(-\tfrac{3}{2} \ln x + \ln m_{2a} + \ln 2\right) \\
&\ge -e^{\gamma x + \ln \ln x + O(1)}.
\end{aligned}$$

◀

Now we go back to prove Lemma 5.1:

Proof of Lemma 5.1. By the integral equation (11) satisfied by f (and symmetry in u about $u = 1/2$), for arbitrary z and a we have

$$f(-z - a) = 2 \int_{u=0}^{1/2} \int_{y \in \mathbb{R}} f(y) f\left(\frac{-z - a - g(u) - (1-u)y}{u}\right) \frac{1}{u} \, dy \, du. \tag{17}$$

Since f is everywhere positive, we can get a lower bound on $f(-z-a)$ by restricting the range of integration in (17). Therefore,

$$f(-z-a) \geq 2\int_{u=\frac{1}{2}-\frac{\epsilon}{2}}^{1/2} \int_{y=-z}^{-z+\epsilon^2} f(y) f\left(\frac{-z-a-g(u)-(1-u)y}{u}\right)\frac{1}{u}\,dy\,du. \tag{18}$$

We claim that in this integral region, we have $\frac{-z-a-g(u)-(1-u)y}{u} \geq -z$, which is equivalent to $y+z \leq \frac{-a-g(u)}{1-u}$. Here is a proof. Observe that when ϵ is small enough and $u \in [\frac{1}{2}-\frac{\epsilon}{2}, \frac{1}{2}]$, we have

$$\begin{aligned}
\frac{-a-g(u)}{1-u} &\geq \frac{g\left(\frac{1}{2}-\epsilon\right)-g\left(\frac{1}{2}-\frac{\epsilon}{2}\right)}{\frac{1}{2}+\frac{\epsilon}{2}} \\
&\geq \frac{\frac{\epsilon}{2}\left|g'\left(\frac{1}{2}-\frac{\epsilon}{2}\right)\right|}{\frac{1}{2}+\frac{\epsilon}{2}} = \frac{\epsilon}{1+\epsilon}\left|2\ln\left(1-\frac{2\epsilon}{1+\epsilon}\right)\right| \\
&\geq \frac{4\epsilon^2}{(1+\epsilon)^2} \geq \epsilon^2.
\end{aligned}$$

Also, in this integral region we have $y+z \leq \epsilon^2$. So we conclude that $y+z \leq \frac{-a-g(u)}{1-u}$.

Next, we claim that $\frac{-z-a-g(u)-(1-u)y}{u} \leq 0$ in this integral region if z is large enough. Here is a proof. Let $\frac{-z-a-g(u)-(1-u)y}{u} = -z+\delta$ with $\delta \geq 0$. Then in the integral region we have $0 \leq y+z = \frac{-a-g(u)-u\delta}{1-u}$. Therefore

$$\begin{aligned}
\delta \leq \frac{-a-g(u)}{u} &\leq \frac{-a-g\left(\frac{1}{2}\right)}{\frac{1}{2}-\frac{\epsilon}{2}} = \frac{2}{1-\epsilon}\left[g\left(\frac{1}{2}-\epsilon\right)-g\left(\frac{1}{2}\right)\right] \\
&\leq \frac{2\epsilon}{1-\epsilon}\left|2\ln\left(1-\frac{4\epsilon}{1+2\epsilon}\right)\right| \\
&\leq 19\epsilon^2,
\end{aligned}$$

where the last inequality can be verified to hold for $\epsilon < 1/10$. That means if we pick z large enough, for example, $z \geq 20\epsilon^2$, then $\frac{-z-a-g(u)-(1-u)y}{u} = -z+\delta$ will be negative. It can also be verified that $a \geq 30\epsilon^2$ for $\epsilon < 1/10$.

Now consider $\epsilon < 1/10$, an integer $k \geq 3$, $z \in [(k-2)a, (k-1)a]$, and $x = z+a \in [(k-1)a, ka]$. Noting $z \geq a \geq 30\epsilon^2 > 20\epsilon^2$, by (18) we have

$$f(-x) \geq 2 \cdot \frac{\epsilon}{2} \cdot m_z^2 \cdot \epsilon^2 \cdot 2 \geq 2\epsilon^3 m_{(k-1)a}^2.$$

Further, for $x \in [0, (k-1)a]$ we have

$$f(-x) \geq m_{(k-1)a} > 2\epsilon^3 m_{(k-1)a}^2$$

since $2\epsilon^3 < 1$ and $m_{(k-1)a} \leq 1$ by definition. Combine these two facts, we can conclude that for $x \in [0, ka]$ we have $f(-x) \geq 2\epsilon^3 m_{(k-1)a}^2$. This implies the recurrence inequality

$$m_{ka} \geq 2\epsilon^3 m_{(k-1)a}^2.$$

The desired inequality follows by iterating:

$$m_{ka} \geq \left(2\epsilon^3\right)^{2^{k-2}-1} m_{2a}^{2^{k-2}} \geq \left(2\epsilon^3 \cdot m_{2a}\right)^{2^{k-2}}. \qquad ◀$$

6 Right Tail Lower Bound on f

Once again we use an iterative approach to derive our right-tail lower bound. The following key lemma is established from a recurrence inequality. Define

$$c := 2[F(1) - F(0)] \in (0, 2)$$

and

$$m_z := \min_{x \in [0,z]} f(x), \quad z \geq 0.$$

▶ **Lemma 6.1.** *Suppose $b \in [0, 1)$ and that $\delta \in (0, 1/2)$ is sufficiently small that $g(\delta) \geq b$. Then for any integer $k \geq 1$ satisfying*

$$2 + (k - 1)b \leq [g(\delta) - b]/\delta$$

we have

$$m_{2+kb} \geq (c\delta)^{k-1} m_3.$$

We delay the proof of Lemma 6.1 in order to show next how the lemma leads us to the desired lower bound in (4) on the right tail of f.

▶ **Proposition 6.2.** *As $x \to \infty$ we have*

$$f(x) \geq \exp[-x \ln x - x \ln \ln x + O(x)].$$

Proof. Given $x \geq 3$ suitably large, we will show next that we can apply Lemma 6.1 for suitably chosen $b > 0$ and δ and $k = \lceil (x - 2)/b \rceil \geq 2$. Then, by the lemma,

$$f(x) \geq m_{2+kb} \geq (c\delta)^{k-1} m_3 \geq (c\delta)^{(x-2)/b} m_3, \tag{19}$$

and we will use (19) to establish the proposition.

We make the same choices of δ and b as in [3, Sec. 4], namely, $\delta = 1/(x \ln x)$ and $b = 1 - (2/\ln x)$. To apply Lemma 6.1, we need to check that $g(\delta) \geq b$ and $2 + (k - 1)b \leq [g(\delta) - b]/\delta$, for the latter of which it is sufficient that $x \leq [g(\delta) - b]/\delta$. Indeed, if x is sufficiently large, then

$$g(\delta) \geq 1 + 3\delta \ln \delta = 1 - \tfrac{3}{x \ln x}(\ln x + \ln \ln x) \geq 1 - \tfrac{4}{x},$$

where the elementary first inequality is (4.1) in [3], and so

$$g(\delta) - b \geq \tfrac{2}{\ln x} - \tfrac{4}{x} \geq \tfrac{1}{\ln x} > 0$$

and

$$\frac{g(\delta) - b}{\delta} \geq \frac{1/\ln x}{1/(x \ln x)} = x.$$

Finally, we use (19) to establish the proposition. Indeed,

$$\begin{aligned} -\ln f(x) &\leq \tfrac{x-2}{b} \ln(\tfrac{1}{c\delta}) - \ln m_3 \\ &\leq \tfrac{x}{1-(2/\ln x)} [\ln(x \ln x) + \ln(\tfrac{1}{c})] - \ln m_3 \\ &= \tfrac{x}{1-(2/\ln x)} \ln(x \ln x) + O(x). \end{aligned}$$

But

$$\begin{aligned}
&\frac{x}{1-(2/\ln x)}\ln(x\ln x)\\
&\quad = x\left[1+\frac{2}{\ln x}+O\left(\frac{1}{(\log x)^2}\right)\right](\ln x+\ln\ln x)\\
&\quad = (x\ln x)\left[1+\frac{2}{\ln x}+O\left(\frac{1}{(\log x)^2}\right)\right]\left(1+\frac{\ln\ln x}{\ln x}\right)\\
&\quad = (x\ln x)\left[1+\frac{\ln\ln x}{\ln x}+\frac{2}{\ln x}+\frac{2\ln\ln x}{(\ln x)^2}+O\left(\frac{1}{(\log x)^2}\right)\right]\\
&\quad = x\ln x+x\ln\ln x+2x+\frac{2x\ln\ln x}{\ln x}+O\left(\frac{x}{\log x}\right)\\
&\quad = x\ln x+x\ln\ln x+O(x).
\end{aligned}$$

So

$$-\ln f(x)\le x\ln x+x\ln\ln x+O(x),$$

as claimed. ◀

Now we go back to prove Lemma 6.1, but first we need two preparatory results.

▶ **Lemma 6.3.** *Suppose $z\ge 2$, $b\ge 0$, and $\delta\in(0,1/2)$ satisfy $g(\delta)\ge b$ and $z\le[g(\delta)-b]/\delta$. Then*

$$f(z+b)\ge c\,\delta\,m_z.$$

Proof. By the integral equation (11) satisfied by f (and symmetry in u about $u=1/2$), for arbitrary z and b we have

$$f(z+b)=2\int_{u=0}^{1/2}\int_{y\in\mathbb{R}}f(y)f\left(\frac{z+b-g(u)-(1-u)y}{u}\right)\frac{1}{u}\,dy\,du.$$

Since f is positive everywhere, a lower bound on $f(z+b)$ can be achieved by shrinking the region of integration:

$$\begin{aligned}
f(z+b)&\ge 2\int_{u=0}^{\delta}\int_{y=0}^{z}f(y)f\left(\frac{z+b-g(u)-(1-u)y}{u}\right)\frac{1}{u}\,dy\,du\\
&\ge 2m_z\int_{u=0}^{\delta}\int_{y=0}^{z}f\left(\frac{z+b-g(u)-(1-u)y}{u}\right)\frac{1}{u}\,dy\,du\\
&= 2m_z\int_{u=0}^{\delta}\int_{\xi=z+\frac{b-g(u)}{u}}^{\frac{z+b-g(u)}{u}}f(\xi)\frac{1}{1-u}\,d\xi\,du.
\end{aligned}\tag{20}$$

The equality comes from a change of variables. We next claim that the integral of integration for ξ contains $(0,z-1)$, and then the desired result follows. Indeed, if $u\in(0,\delta)$ and $\xi\in(0,z-1)$ then

$$\xi<z-1<\tfrac{z-1}{u}\le\tfrac{z+b-g(u)}{u},$$

where the last inequality holds because $b\ge 0$ and $g(u)\le 1$; and, because $g(u)\ge g(\delta)$ and $g(\delta)\ge b$ and $z\le[g(\delta)-b]/\delta$, we have

$$\begin{aligned}
&\xi>0=z+\tfrac{b-g(u)}{u}-[z+\tfrac{b-g(u)}{u}]\ge z+\tfrac{b-g(u)}{u}-[z+\tfrac{b-g(\delta)}{u}]\\
&\quad\ge z+\tfrac{b-g(u)}{u}-[z+\tfrac{b-g(\delta)}{\delta}]\ge z+\tfrac{b-g(u)}{u}.
\end{aligned}$$

◀

▶ **Lemma 6.4.** *Suppose $b \geq 0$ and that $\delta \in (0, 1/2)$ is sufficiently small that $g(\delta) \geq b$. Then for any integer $k \geq 2$ satisfying*

$$2 + (k-1)b \leq [g(\delta) - b]/\delta$$

we have

$$m_{2+kb} \geq c\,\delta\, m_{2+(k-1)b}.$$

Proof. For $y \in [2 + (k-1)b, 2 + kb]$, application of Lemma 6.3 with $z = y - b$ yields

$$f(y) \geq c\,\delta\, m_{y-b} \geq c\,\delta\, m_{2+(k-1)b}.$$

Also, for $y \in [0, 2 + (k-1)b]$ we certainly have

$$f(y) \geq m_{2+(k-1)b} > c\,\delta\, m_{2+(k-1)b}.$$

The result follows. ◀

We are now ready to complete this section by proving Lemma 6.1.

Proof of Lemma 6.1. By iterating the recurrence inequality of Lemma 6.4, it follows that

$$m_{2+kb} \geq (c\,\delta)^{k-1} m_{2+b}.$$

Lemma 6.1 then follows since $b < 1$. ◀

References

1 James Allen Fill and Svante Janson. A characterization of the set of fixed points of the Quicksort transformation. *Electron. Comm. Probab.*, 5:77–84, 2000. doi:10.1214/ECP.v5-1021.

2 James Allen Fill and Svante Janson. Smoothness and decay properties of the limiting quicksort density function. *Mathematics and Computer Science: Algorithms, Trees, Combinatorics and Probabilities, Trends in Mathematics*, pages 53–64, 2000.

3 Svante Janson. On the tails of the limiting quicksort distribution. *Electronic Communications in Probability*, 20, 2015.

4 Charles Knessl and Wojciech Szpankowski. Quicksort algorithm again revisited. *Discrete Mathematics and Theoretical Computer Science*, 3(2), 1999.

5 A. P. Matorin. On inequalities between the maxima of the absolute values of a function and its derivatives on a half-line. *Ukrain. Mat. Ž.*, 7:262–266, 1955.

6 D. S. Mitrinović, J. E. Pečarić, and A. M. Fink. *Inequalities involving functions and their integrals and derivatives*, volume 53 of *Mathematics and its Applications (East European Series)*. Kluwer Academic Publishers Group, Dordrecht, 1991. doi:10.1007/978-94-011-3562-7.

7 Mireille Régnier. A limiting distribution for quicksort. *RAIRO-Theoretical Informatics and Applications*, 23(3):335–343, 1989.

8 Uwe Rösler. A limit theorem for "quicksort". *RAIRO-Theoretical Informatics and Applications*, 25(1):85–100, 1991.

9 Sergei Borisovich Stechkin. Inequalities between the upper bounds of the derivatives of an arbitrary function on the half-line. *Mathematical notes of the Academy of Sciences of the USSR*, 1(6):442–447, 1967.