AofA'12 DMTCS proc. **AQ**, 2012, 339–348

Exact L^2-Distance from the Limit for QuickSort Key Comparisons (Extended Abstract)

Patrick Bindjeme[†] and James Allen Fill

Department of Applied Mathematics and Statistics, The Johns Hopkins University, Baltimore, MD 21218-2682 USA

Using a recursive approach, we obtain a simple exact expression for the L^2-distance from the limit in the classical limit theorem of Régnier (1989) for the number of key comparisons required by `QuickSort`. A previous study by Fill and Janson (2002) using a similar approach found that the d_2-distance is of order between $n^{-1}\log n$ and $n^{-1/2}$, and another by Neininger and Ruschendorf (2002) found that the Zolotarev ζ_3-distance is of exact order $n^{-1}\log n$. Our expression reveals that the L^2-distance is asymptotically equivalent to $(2n^{-1}\ln n)^{1/2}$.

Keywords: QuickSort; key comparisons; limit distribution; L^2-distance

1 Introduction, review of related literature, and summary

We consider the `QuickSort` sorting algorithm of Hoare (1962) applied to an infinite stream of iid (independent and identically distributed) uniform random variables $U_1, U_2, \ldots$. `QuickSort` chooses the first key U_1 as the "pivot", compares each of the other keys to it, and then proceeds recursively to sort both the keys smaller than the pivot and those larger than it. If, for example, the initial round of comparisons finds $U_2 < U_1$, then U_2 is used as the pivot in the recursive call to the algorithm that sorts the keys smaller than U_1 because it is the first element in the sequence $U_1, U_2, \ldots$ which is smaller than U_1. In a natural and obvious way, a realization (requiring infinite time) of the algorithm produces an infinite rooted binary search tree which with probability one has the *completeness* property that each node has two child-nodes.

Essentially the same algorithm can of course be applied to the truncated sequence $U_1, U_2, \ldots, U_n$ for any finite n, where the recursion ends by declaring that a list of size 0 or 1 is already sorted. Let K_n denote the number of key comparisons required by `QuickSort` to sort $U_1, U_2, \ldots, U_n$. Then, with the way we have set things up, all the random variables K_n are defined on a common probability space, and K_n is nondecreasing in n. Indeed, $K_n - K_{n-1}$ is simply the cost of inserting U_n into the usual (finite) binary search tree formed from $U_1, \ldots, U_{n-1}$.

In this framework, Régnier (1989) used martingale techniques to establish the following L^p-limit theorem; she also proved almost sure convergence. We let

$$\mu_n := \mathbf{E}\, K_n.$$

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

Theorem 1.1 (Régnier (1989)) *There exists a random variable T satisfying*

$$Y_n := \frac{K_n - \mu_n}{n+1} \xrightarrow{L^p} T$$

for every finite p.

Rösler (1991) characterized the distribution of Régnier's limiting T as the unique fixed point of a certain distributional transformation, but he also described explicitly how to construct a random variable having the same distribution as T. We will describe his explicit construction in equivalent terms, but first we need two paragraphs of notation.

The nodes of the complete infinite binary search tree are labeled in the natural binary way: the root gets an empty label written ε here, the left (respectively, right) child is labeled 0 (resp., 1), the left child of node 0 is labeled 00, etc. We write $\Theta := \cup_{0 \leq k < \infty} \{0,1\}^k$ for the set of all such labels. If V_θ denotes the key inserted at node $\theta \in \Theta$, let L_θ (resp., R_θ) denote the largest key smaller than V_θ (resp., smallest key larger than V_θ) inserted at any ancestor of θ, with the exceptions $L_\theta := 0$ and $R_\theta := 1$ if the specified ancestor keys do not exist. Further, for each node θ, define

$$\begin{aligned} &\phi_\theta := R_\theta - L_\theta, \qquad U_\theta := \phi_{\theta 0}/\phi_\theta, \\ &G_\theta := \phi_\theta C(U_\theta) = \phi_\theta - 2\phi_\theta \ln \phi_\theta + 2\phi_{\theta 0} \ln \phi_{\theta 0} + 2\phi_{\theta 1} \ln \phi_{\theta 1}, \end{aligned} \tag{1.1}$$

where for $0 < x < 1$ we define

$$C(x) := 1 + 2x \ln x + 2(1-x)\ln(1-x). \tag{1.2}$$

Let $1 \leq p < \infty$. The *d_p-metric* is the metric on the space of all probability distributions with finite pth absolute moment defined by

$$d_p(F_1, F_2) \quad := \quad \inf \|X_1 - X_2\|_p,$$

where we take the infimum of L^p-distances over all pairs of random variables X_1 and X_2 (defined on the same probability space) with respective marginal distributions F_1 and F_2. By the d_p-distance between two random variables we mean the d_p-distance between their distributions.

We are now prepared to state Rösler's main result. NOTE: Here and later results have been adjusted slightly as necessary to utilize the same denominator $n+1$ (rather than n) that Régnier used.

Theorem 1.2 (Rösler (1991)) *For any finite p, the infinite series $Y = \sum_{j=0}^{\infty} \sum_{|\theta|=j} G_\theta$ converges in L^p, and the sequence $Y_n = (K_n - \mu_n)/(n+1)$ converges in the d_p-metric to Y.*

Of course it follows from Theorems 1.1–1.2 that T and Y have the same distribution. **The purpose of the present extended abstract is to show that in fact $T = Y$ and to provide a simple explicit expression for the L^2-distance between Y_n and** Y valid for every n; this is done in Theorem 1.4 below.

We are aware of only two previous studies of the rate of convergence of Y_n to Y, and both of those concern certain distances between *distributions* rather than between *random variables*. The first study, by Fill and Janson (2002), provides upper and lower bounds on $d_p(Y_n, Y)$ for general p; we choose to focus here on d_2.

Theorem 1.3 (Fill and Janson (2002)) *There is a constant $c > 0$ such that for any $n \geq 1$ we have*

$$cn^{-1} \ln n \leq d_2(Y_n, Y) < 2n^{-1/2}.$$

To our knowledge, the gap between the rates $(\log n)/n$ and $n^{-1/2}$ has not been narrowed. Neininger and Rüschendorf (2002) used the Zolotarev ζ_3-metric and found that the correct rate in that metric is $n^{-1} \log n$, but their techniques are not sufficiently sharp to obtain $\zeta_3(Y_n, Y) \sim \tilde{c} n^{-1} \ln n$ for some constant $\tilde{c}$.

In our **main Theorem 1.4**, proved using the same recursive approach as in Fill and Janson (2002), we find not only the lead-order asymptotics for the L^2-distance $\|Y_n - Y\|_2$, but in fact an exact expression for general n. It is interesting to note that the rate $n^{-1/2}(\log n)^{1/2}$ for L^2-convergence is larger even than the upper-bound rate of $n^{-1/2}$ for d_2-convergence from Theorem 1.3.

Theorem 1.4 (main theorem) *For $n \geq 0$ we have*

$$\|Y_n - Y\|_2^2 = (n+1)^{-1} \left(2H_n + 1 + \frac{6}{n+1} \right) - 4 \sum_{k=n+1}^{\infty} k^{-2} = 2 \frac{\ln n}{n} + O\left(\frac{1}{n}\right),$$

where $H_n := \sum_{j=1}^{n} j^{-1}$ is the nth harmonic number and the asymptotic expression holds as $n \to \infty$.

The remainder of this extended abstract is devoted to a proof of Theorem 1.4, which is completed in Section 5.

2 Preliminaries

In this section we provide recursive representations of Y_n (for general n) and Y that will be useful in proving Theorem 1.4. Our first proposition concerns the limit Y and gives a sample-pointwise extension of the very well known [Rösler (1991)] distributional identity satisfied by Y. Recall the notation (1.1) and the definition of Y in Theorem 1.2 as the infinite series $\sum_{j=0}^{\infty} \sum_{|\theta|=j} G_\theta$ in L^2.

Proposition 2.1 *There exist random variables F_θ and H_θ for $\theta \in \Theta$ such that*

(i) *the joint distributions of $(G_\theta : \theta \in \Theta)$, of $(F_\theta : \theta \in \Theta)$, and of $(H_\theta : \theta \in \Theta)$ agree;*

(ii) *$(F_\theta : \theta \in \Theta)$ and $(H_\theta : \theta \in \Theta)$ are independent;*

(iii) *the series*

$$Y^{(0)} := \sum_{j=0}^{\infty} \sum_{|\theta|=j} F_\theta \qquad \text{and} \qquad Y^{(1)} := \sum_{j=0}^{\infty} \sum_{|\theta|=j} H_\theta \tag{2.1}$$

converge in L^2;

(iv) *the random variables $Y^{(0)}$ and $Y^{(1)}$ are independent, each with the same distribution as Y, and*

$$Y = C(U) + UY^{(0)} + \overline{U} Y^{(1)}. \tag{2.2}$$

Here $U := U_1$, with $\overline{U} := 1 - U_1$, and C is defined at (1.2).

Proof: Recall from (1.1) that

$$G_\theta = \phi_\theta - 2\phi_\theta \ln \phi_\theta + 2\phi_{\theta 0} \ln \phi_{\theta 0} + 2\phi_{\theta 1} \ln \phi_{\theta 1}.$$

For $\theta \in \Theta$, define the random variable φ_θ (respectively, ψ_θ) by

$$\varphi_\theta := \phi_{0\theta}/U \qquad (\text{resp., } \psi_\theta := \phi_{1\theta}/\overline{U}).$$

Then U and φ_θ are independent (resp., $\overline{U}$ and ψ_θ are independent), φ_θ and ψ_θ each have the same distribution as ϕ_θ, and

$$G_{0\theta} = UF_\theta \qquad \text{and} \qquad G_{1\theta} = \overline{U} H_\theta,$$

where

$$\begin{aligned} F_\theta &:= \varphi_\theta - 2\varphi_\theta \ln \varphi_\theta + 2\varphi_{\theta 0} \ln \varphi_{\theta 0} + 2\varphi_{\theta 1} \ln \varphi_{\theta 1}, \\ H_\theta &:= \psi_\theta - 2\psi_\theta \ln \psi_\theta + 2\psi_{\theta 0} \ln \psi_{\theta 0} + 2\psi_{\theta 1} \ln \psi_{\theta 1}. \end{aligned}$$

The proposition follows easily from the clear equality

$$\mathcal{L}(F_\theta : \theta \in \Theta) = \mathcal{L}(G_\theta : \theta \in \Theta) = \mathcal{L}(H_\theta : \theta \in \Theta),$$

of joint laws and the fact that

$$\begin{aligned} Y &= \sum_{j=0}^{\infty} \sum_{|\theta|=j} G_\theta = G_\varepsilon + \sum_{j=0}^{\infty} \sum_{|\theta|=j} G_{0\theta} + \sum_{j=0}^{\infty} \sum_{|\theta|=j} G_{1\theta} \\ &= C(U) + U \sum_{j=0}^{\infty} \sum_{|\theta|=j} F_\theta + \overline{U} \sum_{j=0}^{\infty} \sum_{|\theta|=j} H_\theta \\ &= C(U) + UY^{(0)} + \overline{U} Y^{(1)}. \end{aligned}$$

□

We next proceed to provide an analogue [namely, (2.4)] of (2.2) for each Y_n, rather than Y, but first we need a little more notation.

Given $0 \le x < y \le 1$, let $(U_n^{xy})_{n \ge 1}$ be the subsequence of $(U_n)_{n \ge 1}$ that falls in (x, y). The random variable $K_n(x, y)$ is defined to be the (random) number of key comparisons used to sort $U_1^{xy}, \dots, U_n^{xy}$ using `QuickSort`. The distribution of $K_n(x, y)$ of course does not depend on (x, y).

Let $\nu_\theta(n)$ denote the number of keys among $U_1, \dots, U_n$ in the subtree rooted at θ. We now define the random variable

$$Y_{n,\theta} \quad := \quad [K_{\nu_\theta(n)}(L_\theta, R_\theta) - \mu_{\nu_\theta(n)}]/[\nu_\theta(n) + 1], \tag{2.3}$$

with the centering here motivated by the fact that $\mu_{\nu_\theta(n)}$ is the conditional expectation of $K_{\nu_\theta(n)}(L_\theta, R_\theta)$ given $(\nu_\theta(n), L_\theta, R_\theta)$. Then for $n \ge 1$ we have

$$Y_n = \frac{n}{n+1} C_n(\nu_0(n) + 1) + \frac{\nu_0(n) + 1}{n+1} Y_{n,0} + \frac{\nu_1(n) + 1}{n+1} Y_{n,1}, \tag{2.4}$$

where, as in Fill and Janson (2004), for $1 \le i \le n$ we define

$$C_n(i) := \tfrac{1}{n}(n - 1 + \mu_{i-1} + \mu_{n-i} - \mu_n).$$

We note for future reference that the classical divide-and-conquer recurrence for μ_n asserts precisely that

$$\sum_{i=1}^{n} C_n(i) = 0 \tag{2.5}$$

for $n \ge 1$.

It follows from (2.2) and (2.4) that for $n \ge 1$ we have

$$\begin{aligned} Y_n - Y &= \left[\frac{\nu_0(n)+1}{n+1}\, Y_{n,0} - UY^{(0)}\right] + \left[\frac{\nu_1(n)+1}{n+1}\, Y_{n,1} - \overline{U}Y^{(1)}\right] \\ &\qquad + \left[\frac{n}{n+1}\, C_n(\nu_0(n)+1) - C(U)\right] \\ &=: W_1 + W_2 + W_3. \end{aligned} \tag{2.6}$$

Conditionally given U and $\nu_0(n)$, the random variables W_1 and W_2 are independent, each with vanishing mean, and W_3 is constant. Hence

$$\mathbf{E}[(Y_n - Y)^2 \,|\, U, \nu_0(n)] = \mathbf{E}[W_1^2 \,|\, U, \nu_0(n)] + \mathbf{E}[W_2^2 \,|\, U, \nu_0(n)] + W_3^2$$

and thus, taking expectations and using symmetry, for $n \ge 1$ we have

$$a_n^2 := \mathbf{E}(Y_n - Y)^2 = \mathbf{E}\, W_1^2 + \mathbf{E}\, W_2^2 + \mathbf{E}\, W_3^2 = 2\mathbf{E}\, W_1^2 + \mathbf{E}\, W_3^2. \tag{2.7}$$

Note that

$$a_0^2 = \mathbf{E}\, Y^2 = \sigma^2 := 7 - \tfrac{2}{3}\pi^2 \tag{2.8}$$

(for example, Fill and Janson (2004)).

3 Analysis of $\mathbf{E}\, W_1^2$

In this section we analyze $\mathbf{E}\, W_1^2$, producing the following result. Recall the definition of σ^2 at (2.8).

Proposition 3.1 *Let $n \ge 1$. For W_1 defined as at* (2.6)*, we have*

$$\mathbf{E}\, W_1^2 = \frac{1}{n(n+1)^2} \sum_{k=0}^{n-1} (k+1)^2 a_k^2 + \frac{\sigma^2}{6(n+1)}.$$

For that, we first prove the following two lemmas.

Lemma 3.2 *For any $n \ge 1$, we have*

$$\mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1}\right)^2 \left(Y_{n,0} - Y^{(0)}\right)^2\right] \quad = \quad \frac{1}{n} \sum_{k=0}^{n-1} \left(\frac{k+1}{n+1}\right)^2 a_k^2.$$

Lemma 3.3 *For any $n \geq 1$, we have*

$$\mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1} - U\right)^2 \left(Y^{(0)}\right)^2\right] \quad = \quad \frac{\sigma^2}{6(n+1)}.$$

Proof of Lemma 3.2: There is a probabilistic copy $Y^* = (Y_n^*)$ of the stochastic process (Y_n) such that

$$Y_{n,0} \equiv Y^*_{\nu_0(n)}$$

and Y^* and $Y^{(0)}$ are independent of $(U, \nu_0(n))$. This implies

$$\mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1}\right)^2 \left(Y_{n,0} - Y^{(0)}\right)^2\right] \quad = \quad \mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1}\right)^2 \left(Y^*_{\nu_0(n)} - Y^{(0)}\right)^2\right].$$

By conditioning on $\nu_0(n)$, which is uniformly distributed on $\{0, \ldots, n-1\}$, we get

$$\begin{aligned}
\mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1}\right)^2 \left(Y_{n,0} - Y^{(0)}\right)^2\right] &= \mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1}\right)^2 a^2_{\nu_0(n)}\right] \\
&= \frac{1}{n}\sum_{k=0}^{n-1}\left(\frac{k+1}{n+1}\right)^2 a_k^2.
\end{aligned}$$

□

Proof of Lemma 3.3: Conditionally given $\nu_0(n)$ and $Y^{(0)}$, we have that U is distributed as the order statistic of rank $\nu_0(n)+1$ from a sample of size n from the uniform$(0,1)$ distribution, namely, Beta$(\nu_0(n)+1, n-\nu_0(n))$, with expectation $[\nu_0(n)+1]/(n+1)$ and variance $[(\nu_0(n)+1)(n-\nu_0(n))]/[(n+1)^2(n+2)]$. So, using also the independence of $\nu_0(n)$ and $Y^{(0)}$, we find

$$\begin{aligned}
&\mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1} - U\right)^2 \left(Y^{(0)}\right)^2\right] \\
&= \mathbf{E}\left[\frac{(\nu_0(n)+1)(n-\nu_0(n))}{(n+1)^2(n+2)}\left(Y^{(0)}\right)^2\right] = \sigma^2\, \mathbf{E}\left[\frac{(\nu_0(n)+1)(n-\nu_0(n))}{(n+1)^2(n+2)}\right] \\
&= \frac{\sigma^2}{(n+1)^2(n+2)} \times \frac{1}{n}\sum_{k=0}^{n-1}(k+1)(n-k) = \frac{\sigma^2}{n(n+1)^2(n+2)} \times \frac{1}{6}n(n+1)(n+2) \\
&= \frac{\sigma^2}{6(n+1)}.
\end{aligned}$$

□

Proof of Proposition 3.1: We have

$$\begin{aligned}
\mathbf{E}\,W_1^2 &= \mathbf{E}\left[\frac{\nu_0(n)+1}{n+1}\,Y_{n,0} - UY^{(0)}\right]^2 \\
&= \mathbf{E}\left[\frac{\nu_0(n)+1}{n+1}\left(Y_{n,0} - Y^{(0)}\right) + \left(\frac{\nu_0(n)+1}{n+1} - U\right)Y^{(0)}\right]^2 \\
&= \mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1}\right)^2\left(Y_{n,0} - Y^{(0)}\right)^2\right] \\
&\quad + \mathbf{E}\left[\left(\frac{\nu_0(n)+1}{n+1} - U\right)^2 (Y^{(0)})^2\right] \\
&\quad + 2\,\mathbf{E}\left[\frac{\nu_0(n)+1}{n+1}\left(Y_{n,0} - Y^{(0)}\right)\left(\frac{\nu_0(n)+1}{n+1} - U\right)Y^{(0)}\right].
\end{aligned}$$

The result follows from Lemmas 3.2–3.3, and the fact that, conditionally given $(\nu_0(n), Y_{n,0}, Y^{(0)})$, the random variable U is distributed Beta$(\nu_0(n)+1, n-\nu_0(n))$, so that the last expectation in the preceding equation vanishes. □

4 Analysis of $\mathbf{E}\,W_3^2$

In this section we analyze $\mathbf{E}\,W_3^2$, producing the following result.

Proposition 4.1 *For any $n \geq 1$ we have*

$$b_n^2 := \mathbf{E}\,W_3^2 = \frac{\sigma^2 - 7}{3} + \frac{4}{3}\left(\frac{n+2}{n+1}\right)H_n^{(2)} + \frac{4}{3n(n+1)^2}H_n,$$

where $H_n = \sum_{j=1}^n j^{-1}$ is the nth harmonic number and $H_n^{(2)} := \sum_{j=1}^n j^{-2}$ is the nth harmonic number of the second order.

For that, we first prove the following two lemmas.

Lemma 4.2 *For any $1 \leq k \leq n$ we have*

$$D(n,k) := \frac{1}{B(k, n-k+1)}\int_0^1 t^{k-1}(1-t)^{n-k}(\ln t)\,dt = H_{k-1} - H_n,$$

where B is the beta function.

Lemma 4.3 *For any $n \geq 1$ we have*

$$\mathbf{E}[C_n(\nu_0(n)+1)C(U)] = \frac{n}{n+1}\mathbf{E}[C_n(\nu_0(n)+1)]^2.$$

Proof of Lemma 4.2: The result can be proved for each fixed $n \geq 1$ by backwards induction on k and integration by parts, but we give a simpler proof. Recall the defining expression

$$B(\alpha, \beta) = \int_0^1 t^{\alpha-1}(1-t)^{\beta-1}\, dt$$

for the beta function when $\alpha, \beta > 0$. Differentiating with respect to α gives

$$\int_0^1 t^{\alpha-1}(1-t)^{\beta-1}(\ln t)\, dt = B(\alpha, \beta)[\psi(\alpha) - \psi(\alpha+\beta)],$$

where ψ is the classical digamma function, i.e., the logarithmic derivative of the gamma function. But it is well known that $\psi(j) = H_{j-1}$ for positive integers j, so the lemma follows by setting $\alpha = k$ and $\beta = n - k + 1$. $\square$

Proof of Lemma 4.3: We know that $\nu_0(n) + 1 \sim$ unif$\{1, 2, \ldots, n\}$ and that, conditionally given $\nu_0(n)$, the random variable U has the Beta$(\nu_0(n)+1, n-\nu_0(n))$ distribution. So from Lemma 4.2, repeated use of (2.5), and the very well known and easily derived explicit expression

$$\mu_n = 2(n+1)H_n - 4n, \qquad n \geq 0,$$

we have

$$\begin{aligned}
&\mathbf{E}[C_n(\nu_0(n)+1)C(U)] \\
&= \frac{1}{n}\sum_{j=1}^{n} C_n(j)\frac{1}{B(j, n-j+1)}\int_0^1 t^{j-1}(1-t)^{n-j}C(t)\, dt \\
&= \frac{1}{n}\sum_{j=1}^{n} C_n(j)[1 + 2\frac{j}{n+1}(H_j - H_{n+1}) + 2\frac{n-j+1}{n+1}(H_{n-j+1} - H_{n+1})] \\
&= \frac{1}{n(n+1)}\sum_{j=1}^{n} C_n(j)[2jH_j + 2(n-j+1)H_{n-j+1}] \\
&= \frac{1}{n(n+1)}\sum_{j=1}^{n} C_n(j)[2jH_{j-1} - 4(j-1) + 2(n-j+1)H_{n-j} - 4(n-j)] \\
&= \frac{1}{n(n+1)}\sum_{j=1}^{n} C_n(j)[\mu_{j-1} + \mu_{n-j}] = \frac{1}{n(n+1)}\sum_{j=1}^{n} C_n(j)[\mu_{j-1} + \mu_{n-j} - \mu_n] \\
&= \frac{n}{n+1} \times \frac{1}{n}\sum_{j=1}^{n} C_n(j)^2 = \frac{n}{n+1}\mathbf{E}[C_n(\nu_0(n)+1)]^2,
\end{aligned}$$

as desired. $\square$

Proof of Proposition 4.1: It follows from Lemma 4.3 that

$$\begin{aligned} b_n^2 &= \mathbf{E}\left[\frac{n}{n+1}C_n(\nu_0(n)+1) - C(U)\right]^2 \\ &= \left(\frac{n}{n+1}\right)^2 \mathbf{E}\left[C_n(\nu_0(n)+1)\right]^2 - 2\left(\frac{n}{n+1}\right)\mathbf{E}[C_n(\nu_0(n)+1)C(U)] + \mathbf{E}\,C(U)^2 \\ &= \mathbf{E}\,C(U)^2 - \left(\frac{n}{n+1}\right)^2 \mathbf{E}\left[C_n(\nu_0(n)+1)\right]^2. \end{aligned}$$

Knowing that $\mathbf{E}\,C(U)^2 = \sigma^2/3$, and from the proof of Lemma A.5 in Fill and Janson (2004) that

$$\mathbf{E}\left[C_n(\nu_0(n)+1)\right]^2 = \tfrac{7}{3}\left(1+\tfrac{1}{n}\right)^2 - \tfrac{4}{3}\left(1+\tfrac{2}{n}\right)\left(1+\tfrac{1}{n}\right)H_n^{(2)} - \tfrac{4}{3}n^{-3}H_n,$$

we have

$$b_n^2 = \frac{\sigma^2-7}{3} + \frac{4}{3}\left(\frac{n+2}{n+1}\right)H_n^{(2)} + \frac{4}{3n(n+1)^2}H_n,$$

as claimed. □

5 A closed form for a_n^2

In this final section we complete the proof of Theorem 1.4, for which we need one more lemma.

Lemma 5.1 *For $H_n^{(2)} = \sum_{j=1}^n j^{-2}$, the nth harmonic number of the second order, we have*

$$\sum_{j=1}^{n} H_j^{(2)} \quad = \quad (n+1)H_n^{(2)} - H_n$$

for any nonnegative integer n.

The lemma is well known and easily proved.

Proof of main Theorem 1.4: For $n \geq 1$ we have from the decomposition (2.7) and Propositions 3.1 and 4.1 that

$$\begin{aligned} a_n^2 = {} & \frac{2}{n(n+1)^2}\sum_{k=0}^{n-1}(k+1)^2 a_k^2 \\ & + \frac{\sigma^2}{3}\left(\frac{n+2}{n+1}\right) - \frac{7}{3} + \frac{4}{3}\left(\frac{n+2}{n+1}\right)H_n^{(2)} + \frac{4}{3n(n+1)^2}H_n, \end{aligned}$$

and we recall from (2.8) that $a_0^2 = \sigma^2$. Setting $x_n := (n+1)^2 a_n^2$, we have $x_0 = \sigma^2$ and

$$x_n = \frac{2}{n}\sum_{k=0}^{n-1} x_k + c_n \qquad \text{for } n \geq 1,$$

with

$$c_n := \frac{\sigma^2}{3}(n+2)(n+1) - \frac{7}{3}(n+1)^2 + \frac{4}{3}(n+2)(n+1)H_n^{(2)} + \frac{4}{3n}H_n.$$

This is a standard divide-and-conquer recurrence relation for x_n, with solution

$$x_n = (n+1)\left[\sigma^2 + \sum_{k=1}^{n} \frac{kc_k - (k-1)c_{k-1}}{k(k+1)}\right], \qquad n \geq 0.$$

After straightforward computation involving the identity in Lemma 5.1, one finds

$$\begin{aligned} a_n^2 &= (n+1)^{-1}\left(2H_n + 1 + \frac{6}{n+1}\right) + \sigma^2 - 7 + 4H_n^{(2)} \\ &= (n+1)^{-1}\left(2H_n + 1 + \frac{6}{n+1}\right) - 4\sum_{k=n+1}^{\infty} k^{-2} = 2\frac{\ln n}{n} + O\left(\frac{1}{n}\right), \end{aligned}$$

as claimed. □

Acknowledgements

We thank an anonymous referee for helpful suggestions.

References

J. A. Fill and S. Janson. Quicksort asymptotics. *J. Algorithms*, 44(1):4–28, 2002. ISSN 0196-6774. doi: 10.1016/S0196-6774(02)00216-X. URL `http://dx.doi.org/10.1016/S0196-6774(02)00216-X`. Analysis of algorithms.

J. A. Fill and S. Janson. *Quicksort Asymptotics : Appendix*. Unpublished, Available from `http://www.ams.jhu.edu/~fill/`, 2004.

C. A. R. Hoare. Quicksort. *Comput. J.*, 5:10–15, 1962. ISSN 0010-4620.

R. Neininger and L. Rüschendorf. Rates of convergence for Quicksort. *J. Algorithms*, 44(1):51–62, 2002. ISSN 0196-6774. doi: 10.1016/S0196-6774(02)00206-7. URL `http://dx.doi.org/10.1016/S0196-6774(02)00206-7`. Analysis of algorithms.

M. Régnier. A limiting distribution of Quicksort. *RAIRO Informatique Théorique et Applications*, 23: 335–343, 1989.

U. Rösler. A limit theorem for Quicksort. *RAIRO Informatique Théorique et Applications*, 25:85–100, 1991.