The Annals of Probability
2003, Vol. 31, No. 2, 609–629

SECOND PHASE CHANGES IN RANDOM *m*-ARY SEARCH TREES AND GENERALIZED QUICKSORT: CONVERGENCE RATES

BY HSIEN-KUEI HWANG

Academia Sinica

We study the convergence rate to normal limit law for the space requirement of random m-ary search trees. While it is known that the random variable is asymptotically normally distributed for $3 \leq m \leq 26$ and that the limit law does not exist for $m > 26$, we show that the convergence rate is $O(n^{-1/2})$ for $3 \leq m \leq 19$ and is $O(n^{-3(3/2-\alpha)})$, where $4/3 < \alpha < 3/2$ is a parameter depending on m for $20 \leq m \leq 26$. Our approach is based on a refinement to the method of moments and applicable to other recursive random variables; we briefly mention the applications to quicksort proper and the generalized quicksort of Hennequin, where more phase changes are given. These results provide natural, concrete examples for which the Berry–Esseen bounds are not necessarily proportional to the reciprocal of the standard deviation. Local limit theorems are also derived.

1. Introduction. Probabilistic analysis of data structures and algorithms has received increasing recent attention. Roughly, the first goal has been to determine the complexity of the structures or algorithms in terms of simple mathematical functions; the analysis may in turn introduce intriguing random structures as well as challenging probabilistic problems. The problems we study in this paper will be seen to have such a character. We are concerned in this paper with the Berry–Esseen bounds (convergence rates in Kolmogorov distance) for the space requirement of random m-ary search trees, which is shown to exhibit a new "phase change" when m grows. Our method of proof is also applicable to more general search trees and quicksort algorithms for which more "phase changes" are unveiled.

We start from the binary search trees, which are one of the simplest and most fundamental data structures in computer algorithms. A binary search tree is a binary, rooted, labeled tree in which the labels in the left subtrees of any node x are all less than that of x, and those in the right subtrees are all greater than that of x. This property enables one to perform easily queries like "Is the key y in the tree?" Also it is easy to devise algorithms for inserting a new key and for deleting a node in the tree. Although binary search trees have poor performance in the worst-case (e.g., when the tree is a chain of nodes), they are efficient when the tree is constructed from a random sequence; see Mahmoud (1992).

Received June 2001; revised February 2002.
AMS 2000 subject classifications. Primary 60F05; secondary 05C05, 68W40, 68P10.
Key words and phrases. Convergence rates, asymptotic normality, phase change, search trees, quicksort, method of moments, local limit theorems.

Such a data structure is prototypical and admits many different varieties of extensions such as AVL trees, m-ary search trees, quadtrees and k-d trees, on the one hand, and quicksort and its many variants, on the other hand; see Sedgewick (1980), Gonnet (1991), Mahmoud (1992), Devroye (1999). We first describe m-ary search trees, which are the main object of study of this paper; variants of quicksort are briefly mentioned later.

Given a sequence of n keys, an m-ary search tree ($m \geq 2$) is constructed as follows. If $n = 0$ then the tree is empty; if $1 \leq n \leq m - 1$ then the tree consists of only a single internal node holding these keys in increasing order; if $n \geq m$, then the first $m - 1$ keys stay in an internal node (called root node) in increasing order, which are used to direct the remaining keys into the m branches: keys lying between the ith and the $(i + 1)$st keys go to the $(i + 1)$st branch, where $0 \leq i \leq m - 1$ and, for convenience of description, the (imaginary) 0th and the $(m + 1)$st keys are $-\infty$ and $+\infty$, respectively; keys in each subtree are then constructed recursively; see Figure 1 for an illustration and Mahmoud (1992) for more details.

It is visible from Figure 1 that the space requirement (the total number of nodes to store the given keys) depends on the order of the input, and that the number of keys in each node varies from 1 to $m - 1$. Given n keys, it is straightforward to see that the space requirement varies between $n/(m - 1)$ and $mn/(2m - 2)$ [see Mahmoud and Pittel (1989)]. Between these two extremes, what is the "typical behavior" of the storage requirement? To answer this question, we introduce

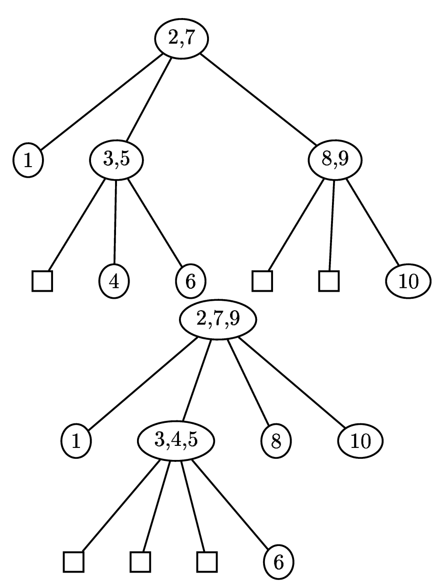

FIG. 1. *The ternary* (*left*) *and quaternary* (*right*) *search trees constructed from the sequence* $\{2, 7, 9, 8, 5, 1, 3, 10, 4, 6\}$.

the usual uniform probability model by assuming that the input is a sequence of n independent and identically distributed random variables with a continuous distribution. Given such a random input, the m-ary search tree constructed from it is called a *random m-ary search tree* (of n keys) and the storage requirement is then a random variable for $m \geq 3$, denoted by X_n. Note that $X_n \equiv n$ for $m = 2$.

In addition to computer algorithms, random search trees also surfaced naturally in several different fields such as evolutionary trees, diffusion models, random fragmentation processes, collision processes; see Aldous (1994), Barlow, Pemantle and Perkins (1997), Ben-Naim, Krapivsky and Majumdar (2001) and the references therein for further information.

By the recursive construction of m-ary search trees and the probability model, $X_0 = 1$, $X_n = 1$ for $1 \leq n \leq m-1$, and

$$X_n \stackrel{d}{=} X_{I_1}^{[1]} + \cdots + X_{I_m}^{[m]} + 1, \qquad n \geq m,$$

where $(X_n^{[1]}), \ldots, (X_n^{[m]}), (I_1, \ldots, I_m)$ are independent and the $(X_n^{[i]})$'s have the same distribution as (X_n). Here [see Mahmoud and Pittel (1984)]

$$P(I_1 = j_1, \ldots, I_m = j_m) = \frac{1}{\binom{n}{m-1}},$$

for all tuples of nonnegative integers $(j_1, \ldots, j_m)$ such that $j_1 + \cdots + j_m = n - m + 1$. [Briefly, there are $\binom{n}{m-1}$ ways of choosing $m-1$ keys for the root node and the m subtrees are independent and equi-distributed.]

Let $P_n(u) := E(e^{X_n u})$. Then the above description translates into

$$P_n(u) = \begin{cases} 1, & \text{if } n = 0, \\ e^u, & \text{if } 1 \leq n \leq m-1, \\ \dfrac{e^u}{\binom{n}{m-1}} \displaystyle\sum_{\substack{j_1+\cdots+j_m=n-m+1 \\ j_1,\ldots,j_m \geq 0}} P_{j_1}(u) \cdots P_{j_m}(u), & \text{if } n \geq m. \end{cases} \tag{1.1}$$

It is known that the limit law of the random variable X_n exhibits a "phase change" at $m = 26$: it is normal for $3 \leq m \leq 26$ and does not exist for $m > 26$; see Mahmoud and Pittel (1989), Lew and Mahmoud (1994), Chern and Hwang (2000) (referred to as CH in the sequel due to frequent citations) for details. Our aim in this paper is to improve the weak convergence in the case of normal limit law by proving the following theorem.

THEOREM 1. *Let $\Phi(x)$ denote the standard normal distribution. Then*

$$\begin{aligned} &\sup_{-\infty < x < \infty} \left| P\left(\frac{X_n - E(X_n)}{\sqrt{\operatorname{Var}(X_n)}} < x\right) - \Phi(x)\right| \\ &\qquad = \begin{cases} O(n^{-1/2}), & \textit{if } 3 \leq m \leq 19, \\ O(n^{-3(3/2-\alpha)}), & \textit{if } 20 \leq m \leq 26, \end{cases} \end{aligned} \tag{1.2}$$

where $4/3 < \alpha < 3/2$ *denotes the real part of the second largest zero*(*s*) (*in real part*) *of the indicial polynomial*

$$z(z+1)\cdots(z+m-2) - m! = 0. \tag{1.3}$$

The approximate values of α and $3(3/2-\alpha)$ are shown in Table 1.

The case when $m = 3$ is proved in CH using a direct analytic approach, but that approach (based on exact solvability of the associated partial differential equation) is not likely to be extended to cover more cases.

The O-terms in (1.2) are, up to implied constants, optimal since they are proportional to $E(X_n - E(X_n))^3/(\mathrm{Var}(X_n))^{3/2}$, and $E(X_n - E(X_n))^3$ is linear for $3 \le m \le 19$ and asymptotic to $n^{3\alpha-3}$ times a periodic function of $\log n$ for $20 \le m \le 26$; see (2.21) and (2.23). We will indeed derive a more precise local limit theorem for X_n; see (2.29).

Technically, while the line $\Re(z) = 3/2$ divides the normal and nonexistence behavior (first phase change), the line $\Re(z) = 4/3$ separates "good" convergence rate to "poor" ones (second phase change).

The result (1.2) not only gives concrete instances for which the Berry–Esseen bound is not proportional to the reciprocal of the standard deviation, but also sheds further light on the change of limit laws at $m = 26$ since the convergence rates are becoming poorer and poorer as m increases from 19 to 26, as shown in Table 1.

Then a natural question is "Why phase change?" We argued in CH that the main cause of the first phase change lies in the periodicity of the second-order term in the asymptotic expansion of $E(X_n)$, which in turn is due to the second largest zeros of the polynomial (1.3). A very rough intuitive interpretation is as follows. Observe first that X_n is degenerate for $n < m$ and that the calculation of X_n for large n involves a large number of terms of X_j's with $j < m$ by recursive decomposition. The contribution of the increasingly small degeneracy then leads to the change of the limit laws when m grows.

We proved in CH the asymptotic normality of X_n for $3 \le m \le 26$ by the method of moments; we extend that method of moments here but instead of

TABLE 1
Numeric values of α *and* $3(3/2-\alpha)$ *for* m *from* 20 *to* 26

m	α	$3(3/2-\alpha)$
20	1.34892881	0.45321354
21	1.38079786	0.35760639
22	1.40936978	0.27189065
23	1.43512896	0.19461309
24	1.45847025	0.12458925
25	1.47971848	0.06084455
26	1.49914326	0.00257020

the "asymptotic transfers" established there (using analytic approach), we will develop more elementary tools for handling inequalities. For other approaches to the asymptotic normality of recursive random variables, see Pittel (1999) (inductive approximation approach), Mahmoud (2002) (urn models), Neininger and Rüschendorf (2001) (contraction method), Devroye (2002) (Stein's method) and the references therein. The essential difference between the method of moments and the approach proposed in this paper is that for the method of moments, we need *asymptotics* for the central moments $E(X_n - E(X_n))^k$ for *each* $k \geq 0$ (*independent of* n), while for the convergence rate, we need *explicit upper bounds* for $E(X_n - E(X_n))^k$ for *all* $k \geq 0$ (*possibly dependent on* n).

Our proof of Theorem 1 is sketched as follows. We start by defining the scaled moment generating function

$$\phi_n(y) := e^{-\sigma_n^2 y^2/2} E\big(e^{(X_n - \mu_n)y}\big) = e^{-\mu_n y - \sigma_n^2 y^2/2} P_n(y),$$

where $\mu_n := E(X_n)$ and $\sigma_n^2 := \mathrm{Var}(X_n)$, which satisfies, by (1.1), $\phi_n(y) = 1$ for $0 \leq n \leq m-1$ and for $n \geq m$,

$$\phi_n(y) = \frac{1}{\binom{n}{m-1}} \sum_{\substack{j_1+\cdots+j_m=n-m+1\\ j_1,\ldots,j_m\geq 0}} \phi_{j_1}(y)\cdots\phi_{j_m}(y) e^{\Delta(\mathbf{j})y + \delta(\mathbf{j})y^2}, \tag{1.4}$$

where

$$\Delta(\mathbf{j}) := 1 + \mu_{j_1} + \cdots + \mu_{j_m} - \mu_n,$$
$$\delta(\mathbf{j}) := \tfrac{1}{2}\big(\sigma_{j_1}^2 + \cdots + \sigma_{j_m}^2 - \sigma_n^2\big).$$

Although the recurrence relation is not explicitly solvable for $m \geq 4$ (see CH), it will suffice to study the "moments" $\phi_{n,k} := \phi_n^{(k)}(0)$. These numbers satisfy the same type of recurrence equations obtained from (1.4) by successive differentiation at $y = 0$. Using these recurrences, we then show that for all $k \geq 0$,

$$|\phi_{n,k}| \leq k! A^k \cdot \begin{cases} n^{k/3}, & \text{if } 3 \leq m \leq 19, \\ n^{k(\alpha-1)}, & \text{if } 20 \leq m \leq 26, \end{cases}$$

where $A > 0$ is a suitably large constant. These estimates are the hard part of the proof and will enable us to derive precise estimate for the characteristic function of $(X_n - \mu_n)/\sigma_n$ when the parameter is small. Another inductive argument is then applied to derive a uniform estimate for the characteristic function. We then conclude Theorem 1 by the Berry–Esseen smoothing inequality [see Berry (1941); Esseen (1945); Petrov (1975)]. The estimates we derived are also strong enough to obtain a local limit theorem (of moderate-deviations type); see (2.29).

This approach is easily extended to recurrences appearing in the analysis of quicksort algorithms, which are the most widely used general-purpose sorting algorithms (for arranging given keys in increasing order). Quicksort was considered to be among the ten best algorithms in the 20th century "with the greatest influence on the development and practice of science and engineering"; see Dongarra and Sullivan (2000). The original quicksort proper invented by Hoare (1962) works, similar to the construction of binary search trees, as follows. Choose first a key from the input as the pivot; partition the input into three parts corresponding to keys whose values are less than, equal to, and larger than the pivot, respectively; then sort recursively the parts with smaller and larger keys until each part contains only one key (which is sorted). This version, although efficient for random input, suffers from several drawbacks such as quadratic cost in the worst case, inefficient when the input is near-sorted, nonstable for equal keys (the relative order of equal keys may not be preserved) and so on. Thus many schemes have been proposed to improve these; see Sedgewick (1980), Gonnet (1991).

Among the many variants of quicksort, we consider here Hennequin's generalized quicksort, where a random sample of $m(t+1)-1$ elements are first chosen and then the $(t+1)$st, $2(t+1)$th, ..., and $(m-1)(t+1)$th smallest elements in this sample are used to partition the remaining elements into m subgroups as in the branching construction of m-ary search trees; the m subgroups are sorted recursively until the file sizes are less than some threshold; then a final run of insertion sort completes the sorting task; see Hennequin (1989, 1991) for further details.

Assume as usual that the input is a sequence of independent and identically distributed random variables with a common continuous distribution. Then the cost measures of Hennequin's generalized quicksort satisfy recurrences of the type

$$Y_n \stackrel{d}{=} Y^{[1]}_{J_1} + \cdots + Y^{[m]}_{J_m} + T_n, \qquad n \geq m(t+1), \tag{1.5}$$

with suitable initial conditions, where $m \geq 2$ and $t \geq 0$ are integers, T_n is some random variable (called "toll function"), $(Y_n^{[1]}), \ldots, (Y_n^{[m]})$, $(J_1, \ldots, J_m)$ are independent and the $(Y_n^{[i]})$'s have the same distribution as (Y_n). Here

$$P(J_1 = j_1, \ldots, J_m = j_m) = \frac{\binom{j_1}{t} \cdots \binom{j_m}{t}}{\binom{n}{m(t+1)-1}},$$

for all tuples of nonnegative integers $(j_1, \ldots, j_m)$ such that $j_1 + \cdots + j_m = n - m + 1$. See Hennequin (1989, 1991) or CH for more information.

For simplicity, we consider the special case when $T_n = 0$ for $n \leq m(t+1) - 1$ and $T_n = 1$ for $n \geq m(t+1)$. Then Y_n denotes the number of partitioning stages used by Hennequin's generalized quicksort [see Hennequin (1989, 1991)]. Phase changes of the limit laws of Y_n have been derived in CH; we will describe the "second phase changes" of Y_n in Section 3. The same phase changes subsist for other cost measures of linear mean (see CH).

Another application to the quicksort proper [$(m, t) = (2, 0)$ in the generalized quicksort] is mentioned in Section 4:

$$Z_n \stackrel{d}{=} Z_{U_n} + Z^*_{n-1-U_n} + T_n, \qquad n \geq 2, \tag{1.6}$$

where U_n is uniformly distributed over $\{0, 1, \dots, n-1\}$. Unlike the preceding cases where either m or t varies, we consider the case when the "toll function" T_n varies and examine the effect of such a variation. The picture of such random variables has been investigated in Devroye (2002), Hwang and Neininger (2002) and is as follows. If the "toll function" is small, say $O(n^{1/2})$, then, under proper conditions, the limit law is normal; if the "toll function" is large, say $\gg n^{1/2}$, then the limit law is nonnormal (under appropriate conditions). This $1/2$-threshold for asymptotic normality is further strengthened by the $1/3$-threshold in which the rate to normality is $O(n^{-1/2})$ if the "toll function" is $\ll n^{1/3}$ and becomes slower for larger "toll functions." As before, the main "determinant" is the order of the third central moment. This consideration will clarify the connection of this order and the convergence rate to normality.

While the practical usefulness of m-ary search trees is limited due to its poor storage utilization $[E(X_n/n) \sim 1/(2/2 + 2/3 + \cdots + 2/m) > 1/(m-1)$; see (2.7)], they naturally introduce several intriguing phenomena and many challenging probabilistic problems. We derived the limit laws in CH by an analytic approach and the convergence rate and local limit theorems in this paper using only elementary tools. Tools are still lacking for, say large deviations problems: what are the rate functions for the large deviation principles for X_n for $m \geq 4$? [The result for $m = 3$ can be derived by the "quasi-power" approximation obtained in CH and the main result in Hwang (1996).] More phase changes are likely to appear.

NOTATION. The symbol $[z^n]f(z)$ represents the coefficient of z^n in the Taylor expansion of $f(z)$. The generic symbols ε, c, K (without subscript) will denote, respectively, suitably small, absolute, and large positive constants whose values may vary from one occurrence to another. For convenience, we also index these symbols with subscripts to denote constants with fixed values.

2. Convergence rate to the normal limit law.

We prove Theorem 1 in this section.

Mean and variance. Define $\mu := 1/(2H_m - 2)$, where $H_n := \sum_{1 \leq j \leq n} 1/j$. The mean μ_n of X_n satisfies [see Mahmoud and Pittel (1989); Mahmoud (1992); CH]

$$\mu_n = \mu n + \begin{cases} O(1), & \text{if } 3 \leq m \leq 13, \\ O(n^{\alpha-1}), & \text{if } m \geq 14, \end{cases} \tag{2.7}$$

since $\alpha - 1 > 0$ for $m \geq 14$; see Table 2.

The variance σ_n^2 of X_n satisfies [see Mahmoud and Pittel (1989); Mahmoud (1992)]

$$\sigma_n^2 = \begin{cases} \sigma^2 n + O(1), & \text{if } 3 \leq m \leq 13, \\ \sigma^2 n + O(n^{2\alpha-2}), & \text{if } 14 \leq m \leq 26, \\ \omega(\log n) n^{2\alpha-2} + o(n^{2\alpha-2}), & \text{if } m > 26, \end{cases} \tag{2.8}$$

TABLE 2
Approximate values of α for $3 \le m \le 20$

m	3	4	5	6	7	8	9	10	11
α	−3	−2.5	−1.5	−0.7682	−0.2663	0.1007	0.3665	0.5685	0.7262
m	12	13	14	15	16	17	18	19	20
α	0.8523	0.9552	1.0406	1.1125	1.1738	1.2267	1.2727	1.3131	1.3489

where $\sigma > 0$ is a constant (see CH) and $\omega(u)$ is a bounded periodic function. Note that the variance is linear for $3 \le m \le 26$ and larger than linear if $m > 26$ (since $2\alpha - 2 > 1$).

Recurrence of $\phi_{n,k}$. From (1.4), we deduce that, for $k \ge 1$,

$$\phi_{n,k} = \begin{cases} 0, & \text{if } 0 \le n \le m-1, \\ \dfrac{m}{\binom{n}{m-1}} \displaystyle\sum_{0 \le j \le n-m+1} \binom{n-1-j}{m-2} \phi_{j,k} + \psi_{n,k}, & \text{if } n \ge m, \end{cases} \tag{2.9}$$

where

$$\begin{aligned} \psi_{n,k} := & \sum_{\substack{i_0+\cdots+i_m+2i_{m+1}=k \\ 0 \le i_1,\ldots,i_m<k}} \frac{k!}{i_0! i_1! \cdots i_m! i_{m+1}!} \\ & \times \frac{1}{\binom{n}{m-1}} \sum_{\substack{j_1+\cdots+j_m=n-m+1 \\ j_1,\ldots,j_m \ge 0}} \phi_{j_1,i_1} \cdots \phi_{j_m,i_m} \Delta(\mathbf{j})^{i_0} \delta(\mathbf{j})^{i_{m+1}}. \end{aligned} \tag{2.10}$$

By definition,

$$\phi_{n,0} = 1 \quad \text{and} \quad \phi_{n,1} = \phi_{n,2} = 0. \tag{2.11}$$

Also by (2.7) and (2.8), we have

$$\begin{aligned} |\Delta(\mathbf{j})| &\le K_1(n^{\alpha-1} \vee 1), \\ |\delta(\mathbf{j})| &\le K_2(n^{2\alpha-2} \vee 1), \end{aligned} \tag{2.12}$$

uniformly for any tuples $\mathbf{j} = (j_1, \ldots, j_m)$, where $K_1, K_2 > 0$ are constants independent of n and $\mathbf{j}$.

We need tools for handling asymptotics of the recurrence

$$a_n = \frac{m}{\binom{n}{m-1}} \sum_{0 \le j \le n-m+1} \binom{n-1-j}{m-2} a_j + b_n, \qquad n \ge m, \tag{2.13}$$

with suitable initial conditions. To avoid ambiguity in the following discussions, we may, by modifying the values of b_n if necessary, take $a_n := b_n$ for $n < m$.

Asymptotic transfers for the recurrence (2.13). We bridge the asymptotics of b_n to that of a_n using the following result.

PROPOSITION 1. *Assume that a_n satisfies* (2.13). (i) *The conditions*

$$b_n = o(n) \quad \text{and} \quad \sum_n b_n n^{-2} < \infty, \tag{2.14}$$

are both necessary and sufficient for $a_n \sim c_0 n$, where

$$c_0 := \frac{1}{H_m - 1} \sum_{j \geq 0} \frac{b_j}{(j+1)(j+2)};$$

(ii) *if $|b_n| \leq c_1 n^{\upsilon}$, where $\upsilon > 1$, then*

$$|a_n| \leq \frac{K c_1}{1 - m!/((\upsilon+1)\cdots(\upsilon+m-1))} n^{\upsilon}, \tag{2.15}$$

uniformly in υ, where $K > 1$ is a constant independent of υ.

Unlike the method of moments used in CH, we need in (ii) more explicit upper bounds instead of asymptotic equivalent and o-estimate.

We first prove a lemma for handling the case of small υ.

LEMMA 1. *Assume that a_n satisfies* (2.13). *If $|b_n| \leq c_2 \binom{n+\upsilon}{n}$ for $n \geq 0$, where $\upsilon > 1$, then*

$$|a_n| \leq \frac{c_2}{1 - m!/((\upsilon+1)\cdots(\upsilon+m-1))} \binom{n+\upsilon}{n}. \tag{2.16}$$

PROOF. Observe first that (2.16) holds for $n < m$ by definition. Assume that $|a_j| \leq K_3 \binom{j+\upsilon}{j}$ for $0 \leq j \leq n-1$, $n > m$. Then

$$\begin{aligned}
|a_n| &\leq \frac{K_3 m}{\binom{n}{m-1}} \sum_{0 \leq j \leq n-m+1} \binom{n-1-j}{m-2} \binom{j+\upsilon}{j} + c_2 \binom{n+\upsilon}{n} \\
&= \frac{K_3 m}{\binom{n}{m-1}} [z^{n-1}] \frac{z^{m-2}}{(1-z)^{m-1}} \frac{1}{(1-z)^{\upsilon+1}} + c_2 \binom{n+\upsilon}{n} \\
&= K_3 m! \frac{\Gamma(\upsilon+1)}{\Gamma(\upsilon+m)} \binom{n+\upsilon}{n} + c_2 \binom{n+\upsilon}{n} \\
&\leq K_3 \binom{n+\upsilon}{n},
\end{aligned}$$

solving the last inequality gives

$$K_3 \geq \frac{c_2}{1 - m!\Gamma(\upsilon+1)/(\Gamma(\upsilon+m))} = \frac{c_2}{1 - m!/((\upsilon+1)\cdots(\upsilon+m-1))}.$$

This completes the induction. □

LEMMA 2. *If* $x, y \ge 0$, *then*

$$n^{-1} \sum_{0 \le j \le n-1} j^x (n-1-j)^y \le 2n^{x+y} \frac{\Gamma(x+1)\Gamma(y+1)}{\Gamma(x+y+2)}; \tag{2.17}$$

more generally,

$$\frac{1}{\binom{n}{\ell-1}} \sum_{\substack{j_1+\cdots+j_\ell=n-\ell+1 \\ j_1,\ldots,j_m \ge 0}} j_1^{x_1} \cdots j_\ell^{x_\ell} \le 2^{\ell-1}(\ell-1)! n^{x_1+\cdots+x_\ell} \frac{\Gamma(x_1+1)\cdots\Gamma(x_\ell+1)}{\Gamma(x_1+\cdots+x_\ell+\ell)}, \tag{2.18}$$

for $x_1, \ldots, x_\ell \ge 0$ *and* $\ell \ge 2$.

PROOF. First write the sum as a Stieltjes integral,

$$n^{-1} \int_0^{n-1} \upsilon^x (n-1-\upsilon)^y \, d\lfloor \upsilon \rfloor,$$

and then crudely bound the integral by

$$2n^{-1} \int_0^n \upsilon^x (n-\upsilon)^y \, d\upsilon;$$

making a change of variables yields a beta integral and thus (2.17).

The general version (2.18) follows by an induction on ℓ. $\square$

PROOF OF PROPOSITION 1. Case (i) is proved in CH. We prove case (ii).

Consider first the case when υ is small, say $\upsilon \le m+1$. Then by the asymptotic formula

$$\binom{n+\upsilon}{n} = \frac{n^\upsilon}{\Gamma(\upsilon+1)}\left(1 + O(n^{-1}|\upsilon|^2)\right),$$

we deduce that

$$c_3 \Gamma(\upsilon+1) \binom{n+\upsilon}{n} \le n^\upsilon \le c_4 \Gamma(\upsilon+1) \binom{n+\upsilon}{n},$$

uniformly for $1 \le \upsilon \le m+1$, for some constants $0 < c_3 < c_4 < \infty$. Thus

$$|b_n| \le c_1 n^\upsilon \le c_1 c_4 \Gamma(\upsilon+1) \binom{n+\upsilon}{n},$$

and we apply Lemma 1, obtaining

$$|a_n| \le \frac{c_1 c_4 \Gamma(\upsilon+1)}{1 - m!/((\upsilon+1)\cdots(\upsilon+m-1))}\binom{n+\upsilon}{n}$$
$$\le \frac{K c_1}{1 - m!/((\upsilon+1)\cdots(\upsilon+m-1))} n^{\upsilon},$$

where $K = c_4/c_3$.

For $\upsilon \ge m+1$, we use a different argument. Assume that $|a_j| \le K_4 j^{\upsilon}$ for $1 \le j \le n-1$. Then by induction, we have

$$\begin{aligned}
|a_n| &\le \frac{m}{\binom{n}{m-1}} \sum_{0\le j\le n-m+1} \binom{n-1-j}{m-2} K_4 j^{\upsilon} + c_1 n^{\upsilon} \\
&\le K_4 \frac{m(m-1)}{n} \sum_{0\le j\le n-m+1} \left(1 - \frac{j}{n-1}\right)^{m-2} j^{\upsilon} + c_1 n^{\upsilon} \\
&\le 2K_4 \frac{m(m-1)}{n} \int_0^n \left(1 - \frac{x}{n}\right)^{m-2} x^{\upsilon}\, dx + c_1 n^{\upsilon} \\
&= 2K_4 m! \frac{\Gamma(\upsilon+1)}{\Gamma(\upsilon+m)} n^{\upsilon} + c_1 n^{\upsilon} \\
&\le K_4 n^{\upsilon},
\end{aligned}$$

where we used (2.17). Solving the last inequality for K_4 gives

$$K_4 \ge \frac{c_1}{1 - 2m!\Gamma(\upsilon+1)/(\Gamma(\upsilon+m))} = \frac{c_1}{1 - 2m!/((\upsilon+1)\cdots(\upsilon+m-1))}.$$

Since $\upsilon \ge m+1$, the inequality $(\upsilon+1)\cdots(\upsilon+m-1) > 2m!$ holds for $m \ge 3$, and thus the denominator is bounded away from zero for $m \ge 3$. By suitably tuning the constants involved if needed, we deduce (2.16). $\square$

LEMMA 3. *Define*

$$S_m(k) := \sum_{\substack{i_1+\cdots+i_m=k\\ i_1,\ldots,i_m\ge 0}} \frac{\Gamma(i_1\bar{\alpha}+1)\cdots\Gamma(i_m\bar{\alpha}+1)}{\Gamma(k\bar{\alpha}+m)}, \qquad m \ge 2;\ k \ge 0. \tag{2.19}$$

Then for $k \ge 0$,

$$S_m(k) \le \frac{K_5^{m-1}}{(k\bar{\alpha}+1)\cdots(k\bar{\alpha}+m-1)}, \qquad m \ge 2. \tag{2.20}$$

PROOF. Consider first the case $m = 2$. By interchanging integration and summation and by summing the integrand, we obtain

$$
\begin{aligned}
S_2(k) &= \sum_{0\le j\le k} \int_0^1 x^{j\bar\alpha}(1-x)^{(k-j)\bar\alpha}\,dx \\
&= 2\int_0^{1/2} \frac{(1-x)^{(k+1)\bar\alpha} - x^{(k+1)\bar\alpha}}{(1-x)^{\bar\alpha} - x^{\bar\alpha}}\,dx \\
&\le \bigl(2+o(1)\bigr)\int_0^{\infty} e^{-k\bar\alpha x}\,dx \\
&\le \frac{K_5}{k\bar\alpha+1}.
\end{aligned}
$$

By induction, for $m \ge 3$,

$$
\begin{aligned}
S_m(k) &= \sum_{0\le j\le k} \frac{\Gamma(j\bar\alpha+1)\Gamma((k-j)\bar\alpha+m-1)}{\Gamma(k\bar\alpha+m)} S_{m-1}(k-j) \\
&\le \frac{K_5^{m-2}}{(k\bar\alpha+2)\cdots(k\bar\alpha+m-1)} \sum_{0\le j\le k} \frac{\Gamma(j\bar\alpha+1)\Gamma((k-j)\bar\alpha+1)}{\Gamma(k\bar\alpha+2)} \\
&\le \frac{K_5^{m-1}}{(k\bar\alpha+1)\cdots(k\bar\alpha+m-1)}.
\end{aligned}
$$

This proves (2.20). □

Estimate of $\phi_{n,3}$. We first determine the order of $\phi_{n,3} = E(X_n - \mu_n)^3$, which plays the determinant role in the rate (1.2).

By the definition of $\psi_{n,3}$ by (2.10) and (2.11), (2.12), we obtain

$$
\begin{aligned}
\psi_{n,3} &= \frac{1}{\binom{n}{m-1}} \sum_{\substack{j_1+\cdots+j_m=n-m+1\\ j_1,\dots,j_m\ge 0}} \bigl(6\Delta(\mathbf{j})\delta(\mathbf{j}) + \Delta(\mathbf{j})^3\bigr) \\
&= O(n^{3\alpha-3}+1);
\end{aligned}
$$

it follows by Proposition 1 that

$$\phi_{n,3} \sim c_5 n, \tag{2.21}$$

for $3 \le m \le 19$ and

$$\phi_{n,3} = O(n^{3\alpha-3}),$$

if $20 \le m \le 26$, since $3\alpha - 3 > 1$ for $m > 19$. For simplicity, define

$$\bar\alpha := \begin{cases} 1/3, & \text{if } 3 \le m \le 19,\\ \alpha - 1, & \text{if } 20 \le m \le 26,\end{cases}$$

so that we can write

$$|\phi_{n,3}| \le K_6 n^{3\bar{\alpha}}, \qquad 3 \le m \le 26. \tag{2.22}$$

Note that by refining the analytic approach given in CH, we can show that, for $20 \le m \le 26$,

$$\phi_{n,3} \sim \varpi(\log n) n^{3\alpha - 3}, \tag{2.23}$$

where $\varpi(u)$ is a continuous, periodic function of bounded fluctuation; the details being laborious and uninteresting are omitted here. [Roughly, we first write

$$\begin{aligned} E(X_n - \mu_n)^3 &= E\big(X_n - \mu(n+1)\big)^3 - 3E\big(X_n - \mu(n+1)\big)^2\big(\mu_n - \mu(n+1)\big) \\ &\quad + 2\big(\mu_n - \mu(n+1)\big)^3, \end{aligned}$$

and then apply the same arguments as in CH to derive more precise approximations for $E(X_n - \mu(n+1))^3$.]

An upper bound for $\phi_{n,k}$. We now proceed by induction to show that

$$|\phi_{n,k}| \le k! A^k n^{k\bar{\alpha}}, \tag{2.24}$$

where $A > 0$ is a sufficiently large constant to be specified later. The inequality (2.24) holds with $A > (K_6/6)^{1/3}$ for $0 \le k \le 3$ by (2.11) and (2.22).

By (2.12), (2.22) and induction, we have

$$\begin{aligned} |\psi_{n,k}| \le k! \sum_{\substack{i_0 + \cdots + i_m + 2i_{m+1} = k \\ 0 \le i_1, \ldots, i_m < k}} \frac{K_1^{i_0} K_2^{i_{m+1}}}{i_0! i_{m+1}!} A^{i_1 + \cdots + i_m} n^{i_0 \bar{\alpha} + 2 i_{m+1} \bar{\alpha}} \\ \times \frac{1}{\binom{n}{m-1}} \sum_{\substack{j_1 + \cdots + j_m = n - m + 1 \\ j_1, \ldots, j_m \ge 0}} j_1^{i_1 \bar{\alpha}} \cdots j_m^{i_m \bar{\alpha}}. \end{aligned}$$

By applying (2.18), we obtain

$$\begin{aligned} |\psi_{n,k}| &\le 2^m (m-1)! k! n^{k\bar{\alpha}} \sum_{\substack{i_0 + \cdots + i_m + 2i_{m+1} = k \\ 0 \le i_1, \ldots, i_m < k}} \frac{K_1^{i_0} K_2^{i_{m+1}}}{i_0! i_{m+1}!} A^{i_1 + \cdots + i_m} \\ &\qquad \times \frac{\Gamma(i_1 \bar{\alpha} + 1) \cdots \Gamma(i_m \bar{\alpha} + 1)}{\Gamma((i_1 + \cdots + i_m)\bar{\alpha} + m)} \\ &\le 2^m (m-1)! e^{K_1 + K_2} k! n^{k\bar{\alpha}} \sum_{\substack{i_1 + \cdots + i_m \le k \\ 0 \le i_1, \ldots, i_m < k}} A^{i_1 + \cdots + i_m} \frac{\Gamma(i_1 \bar{\alpha} + 1) \cdots \Gamma(i_m \bar{\alpha} + 1)}{\Gamma((i_1 + \cdots + i_m)\bar{\alpha} + m)} \\ &\le 2^m (m-1)! e^{K_1 + K_2} k! n^{k\bar{\alpha}} \sum_{0 \le p \le k} A^p S_m(p), \end{aligned}$$

where $S_m(p)$ is defined in (2.19). By (2.20), we deduce that

$$\begin{aligned}\sum_{0\le p\le k} A^p S_m(p) &\le K_5^{m-1} \sum_{0\le p\le k} \frac{A^p}{(p\bar{\alpha}+1)\cdots(p\bar{\alpha}+m-1)}\\ &\le K \frac{A^k}{(k\bar{\alpha}+1)\cdots(k\bar{\alpha}+m-1)}.\end{aligned}$$

It follows that

$$|\psi_{n,k}| \le \frac{K_6}{(k\bar{\alpha}+1)\cdots(k\bar{\alpha}+m-1)} k!A^k n^{k\bar{\alpha}}.$$

Transferring this bound to $\phi_{n,k}$ via (2.15), we obtain

$$|\phi_{n,k}| \le \frac{K_6}{(k\bar{\alpha}+1)\cdots(k\bar{\alpha}+m-1)-m!} k!A^k n^{k\bar{\alpha}}. \tag{2.25}$$

Thus if

$$\frac{K_6}{(k\bar{\alpha}+1)\cdots(k\bar{\alpha}+m-1)-m!} \le 1,$$

for $k > k_0$, then (2.24) holds for $k > k_0$. It remains to tune the value of A so that (2.24) holds for $4 \le k \le k_0$. Define

$$A_r := \left(\frac{K_6}{6}\right)^{1/3} \prod_{4\le j\le r} \left(\frac{K_6}{(j\bar{\alpha}+1)\cdots(j\bar{\alpha}+m-1)-m!}\right)^{1/j}, \qquad 4 \le r \le k_0.$$

Observe that (2.25) relies only on A_j, $j < k$. Thus (2.25) can be written as

$$|\phi_{n,r}| \le \frac{K_6}{(r\bar{\alpha}+1)\cdots(r\bar{\alpha}+m-1)-m!} r!A_{r-1}^r n^{r\bar{\alpha}} = r!A_r^r n^{r\bar{\alpha}}.$$

The proof is complete by taking $A = A_{k_0}$. $\square$

An estimate for the difference of the characteristic functions. Consider now the characteristic function

$$\varphi_n(y) := E\left(e^{(X_n-\mu_n)iy/\sigma_n}\right) = e^{-\mu_n iy/\sigma_n} P_n(iy/\sigma_n) = e^{-y^2/2}\phi_n(iy/\sigma_n).$$

By (2.24),

$$\begin{aligned}\left|\varphi_n(y) - e^{-y^2/2}\right| &= e^{-y^2/2}\,|\phi_n(iy/\sigma_n) - 1|\\ &\le e^{-y^2/2} \sum_{k\ge 3} \frac{|\phi_{n,k}|}{k!\sigma_n^k}|y|^k\\ &\le e^{-y^2/2} \sum_{k\ge 3} \left(An^{\bar{\alpha}}\sigma_n^{-1}|y|\right)^k\\ &\le Ke^{-y^2/2}|y|^3 n^{-3/2+3\bar{\alpha}},\end{aligned} \tag{2.26}$$

if $|y| \le \varepsilon_0 n^{1/2-\bar{\alpha}}$. We need another estimate for $|\varphi_n(y)|$ for larger $|y|$.

Note that from (2.26) we have

$$|\varphi_n(y)| \le e^{-y^2/2}\bigl(1 + K|y|^3 n^{-3/2+3\bar{\alpha}}\bigr)$$
$$\le e^{-y^2/2+K|y|^3 n^{-3/2+3\bar{\alpha}}}$$

for $|y| \le \varepsilon_0 n^{1/2-\bar{\alpha}}$. In terms of $P_n(iu)$, this yields

$$|P_n(iu)| \le e^{-\sigma_n^2 u^2/2+K|u|^3 n^{3\bar{\alpha}}}$$

for $|u| \le \varepsilon_1 n^{-\bar{\alpha}}$, $u \in \mathbb{R}$.

By definition, $|P_n(iu)| = 1$ for $0 \le n \le m$ and

$$P_n(iu) = \sum_{1\le j\le n-m+1} \frac{\binom{m}{j}\binom{n-m}{j-1}}{\binom{n}{m-1}} e^{i(j+1)u}, \qquad m \le n \le 2m-2,$$

for real u. Thus

$$|P_n(iu)| \le e^{-\varepsilon_2 (n+K_8)u^2}, \tag{2.27}$$

for $|u| \le \varepsilon_1 n^{-\bar{\alpha}}$ and $n \ge m+1$, where $K_8 > 1$ is a suitably chosen constant to be specified later.

A uniform estimate for the characteristic function. We now show by induction that the same estimate (2.27) holds for $|u| \le \varepsilon_3$, where $0 < \varepsilon_3 < \pi$ is sufficiently small.

Take $\varepsilon_3 := \varepsilon_1 n_0^{-\bar{\alpha}}$, where n_0 is a large constant. Then, if $|u| \le \varepsilon_1 n^{-\bar{\alpha}}$, then $|u| \le \varepsilon_3$ for $n \le n_0$, so that (2.27) holds for $m+1 \le n \le n_0$ and $|u| \le \varepsilon_3$. By induction using (1.1),

$$\begin{aligned}
&|P_n(iu)| \\
&\quad\le \frac{1}{\binom{n}{m-1}} \sum_{\substack{j_1+\cdots+j_m=n-m+1\\ j_1,\dots,j_m\ge 0}} |P_{j_1}(iu)|\cdots|P_{j_m}(iu)| \\
&\quad\le \frac{1}{\binom{n}{m-1}} [z^{n-m+1}]\Biggl(1+z+\cdots+z^m + \sum_{j\ge m+1} e^{-\varepsilon_2 (j+K_8)u^2} z^j\Biggr)^m \\
&\quad= \frac{1}{\binom{n}{m-1}} [z^{n-m+1}]\Biggl(\frac{e^{-\varepsilon_2 K_8 u^2}}{1-e^{-\varepsilon_2 u^2} z} + \sum_{0\le j\le m} \bigl(1-e^{-\varepsilon_2 (j+K_8)u^2}\bigr) z^j\Biggr)^m \\
&\quad= \frac{e^{-\varepsilon_2 (n-m+1)u^2}}{\binom{n}{m-1}} [z^{n-m+1}]\Biggl(\frac{e^{-\varepsilon_2 K_8 u^2}}{1-z} + \sum_{0\le j\le m} \bigl(1-e^{-\varepsilon_2 (j+K_8)u^2}\bigr) e^{\varepsilon_2 j u^2} z^j\Biggr)^m \\
&\quad= e^{-\varepsilon_2 (n+K_8)u^2}\bigl(e^{-\varepsilon_2 (m-1)(K_8-1)u^2} + R_n(u)\bigr)
\end{aligned}$$

for $n > n_0$, where we used the relation $[z^k]f(xz) = x^k[z^k]f(z)$ and

$$R_n(u) := \frac{1}{\binom{n}{m-1}} \sum_{0\le \ell<m} \binom{m}{\ell} e^{-\varepsilon_2(\ell K_8 - K_8 - m+1)u^2} [z^{n-m+1}](1-z)^{-\ell} \times \left(\sum_{0\le j\le m} (1-e^{-\varepsilon_2(j+K_8)u^2}) e^{\varepsilon_2 j u^2} z^j \right)^{m-\ell}.$$

Since both m and u are finite and for $\ell < m$, $k \ge 0$,

$$[z^{n-m+1}](1-z)^{-\ell} z^k = \binom{n-m+\ell-k}{\ell-1} \le \binom{n-1}{m-2} = \frac{m-1}{n}\binom{n}{m-1},$$

we have $R_n(u) = O(n^{-1})$; on the other hand, for $\varepsilon_1 n^{-\bar\alpha} \le |u| \le \varepsilon_3$,

$$u^2 \ge \varepsilon_1^2 n^{-2\bar\alpha} \gg n^{-1}.$$

Thus we first take n_0 so large that

$$\max_{|u|\le\pi} |R_n(u)| \le K_9/n,$$

for $n \ge n_0$; then we take K_8 such that the inequality

$$e^{-\varepsilon_2(m-1)(K_8-1)u^2} \le 1 - 2K_9/n$$

holds for $n \ge n_0$ and $\varepsilon_1 n^{-\bar\alpha} \le |u| \le \varepsilon_3$. This completes the induction and proves (2.27) for $|u| \le \varepsilon_3$.

We now extend the range from $|u| \le \varepsilon_3$ to $|u| \le \pi$. The same argument as above applies provided that we can show that

$$|P_n(iu)| \le e^{-\varepsilon_2(n+K_8)u^2}, \qquad \varepsilon_3 \le |u| \le \pi,$$

for $m+1 \le n \le n_0$. This follows easily from (i) the span of X_n is 1 (by induction); and (ii)

$$|P_n(iu)| \le 1 - \varepsilon_4,$$

for $\varepsilon_3 \le |u| \le \pi$. (Indeed, we need only tune the value of K_8 if needed.)

In terms of the characteristic function $\varphi_n(y)$, we have

$$|\varphi_n(y)| \le e^{-\varepsilon y^2}, \tag{2.28}$$

uniformly for $|y| \le \pi\sigma_n$.

Berry–Esseen smoothing inequality. We now apply the Berry–Esseen smoothing inequality [see Petrov (1975)], which states for our problem that

$$\sup_{-\infty<x<\infty}\left|P\left(\frac{X_n-E(X_n)}{\sqrt{\operatorname{Var}(X_n)}}<x\right)-\Phi(x)\right| = O\left(T^{-1}+\int_{-T}^{T}\left|\frac{\varphi_n(y)-e^{-y^2/2}}{y}\right|dy\right),$$

where T is taken to be $\varepsilon n^{3/2-3\bar{\alpha}}$. By the two estimates (2.26) and (2.28), we easily have

$$\begin{aligned}
&\int_{-T}^{T}\left|\frac{\varphi_n(y)-e^{-y^2/2}}{y}\right|dy\\
&\quad = O\left(n^{-3/2+3\bar{\alpha}}\int_{-\varepsilon_0 n^{1/2-\bar{\alpha}}}^{\varepsilon_0 n^{1/2-\bar{\alpha}}} e^{-y^2/2}y^2\,dy+\int_{\varepsilon_0 n^{1/2-\bar{\alpha}}}^{\varepsilon n^{3/2-3\bar{\alpha}}}\frac{e^{-Ky^2}+e^{-y^2/2}}{y}\,dy\right)\\
&\quad = O\left(n^{-3/2+3\bar{\alpha}}\right)+O\left(n^{-1/2+\bar{\alpha}}e^{-\varepsilon n^{1-2\bar{\alpha}}}\right).
\end{aligned}$$

This proves Theorem 1. □

Local limit theorem. By the inversion formula

$$\begin{aligned}
P(X_n=k) &= \frac{1}{2\pi}\int_{-\pi}^{\pi} e^{-iky}P_n(iy)\,dy\\
&= \frac{1}{2\pi\sigma_n}\int_{-\pi\sigma_n}^{\pi\sigma_n} e^{-ixy}\left(1+O(|y|\sigma_n^{-1})\right)\varphi_n(y)\,dy,
\end{aligned}$$

where $k=\lfloor\mu_n+x\sigma_n\rfloor$, we deduce, by splitting the integral similarly as above, the following local limit theorem.

THEOREM 2. *Uniformly for* $x=o(n^{1/2-\bar{\alpha}})$,

$$P(X_n=\lfloor\mu_n+x\sigma_n\rfloor)=\frac{e^{-x^2/2}}{\sqrt{2\pi}\,\sigma_n}\left(1+O\left((1+|x|^3)n^{-3/2+3\bar{\alpha}}\right)\right). \tag{2.29}$$

3. Second phase changes in generalized quicksort. Let $m\ge 2$ and $t\ge 0$ be two fixed integers. We consider in this section Y_n, the number of partitioning stages used by the generalized quicksort of Hennequin [see Hennequin (1989, 1991)]. Recall that, by (1.5), $Q_n(u):=E(e^{Y_n u})$ satisfies the recurrence

$$Q_n(u)=\begin{cases}1, & \text{if } 0\le n\le m(t+1)-1,\\[4pt] e^u\displaystyle\sum_{\substack{j_1+\cdots+j_m=n-m+1\\ j_1,\dots,j_m\ge 0}}\frac{\binom{j_1}{t}\cdots\binom{j_m}{t}}{\binom{n}{m(t+1)-1}}Q_{j_1}(u)\cdots Q_{j_m}(u), & \\ & \text{if } n\ge m(t+1).\end{cases}$$

TABLE 3
All pairs of integers (m, t) for which Y_n are asymptotically normally distributed

m	2	3	4	5	6	7	8, 9	10, ..., 13	14, ..., 26
t	1, ..., 58	0, ..., 19	0, ..., 10	0, ..., 6	0, ..., 4	0, ..., 3	0, 1, 2	0, 1	0

The thresholds separating normal limit law from nonexistence of limit law for Y_n are derived in CH and are repeated in Table 3.

Define the set $\mathcal{S}_1$ of integer pairs (m, t) by Table 4.

TABLE 4
All pairs of integers (m, t) for which the Berry–Essen bounds are $O(n^{-1/2})$

m	2	3	4	5	6	7	8, 9, 10	11, ..., 19
t	1, ..., 43	0, ..., 14	0, ..., 7	0, ..., 4	0, ..., 3	0, 1, 2	0, 1	0

Define the set $\mathcal{S}_2$ of integer pairs (m, t) by Table 5.

TABLE 5
All pairs of integers (m, t) for which the Berry–Essen bounds are of the form $O(n^{-3(3/2-a)})$, where $4/3 < a < 3/2$. Note that there is no such pair when $m = 10, 14, \ldots, 19$

m	2	3	4	5	6	7	8, 9	11, 12, 13	20, ..., 26
t	44, ..., 58	15, ..., 19	8, 9, 10	5, 6	4	3	2	1	0

THEOREM 3.

$$\sup_{-\infty<x<\infty} \left| P\left(\frac{Y_n - E(Y_n)}{\sqrt{\operatorname{Var}(Y_n)}} < x \right) - \Phi(x) \right| = \begin{cases} O(n^{-1/2}), & \text{if } (m,t) \in \mathcal{S}_1, \\ O(n^{-3(3/2-a)}), & \text{if } (m,t) \in \mathcal{S}_2, \end{cases} \tag{3.30}$$

where $4/3 < a < 3/2$ denotes the real part of the second largest zero(s) (in real part) of the indicial polynomial

$$z(z+1)\cdots(z+m(t+1)-2) - m\frac{(m(t+1)-1)!}{t!} z\cdots(z+t-1) = 0. \tag{3.31}$$

The proof follows the same line of arguments used for X_n and is omitted here; see CH for estimates and tools needed.

4. Second phase change in quicksort recurrence. We consider in this section random variables Z_n associated with the cost of quicksort proper, namely (1.6), which, defining $R_n(u) := E(e^{Z_n u})$, translates into the recurrence

$$R_n(u) = \frac{1}{n} \sum_{0 \le j < n} R_j(u) R_{n-1-j}(u) W_{n,j}(u), \qquad n \ge 1,$$

with $R_0(u) := 1$, where $n^{-1} \sum_{0 \le j < n} W_{n,j}(u) := E(e^{T_n u})$, T_n being the "toll cost" used to divide the original problem into two smaller problems. If $E(T_n) = O(\sqrt{n} L(n))$, where $L(n)$ is slowly varying, then under suitable conditions on T_n, Z_n is asymptotically normally distributed; see Devroye (2002), Hwang and Neininger (2002). We derive convergence rates in this section under slightly stronger conditions. Note that we allow explicit dependence of T_n on the rank of the "partitioning key" j.

THEOREM 4. *If*

$$|E(T_n^k)| \le \tau_k n^{k\beta}, \qquad k = 1, 2, \ldots,$$

where $0 \le \beta < 1/2$ *and the sequence* $\{\tau_k\}$ *satisfies* $\sum_k \tau_k \varepsilon^k / k! < \infty$ *for some* $\varepsilon > 0$, *then*

$$\sup_{-\infty < x < \infty} \left| P\left(\frac{Z_n - E(Z_n)}{\sqrt{\operatorname{Var}(Z_n)}} < x \right) - \Phi(x) \right|$$
$$= \begin{cases} O(n^{-1/2}), & \text{if } 0 \le \beta < 1/3, \\ O(n^{-1/2} \log n), & \text{if } \beta = 1/3, \\ O\big(n^{-3(1/2-\beta)}\big), & \text{if } 1/3 < \beta < 1/2. \end{cases}$$

The same method of proof extends to the case $\beta = 1/2$ for which the convergence rate drops from polynomial to logarithmic; we content ourselves with the current version for simplicity. Since the proof does not require any new argument, we omit it.

While our approach applies well to the normal range, it is unclear how it can apply to the range when the limit law exists and is not normal [roughly, when $E(T_n) \gg n^{1/2}$]. The simplest concrete example is $T_n \equiv n - 1$; in this case Z_n is the total number of comparisons used by the quicksort proper. A convergence rate for Kolmogorov distance of order $n^{-1/2+\varepsilon}$ was recently derived by Fill and Janson (2002). It is generally conjectured that the true rate should be of order $n^{-1} \log n$, but no proof has yet been found.

Acknowledgments. Thanks to the School of Computer Science, McGill University for its hospitality and support. The author also thanks Zhi-Dong Bai, the Editor and the two referees for helpful comments and suggestions, and

Bob Sedgewick for suggesting the intuitive interpretation of the phase change of limit laws.

REFERENCES

ALDOUS, D. (1996). Probability distributions on cladograms. In *Random Discrete Structures* (D. Aldous and R. Pemantle, eds.) 1–18. Springer, New York.

BARLOW, M. T., PEMANTLE, R. and PERKINS, E. A. (1997). Diffusion-limited aggregation on a tree. *Probab. Theory Related Fields* **107** 1–60.

BEN-NAIM, E., KRAPIVSKY, P. L. and MAJUMDAR, S. N. (2001). Extremal properties of random trees. *Phys. Rev. E* **64**.

BERRY, A. C. (1941). The accuracy of the Gaussian approximation to the sum of independent variates. *Trans. Amer. Math. Soc.* **49** 122–136.

CHERN, H.-H. and HWANG, H.-K. (2001). Phase changes in random m-ary search trees and generalized quicksort. *Random Structures Algorithms* **19** 316–358.

DEVROYE, L. (1999). Universal limit laws for depths in random trees. *SIAM J. Comput.* **28** 409–432.

DEVROYE, L. (2002). Limit laws for sums of functions of subtrees of random binary search trees. *SIAM J. Comput.* **32** 152–171.

DONGARRA, J. and SULLIVAN, S. (2000). The top 10 algorithms. *Computing in Science and Engineering* January/February 22–23.

ESSEEN, C.-G. (1945). Fourier analysis of distribution functions. A mathematical study of the Laplace–Gaussian law. *Acta Math.* **77** 1–125.

FILL, J. A. and JANSON, S. (2002). Quicksort asymptotics. *J. Algorithms* **44** 4–28.

GONNET, G. H. (1991). *Handbook of Algorithms and Data Structures* (*in Pascal and C*). Addison-Wesley, Reading, MA.

HENNEQUIN, P. (1989). Combinatorial analysis of quicksort algorithm. *RAIRO Inform. Théor. Appl.* **23** 317–333.

HENNEQUIN, P. (1991). *Analyse en moyenne d'algorithme, tri rapide et arbres de recherche*. Ph.D. thesis, Ecole Polytechnique. Available at pauillac.inria.fr/algo/AofA/Research/src/Hennequin.PhD.html.

HOARE, C. A. R. (1962). Quicksort. *Comput. J.* **5** 10–15.

HWANG, H.-K. (1996). Large deviations for combinatorial distributions. I. Central limit theorems. *Ann. Appl. Probab.* **6** 297–319.

HWANG, H.-K. and NEININGER, R. (2002). Phase changes of the limit laws in the quicksort recurrence under varying toll functions. *SIAM J. Comput.* **31** 1687–1722.

LEW, W. and MAHMOUD, H. M. (1994). The joint distribution of elastic buckets in multiway search trees. *SIAM J. Comput.* **23** 1050–1074.

MAHMOUD, H. M. (1992). *Evolution of Random Search Trees*. Wiley, New York.

MAHMOUD, H. M. (2002). The size of random bucket trees via urn models. *Acta Inform.* **38** 813–383.

MAHMOUD, H. M. and PITTEL, B. (1984). On the most probable shape of a search tree grown from random permutations. *SIAM Journal on Algebraic and Discrete Methods* **1** 69–81.

MAHMOUD, H. M. and PITTEL, B. (1989). Analysis of the space of search trees under the random insertion algorithm. *J. Algorithms* **10** 52–75.

NEININGER, R. and RÜSCHENDORF, L. (2001). A general contraction theorem and asymptotic normality in combinatorial structures. Available at www.stochastik.uni-freiburg.de/homepages/neininger.

PETROV, V. V. (1975). *Sums of Independent Random Variables*. Springer, New York.

PITTEL, B. (1999). Normal convergence problem? Two moments and a recurrence may be the clues. *Ann. Applied Probab.* **9** 1260–1302.
SEDGEWICK, R. (1980). *Quicksort*. Garland, New York.

INSTITUTE OF STATISTICAL SCIENCE
ACADEMIA SINICA
TAIPEI 115
TAIWAN
E-MAIL: hkhwang@stat.sinica.edu.tw
WEB: http://algo.stat.sinica.edu.tw