Electronic Journal of Probability

Electron. J. Probab. **24** (2019), no. 67, 1–13.
ISSN: 1083-6489 https://doi.org/10.1214/19-EJP331

QuickSort: improved right-tail asymptotics for the limiting distribution, and large deviations*

James Allen Fill† Wei-Chun Hung†

Abstract

We substantially refine asymptotic logarithmic upper bounds produced by Svante Janson (2015) on the right tail of the limiting `QuickSort` distribution function F and by Fill and Hung (2018) on the right tails of the corresponding density f and of the absolute derivatives of f of each order. For example, we establish an upper bound on $\log[1 - F(x)]$ that matches conjectured asymptotics of Knessl and Szpankowski (1999) through terms of order $(\log x)^2$; the corresponding order for the Janson (2015) bound is the lead order, $x \log x$.

Using the refined asymptotic bounds on F, we derive right-tail large deviation (LD) results for the distribution of the number of comparisons required by `QuickSort` that substantially sharpen the two-sided LD results of McDiarmid and Hayward (1996).

Keywords: `QuickSort`; asymptotic bounds; tails of distributions; large deviations; moment generating functions; Chernoff bounds.
AMS MSC 2010: Primary 68P10, Secondary 60E05; 60C05.
Submitted to EJP on January 19, 2019, final version accepted on June 7, 2019.
Supersedes arXiv:1903.07775.

1 Introduction

To set the stage, and for the reader's convenience, we repeat here relevant portions of Section 1 of Fill and Hung [2]. 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

*Research of both authors supported by the Acheson J. Duncan Fund for the Advancement of Research in Statistics.

†The Johns Hopkins University, Department of Applied Mathematics and Statistics, Baltimore, MD, USA.
E-mail: jimfill@jhu.edu,whung6@jhu.edu

independent. It is well known that

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

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:

$$Z_n = \frac{X_n - \mu_n}{n}. \tag{1.1}$$

Using the Wasserstein d_2-metric, Rösler [12] proved that Z_n converges to Z weakly as $n \to \infty$. Using a martingale argument, Régnier [11] 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

$$Z \stackrel{\mathcal{L}}{=} UZ + (1-U)Z^* + g(U). \tag{1.2}$$

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 [12] under the additional condition that $\operatorname{Var} Z < \infty$, and later in full by Fill and Janson [4].

Fill and Janson [5] 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.

Janson [7] studied logarithmic asymptotics in both tails for the corresponding distribution function F, and Fill and Hung [2] did the same for f and each of its derivatives. For right tails, all these results can be summarized in the following theorem. We let $\overline{F}(x) := 1 - F(x)$, and for a function $h : \mathbb{R} \to \mathbb{R}$ we write

$$\|h\|_x := \sup_{t \ge x} |h(t)|. \tag{1.3}$$

Theorem 1.1 ([7], Thm. 1.1; [2], Thms. 1.1–1.2)**.**
(a) *As $x \to \infty$, the limiting* `QuickSort` *density function f satisfies*

$$\exp[-x\ln x - x \ln\ln x + O(x)] \le f(x) \le \exp[-x \ln x + O(x)]. \tag{1.4}$$

(b) *Given an integer $k \ge 0$, as $x \to \infty$ the k^{th} derivative of the limiting* `QuickSort` *distribution function F satisfies*

$$\exp[-x\ln x - (k \vee 1)x \ln\ln x + O(x)] \le \|\overline{F}^{(k)}\|_x \le \exp[-x \ln x + O(x)]. \tag{1.5}$$

As discussed in [7, Section 1] and in [2, Remark 1.3(b)], non-rigorous arguments of Knessl and Szpankowski [8] suggest very refined asymptotics, which to three logarithmic terms assert that for each $k \ge 0$ we have

$$\overline{F}^{(k)}(x) = \exp[-x\ln x - x\ln\ln x + (1+\ln 2)x + o(x)] \tag{1.6}$$

as $x \to \infty$ (and hence that the same asymptotics hold for $\|\overline{F}^{(k)}\|_x$). Note that for $k = 0, 1$ these expansions match the lower bounds on f and $\overline{F}$ in Theorem 1.1 to two logarithmic terms.

In an earlier extended-abstract version [3] of this paper, we refined the upper bounds of Theorem 1.1 to match (1.6), and we were also able to improve the lower bound in (1.5) to match (1.6) to two terms. Here is the main theorem of [3]:

Theorem 1.2 ([3], Thm. 1.2)**.** (a) *As $x \to \infty$, the limiting* `QuickSort` *density function f satisfies*

$$\begin{aligned}\exp[-x \ln x - x \ln \ln x + O(x)] &\le f(x) &(1.7)\\ &\le \exp[-x \ln x - x \ln \ln x + (1 + \ln 2)x + o(x)]. &(1.8)\end{aligned}$$

(b) *Given an integer $k \ge 0$, as $x \to \infty$ the k^{th} derivative of the limiting* `QuickSort` *distribution function F satisfies*

$$\begin{aligned}\exp[-x \ln x - x \ln \ln x + O(x)] &\le \|\overline{F}^{(k)}\|_x &(1.9)\\ &\le \exp[-x \ln x - x \ln \ln x + (1 + \ln 2)x + o(x)]. &(1.10)\end{aligned}$$

In this paper we substantially refine the upper bound

$$\overline{F}(x) \le \exp[-x \ln x - x \ln \ln x + (1 + \ln 2)x + o(x)] \tag{1.11}$$

of Theorem 1.2(b) with $k = 0$; we also improve the upper bounds for $k \ge 1$, though not as dramatically.

Let

$$J(t) := \int_{s=1}^{t} \frac{2e^s}{s}\, ds, \quad t \ge 1. \tag{1.12}$$

It is elementary using integration by parts that $J(t)$ has the (divergent) asymptotic expansion

$$J(t) \sim 2t^{-1}e^t \sum_{j=0}^{\infty} j!\, t^{-j}. \tag{1.13}$$

Here is the main theorem of this paper:

Theorem 1.3. *For $x \ge 2e$, let $w \equiv w(x)$ denote the unique real solution satisfying $w \ge 1$ to*

$$x = 2w^{-1}e^w.$$

(a) *As $x \to \infty$, the limiting* `QuickSort` *distribution function F satisfies*

$$\begin{aligned}\overline{F}(x) &\le \exp[-xw + J(w) - w^2 + O(\log x)]\\ &= \exp[-2e^w + J(w) - w^2 + O(w)].\end{aligned}$$

(b) *Given an integer $k \ge 1$, as $x \to \infty$ the k^{th} derivative of the limiting* `QuickSort` *distribution function F satisfies*

$$\|\overline{F}^{(k)}\|_x \le \exp[-xw + J(w) + O(\sqrt{x \log x})]. \tag{1.14}$$

Remark 1.4. (a) We aid the reader in gauging the approximate sizes of the various terms in the bounds appearing in Theorem 1.3. It is routine to check that, as noted by Knessl and Szpankowski [8, eq. (20)],

$$w = \ln(x/2) + \ln \ln(x/2) + (1 + o(1))\frac{\ln \ln(x/2)}{\ln(x/2)} \tag{1.15}$$

as $x \to \infty$. Thus, by (1.13), we have the asymptotic equivalence

$$J(w) \sim 2w^{-1}e^w = x. \tag{1.16}$$

From (1.15)–(1.16) it's easy to see that Theorem 1.3 does indeed strengthen the upper bounds in Theorem 1.2. Inclusion of the term $J(w)$ in the bounds of Theorem 1.3 enables us effectively to bypass the entire infinite asymptotic expansion (1.13).

(b) Using non-rigorous methods, Knessl and Szpankowski [8, see esp. their eq. (18)] derive the following exact asymptotics for $\overline{F}(x)$ as $x \to \infty$:

$$\overline{F}(x) = \exp\left[-xw + J(w) - w^2 - (\alpha + \tfrac{1}{2})w - \tfrac{3}{2}\ln w + C - \ln(2\sqrt{\pi}) + o(1)\right] \tag{1.17}$$

for some (unspecified) constant C, with $\alpha := 2\ln 2 + 2\gamma - 1$, where γ denotes the Euler–Mascheroni constant. Hence the bound of Theorem 1.3(a) on $\ln \overline{F}(x)$ matches the conjectured asymptotics to within an additive term $O(w) = O(\log x)$.

(c) In their notation, the non-rigorously derived eq. (88) of [8] should read

$$P(y) \sim \frac{C_*}{\sqrt{2\pi}} \frac{1}{\sqrt{y}\, w_* \sqrt{1 - (1/w_*)}} \exp\left[-yw_* + \int_1^{w_*} \frac{2e^u}{u}\, du - w_*^2 - \alpha w_*\right],$$

recalling $\alpha = 2\gamma + 2\ln 2 - 1$. Ignoring the factor $\sqrt{1-(1/w^*)}$ which ~ 1, this result in our notation is

$$\begin{aligned} f(x) &\sim (2\pi \times 2w^{-1}e^{w})^{-1/2} e^{-xw}\psi(w) \\ &\sim (2\pi x)^{-1/2} \exp[-xw + J(w) - w^2 - \alpha w - \ln w + C], \end{aligned} \tag{1.18}$$

where ψ is the moment generating function corresponding to f and C is the same constant as at (1.17). [They derive their (88) by the "standard saddle point approximation" from the moment generating function expansion (2.4) recalled in Remark 2.2 below, and they derive (1.17) by integrating (1.18).] Hence the bound of Theorem 1.3(b) on $\ln f(x)$ matches the conjectured asymptotics (1.18) to within an additive term $O(\sqrt{x \log x})$.

We prove Theorem 1.3 in Section 2. In Section 3 we use our refined asymptotic bounds on F to derive right-tail large deviation results for the distribution of the number of comparisons required by `QuickSort` that sharpen somewhat the two-sided large-deviation results of McDiarmid and Hayward [9].

We conclude this section by repeating from [3] an open problem concerning *left*-tail behavior.

Open Problem. With $\underline{F}(x) := F(-x)$ can the *lower* bounds as $x \to \infty$ in the left-tail results

$$\exp\left[-e^{\Gamma x + \ln\ln x + O(1)}\right] \le f(-x) \le \exp\left[-e^{\Gamma x + O(1)}\right], \tag{1.19}$$

$$\exp\left[-e^{\Gamma x + \ln\ln x + O(1)}\right] \le \|\underline{F}^{(k)}\|_x \le \exp\left[-e^{\Gamma x + O(1)}\right] \tag{1.20}$$

of [7] and [2] be improved to match the asymptotics

$$\underline{F}^{(k)}(x) = \exp\left[-e^{\Gamma x + O(1)}\right]$$

suggested by Knessl and Szpankowski [8] (and known rigorously [7, 2] for *upper* bounds), where $\Gamma := (2 - \frac{1}{\ln 2})^{-1}$?

2 Proof of the main Theorem 1.3

In Section 2.1 we bound the moment generating function (mgf) ψ of Z. In Section 2.2 we prove Theorem 1.3(a) by combining the Chernoff bound

$$\overline{F}(x) = \mathbb{P}(Z \ge x) \le e^{-tx}\psi(t),$$

for judicious choice of $t \equiv t(x) > 0$, with our bound on ψ. In Section 2.3 we prove Theorem 1.3(b).

2.1 A bound on the moment generating function of Z

Let ψ denote the mgf of Z. It was shown by Rösler [12] that ψ is everywhere finite. In this subsection we establish a bound on $\psi(t)$ which (for large t) improves on that of [3, Lemma 2.1], which asserts that for every $\epsilon > 0$ there exists $a \equiv a(\epsilon) \geq 0$ such that the mgf ψ of Z satisfies

$$\psi(t) \leq \exp[(2+\epsilon)t^{-1}e^t + at] \tag{2.1}$$

for every $t > 0$. The bound (2.1) in turn improved the one obtained in the proof of [7, Lemma 6.1], namely, that there exists $a \geq 0$ such that

$$\psi(t) \leq \exp(e^t + at) \quad \text{for every } t \geq 0. \tag{2.2}$$

Recalling the definition (1.12) of $J(t)$, we next state our bound on $\psi(t)$ which, according to (1.13), does indeed improve on (2.1) for large t.

Proposition 2.1. *There exists a constant $a \geq 0$ such that the moment generating function ψ of Z satisfies*

$$\psi(t) \leq \exp[J(t) - t^2 + at] \tag{2.3}$$

for every $t \geq 1$.

We postpone the proof of Proposition 2.1 for a preliminary remark.

Remark 2.2. Using non-rigorous methods, Knessl and Szpankowski [8] derive that as $t \to \infty$ the mgf ψ satisfies

$$\psi(t) = \exp[J(t) - t^2 - \alpha t - \ln t + C + o(1)] \tag{2.4}$$

as $t \to \infty$ for the same (unspecified) constant C as at (1.17), with $\alpha = 2\ln 2 + 2\gamma - 1$; see their equation (71) (we have corrected a misplaced-right-parenthesis typo). If (2.4) is true, then our bound on $\ln \psi(t)$ agrees with the truth to within $O(t)$, whereas the bound (2.1) (for fixed ϵ) exceeds the true value by $(1+o(1))\epsilon t^{-1}e^t$. Thus our bound (2.3) comes substantially closer to the apparent truth than does (2.1).

The proof of Proposition 2.1 will require the following lemma. Recall from Remark 2.2 that $\alpha = 2\ln 2 + 2\gamma - 1$, and define

$$\widehat{\psi}(t) := \begin{cases} (1-e^{-t/2})\exp[J(t) - t^2 - \alpha t - \ln t] & \text{if } t > 1 \\ 1 & \text{otherwise.} \end{cases}$$

Lemma 2.3. *For all sufficiently large t we have the strict inequality*

$$2\int_{u=0}^{1/2} \widehat{\psi}(ut)\widehat{\psi}((1-u)t)\exp[tg(u)]\,du < \widehat{\psi}(t).$$

Proof. Call the left side of this inequality $\lambda(t)$. To handle $\lambda(t)$, we begin by changing the variable of integration from u to η, where $u = \frac{1}{2}e^{-t}\eta$:

$$\begin{aligned} \lambda(t) &= e^{-t}\int_{\eta=0}^{e^t} \widehat{\psi}\left(\frac{1}{2}te^{-t}\eta\right)\widehat{\psi}\left(t - \frac{1}{2}te^{-t}\eta\right)\exp\left[tg\left(\frac{1}{2}e^{-t}\eta\right)\right]\,d\eta \\ &= \int_{\eta=0}^{e^t} \widehat{\psi}\left(\frac{1}{2}te^{-t}\eta\right)\widehat{\psi}\left(t - \frac{1}{2}te^{-t}\eta\right)\exp\left[2t\phi\left(\frac{1}{2}e^{-t}\eta\right)\right]\,d\eta \end{aligned}$$

with $\phi(u) := u\ln u + (1-u)\ln(1-u) \leq 0$.

We next show that the contribution to $\int_{\eta=0}^{e^t}$ here from $\int_{\eta=e^{t/10}}^{e^t}$ is effectively quite negligible. To see this, we consider the integrand in two cases. Before breaking into cases, observe that the second argument for $\widehat{\psi}$ is at least $t/2$, which exceeds 1 if (as we may suppose) $t > 2$. For the first case, suppose that the first argument for $\widehat{\psi}$ also exceeds 1. In this case we need to treat the sum of the J-values at these arguments. But, using the increasingness of $2s^{-1}e^s$ for $s \geq 1$, we see that if $a, b \geq 1$ and $a + b = t$, then

$$\begin{aligned} J(a) + J(b) &= \int_{s=1}^{a} 2s^{-1}e^s\,ds + \int_{s=1}^{b} 2s^{-1}e^s\,ds \\ &\leq \int_{s=1}^{a} 2s^{-1}e^s\,ds + \int_{s=a}^{a+b-1} 2s^{-1}e^s\,ds = J(t-1) \end{aligned}$$

and therefore

$$\begin{aligned} J(a) + J(b) - J(t) &\leq -[J(t) - J(t-1)] = -\int_{s=t-1}^{t} 2s^{-1}e^s\,ds \\ &\leq -2(t-1)^{-1}e^{t-1} = -(1+o(1))\,2e^{-1}\,t^{-1}e^t. \end{aligned}$$

For the second case, suppose that the first argument for $\widehat{\psi}$ does not exceed 1. In this case we need to treat $J(t - \frac{1}{2}te^{-t}\eta) \leq J(t - \frac{1}{2}te^{-9t/10})$. In this case, observe that

$$\begin{aligned} J(t - \tfrac{1}{2}te^{-9t/10}) - J(t) &\leq (\tfrac{1}{2}te^{-9t/10}) \cdot -2(t - \tfrac{1}{2}te^{-9t/10})^{-1}\exp[t - \tfrac{1}{2}te^{-9t/10}] \\ &= -(1+o(1))e^{t/10}. \end{aligned}$$

The minor contribution $\int_{\eta=e^{t/10}}^{e^t}$ is thus bounded between 0 and

$$\begin{aligned} &(e^t - e^{t/10}) \times \exp[J(t) - (1+o(1))e^{t/10} + O(t^2)] \times 1 \\ &\quad= \exp[J(t) - (1+o(1))e^{t/10} + O(t^2)] \\ &\quad= \exp[-(1+o(1))e^{t/10}]\widehat{\psi}(t). \end{aligned}$$

For the major contribution $\int_{\eta=0}^{e^{t/10}}$, we can use simple expansions for the first and third factors in the integrand, because $0 \leq \frac{1}{2}te^{-t}\eta \leq \frac{1}{2}te^{-9t/10} = o(1)$:

$$\begin{aligned} \widehat{\psi}\left(\tfrac{1}{2}te^{-t}\eta\right) &= 1, \\ \phi\left(\tfrac{1}{2}e^{-t}\eta\right) &= \tfrac{1}{2}e^{-t}\eta(-t + \ln\eta - \ln 2) - \tfrac{1}{2}e^{-t}\eta + O(e^{-2t}\eta^2). \end{aligned}$$

We also use an expansion for $J(t - \frac{1}{2}te^{-t}\eta)$ appearing in the second factor in the integrand:

$$\begin{aligned} J(t - \tfrac{1}{2}te^{-t}\eta) - J(t) &= -\tfrac{1}{2}te^{-t}\eta J'(t) + \tfrac{1}{8}t^2e^{-2t}\eta^2 J''(t) + O(t^2e^{-2t}\eta^3) \\ &= -\eta + \tfrac{1}{4}(t-1)e^{-t}\eta^2 + O(t^2e^{-2t}\eta^3). \end{aligned}$$

Thus, abbreviating $t - \frac{1}{2}te^{-t}\eta$ as $t_1 \equiv t_1(t,\eta)$, the major contribution to $\lambda(t)$ equals

$$\exp[J(t)]I(t),$$

where $I(t)$ is the integral

$$\begin{aligned} I(t) := \int_{\eta=0}^{e^{t/10}} & e^{-\eta}(1 - e^{-t_1/2})\exp\left[\tfrac{1}{4}(t-1)e^{-t}\eta^2 + O(t^2e^{-2t}\eta^3)\right. \\ &\left. + te^{-t}\eta(-t + \ln\eta - \ln 2) - te^{-t}\eta + O(te^{-2t}\eta^2) - t_1^2 - \alpha t_1 - \ln t_1\right] d\eta. \end{aligned}$$

We now use the following additional expansions:

$$\begin{aligned} t_1^2 &= t^2 - t^2e^{-t}\eta + O(t^2e^{-2t}\eta^2), \\ \ln t_1 &= \ln(t - \tfrac{1}{2}te^{-t}\eta) = \ln t - \tfrac{1}{2}e^{-t}\eta + O(e^{-2t}\eta^2), \\ e^{-t_1/2} &= e^{-t/2}[1 + \tfrac{1}{4}te^{-t}\eta + O(t^2e^{-2t}\eta^2)]. \end{aligned}$$

Further we can expand the factor $\exp[\cdot]$ appearing in $I(t)$ as $1+\cdot+O(\cdot^2)$, because $\cdot = o(1)$ uniformly throughout the range of integration.

Calculus now gives

$$\begin{aligned} I(t) &= (1-e^{-t/2})\exp[-t^2-\alpha t-\ln t] \\ &\quad \times \Big[O(t^4 e^{-2t}) + \int_{\eta=0}^{\infty} e^{-\eta}(1-\tfrac{1}{4}te^{-3t/2}\eta+\tfrac{1}{4}(t-1)e^{-t}\eta^2 \\ &\qquad + te^{-t}\eta(-t+\ln\eta-\ln 2) - te^{-t}\eta + t^2e^{-t}\eta + \tfrac{1}{2}\alpha te^{-t}\eta + \tfrac{1}{2}e^{-t}\eta)\,d\eta\Big] \\ &= (1-e^{-t/2})\exp[-t^2-\alpha t-\ln t]\times[1-\tfrac{1}{4}te^{-3t/2}+O(t^4e^{-2t})]. \end{aligned}$$

We conclude for sufficiently large t that

$$\lambda(t) = \widehat{\psi}(t)[1-\tfrac{1}{4}te^{-3t/2}+O(t^4e^{-2t})] < \widehat{\psi}(t). \qquad \square$$

Remark 2.4. If we change the factor $(1-e^{-t/2})$ in the definition of $\widehat{\psi}$ to $(1+e^{-t/2})$, then a similar proof shows that the reverse strict inequality holds in Lemma 2.3. In fact, the proof becomes a bit simpler, since the minor contribution can simply be bounded below by 0.

Proof of Proposition 2.1. We carry out the proof by showing that there exists $a' \ge 0$ such that

$$\psi(t) \le e^{a't}\widehat{\psi}(t) \tag{2.5}$$

for every $t > 0$.

To begin, we compare asymptotics of $\psi(t)$ and $\widehat{\psi}(t)$ as $t \to 0$. Because Z has zero mean and finite variance, we have $\psi(t) = 1 + O(t^2)$. On the other hand, $\widehat{\psi}(t) = 1$ for all $0 < t \le 1$. We can thus choose $t_1 > 0$ and $a'' > 0$ such that (2.5) holds for $t \in [0, t_1]$ and any $a' \ge a''$.

Let $t_2 > 1$ be such that the strict inequality in Lemma 2.3 holds for all $t \ge t_2$, and choose $a' \ge a''$ so that (2.5) holds for $t \in [t_1, t_2]$. Assuming for the sake of contradiction that (2.5) fails for some $t > 0$, let $T := \inf\{t > 0 : \text{(2.5) fails}\}$. Then $T \ge t_2$, and continuity gives

$$\psi(T) = e^{a'T}\widehat{\psi}(T).$$

Further, if $0 < u < 1$, then (2.5) holds for $t = uT$ and $t = (1-u)T$, and thus, using our standard integral equation for ψ, we have

$$\psi(T) \le e^{a'T} \times 2\int_{u=0}^{1/2} \widehat{\psi}(uT)\widehat{\psi}((1-u)T)\exp[tg(u)]\,du,$$

which is strictly smaller than $e^{a'T}\widehat{\psi}(T)$ by applying Lemma 2.3 with $t = T \ge t_2$. The resulting strict inequality $\psi(T) < e^{a'T}\widehat{\psi}(T)$ contradicts the definition of T. Hence (2.5) holds for all $t \ge 0$. $\square$

Remark 2.5. Using Remark 2.4 just as Lemma 2.3 is used in the proof of Proposition 2.1, we have the following reverse of Proposition 2.1: There exists a constant $a \ge 0$ such that the mgf ψ of Z satisfies

$$\psi(t) \ge \exp[J(t) - t^2 - at] \tag{2.6}$$

for every $t \ge 1$.

Remark 2.6. (a) Unfortunately, due to the need to handle small values of t in the proofs of Proposition 2.1 and Remark 2.5, we sacrifice the information in the linear term of $\ln\psi(t)$ that Remark 2.2 and Lemma 2.3 strongly suggest. Thus any further progress on asymptotic determination of ψ would have to employ a technique different from the one used to derive (2.2), (2.1), and (2.3).

(b) The extent to which we are able to make rigorous the claim (2.4) and thereby, in particular, identify the linear term in $\ln \psi(t)$ is the following. If

$$\psi(t) = \exp[J(t) + K(t)]$$

where we *assume* $K'(t) = O(t^{b_1})$ and $K''(t) = O(t^{b_2})$ for some b_1 and b_2 [just as we now know rigorously that $K(t) \sim -t^2 = O(t^2)$], then we must have

$$K(t) = -t^2 - \alpha t - \ln t + C + O(t^b e^{-t})$$

for some constant C, where $b := \max\{4, 2 + 2b_1, 2 + b_2\}$. (Aside: It is natural to *assume further* that $b_1 = 1$ and $b_2 = 0$, in which case $b = 4$.) The proof of this assertion is quite similar to the proof of Lemma 2.3 and is omitted.

2.2 Proof of improved asymptotic upper bound on $\overline{F}$

Proof of Theorem 1.3(a). Choose $t = w$, apply the Chernoff bound

$$\overline{F}(x) = \mathbb{P}(Z \geq x) \leq e^{-tx}\psi(t),$$

and utilize Proposition 2.1 to establish Theorem 1.3(a). □

Remark 2.7. (a) For large x, the optimal choice of t for the Chernoff bound combined with (2.3) is not $t = w$, but rather the larger $\widetilde{w} \equiv \widetilde{w}(x)$ of the two positive real solutions to

$$x = 2(\widetilde{w}^{-1}e^{\widetilde{w}} - \widetilde{w}) + a.$$

But the resulting improvement in the bound on $\ln \overline{F}(x)$ not only is subsumed by the error bound $O(\log x)$ but in fact is asymptotically equivalent to $2x^{-1}(\log x)^2 = o(1)$ and so is negligible even as concerns estimating $\overline{F}(x)$ to within a factor $1 + o(1)$.

Here is a proof. Use of $t = w$ vs. $t = \widetilde{w}$ gives the larger expression

$$-xw + J(w) - w^2 + aw$$

vs.

$$-x\widetilde{w} + J(\widetilde{w}) - \widetilde{w}^2 + a\widetilde{w};$$

the increase is

$$\Delta \equiv \Delta(x) := x(\widetilde{w} - w) - [J(\widetilde{w}) - J(w)] + (\widetilde{w}^2 - w^2) - a(\widetilde{w} - w).$$

Using Taylor's theorem, we write

$$\begin{aligned} J(\widetilde{w}) - J(w) &= 2w^{-1}e^w(\widetilde{w} - w) + t^{-1}e^t(1 - t^{-1})(\widetilde{w} - w)^2 \\ &= x(\widetilde{w} - w) + (1 + o(1))\tfrac{1}{2}x(\widetilde{w} - w)^2 \end{aligned}$$

where t belongs to $(w, \widetilde{w})$, and we also note

$$\widetilde{w}^2 - w^2 = 2(\widetilde{w} - w)(\widetilde{w} + w) \sim 2(\widetilde{w} - w)\ln x.$$

Thus

$$\Delta = -(1 + o(1))\tfrac{1}{2}x(\widetilde{w} - w)^2 + (1 + o(1))2(\widetilde{w} - w)\ln x.$$

It remains to estimate $\widetilde{w} - w$. We have

$$1 = \frac{x}{x} = \frac{2(\widetilde{w}^{-1}e^{\widetilde{w}} - \widetilde{w}) + a}{2w^{-1}e^w} = \frac{w}{\widetilde{w}}e^{\widetilde{w} - w} - \frac{2\widetilde{w} - a}{x}.$$

Write this as

$$\frac{w}{\widetilde{w}} e^{\widetilde{w}-w} = 1 + \frac{2\widetilde{w} - a}{x}$$

and take logs. Note

$$\ln\left(\frac{w}{\widetilde{w}} e^{\widetilde{w}-w}\right) = -\ln\left(1 + \frac{\widetilde{w} - w}{w}\right) + \widetilde{w} - w \sim \widetilde{w} - w$$

and

$$\ln\left(1 + \frac{2\widetilde{w} - a}{x}\right) \sim \frac{2\ln x}{x}.$$

Thus

$$\widetilde{w} - w \sim 2x^{-1} \ln x.$$

It now follows that

$$\Delta = -(1 + o(1))2x^{-1}(\ln x)^2 + (1 + o(1))4x^{-1}(\ln x)^2 \sim 2x^{-1}(\ln x)^2,$$

as claimed.

(b) If we grant the truth of (2.4), the following upper bound on $\overline{F}(x)$ resulting from use of a Chernoff inequality with $t = w$ together with (2.4) still does not completely match (1.17):

$$\begin{aligned}\overline{F}(x) &\le \exp[-xw + J(w) - w^2 - \alpha w - \ln w + C + o(1)] \\ &= 2\sqrt{\pi}\, w^{1/2} e^{w/2} \times \text{RHS(1.17)} \sim (2\pi x)^{1/2} \times \text{RHS(1.17)}.\end{aligned}$$

Further, use of the *exactly* optimal t [ignoring the $o(1)$ remainder term in (2.4)] gives a bound that is still asymptotically $(2\pi x)^{1/2} \times$ RHS(1.17). Thus if the asymptotic inequality $\overline{F}(x) \le$ RHS(1.17) is ever to be established rigorously, it would have to involve some technique (such as a rigorization of the saddle-point arguments used in [8]) we have not used; Chernoff bounds are insufficient.

2.3 Proof of improved asymptotic upper bounds on absolute values of derivatives of F

Using the improved right-tail upper bound of the distribution function in Theorem 1.3(a), we are now able to establish Theorem 1.3(b).

Proof of Theorem 1.3(b). The bound (1.14) holds for $k = 0$ because it is cruder than the bound of Theorem 1.3(a). The bound (1.14) for general values of k then follows inductively using Proposition 6.1 of [2], according to which

$$\limsup_{x\to\infty} r(x)^{-1}\left(\ln \|\overline{F}^{(k+1)}\|_x - \ln \|\overline{F}^{(k)}\|_x\right) \le 0$$

provided $r(x) = \omega(\sqrt{x \log x})$ as $x \to \infty$. □

3 Large deviations for QuickSort

With some improvements, this section repeats Section 3 of [3].

McDiarmid and Hayward [9] study large deviations for the variant of `QuickSort` in which the pivot (that is, the initial partitioning key) is chosen as the median of $2t + 1$ keys chosen uniformly at random without replacement from among all the keys. The case $t = 0$ is the classical `QuickSort` algorithm of our ongoing limited focus in this paper. Restated equivalently in terms of the random variable Z_n in (1.1) (as straightforward calculation reveals), the following is their main theorem for classical `QuickSort`.

Theorem 3.1 ([9])**.** *Let x_n satisfy*

$$\frac{\mu_n}{n \ln n} < x_n \leq \frac{\mu_n}{n}. \tag{3.1}$$

Then as $n \to \infty$ we have

$$\mathbb{P}(|Z_n| > x_n) = \exp\{-x_n[\ln x_n + O(\log\log\log n)]\}. \tag{3.2}$$

Observe that (3.1) is roughly equivalent to the condition that x_n lie between 2 and $2 \ln n$, and rather trivially the range can be extended to $1 < x_n \leq \mu_n/n$. But notice also that if $x_n = (\ln \ln n)^{c_n}$ with c_n nondecreasing (say), then (3.2) provides a nontrivial upper bound on $\mathbb{P}(|Z_n| > x_n)$ if and only if $c_n \to \infty$.

McDiarmid and Hayward require a fairly involved proof utilizing primarily the method of bounded differences pioneered by McDiarmid [10] to establish the $\leq$ half of (3.2). The $\geq$ half is proven by establishing (by means of another substantial argument) the right-tail lower bound

$$\mathbb{P}(Z_n > x_n) \geq \exp\{-x_n[\ln x_n + O(\log\log\log n)]\}, \tag{3.3}$$

again assuming (3.1) (see [9, Lemma 2.9]). It follows from (3.2)–(3.3) that we have the right-tail large deviation result that

$$\mathbb{P}(Z_n > x_n) = \exp\{-x_n[\ln x_n + O(\log\log\log n)]\}. \tag{3.4}$$

The main point of this section [see Theorem 3.3(b)–(d)] is to note that (3.4) can be refined, for deviations not allowed to be quite as large as those permitted by Theorem 3.1, rather effortlessly by combining our upper bound [Theorem 1.3(a)] and lower bound [Theorem 1.2(b), with $k = 0$] on the right tail of F with the following bound on Kolmogorov–Smirnov distance between the distributions of Z_n and Z (see [6, Section 5]):

Lemma 3.2 ([6])**.** *We have*

$$\sup_x |\mathbb{P}(Z_n > x) - \mathbb{P}(Z > x)| \leq \exp\left[-\tfrac{1}{2}\ln n + O\left((\log n)^{1/2}\right)\right].$$

We state next our right-tail large-deviations theorem for `QuickSort`. With the additional indicated restriction on the growth of x_n (which allows for x_n nearly as large as $\frac{1}{2}\frac{\ln n}{\ln \ln n}$), parts (b)–(c) strictly refine (3.3) and the asymptotic upper bound on $\mathbb{P}(Z_n > x_n)$ implied by (3.4). The left-hand endpoint of the interval I_n in Theorem 3.3 is chosen as $c > 1$ simply to ensure that $\sup\{-\ln \ln x : x \in I_n\} < \infty$.

Theorem 3.3. *Let (ω_n) be any sequence diverging to $+\infty$ as $n \to \infty$ and let $c > 1$. For integer $n \geq 3$, consider the interval $I_n := \left[c, \frac{1}{2}\frac{\ln n}{\ln \ln n}\left(1 - \frac{\omega_n}{\ln \ln n}\right)\right]$.*
(a) *Uniformly for $x \in I_n$ we have*

$$\mathbb{P}(Z_n > x) = (1 + o(1))\mathbb{P}(Z > x) \quad \text{as } n \to \infty. \tag{3.5}$$

(b) *If $x_n \in I_n$ for all large n, then*

$$\mathbb{P}(Z_n > x_n) \geq \exp[-x_n \ln x_n - x_n \ln \ln x_n + O(x_n)]. \tag{3.6}$$

(c) *If $x_n \in I_n$ for all large n and $x_n \to \infty$, then*

$$\begin{aligned} \mathbb{P}(Z_n > x_n) &\leq \exp[-x_n w_n + J(w_n) - w_n^2 + O(\log x_n)] \qquad (3.7)\\ &= \exp[-x_n \ln x_n - x_n \ln \ln x_n + (1 + \ln 2)x_n + o(x_n)], \qquad (3.8) \end{aligned}$$

where w_n is the larger of the two real solutions to $x_n = 2w_n^{-1}e^{w_n}$.
(d) *If $x_n \in I_n$ for all large n, then*

$$\mathbb{P}(Z_n > x_n) = \exp[-x_n \ln x_n - x_n \ln \ln x_n + O(x_n)]. \tag{3.9}$$

Proof. Parts (b)–(c) follow immediately from part (a) and Theorem 1.3(a), and part (d) by combining parts (b)–(c). So we need only prove part (a), for which by Lemma 3.2 it is sufficient to prove that

$$\exp\left[-\tfrac{1}{2}\ln n + O\left((\log n)^{1/2}\right)\right] \le o(\mathbb{P}(Z > x_n))$$

with $x_n \equiv \frac{1}{2}\frac{\ln n}{\ln\ln n}\left(1 - \frac{\omega_n}{\ln\ln n}\right)$; this assertion decreases in strength as the choice of ω_n is increased, so we may assume that $\omega_n = o(\log\log n)$. Since, by Theorem 1.2(b), we have

$$\mathbb{P}(Z > x_n) \ge \exp[-x_n \ln x_n - x_n \ln\ln x_n + O(x_n)],$$

it suffices to show that for any constant $C < \infty$ we have

$$-\tfrac{1}{2}\ln n + C(\ln n)^{1/2} + x_n \ln x_n + x_n \ln\ln x_n + C x_n \to -\infty.$$

But, writing L for $\ln$ and L_k for the kth iterate of L, and abbreviating $\alpha_n := 1 - \frac{\omega_n}{\mathrm{L}_2 n}$, this follows from the observation that, for n large,

$$\begin{aligned}
&x_n(\mathrm{L}\,x_n + \mathrm{L}_2\,x_n + C)\\
&= \frac{1}{2}\frac{\mathrm{L}\,n}{\mathrm{L}_2\,n}\alpha_n[(\mathrm{L}_2\,n - \mathrm{L}_3\,n - \mathrm{L}\,2 + \mathrm{L}\,\alpha_n) + \mathrm{L}(\mathrm{L}_2\,n - \mathrm{L}_3\,n - \mathrm{L}\,2 + \mathrm{L}\,\alpha_n) + C]\\
&= \frac{1}{2}\frac{\mathrm{L}\,n}{\mathrm{L}_2\,n}\alpha_n\left[\mathrm{L}_2\,n + C - \mathrm{L}\,2 + \mathrm{L}\,\alpha_n + \mathrm{L}\left(1 - \frac{\mathrm{L}_3\,n + \mathrm{L}\,2 - \mathrm{L}\,\alpha_n}{\mathrm{L}_2\,n}\right)\right]\\
&= \frac{1}{2}\frac{\mathrm{L}\,n}{\mathrm{L}_2\,n}\alpha_n\left[\mathrm{L}_2\,n + C - \mathrm{L}\,2 + \mathrm{L}\,\alpha_n - (1+o(1))\frac{\mathrm{L}_3\,n}{\mathrm{L}_2\,n}\right]\\
&= \frac{1}{2}\frac{\mathrm{L}\,n}{\mathrm{L}_2\,n}\alpha_n\left[\mathrm{L}_2\,n + C - \mathrm{L}\,2 + o(1)\right]\\
&= \left(\frac{1}{2}\mathrm{L}\,n\right)\alpha_n\left[1 + \frac{C - \mathrm{L}\,2 + o(1)}{\mathrm{L}_2\,n}\right] = \frac{1}{2}\mathrm{L}\,n - (1+o(1))\omega_n\frac{\mathrm{L}\,n}{2\,\mathrm{L}_2\,n}.
\end{aligned}$$

□

For completeness we next present a left-tail analogue of Theorem 3.3 [but, for brevity, only parts (b)–(c) thereof]. Theorem 3.4 follows in similar fashion using the case $k = 0$ of (1.20) in place of Theorem 1.2(b). No such left-tail large-deviation result is found in [9]. Recall $\Gamma := (2 - \frac{1}{\ln 2})^{-1}$ and the notation L_k used in the proof of Theorem 3.3.

Theorem 3.4. *If* $1 < x_n \le \Gamma^{-1}(\mathrm{L}_2\,n - \mathrm{L}_4\,n - \omega_n)$ *with* $\omega_n \to \infty$*, then*

$$\exp\left[-e^{\Gamma x_n + \mathrm{L}_2\,x_n + O(1)}\right] \le \mathbb{P}(Z_n \le -x_n) \le \left[-e^{\Gamma x_n + O(1)}\right].$$

Remark 3.5. The upper bound in Theorem 3.4 requires only the weaker restriction

$$-M \le x_n \le \Gamma^{-1}(\mathrm{L}_2\,n - \omega_n)$$

with $M < \infty$ and $\omega_n \to \infty$.

Remark 3.6. If we let $N := n + 1$ and study the slight modification $\widehat{Z}_n := (X_n - \mu_n)/N = [n/(n+1)]Z_n$ instead of (1.1), then large deviation upper bounds based on tail estimates of the limiting F have broader applicability and are easier to derive, too. The reason is that (i) both Theorem 1.3(a) and the upper bound for $k = 0$ in (1.20) have been derived by establishing an upper bound on the limiting mgf ψ and using a Chernoff bound, and (ii) according to [6, Theorem 7.1], ψ majorizes the moment generating function $\widehat{\psi}_n$ of $\widehat{Z}_n$ for every n. It follows immediately (with w defined in the now-familiar way in terms of x) that $\mathbb{P}(\widehat{Z}_n > x)$ (respectively, $\mathbb{P}(\widehat{Z}_n \le -x)$) is bounded above uniformly in n by

$$\exp[-xw + J(w) - w^2 + O(\log x)] \tag{3.10}$$

$$= \exp[-x\ln x - x\ln\ln x + (1 + \ln 2)x + o(x)] \tag{3.11}$$

(resp., by $\exp\left[-e^{\gamma x+O(1)}\right]$) as $x \to \infty$; there is *no restriction at all* on how large x can be in terms of n.

Here are examples of *very* large values of x for which the tail probabilities are nonzero and the aforementioned bounds still match logarithmic asymptotics to lead order of magnitude, albeit not to lead-order term. Let $\lg$ denote binary log. The largest possible value of X_n is $\binom{n}{2}$ (corresponding to any binary search tree which is a path), which occurs with probability $2^{n-1}/n!$. The smallest possible value (supposing, for simplicity, that $n = 2^k - 1$ for integer k) is $(k-2)2^k + 2 = N(\lg N - 2) + 2$ (corresponding to the perfect tree, in the terminology of [1, Section 3]); according to [1, Proposition 4.1], this value occurs with probability $\exp[-s(1)N + s(N+1)]$, where

$$s(\nu) := \sum_{j=1}^{\infty} 2^{-j} \ln(2^j \nu - 1).$$

Correspondingly, the largest possible value of $\widehat{Z}_n$ is

$$\lambda_n := \tfrac{n(n+7)}{2(n+1)} - 2H_n = (1 + o(1))\tfrac{1}{2}N,$$

and the smallest is $-\sigma_n$, with

$$\sigma_n := -2H_N - \lg N - 2 = (2 - \tfrac{1}{\ln 2})\ln N + O(1).$$

The bound (3.11) on $\mathbb{P}(\widehat{Z}_n > \lambda_n)$ is in fact also (by the same proof) a bound on the larger probability $\mathbb{P}(\widehat{Z}_n \ge \lambda_n)$, and equals

$$\exp\left\{-\tfrac{1}{2}N[\ln N + \ln\ln N - (2\ln 2 + 1) + o(1)]\right\},$$

whereas (using Stirling's formula) the truth is

$$\mathbb{P}(\widehat{Z}_n \ge \lambda_n) = \exp[-N\ln N + (1 + \ln 2)N + O(\log N)].$$

The bound on $\mathbb{P}(\widehat{Z}_n \le -\sigma_n)$ equals

$$\exp\left[-e^{\ln N + O(1)}\right] = \exp[-\Omega(N)],$$

whereas (by [1, Proposition 4.1 and Table 1]) the truth is

$$\mathbb{P}(\widehat{Z}_n \le -\sigma_n) = \exp[-s(1)N + O(\log N)]$$

and (rounded to seven decimal places) $s(1) = 0.9457553$.

References

[1] Robert P. Dobrow and James Allen Fill, *Multiway trees of maximum and minimum probability under the random permutation model*, Combinatorics, Probability and Computing **5** (1996), no. 4, 351–371. MR-1426429

[2] James Allen Fill and Wei-Chun Hung, *On the tails of the limiting `QuickSort` density*, Electron. Commun. Probab. **24** (2019), Paper No. 7, 11. MR-3916339

[3] James Allen Fill and Wei-Chun Hung, *QuickSort: improved right-tail asymptotics for the limiting distribution, and large deviations (extended abstract)*, 2019 Proceedings of the Sixteenth Workshop on Analytic Algorithmics and Combinatorics (ANALCO), SIAM, Philadelphia, PA, 2019, pp. 87–93. MR-3909444

[4] James Allen Fill and Svante Janson, *A characterization of the set of fixed points of the Quicksort transformation*, Electronic Communications in Probability **5** (2000), 77–84. MR-1781841

[5] James Allen Fill and Svante Janson, *Smoothness and decay properties of the limiting Quicksort density function*, Mathematics and computer science (Versailles, 2000), Trends Math., Birkhäuser, Basel, 2000, pp. 53–64. MR-1798287

[6] James Allen Fill and Svante Janson, *Quicksort asymptotics*, Journal of Algorithms. Cognition, Informatics and Logic **44** (2002), no. 1, 4–28, Analysis of algorithms. MR-1932675

[7] Svante Janson, *On the tails of the limiting Quicksort distribution*, Electron. Commun. Probab. **20** (2015), no. 81, 7. MR-3434198

[8] Charles Knessl and Wojciech Szpankowski, *Quicksort algorithm again revisited*, Discrete Mathematics & Theoretical Computer Science. DMTCS. An Electronic Journal **3** (1999), no. 2, 43–64. MR-1695194

[9] C. J. H. McDiarmid and R. B. Hayward, *Large deviations for Quicksort*, Journal of Algorithms. Cognition, Informatics and Logic **21** (1996), no. 3, 476–507. MR-1417660

[10] Colin McDiarmid, *On the method of bounded differences*, Surveys in combinatorics, 1989 (Norwich, 1989), London Math. Soc. Lecture Note Ser., vol. 141, Cambridge Univ. Press, Cambridge, 1989, pp. 148–188. MR-1036755

[11] Mireille Régnier, *A limiting distribution for quicksort*, RAIRO Inform. Théor. Appl. **23** (1989), no. 3, 335–343. MR-1020478

[12] Uwe Rösler, *A limit theorem for "Quicksort"*, RAIRO Inform. Théor. Appl. **25** (1991), no. 1, 85–100. MR-1104413