JOURNAL OF ALGORITHMS **21**, 476-507 (1996)
ARTICLE NO. 0055

Large Deviations for Quicksort*

C.J.H. McDiarmid†

Department of Statistics, Oxford University, Oxford OX1 3TG, United Kingdom

AND

R.B. Hayward‡

Department of Mathematics and Computer Science, University of Lethbridge, 4401 University Drive W., Lethbridge, Alberta, Canada T1K 3M4

Received March 22, 1993

Let Q_n be the random number of comparisons made by quicksort in sorting n distinct keys, when we assume that all $n!$ possible orderings are equally likely. Known results concerning moments for Q_n do not show how rare it is for Q_n to make large deviations from its mean. Here we give a good approximation to the probability of such a large deviation, and find that this probability is quite small. As well as the basic quicksort we consider the variant in which the partitioning key is chosen as the median of $(2t+1)$ keys.

1 Introduction

In the short history of computer science, Hoare's quicksort has emerged as one of the classic algorithms. There are several reasons for this.

First, the algorithm is efficient. It is a $\Theta(n \log n)$ expected time sorting algorithm (where n is the number of keys to be sorted). Arguably one of the best general purpose computer sorting algorithms from a point of view

* The second author gratefully acknowledges the support of N.S.E.R.C. and the Alexander von Humboldt-Stiftung. Some work was carried out while at the Institute for Discrete Mathematics in Bonn, and the Department of Computing and Information Science at Queen's University, Kingston. A shortened version of this paper appeared in the proceedings of the Third Annual ACM-SIAM Symposium on Discrete Algorithms [MH92].

† E-mail: cmcd%statistics.oxford.ac.uk@prg.oxford.ac.uk

‡ E-mail: hayward@ultrix.uleth.ca

of space and time efficiency, it is the basis for example of the Unix "sort" feature, which uses the variant of quicksort in which the partitioning key is chosen as the median of three keys.

Second, the algorithm embodies two paradigms that are today considered key ideas in algorithm design, namely divide and conquer, and randomization.

Third, since the introduction of quicksort [Hoa61], an extensive body of literature has been published that is based on the design and analysis of many variants of the original algorithm. (Many of the commonly considered variants were in fact anticipated by Hoare in [Hoa62]. For a discussion of this, see [Sed80].) Indeed, the study of quicksort and its variants has become a model for the analysis of algorithms in general. Examples in point are the work by Knuth [Knu73] and by Sedgewick [Sed80]. (For recent relevant results on the analysis of quicksort, see [Hen89], [Rég89], [Rös91] and [Mah92].)

The aim of this paper is to establish a new and rather precise result concerning quicksort's typical performance. We establish much tighter bounds than have been shown previously on

> *the probability that the number of comparisons of a random execution of quicksort will have a large deviation from the expected number of comparisons.*

Our proofs involve the recently popularised combinatorial approach known as the "method of bounded differences".

Before introducing the definitions we need in order to state our theorems precisely, let us discuss the two variants of quicksort we shall refer to in this paper. First, by "basic" quicksort (and unless otherwise stated, that is the version we will be referring to) we mean the original, unadorned version:

> *A partitioning key is selected at random from the list of unsorted keys, and used to partition the keys into two sublists. The algorithm is called recursively on remaining unsorted sublists, until sublists have size one or zero.*

One common variant is to use a different sorting algorithm (usually insertion sort) for lists whose size is not greater than a certain threshold value M. We refer to this variant as "cutting at length M". Another common variant is to select the partitioning key as the median of $2t+1$ keys selected from the list of unsorted keys. (Observe that basic quicksort can be considered as the "median of 1" version of this variant.) A comprehensive survey and comparative analysis of common variants is given in [Sed80].

We now introduce some notation. Let Q_n be the number of key comparisons made when (basic) quicksort sorts n keys. We make the usual assumption that the n keys are distinct and that all $n!$ linear orders are equally likely. (Alternatively, our results apply to a suitably randomised version of quicksort acting on any list of n distinct keys.) We further assume that the partitioning phase of the algorithm is performed so that the resulting sublists are also "random". Some care must be taken to ensure this, but it is not difficult to do. See [Sed80] for a description of such a partitioning algorithm, and a proof that the randomness of the sublists is preserved.

A straightforward and well known analysis (in fact, so well known it is likely to appear in any undergraduate algorithms course!) shows that the expected number $q_n = E[Q_n]$ of key comparisons satisfies $q_0 = 0$ and for $n \geq 1$

$$q_n \;=\; n-1 \;+\; \frac{1}{n}\sum_{j=1}^{n}(q_{j-1}+q_{n-j})\,. \tag{1}$$

Define H_n so that $H_0 = 0$ and $H_t = \sum_1^t 1/j$ for all integers $t \geq 1$. It follows easily from Equation 1 that $q_n = 2(n+1)H_n - 4n$. Since $H_n = \ln n + \gamma + O(1/n)$ as $n \to \infty$, (where γ is Euler's constant, namely $0.5772156649\ldots$), it follows that as $n \to \infty$,

$$q_n \;=\; 2n\ln n \;-\; (4-2\gamma)n \;+\; 2\ln n \;+\; O(1)\,.$$

Of course one attaches more credibility to an average case result if it is known that there is a strong concentration of probability around the mean. In particular, given $\varepsilon > 0$, what bounds can be placed on the quantity

$$\Pr\left[\,\left|\,\frac{Q_n}{q_n} - 1\,\right| > \varepsilon\,\right] \quad ?$$

The point of this paper is to establish tight bounds for the above probability. There are some previous results of this form. For instance, from the fact that the variance of Q_n is $\Theta(n^2)$ (see [Knu73] or [Sed80]), it follows using Chebyshev's inequality that for $\varepsilon > 0$

$$\Pr\left[\,\left|\,\frac{Q_n}{q_n} - 1\,\right| > \varepsilon\,\right] \;=\; O((\varepsilon\ln n)^{-2})\,.$$

Recently Hennequin [Hen89] used Chebyshev's inequality with fourth moments to show that for $\varepsilon > 0$ the above probability is $O((\varepsilon\ln n)^{-4})$. Even more recently Rösler [Rös91] improved these upper bounds dramatically , and showed that for fixed ε this probability is $O(n^{-k})$ for any fixed k.

We shall see that the probability of such deviations is even smaller than this. For fixed $\varepsilon > 0$ we shall show that

$$\Pr\left[\,\left|\,\frac{Q_n}{q_n} - 1\,\right| > \varepsilon\,\right] \;=\; n^{-(2+o(1))\,\varepsilon\ln^{(2)} n} \tag{2}$$

as $n \to \infty$. Note that here we use $\ln n$ to denote $\log_e n$. Also, we use the notation $\ln^{(k)} n$, where $\ln^{(1)} n = \ln n$ and $\ln^{(k+1)} n = \ln(\ln^{(k)} n)$. Our main result for basic quicksort is a more precise version of (2) above.

Theorem 1. *Let* $\varepsilon = \varepsilon(n)$ *satisfy* $1/\ln n < \varepsilon \le 1$. *Then as* $n \to \infty$,

$$\Pr\left[\left|\frac{Q_n}{q_n} - 1\right| > \varepsilon\right] = n^{-2\varepsilon(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n))}.$$

This result is quite precise for $\varepsilon > 0$ fixed or tending to 0 very slowly, but says little when $\varepsilon = O(\ln^{(2)} n/\ln n)$.

Theorem 1 is stated above for the most basic form of quicksort. However, the difference between the number of comparisons required by this version of quicksort and the corresponding number for the more efficient "cutting at length M" variant is only $O(n)$. This small term does not affect our results, and the theorem we have just stated (as well as the one that is to follow) also holds for the "cutting at length M" variant.

More care must be taken, however, when considering the "median of $(2t+1)$" variants of quicksorts. Recall that these are the variants in which the partitioning key is chosen as the median of $(2t+1)$ keys. Here t is a fixed non-negative integer, $t = 0$ corresponds to basic quicksort, and $t = 1$ is perhaps most common in practice.

We need more notation before stating our second and final theorem. Let $Q_n^{(t)}$ be the random number of comparisons taken to sort a random list of length n, and let $q_n^{(t)} = E\left[Q_n^{(t)}\right]$. Thus $Q_n^{(0)}$ is Q_n. For $j = 1, 2, \ldots$ let $\kappa_j = (H_{2j+2} - H_{j+1})^{-1}$. Thus for example $\kappa_0 = 2$ and $\kappa_1 = 12/7$. It is well known ([Van70] and [Sed80]) that

$$q_n^{(t)} = \kappa_t\, n \ln n + O(n) \qquad \text{as } n \to \infty.$$

We can now state the "median of $(2t+1)$" extension of our first theorem:

Theorem 2. *Let* $\varepsilon = \varepsilon(n)$ *satisfy* $1/\ln n < \varepsilon \le 1$. *Then as* $n \to \infty$,

$$\Pr\left[\left|\frac{Q_n^{(t)}}{q_n^{(t)}} - 1\right| > \varepsilon\right] = n^{-(t+1)\kappa_t\varepsilon(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n))}.$$

We close this introductory section by mentioning that our results for basic quicksort apply to the analysis of binary search trees, via a well known correspondence which we now describe. An execution of quicksort is commonly associated with a so-called partition tree, namely the binary tree whose root is the initial partitioning key, and whose two subtrees are the trees of the recursive calls. Given a fixed number n of keys, there is a natural one-to-one correspondence between partition trees and binary search trees (binary trees in which each left/right child is lesser/greater than its parent). Thus the distribution of the number of key comparisons of a quicksort execution is the same as the distribution of the number of key comparisons performed in constructing a binary search tree by repeated insertion, when all $n!$ initial orders are equally likely.

Thus, for example, Theorem 1 can be interpreted as a result on binary search trees, by letting Q_n be the number of key comparisons made when a binary search tree on n keys is constructed.

2 Basic quicksort

The purpose of this section is to prove Theorem 1. In §3 we will prove Theorem 2. In this section we present the proofs of the lemmas in detail. As the proofs of the corresponding lemmas in §3 are often similar, the arguments there will be brisker.

We shall work in the simple probability space $(\Omega, \mathcal{F}, \mathrm{Pr})$ defined as follows (for a given positive integer n). The set Ω is the set of all $n!$ permutations π of $\{1, \ldots, n\}$; the σ-field $\mathcal{F}$ of events is the set $\mathcal{P}(\Omega)$ of all subsets of Ω; and the probability measure Pr is the uniform measure, so that $\mathrm{Pr}[\,A\,] = |A|/n!$ for $A \subseteq \Omega$. (When we prove the lemmas on list lengths, lemmas 1, 10, and 11, we shall need briefly to consider a richer probability space.)

2.1 List lengths in the partition tree

A standard technique in the analysis of randomised quicksort is to associate with an execution of quicksort a binary tree whose nodes contain the sublists obtained by the algorithm, the root corresponding to the original unsorted list, and the children of any node being the sublists obtained by the splitting of the parent node. We now describe this correspondence more precisely.

Consider the infinite binary tree, with nodes numbered 1,2,3,... level by level and left to right in the usual manner. (So for instance, the path from the root down the left side consists of nodes 1,2,4,8,...). Each execution of (basic) quicksort yields a labelling of a subtree of this tree, corresponding to the recursive structure of quicksort. The root, node 1, is labelled with the unsorted list of n keys, and its "list length" L_1 is n. A partitioning key is chosen, and after partitioning an (unsorted) list of those keys less than the partitioning key forms the label for the left child (node 2) which then acts like the root of a new tree. Similarly, those keys greater than the partitioning key are sent to the right child of the root.

For each $j = 1, 2, \ldots$ let L_j be the length of the list to be sorted at node j. Thus $L_1 = n$ and exactly n of the L_j are non-zero. Our aim in this section is to show that as we move down the tree the list lengths shorten suitably with high probability. Let M_k^n be the maximum value of the list length L_j over the 2^k nodes j at depth k, that is

$$M_k^n \;=\; \max\{L_{2^k+i} : i = 0, 1, \ldots, 2^k - 1\} \,.$$

The following lemma is essentially contained in Lemma 3.1 of [Dev86]. We give a short proof here for completeness.

Lemma 1. *For any* $0 < \alpha < 1$ *and any integer* $k \geq \ln(1/\alpha)$

$$\Pr[\, M_k^n \geq \alpha n \,] \;\leq\; \alpha \left(\frac{2e\ln(1/\alpha)}{k} \right)^k \,.$$

Proof. The key observation is that we can obtain the exact joint distribution of $(L_1, L_2, \ldots)$ as follows. Let the random variables $X_1, X_2, \ldots$ be independent with each uniformly distributed on the interval (0,1). Define random variables $\tilde{L}_1, \tilde{L}_2, \ldots$ as follows. Let $\tilde{L}_1 = n$ and for $i \geq 1$ let $\tilde{L}_{2i} = \lfloor X_i \tilde{L}_i \rfloor$ and $\tilde{L}_{2i+1} = \lfloor (1 - X_i)\tilde{L}_i \rfloor$. Then it is easily seen that $(L_1, L_2, \ldots)$ and $(\tilde{L}_1, \tilde{L}_2, \ldots)$ have the same joint distribution. Also, let $\tilde{M}_k^n$ be the maximum value of $\tilde{L}_j$ over the nodes j at depth k. Then it follows that M_k^n and $\tilde{M}_k^n$ have the same distribution.

Now define further random variables $Z_1, Z_2, \ldots$ from $X_1, X_2, \ldots$ as follows. Let $Z_1 = 1$ and for $i \geq 1$ let $Z_{2i} = X_i Z_i$ and $Z_{2i+1} = (1 - X_i) Z_i$. Then we have $\tilde{L}_i \leq n Z_i$ for each $i = 1, 2, \ldots$. Let Z_k^* be the maximum value of Z_j over the 2^k nodes j at depth k. Then

$$\tilde{M}_k^n \leq n Z_k^* \,.$$

Now the conclusion of the lemma follows from a series of routine probability inequalities and arguments involving Z_k^*. We have

$$\begin{aligned}
\Pr[\, M_k^n \geq \alpha n \,] \quad &= \Pr\left[\, \tilde{M}_k^n \geq \alpha n \,\right] \\
&\leq \Pr[\, Z_k^* \geq \alpha \,] \\
&\leq \sum_{j=2^k}^{2^{k+1}-1} \Pr[\, Z_j \geq \alpha \,] \\
&= 2^k \Pr\left[\prod_{i=1}^{k} X_i \geq \alpha \right] ,
\end{aligned}$$

since each random variable Z_j at depth k is the product of k independent random variables uniformly distributed on $(0,1)$. Thus

$$\begin{aligned}
\Pr\left[\, M_k^n \geq \alpha n \,\right] &\leq 2^k \Pr\left[\prod_{i=1}^{k} X_i \geq \alpha\right] \\
&\leq 2^k \alpha^{-s} E\left[\prod_{i=1}^{k} X_i^s\right] \quad \text{for any } \ s>0\,, \text{ by the Markov inequality} \\
&= 2^k \alpha^{-s} (E\,[X^s])^k \\
&= 2^k \alpha^{-s} (s+1)^{-k} \\
&= \alpha \left(\frac{2e\ln(1/\alpha)}{k}\right)^k \quad \text{upon taking } \ s+1 = \frac{k}{\ln(1/\alpha)} \\
&\qquad\qquad\qquad\qquad \text{to minimise the bound.}
\end{aligned}$$

2.2 The bounded differences approach

In this section we shall use the idea of the method of “bounded differences” [McD89] to establish some necessary lemmas.

We shall be interested in the comparisons performed by quicksort on each of the levels of the partition tree. (Recall that the root is at level 0.) For $k = 0,1,2,\ldots$ let $H_k = H_k(\pi)$ be the random “history” of the comparisons performed at level k, namely, a record of each comparison, and its outcome, made at level k. Thus the vector $(H_0, H_1, \ldots, H_{k-1})$ records the entire history of the process for the first k levels: we call this the *k-history* $\underline{H}^{(k)}$.

It is convenient to introduce some more notation. Let k be a positive integer and let $\underline{h}$ be a (fixed) possible k-history.

Let $\Omega_{\underline{h}}$ denote the event that we observe this particular history, that is, $\Omega_{\underline{h}} = \{\pi \in \Omega \ : \ H^{(k)}(\pi) = \underline{h}\}$. Note that the list lengths at depth k are determined by $\underline{h}$, that is, they are the same for all $\pi \in \Omega_{\underline{h}}$. We shall be interested in the case when their maximum size M_k^n is not too large.

Given that the event $\Omega_{\underline{h}}$ occurs, we are now concerned with the probability space $(\Omega_{\underline{h}}, \mathcal{P}(\Omega_{\underline{h}}), \Pr_{\underline{h}})$, where $\Pr_{\underline{h}}$ is the uniform probability measure on $\Omega_{\underline{h}}$. Let us use $E_{\underline{h}}$ for expectation in this space.

Also, given that $\Omega_{\underline{h}}$ occurs, our “base point” is $E_{\underline{h}}[Q_n]$. Suppose that we run the process on down to some level $k'-1$, where $k' > k$. The random variable $E_{\underline{h}}[Q_n|\underline{H}^{(k')}]$ depends on the random extension $\underline{H}^{(k')}$ of $\underline{h}$ observed. [Recall that by definition $E_{\underline{h}}[Q_n|\underline{H}^{(k')}](\pi) = E_{\underline{h}}[Q_n|\underline{H}^{(k')} = \underline{h}']$ for $\pi \in \Omega_{\underline{h}}$, where $\underline{h}' = \underline{H}^{(k')}(\pi)$: another notation for this is $E_{\underline{h}}[Q_n|\mathcal{F}]$ where $\mathcal{F}$ is the σ-field generated by $\underline{H}^{(k')}$ restricted to $\Omega_{\underline{h}}$.] Thus $E_{\underline{h}}[Q_n|\underline{H}^{(k')}]$

will deviate from $E_{\underline{h}}[Q_n]$ depending on whether there were good or bad splits after $\underline{h}$. We want to show that these deviations are very unlikely to be large.

If h is any level 0 history, so that h tells us the rank r of the initial splitting element, then

$$E_h[Q_n] - q_n \;=\; n - 1 + q_{r-1} + q_{n-r} - q_n \;.$$

The key property of (basic) quicksort that makes our proofs work is given in the following lemma.

Lemma 2. *Let n be a positive integer and let*

$$A_n \;=\; \{n - 1 + q_{r-1} + q_{n-r} - q_n \;:\; r = 1, 2, \ldots, n\}\;.$$

Then
(a) $|\, x \,| \leq n$ *for all* $x \in A_n$, *and*
(b) $\max(A_n) - \min(A_n) \leq (2 \ln 2) n$.

Proof. This follows from straightforward manipulations of the closed form solution to Equation 1 that we have described earlier. The details of this proof are given in an Appendix.

Lemma 3. *Let n and k be positive integers and let $\underline{h}$ be any possible k-history for Q_n. Then*

$$|\; E_{\underline{h}}[Q_n] - q_n \;| \;\leq\; kn\;.$$

Proof. If $k = 1$ the result is just lemma 2(a). Now let $k \geq 2$, and under a given k-history $\underline{h}$ let the list lengths at level k be $l_1, l_2, \ldots, l_{2^k}$. Then

$$\begin{aligned}
&|\; E_{\underline{h}}[Q_n | H_k = h] \;-\; E_h[Q_n] \;| \\
&\quad = \left| \sum_{j=1}^{2^k} \left(E\left[Q_{l_j} | H_0 = h(j) \right] \;-\; q_{l_j} \right) \right| && \text{for suitable level-0 histories } h(j) \text{ obtained from } h, \\
&\quad \leq \sum_{j=1}^{2^k} l_j && \text{by the case } k = 1 \\
&\quad \leq n\;.
\end{aligned}$$

The lemma now follows by induction on k.

We shall use Lemma 3 when considering levels near the top of the partition tree. For further down the tree we use another inequality, again based on Lemma 2. First we need two preliminary lemmas taken (essentially) from [McD89].

Lemma 4. (see lemma 5.8 in [McD89]) *Let X be a random variable such that $E[X]=0$, and $a \le X \le b$ for some constants a and b. Then for any s,*

$$E[\exp\{\, sX \,\}] \;\le\; \exp\left\{\, s^2(b-a)^2/8 \,\right\}. \qquad \square$$

The second preliminary lemma is a variant of Theorem 6.7 of [McD89]. It is our 'bounded differences' inequality, though 'constrained differences' might be a more appropriate phrase here.

Recall that a *filter* in a probability space $(\Omega, \mathcal{F}, \Pr)$ is an increasing sequence $\mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots$ of σ-fields contained in $\mathcal{F}$. Given a filter, a sequence $X_0, X_1, \ldots$ of integrable random variables is a *martingale* if $E[X_{t+1} \mid \mathcal{F}_t] = X_t$ for each t. For an introduction to the theory of martingales, see for example [Nev75] or [Wil91].

Lemma 5. *Let $(\Phi, \Omega) = \mathcal{F}_0 \subseteq \mathcal{F}_1 \subseteq \ldots \subseteq \mathcal{F}_n$ be a filter, let X be an integrable random variable, and let $X_0, X_1, \ldots, X_n$ be the martingale obtained by setting $X_k = E[X \mid \mathcal{F}_k]$. Suppose that for each $k = 1, \ldots, n$ there is a constant c_k such that for any s we have*

$$E[\exp\{\, s(X_k - X_{k-1}) \,\} \mid \mathcal{F}_{k-1}] \;\le\; \exp\left\{\, c_k^2 s^2 \,\right\}.$$

Then for any $t > 0$

$$\Pr[\,|\, X_n - X_0 \,| \;\ge\; t\,] \quad \le \quad 2\exp\left\{\, -t^2/4\sum_{k=1}^{n} c_k^2 \,\right\}.$$

Proof. We follow the lines of the proof of Theorem 6.7 of [McD89]. Let $Y_k = X_k - X_{k-1}$ and $S_k = \sum_{j=1}^{k} Y_j = X_k - X_0$. For any $s > 0$,

$$\begin{aligned}
\Pr[\,X - E[X] \ge t\,] &= \Pr[\,S_n \ge t\,] \\
&\le e^{-st} E[\exp\{\, sS_n \,\}] \\
&= e^{-st} E[\exp\{\, sS_{n-1} \,\} E[\exp\{\, sY_n \,\} \mid \mathcal{F}_{n-1}]] \\
&\le e^{-st} E[\exp\{\, sS_{n-1} \,\}] \exp\left\{\, c_n^2 s^2 \,\right\} \\
&\le e^{-st} \exp\left\{\, s^2 \sum c_k^2 \,\right\} \qquad \text{by iterating.}
\end{aligned}$$

Now set $s = t/2\sum c_k^2$ to obtain

$$\Pr[\,X - E\,[X] \geq t\,] \;\leq\; \exp\left\{\, -t^2/4\sum c_k^2 \,\right\}\,.$$

To obtain the same bound for $\Pr[\,X - E\,[X] \leq -t\,]$, just replace X by $-X$.

Lemma 6. *Let $k_1 < k_2$ be positive integers, let $\alpha > 0$, and let $\underline{h}$ be a possible k_1-history for which $M_{k_1}^n \leq \alpha n$. Then for any $t > 0$*

$$\Pr_{\underline{h}}\left[\left|\; E_h[Q_n|\underline{H}^{(k_2)}] \;-\; E_{\underline{h}}[Q_n] \;\right| \geq t\right] \leq 2\exp\left\{\, -\frac{t^2}{2(\ln 2)^2(k_2-k_1)\alpha n^2} \,\right\}.$$

Proof. We shall set up things to use Lemma 5. We work in the probability space $(\Omega_{\underline{h}}, \mathcal{P}(\Omega_{\underline{h}}), \Pr_{\underline{h}})$. The filter is defined by letting $\mathcal{F}_k$ be the σ-field generated by the histories $H_{k_1}, \ldots, H_{k_1+k-1}$. Thus the filter corresponds to the increasingly refined partitions of $\Omega_{\underline{h}}$ obtained from all the different possible extensions of the k_1-history $\underline{h}$.

Naturally we take X to be Q_n (restricted to $\Omega_{\underline{h}}$). Thus in the notation of Lemma 5 we have $X_0 = E_{\underline{h}}[Q_n]$, and $X_{k_2-k_1} = E_{\underline{h}}[Q_n|\underline{H}^{(k_2)}]$. We claim that for any $k = 1, 2, \ldots$ and any s we have

$$E_{\underline{h}}\,[\exp\{\; s(X_k - X_{k-1})\;\} \mid \mathcal{F}_{k-1}] \;\leq\; \exp\left\{\; c_k^2 s^2 \;\right\} \tag{$*$}$$

where each $c_k^2 = \dfrac{1}{2}(\ln 2)^2\alpha n^2$. Once this claim is established, the lemma will follow immediately from Lemma 5.

Let us then prove the claim $(*)$. Let k be a non-negative integer, let $j = k_1 + k$, and let $\underline{h}'$ be a possible j-history extending $\underline{h}$. We condition on the event $\underline{H}^{(j)} = \underline{h}'$ and use notation as before. Define the random variable T on $\Omega_{\underline{h}'}$ by setting

$$T = E_{\underline{h}'}[Q_n \mid H_j] \;-\; E_{\underline{h}'}[Q_n]\,.$$

The claim $(*)$ is equivalent to showing that for any s

$$E_{\underline{h}'}[\exp\{\; sT\;\}] \;\leq\; \exp\left\{\; s^2(\ln 2)^2\alpha n^2/2 \;\right\}\,. \tag{$**$}$$

It remains then to prove $(**)$. As we observed earlier, given that $\underline{H}^{(j)} = \underline{h}'$, all the list lengths at depth j are determined, say as $l_1, \ldots, l_{2^j}$. Also we have

$$\sum_i l_i^2 \leq \alpha n^2\;\;,$$

since each $l_i \leq \alpha n$ and $\sum_i l_i \leq n$.

Now T has the distribution of the sum of 2^j independent random variables T_{l_i} for $i = 1, \ldots, 2^j$, where T_{l_i} is uniformly distributed on the set A_{l_i} (defined in Lemma 2). Hence, by Lemmas 2(b) and 4, for any s

$$
\begin{aligned}
E_{\underline{h}'}\left[\exp\{\, sT \,\}\right] &= \prod_i E\left[\exp\{\, sT_{l_i} \,\}\right] \\
&\leq \exp\left\{\, s^2 (2\ln 2)^2 \sum_i l_i^2/8 \,\right\} \\
&\leq \exp\left\{\, s^2 (\ln 2)^2 \alpha n^2/2 \,\right\},
\end{aligned}
$$

as required.

2.3 Upper bound

We now need only assemble the pieces from §2.1 and §2.2 to give a non-asymptotic upper bound for the probability of a large deviation − this is Lemma 7 below − and then choose appropriate values for the parameters to yield the upper bound in Theorem 1.

Lemma 7. *Let n, k_1 and s be positive integers. Then for any real α with $0 < \alpha \leq 1$ and positive integer k_2 such that $\ln(1/\alpha) \leq k_1$, $k_2 > k_1$, $k_2 \geq \ln(n/2)$ we have*

$$
\begin{aligned}
\Pr\left[\,|\, Q_n - q_n \,| \geq\; k_1 n + s\,\right] \;\leq\; & \frac{2}{n}\left(\frac{2e\ln(n/2)}{k_2}\right)^{k_2} + \alpha\left(\frac{2e\ln(1/\alpha)}{k_1}\right)^{k_1} \\
& + 2\exp\left\{\, \frac{-s^2}{2(k_2 - k_1)\alpha n^2} \,\right\}.
\end{aligned}
$$

Proof. Let R_n be the random variable $E\left[Q_n \mid \underline{H}^{(k_2)}\right]$. Recall that a k_1-history $\underline{h}^{(k_1)}$ determines $M^n_{k_1}$. Let $\mathcal{H}$ be the set of k_1-histories $\underline{h}^{(k_1)}$ for which $M^n_{k_1} \leq \alpha n$. Then

$$
\begin{aligned}
&\Pr\left[\,|\, Q_n - q_n \,| \;\geq\; k_1 n + s\,\right] \\
&\leq \Pr\left[\, Q_n \neq R_n \,\right] \;+\; \Pr\left[\underline{H}^{(k_1)} \notin \mathcal{H}\right] \\
&\quad + \Pr\left[\,|\, R_n - q_n \,| \geq k_1 n + s \quad \text{and} \quad \underline{H}^{(k_1)} \in \mathcal{H}\right] \\
&= \Pr\left[\, Q_n \neq R_n \,\right] \;+\; \Pr\left[\underline{H}^{(k_1)} \notin \mathcal{H}\right] \\
&\quad + \sum_{\underline{h} \in \mathcal{H}} \left(\Pr_{\underline{h}}\left[\,|\, R_n - q_n \,| \geq k_1 n + s\,\right] \;\times \Pr\left[\underline{H}^{(k_1)} = \underline{h}\right]\right)
\end{aligned}
$$

$$
\begin{aligned}
&\leq \Pr\left[\, M^n_{k_2} \geq 2 \,\right] \;+\; \Pr\left[\, M^n_{k_1} > \alpha n \,\right] \\
&\quad + \sum_{\underline{h} \in \mathcal{H}} \Pr_{\underline{h}}\left[\, (\mid R_n - E_{\underline{h}}[R_n] \mid \geq s) \,\right] \;\times\; \Pr\left[\underline{H}^{(k_1)} = \underline{h} \right]
\end{aligned}
$$

since $\mid E_{\underline{h}}[R_n] - q_n \mid \;\leq\; k_1 n$ by Lemma 3. The result now follows from Lemmas 1 and 6.

Lemma 8. *Let* $\varepsilon = \varepsilon(n)$ *satisfy* $0 < \varepsilon \leq 1$. *Then as* $n \to \infty$,

$$
\Pr\left[\left| \frac{Q_n}{q_n} - 1 \right| > \varepsilon \right] \;\leq\; \exp\left\{ \, -2\varepsilon \ln n (\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n)) \, \right\} \;.
$$

Proof. It suffices to assume that $\varepsilon(n) \;\geq\; \varepsilon_0(n) \;=\; 2\ln^{(2)} n / \ln n$, for if $\varepsilon < \varepsilon_0$ then the upper bound can take the value 1.

We need to choose our parameters appropriately. Let $s = s(n)$ and $k_1 = k_1(n)$ be integers with

$$
s \;=\; \left\lceil \frac{\varepsilon n \ln n}{\ln^{(2)} n} \right\rceil \;,
$$

and

$$
k_1 \;=\; \lfloor 2\varepsilon \ln n - 2s/n \rfloor \;=\; 2\varepsilon \ln n \left(1 + O(\frac{1}{\ln^{(2)}}) \right) \;.
$$

Observe that

$$
\begin{aligned}
k_1 n + s &= \lfloor 2\varepsilon n \ln n - s \rfloor \\
&\leq \varepsilon q_n \qquad \text{for } n \text{ sufficiently large} \;;
\end{aligned}
$$

and then

$$
\Pr\left[\left| \frac{Q_n}{q_n} - 1 \right| > \varepsilon \right] \;\leq\; \Pr\left[\, \mid Q_n - q_n \mid \geq k_1 n + s \,\right] \;.
$$

Next let

$$
\alpha \;=\; \alpha(n) \;=\; \varepsilon^2 (\ln^{(2)} n)^{-5}
$$

and let

$$
k_2 \;=\; k_2(n) \;=\; \left\lceil (\ln n)(\ln^{(2)} n) \right\rceil .
$$

Note that $k_1 \geq \ln(1/\alpha)$ for n sufficiently large, since

$$
\frac{\ln(1/\alpha)}{k_1} \;\sim\; \frac{2\ln(1/\varepsilon) + 10\ln^{(3)} n}{2\varepsilon \ln n} \;=\; \frac{(1/\varepsilon)\ln(1/\varepsilon)}{\ln n} + o(1) \;\leq\; \frac{1}{2} + o(1) \;.
$$

It remains only to check that each of the three terms in the right hand side of the inequality in Lemma 7 is suitably small. The first term is

$$\frac{2}{n}\left(\frac{2e\ln(n/2)}{k_2}\right)^{k_2} \le \exp\left\{ -k_2 \ln^{(3)} n \right\} ,$$

which is very small. The second term is

$$\begin{aligned}\alpha\left(\frac{2e\ln(1/\alpha)}{k_1}\right)^{k_1} &\le \exp\left\{ -k_1(\ln k_1 - \ln^{(2)}(1/\alpha) + O(1)) \right\} \\ &= \exp\left\{ -2\varepsilon \ln n \left(\ln^{(2)} n - \ln(\frac{1}{\varepsilon}) - \ln^{(2)}(\frac{2}{\varepsilon}) - \ln^{(4)} n + O(1)\right) \right\} , \\ &= \exp\left\{ -2\varepsilon \ln n \left(\ln^{(2)} n - \ln(\frac{1}{\varepsilon}) + O(\ln^{(3)} n)\right) \right\} ,\end{aligned}$$

which is as required. The third and final term is

$$\begin{aligned}&2\exp\left\{ \frac{-2s^2}{(k_2-k_1)(2\ln 2)^2\alpha n^2} \right\} \\ &\quad\le\ 2\exp\left\{ -\frac{(1+o(1))}{2(\ln 2)^2}\frac{n^2(\ln n)^2/(\ln^{(2)} n)^2}{(\ln n)(\ln^{(2)} n)(\ln^{(2)} n)^{-5}n^2} \right\} \\ &\quad\le\ 2\exp\left\{ -\frac{(1+o(1))}{2(\ln 2)^2}\ln n(\ln^{(2)} n)^2 \right\} ,\end{aligned}$$

which is very small.

2.4 The lower bound

In this section we shall prove the lower bound part of Theorem 1. We shall see that a few "bad splits" near the top of the partition tree can account for the probability of large deviations.

Lemma 9. *Let $\varepsilon = \varepsilon(n)$ satisfy $1/\ln n \le \varepsilon \le 1$. Then as $n \to \infty$,*

$$\Pr\left[\frac{Q_n}{q_n} - 1 > \varepsilon\right] \ge \exp\left\{ -2\varepsilon \ln n \left(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n)\right) \right\} .$$

Proof. The proof for the median of $(2t+1)$ quicksort will be very similar to the proof below, so it is convenient to introduce $\kappa_0 = 2$ into our discussion. As in the proof of Lemma 8, we need to choose the appropriate parameters carefully.

To begin with, let us assume that

$$\varepsilon(n)\ \ge\ 9\ln^{(2)} n/\ln n\ . \qquad (*)$$

Let

$$\mu \;=\; \mu(n) \;=\; 3\ln^{(3)} n/\ln^{(2)} n$$

and let

$$\kappa \;=\; \kappa_0 + \mu,\;\; \lambda \;=\; \mu/3 \;\text{and}\;\; \delta \;=\; \mu/4\,.$$

Note that

$$\kappa(1-\kappa\lambda/2) \;\geq\; \kappa_0+\delta \quad \text{for } \; n \; \text{ sufficiently large.}$$

Now let

$$\begin{aligned} k \;&=\; k(n) = \lfloor \kappa\varepsilon\ln n\rfloor\,, \\ l \;&=\; l(n) = \lfloor (\lambda n)/(\varepsilon\ln n)\rfloor\,, \\ J \;&=\; J(n) = \{2^j+1 : \, j=0,1,\ldots,k-1\}\cup\{2^k\}\,, \end{aligned}$$

and let $\mathcal{L}=\mathcal{L}(n)$ be the set of vectors $(l_j : j\in J)$ of non-negative integers l_j such that $l_j \leq l$ for each $j\in J\backslash\{2^k\}$.

For each $\underline{l}$ in $\mathcal{L}$ let $A(\underline{l})$ be the event that $L_j = l_j$ for each j in J. Finally let A be the union of the events $A(\underline{l})$ for $\underline{l}$ in $\mathcal{L}$. Note that

$$\begin{aligned} \Pr[\,A\,] &\geq \left(\frac{l+1}{n}\right)^k \\ &\geq \left(\frac{\lambda}{\varepsilon\ln n}\right)^k \\ &= \exp\left\{\; -(\kappa_0+\mu)\varepsilon\ln n\;\left(\ln^{(2)} n \;-\ln(\frac{1}{\varepsilon})+\ln(\frac{1}{\lambda})\;\right)\;\right\} \\ &= \exp\left\{\; -\kappa_0\varepsilon\ln n\left(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n)\right)\;\right\}\;. \end{aligned}$$

Let Q'_n be the number of comparisons corresponding to partitioning the leftmost nodes at depth at most $k-1$ (these are the parents of the nodes in the set J) and let $Q''_n = Q_n - Q'_n$. If the event A occurs then

$$\begin{aligned} Q'_n &\geq \sum_{j=0}^{k-1}(n-j(l+1)-1) \\ &= k(n-1-\frac{1}{2}(k-1)(l+1)) \\ &= \kappa\varepsilon(1-\kappa\lambda/2)n\ln n \;+\; O(n) \\ &\geq (\kappa_0+\delta)\varepsilon n\ln n \;+\; O(n)\;. \end{aligned}$$

Now let $\underline{l}\in\mathcal{L}$ be such that $\Pr[\,A(\underline{l})\,]>0$. Observe that $\sum_{j\in J} l_j = n-k$ at least once $k(l+1) < n$. Conditional on $A(\underline{l})$, Q''_n is distributed like

$\sum_{j\in J} Q_{l_j}$, where these $k+1$ random variables Q_{l_j} are independent. Hence

$$\begin{aligned}
E[Q''_n \mid A(\underline{l})] &= \sum_{j\in J} E[Q_{l_j}] \\
&\geq \sum_{j\in J} (\kappa_0\, l_j \ln l_j \;-\; 4 l_j) \\
&\geq \kappa_0\, l_{2^k}\ \ln l_{2^k} \;+\; \kappa_0 k\, \hat{l} \ln \hat{l} \;-\; 4n
\end{aligned}$$

where $\hat{l} = \frac{1}{k}\sum_{j\in J\setminus\{2^k\}} l_j$. But $l_{2^k} \;=\; n - k(\hat{l}+1)$ and so

$$\begin{aligned}
l_{2^k} \ln l_{2^k} &\geq (n - k(\hat{l}+1))(\ln n + \ln(1 - \frac{k(l+1)}{n})) \\
&= n \ln n - k\hat{l} \ln n + O((\ln n)^2)\ .
\end{aligned}$$

Hence for n sufficiently large

$$\begin{aligned}
E[Q''_n \mid A(\underline{l})] &\geq \kappa_0 n \ln n - \kappa_0 k\hat{l} \ln(n/\hat{l}) - 5n \\
&\geq \kappa_0 n \ln n - \kappa_0 k l \ln(n/l) - 5n \\
&\geq \kappa_0 n \ln n - \kappa_0 \kappa \lambda n \ln\left(\frac{\varepsilon \ln n}{\lambda}\right) - 6n \\
&\geq \kappa_0 n \ln n - (4 + o(1)) n \ln^{(3)} n\ .
\end{aligned}$$

Also

$$\text{var}\ \ (Q''_n \mid A(\underline{l})) \;=\; \sum_{j\in J} \text{var}\ \ Q_{l_j} \;=\; O(\sum_{j\in J} l_j^2) \;=\; O(n^2)\ .$$

Here we are using the fact that $\text{var}(Q_n) = O(n^2)$, as noted in §1. Hence, by Chebyshev's inequality

$$\Pr\left[Q''_n \geq\ \kappa_0 n \ln n - 5n \ln^{(3)} n \mid A(\underline{l}) \right] \to 1\ \text{ as }\ \ n \to \infty\ ,$$

and this convergence is uniform over $\underline{l}$ in $\mathcal{L}$.

Now, for n sufficiently large,

$$(\kappa_0 + \delta)\varepsilon n \ln n + \kappa_0 n \ln n - 6n \ln^{(3)} n \;>\; (1+\varepsilon) q_n\ .$$

Hence we have

$$\begin{aligned}
\Pr[\, Q_n > (1+\varepsilon) q_n \,] \quad &\geq \Pr[\, Q_n > (1+\varepsilon) q_n \mid A \,]\ \Pr[\, A \,] \\
&= \sum_{\underline{l}\in\mathcal{L}} \Pr[\, Q_n > (1+\varepsilon) q_n \mid A(\underline{l}) \,]\ \Pr[\, A(\underline{l}) \,] \\
&\geq \sum_{\underline{l}\in\mathcal{L}} \Pr\left[Q''_n \geq \kappa_0 n \ln n - 5n \ln^{(3)} n \mid A(\underline{l}) \right] \Pr[\, A(\underline{l}) \,] \\
&= (1 + o(1))\ \ \Pr[\, A \,] \\
&\geq \exp\left\{\ -\kappa_0 \varepsilon \ln n \left(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n) \right)\ \right\}
\end{aligned}$$

as required. This completes the proof of Lemma 9 for the case $(*)$.

Now suppose that

$$\frac{1}{\ln n} \;\leq\; \varepsilon(n) \;\leq\; \frac{9\ln^{(2)} n}{\ln n}\,. \qquad (**)$$

Thus we can write $\varepsilon(n) = f(n)/\ln(n)$, where $1 \leq f(n) \leq 9\ln^{(2)} n$, and so

$$\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n) = \ln^{(2)} n - \ln^{(2)} n + \ln(f(n)) + O(\ln^{(3)} n)) = O(\ln^{(3)} n).$$

Hence to prove the lemma for $(**)$, it suffices to prove that

$$\Pr\left[\frac{Q_n}{q_n} - 1 > \varepsilon\right] \;\geq\; n^{-c\varepsilon\ln^{(3)} n}$$

for some constant c. We shall in fact show that for n sufficiently large,

$$\Pr\left[\frac{Q_n}{q_n} - 1 > \varepsilon\right] \;\geq\; n^{-(2c+2)\varepsilon\ln^{(3)} n}$$

where $c = 12$. We choose parameters as before, except that now $k = k(n) = \lfloor c\varepsilon\ln n\rfloor$. Arguing much as before, we find that for n sufficiently large,

$$\begin{aligned}
\Pr\left[\,A\,\right] &\geq n^{-(2c+1)\varepsilon\ln^{(3)} n}\;; \\
Q'_n &\geq (c-1)\varepsilon n\ln n\;; \quad \text{and} \\
E\left[Q''_n \mid A(\underline{l})\right] &\geq \kappa_0 n\ln n \;-\; \kappa_0 c\lambda n\ln\left(\frac{\varepsilon\ln n}{\lambda}\right) \;-\; 6n\;. \\
&\geq \kappa_0 n\ln n - 7n\;.
\end{aligned}$$

But $\mathrm{var}(Q''_n \mid A(\underline{l})) \leq n^2$ (for n sufficiently large), since $\mathrm{var}(Q_n) \sim (7 - 2\pi^2/3)n^2$ as $n \to \infty$, and $7 - 2\pi^2/3 \simeq 0.42 < 1$ (see [Knu73]). Hence by Chebyshev's inequality, with probability $\geq 1/2$, $Q''_n \;\geq\; \kappa_0 n\ln n \,-9n$. Further (for n sufficiently large)

$$\kappa_0 n\ln n \;+\; (c-1)\varepsilon n\ln n \;-\; 9n \;\geq\; (1+\varepsilon)q_n\;.$$

Now we complete the proof much as in the case of $(*)$.

3 Median-of-$(2t+1)$ quicksort

In this section we shall prove Theorem 2. Many of the ideas and techniques of this section are similar to those explained in the previous section. Let t be a fixed non-negative integer throughout.

3.1 List lengths in the partition tree

We argue much as in §2.1 though the details are more complicated here.

Lemma 10. *Let n and k be positive integers, let $0 < \alpha < 1$ and suppose that $(2t+1)2tk < \alpha n$.*

Let $X_1, X_2, \ldots, X_k$ be independent random variables each distributed as the median of $2t+1$ independent random variables uniformly distributed on [0,1]. Then

$$\Pr[\, M_k^n \geq \alpha n \,] \leq 2^k \left(1 - \frac{(2t+1)2tk}{\alpha n}\right)^{-1} \Pr\left[\prod_{i=1}^{k} X_i \geq \alpha\right] .$$

Proof. Let $U_j^{(i)}$ for positive integers i and j be independent random variables each uniformly distributed on [0,1]. For each $i = 1, \ldots, k$ we may take X_i to be the median of $U_1^{(i)}, \ldots, U_{2t+1}^{(i)}$.

Next we define a decreasing sequence N_0, N_1, …, N_k of random variables corresponding to the list lengths L_0, L_2, …, L_{2^k}. Let $N_0 = n$. For each $i = 1, \ldots, k$ do the following. If $N_{i-1} < \alpha n$ then set $N_i = \ldots = N_k = 0$ and stop. If $N_{i-1} \geq \alpha n$ then consider $U_1^{(i)}, U_2^{(i)}, \ldots$ in turn until we obtain $2t+1$ distinct numbers $\lfloor U_j^{(i)} N_{i-1} \rfloor$. Then let N_i be one less than the median of these $2t+1$ numbers. The key observation is that

$$\Pr[\, L_{2^k} \geq \alpha n \,] \;=\; \Pr[\, N_k \geq \alpha n \,] \;.$$

Let A be the event that for each $i = 1, \ldots, k$ with $N_{i-1} \geq \alpha n$ the first $2t+1$ numbers $\lfloor U_1^{(i)} N_{i-1} \rfloor, \ldots, \lfloor U_{2t+1}^{(i)} N_{i-1} \rfloor$ are distinct. Then

$$\Pr[\, N_k \geq \alpha n \text{ and } A \,] \;\leq\; \Pr\left[\prod_{i=1}^{k} X_i \geq \alpha\right]$$

since clearly $N_k \leq n \prod_{i=1}^{k} X_i$ on A. Also

$$\begin{aligned}\Pr[\, A \mid N_k \geq \alpha n \,] &\geq \left(1 - \binom{2t+1}{2} \frac{2}{\alpha n}\right)^k \\ &\geq 1 - \frac{(2t+1)(2t)k}{\alpha n} \,.\end{aligned}$$

Now the desired conclusion follows from routine probability inequalities. In particular,

$$\begin{aligned}\Pr[\, M_k^n \geq \alpha n \,] &\leq 2^k \ \Pr[\, L_{2^k} \geq \alpha n \,] \\ &= 2^k \ \Pr[\, N_k \geq \alpha n \,]\end{aligned}$$

$$= 2^k \ \Pr[\, N_k \geq \alpha n \quad \text{and} \quad A\,] \Big/ \Pr[\, A \mid N_k \geq \alpha n\,]$$
$$\leq 2^k \ \Pr\left[\prod_{i=1}^{k} X_i \geq \alpha\right] \Big/ \left(1 - \frac{(2t+1)2tk}{\alpha n}\right) .$$

Lemma 11. *Let* $0 < \alpha < 1$ *and let* n *and* k *be positive integers such that*

$$k \ > \ \frac{2t+1}{t+1} \ \ln \frac{1}{\alpha} \, .$$

Let X *be distributed like the median of* $2t+1$ *random variables each uniform on [0,1], and let* $X_1, \ldots, X_k$ *be independent random variables, each distributed like* X*. Then*

$$\Pr\left[\prod_{i=1}^{k} X_i \geq \alpha\right] \ \leq \ \alpha^{2t+1} \left(\frac{(2t+1)e \ln(1/\alpha)}{(t+1)k}\right)^{(t+1)k} .$$

Proof. It is well known and routine to check that X has the β distribution with parameters $t+1$, $t+1$; which has probability density function

$$f(x) \ = \ \frac{\Gamma(2t+2)}{(\Gamma(t+1))^2} x^t (1-x)^t \qquad \text{for} \quad 0 < x < 1 \, .$$

Thus, for $s > -(t+1)$

$$\begin{aligned}
E\,[X^s] &= \int_0^1 x^s f(x)\, dx \\
&= \frac{\Gamma(2t+2)}{(\Gamma(t+1))^2} \ \int_0^1 x^{s+t}(1-x)^t \, dt \\
&= \frac{\Gamma(2t+2)}{\Gamma(t+1)} \ \frac{\Gamma(s+t+1)}{\Gamma(s+2t+2)} \\
&= \prod_{i=0}^{t} \left(\frac{2t+1-i}{s+2t+1-i}\right) .
\end{aligned}$$

Now we argue as in the proof of Lemma 1. For any $s > 0$

$$\begin{aligned}
\Pr\left[\prod_{i=1}^{k} X_i \geq \alpha\right] &= \Pr\left[\prod_{i=1}^{k} X_i^s \geq \alpha^s\right] \\
&\leq \alpha^{-s} (E\,[X^s])^k \\
&= \alpha^{-s} \left[\prod_{i=0}^{t} \left(\frac{2t+1-i}{s+2t+1-i}\right)\right]^k \\
&\leq \alpha^{-s} \left(\frac{2t+1}{s+2t+1}\right)^{(t+1)k} .
\end{aligned}$$

Now we choose $s > 0$ to minimise this bound b say. Note that

$$\begin{aligned}\frac{d}{ds}\ln(b) &= \frac{d}{ds}\left(s\ln\frac{1}{\alpha} - (t+1)k\ \ln(s+2t+1)\right) \\ &= \ln\frac{1}{\alpha} - \frac{(t+1)k}{s+2t+1}\,,\end{aligned}$$

and so we take

$$s \;=\; \frac{(t+1)k}{\ln(1/\alpha)} - 2t - 1 \;>\; 0\,.$$

Then

$$\alpha^{-s} \;=\; \alpha^{2t+1}\, e^{(t+1)k}$$

and

$$\frac{2t+1}{s+2t+1} \;=\; \frac{(2t+1)\ln(1/\alpha)}{(t+1)k}$$

and so we obtain the required bound.

3.2 The bounded differences approach

Our first result here is the key property for median-of-$(2t+1)$ quicksort and corresponds to lemma 2 for basic quicksort. That result was non-asymptotic and proved from first principles: here we are not so lucky and must use asymptotic results for $q_n^{(t)}$.

If h is any level 0 history, so that h tells us the rank r of the splitting element, then

$$E_{\underline{h}}[Q_n^{(t)}] - q_n^{(t)} \;=\; n - 1 + q_{r-1}^{(t)} + q_{n-r}^{(t)} - q_n^{(t)}\,.$$

Lemma 12. *For each positive integer n let*

$$A_n \;=\; \{n - 1 \;+\; q_{r-1}^{(t)} \;+\; q_{n-r}^{(t)} \;-\; q_n^{(t)} \;:\; r = 1, 2, \ldots, n\}\,.$$

Then there is an $\eta > 0$ and a function $g(n) > 0$ with $g(n) = O(n^{1-\eta})$ such that for each positive integer n

$$-(\kappa_t \ln 2\; - 1)\, n \;-\; g(n) \;<\; x \;<\; n \;+\; g(n) \qquad \text{for all } x \in A_n\,.$$

Proof. It is known (see [Hen89]) that for some constants $\beta = \beta(t)$ and $\eta = \eta(t)$ with $0 < \eta < 1$ we have

$$q_n^{(t)} \;=\; \kappa_t n \ln n \;+\; \beta n \;+\; \gamma(n)\,,$$

where $\gamma(n) = O(n^{1-\eta})$. Suppose that $\mid \gamma(n) \mid \leq cn^{1-\gamma}$. Let

$$f(n) \;=\; \kappa_t n \ln n \;+\; \beta n\,.$$

It is easy to check that

$$f(r-1)+f(n-r)-f(n)\quad\begin{cases}\le\ O(1)\\[2ex] \ge\ -\kappa_t(n\,\ln 2\,+\,\ln n)\,+\,O(1)\,.\end{cases}$$

It follows that

$$q^{(t)}_{r-1}\;+\;q^{(t)}_{n-r}\;-\;q^{(t)}_{n}\quad\begin{cases}\le 3\max\limits_{1\le k\le n}\mid\gamma(k)\mid\;+\;O(1)\\[2ex] \ge -\kappa_t(n\,\ln 2\;+\;\ln n)\\[2ex] \qquad -3\max\limits_{1\le k\le n}\mid\gamma(k)\mid\;+\;O(1)\,.\end{cases}$$

Thus there is a suitable function $g(n)$ with

$$g(n)\;=\;3cn^{1-\gamma}\,+\,2\,\ln n\,+\,O(1)\,.\qquad\square$$

Consider the variant of median of (2t+1) quicksort which cuts lists at length $\ln(n)$ say. Let $\tilde{Q}_n$ be the corresponding number of comparisons and let $\tilde{q}_n=E\left[\tilde{Q}_n\right]$. Note that

$$\left|\;Q_n^{(t)}-\tilde{Q}_n\;\right|\;=\;O(n\ln^{(2)}n)\,.$$

Lemma 13. *There exist $c,\eta>0$ such that the following holds. Let n and k be positive integers and let $\underline{h}$ be a possible k-history for $\tilde{Q}_n$. Then*

$$\left|\;E_{\underline{h}}[\tilde{Q}_n]\;-\;\tilde{q}_n\;\right|\;\le\;\left(1+\;c(\ln n)^{-\eta}\right)\;kn\,.$$

Proof. This follows from Lemma 12 much as Lemma 3 followed from Lemma 2(a).

Choose $c,\eta>0$ such that $g(n)\le cn^{1-\eta}$ in Lemma 12. Note that since $\kappa_t\ln 2\le 2\ln 2\le 2$, Lemma 12 shows that $|x|\le n+g(n)$ for all $x\in A_n$. We now see that the case $k=1$ follows immediately from Lemma 12.

Now let $k\ge 1$, let $\underline{h}$ be a possible k-history with corresponding list lengths $l_1,\ldots,l_{2^k}$ at level k, and let h be a possible level k history extending $\underline{h}$. Then

$$\begin{aligned}&\left|\;E_{\underline{h}}\left[\tilde{Q}_n\mid H_k=h\right]\;-\;E_{\underline{h}}\left[\tilde{Q}_n\right]\;\right|\\ &=\left|\;\sum_{i=1}^{2^k}\left(E\left[\tilde{Q}_{l_i}\mid H_0=h(i)\right]\;-\;\tilde{q}_{l_i}\right)\;\right|\qquad\text{for suitable depth-0 histories } h(i) \text{ obtained from } h\end{aligned}$$

$$\le \sum_i \left\{ l_i + g(l_i) \ : \ i = 1, \ldots, 2^k \text{ with } l_i \ge \ln n \right\}$$
$$\le n + c \sum_i \left\{ l_i^{1-\eta} \ : \ i = 1, \ldots, 2^k \text{ with } l_i \ge \ln n \right\}$$
$$\le n + cn(\ln n)^{-\eta} \ .$$

In the last step we used the fact that if $a > 0$, $x_1, x_2, \ldots \ge a$ and $\sum_i x_i \le n$ then $\sum_i x_i^{1-\eta} \le na^{-\eta}$.

The lemma now follows by induction on k.

3.3 An upper bound

In this section we shall prove the upper bound part of Theorem 2, following the pattern of the proofs for the case of basic quicksort.

Lemma 14. *Let $\varepsilon = \varepsilon(n)$ satisfy $0 < \varepsilon \le 1$. Then*

$$\Pr\left[\left| \frac{Q_n^{(t)}}{q_n^{(t)}} - 1 \right| > \varepsilon \right] \le \exp\left\{ -(t+1)\kappa_t \varepsilon \ln n \left(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n) \right) \right\}.$$

Proof. Consider $\tilde{Q}_n$ as in Lemma 13 above. Define s, k_1, α, k_2, R_n, $\mathcal{H}$ as in the proof of Lemma 8, except with the term $2\varepsilon \ln n$ in the definition of k_1 replaced by $\kappa_t\ \varepsilon\ \ln n$, and referring to $\tilde{Q}_n$. (Note that $\kappa_0 = 2$.) Now the proof of this lemma is similar to the proofs of Lemmas 7 and 8. As there, it suffices to assume that $\varepsilon(n) \ge \varepsilon_0(n)$. In particular, for n sufficiently large,

$$\begin{aligned}
\Pr&\left[\left| Q_n^{(t)} - q_n^{(t)} \right| > \varepsilon\, q_n^{(t)} \right] \\
&\le \Pr\left[\left| \tilde{Q}_n - \tilde{q}_n \right| > k_1\, n + s \right] \\
&\le \Pr\left[\tilde{Q}_n \ne R_n \right] + \Pr\left[\underline{H}^{(k_1)} \notin \mathcal{H} \right] \\
&\quad + \Pr\left[\left| \tilde{Q}_n - \tilde{q}_n \right| > k_1 n + s \text{ and } \underline{H}^{(k_1)} \in \mathcal{H} \right] \\
&\le \Pr\left[M_{k_2}^n > \ln n \right] + \Pr\left[M_{k_1}^n > \alpha\, n \right] \\
&\quad + \sum_{\underline{h} \in \mathcal{H}} \left(\Pr_{\underline{h}}\left[\, | R_n - \tilde{q}_n | > k_1 n + s \,\right] \times \Pr\left[\underline{H}^{(k_1)} = \underline{h} \right] \right) .
\end{aligned}$$

Now consider the three terms in this bound. First we shall show that $\Pr\left[M_{k_2}^n > \ln n \right]$ is very small. Let $k_3 = \lfloor k_2/2 \rfloor$. Then by Lemmas 10 and 11

$$\begin{aligned}
\Pr\left[M_{k_2}^n > \ln n \right] &\le \Pr\left[M_{k_2}^n > 1 \right] \\
&\le \Pr\left[M_{k_3}^n > k_3 \right] \\
&\le \exp\left\{ -(t+1)\ k_3 \left(\ln^{(3)} n + O(1) \right) \right\} ,
\end{aligned}$$

which is very small.

Next consider $\Pr\left[M^n_{k_1} \ > \ \alpha\, n \right]$. By Lemmas 10 and 11 this is at most

$$\exp\left\{ \ -(t+1)\ k_1 \ \left(\ln k_1 \ - \ \ln^{(2)}(1/\alpha) \ + \ O(1)\right) \right\}$$
$$= \exp\left\{ \ -(t+1)\kappa_t \varepsilon \ln n \left(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n)\right) \right\} \ ,$$

which is as required.

Finally, for the third term, consider a k_1-history $\underline{h}$ in $\mathcal{H}$. Note that by Lemma 13

$$|\ E_{\underline{h}}\,[R_n] \ - \ \tilde{q}_n \ | \ \le \ (1+(\ln n)^{-\eta}) \ \ k_1 n \ .$$

Hence

$$\Pr_{\underline{h}}\,[\,|\ R_n - \tilde{q}_n \ | \ > \ k_1 n + s\,]$$
$$\le \Pr_{\underline{h}}\,[\,|\ R_n - E_{\underline{h}}\,[R_n]\ | \ > s - k_1 n (\ln n)^{-\eta}\,] \ .$$

But $s - k_1 n(\ln n)^{-\eta} = (1+o(1))s$. Now we use Lemma 5 as in the proof of Lemma 6. We find

$$\Pr_{\underline{h}}\,[\,|\ R_n - \tilde{q}_n \ | \ > \ k_1 n + s\,] \le 2\exp\left\{ -\frac{(2+o(1))s^2}{(k_2-k_1)(\kappa_t \ln 2)^2 \alpha n^2} \right\}$$
$$\le \exp\left\{ -\frac{(2+o(1))}{(\kappa_t \ln 2)^2} (\ln n) \left(\ln^{(2)} n\right)^2 \right\}$$

which is very small.

3.4 Lower bound – general case

In this section we shall prove the lower bound part of the proof of Theorem 2. We have set up the proof of Lemma 9 above so that we can follow it closely.

Lemma 15. *Let $\varepsilon = \varepsilon(n)$ satisfy $1/\ln n < \varepsilon \le 1$. Then, as $n \to \infty$,*

$$\Pr\left[\left(Q^{(t)}_n > (1+\varepsilon) q^{(t)}_n\right) \right] \ge \exp\left\{ \ -(t+1)\kappa_t \varepsilon \ln n \left(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)})\right) \right\}.$$

Proof. In the proof of Lemma 9 replace all references to Q, q, κ_0 by $Q^{(t)}$, $q^{(t)}$, κ_t respectively. Note that $\kappa_t \le \kappa_0 = 2$. Also, $\mathrm{var}(Q^{(t)}_n)/n^2 \to \sigma^2_t \le \sigma^2_0$ as $n \to \infty$ (see [Hen89]). The only parts that need to be changed concern $\Pr\,[\,A\,]$.

Note that $L_3 \le l$ if and only if at least $t+1$ of the chosen $2t+1$ keys are amongst the last $l+1$ of the n possible keys. So certainly $L_3 \le l$ if the first $t+1$ keys chosen are amongst the last $l+1$ of the n possible keys. Thus

$$\Pr[\,L_3 \le l\,] \;\ge\; \left(\frac{l-t}{n}\right)^{t+1} .$$

Considering the other members j of $J - \{2^k\}$ similarly, we see that

$$\begin{aligned} \Pr[\,A\,] &\ge \left(\frac{l-t}{n}\right)^{(t+1)k} \\ &= \exp\Big\{ -(t+1)\kappa_t \varepsilon \ln n \left(\ln^{(2)} n - \ln(1/\varepsilon) + O(\ln^{(3)} n)\right) \Big\} . \end{aligned}$$

4 Appendix 1: A lemma

In this section we prove in detail the following lemma, which describes how the expected number of comparisons in basic quicksort can change, given the knowledge of the first partition. Recall that $q_n = E[Q_n]$. For positive integers n and for $1 \le k \le n$ define

$$\gamma_{n,k} \;=\; q_n - (q_{k-1} + q_{n-k}) ,$$

and for positive integers n define

$$A_n \;=\; \{n-1 \;-\; \gamma_{n,k} \;:\; 1 \le k \le n\} .$$

Lemma 16. *(a) For any integers n, k with $1 \le k \le n/2$,*

$$q_{k-1} + q_{n-k} \;\ge\; q_k + q_{n-k-1} .$$

(b) For any integers n, k with $1 \le k \le n$,

$$q_{n-1} \;\ge\; q_{k-1} + q_{n-k} .$$

(c) For any integer $n \ge 1$,

$$q_{n-1} \;-\; (q_{\lfloor (n-1)/2 \rfloor} + q_{\lceil (n-1)/2 \rceil}) \;\le\; (2\ln 2) n .$$

(d) For any integer $n \ge 1$,

$$\max(A_n) - \min(A_n) \;\le\; (2\ln 2) n .$$

(e) For any integer $n \ge 1$,

$$-(n-1) \;\le\; \min(A_n) \;\le\; \max(A_n) \;\le\; n-1 .$$

Proof. The proof follows from routine arithmetic manipulations of Equation 1. A closed form solution to this equation (see [Knu73] or [Sed80]) is given by $q_0 = 0$, and for $n \geq 1$,

$$q_n \;=\; 2(n+1)\sum_{j=2}^{n} \frac{j-1}{j(j+1)} \,.$$

By rewriting the summand in the above equation as $\dfrac{2j}{j+1} - \dfrac{1}{j}$, it follows that for all $n \geq 0$,

$$q_n \;=\; 2(n+1)H_n \;-4n \,.$$

It is convenient to note here that if n is odd ($n \geq 1$) then

$$H_{n-1} - H_{(n-1)/2} \;\leq\; \int_{(n-1)/2}^{n-1} \frac{dx}{x} \;=\; \ln 2$$

and if n is even ($n \geq 2$) then

$$H_n - H_{n/2} \;\leq\; \int_{n/2}^{n} \frac{dx}{x} \;=\; \ln 2 \;.$$

To prove part (a) of the lemma, we wish to show that $q_{n-k} - q_{n-(k-1)} \geq q_k - q_{k-1}$. It suffices to show that $q_t - q_{t-1}$ is a non-decreasing function of t. But it follows from (*) that $q_t - q_{t-1} = 2(H_t - 1)$, so the result follows.

Now prove part (b). For $1 \leq k \leq n$,

$$q_{k-1} + q_{n-k} \leq q_0 + q_{n-1} \;\; \leq q_{n-1} \;\text{ by lemma 16(a)} \,.$$

Now prove part (c) of the lemma. For integers $n \geq 1$ let

$$\delta_n = q_{n-1} - (q_{\lfloor (n-1)/2 \rfloor} + q_{\lceil (n-1)/2 \rceil}) \,.$$

We wish to show that $\delta_n/2 \leq n \ln 2$. There are two cases to consider, depending on the parity of n. First suppose that n is odd. Then $\lfloor (n-1)/(2) \rfloor = \lceil (n-1)/(2) \rceil = (n-1)/(2)$ and so

$$\begin{aligned}
\delta_n/2 &= \frac{1}{2}(q_{n-1} \;-\; 2q_{(n-1)/2}) \\
&= nH_{n-1} - (n+1)H_{(n-1)/2} \\
&\leq n(H_{n-1} - H_{(n-1)/2}) \\
&\leq n \ln 2 \,.
\end{aligned}$$

Now suppose that n is even and at least 2. Then $\lfloor (n-1)/2 \rfloor = (n-2)/2$ and $\lceil (n-1)/2 \rceil = n/2$. Hence

$$\begin{aligned}\delta_n/2 &= \frac{1}{2}(q_{n-1} \; - \; q_{n/2} \; - \; q_{(n-2)/2}) \\ &= nH_{n-1} - (n+1)H_{n/2} + 1 \\ &\le n(H_{n-1} - H_{n/2}) \\ &\le n \ln 2 \,.\end{aligned}$$

Now prove (d). Observe that

$$\max(A_n) \; - \; \min(A_n) = \max(\gamma_{n,k}) \; - \; \min(\gamma_{n,k}) \,.$$

Define

$$q_n - (q_{\lfloor (n-1)/2 \rfloor} + q_{\lceil (n-1)/2 \rceil}) \;\; = \;\; \mu_n \,.$$

Using lemma 16(a),

$$\max_k(\gamma_{n,k}) \;\; = \;\; \mu_n$$

and

$$\min_k(\gamma_{n,k}) = q_n - (q_{n-1} + q_0) = q_n - q_{n-1} \,.$$

Thus

$$\begin{aligned}\max(A_n) \; - \; \min(A_n) &= q_n - (q_{\lfloor (n-1)/2 \rfloor} + q_{\lceil (n-1)/2 \rceil}) \; - \; (q_n - q_{n-1}) \\ &= \delta_n\end{aligned}$$

and now the conclusion follows by lemma 16(c).

Finally, we prove (e). Note that the upper bound follows by observing that $\gamma_{n,k}$ is non-negative. To establish the lower bound, observe that

$$\begin{aligned}\min(A_n) &= n - 1 \; - \; \max(\gamma_{n,k}) \\ &= n - 1 \; - \; \mu_n \,.\end{aligned}$$

To complete the proof, we wish to show that

$$-(n-1) \;\; \le \;\; n - 1 \; - \; \mu_n \,,$$

namely, that

$$\mu_n \;\; \le \;\; 2n - 2 \qquad \text{for} \quad n \ge 1 \,. \tag{$**$}$$

Observe that for all $n \ge 1$,

$$\mu_n = \delta_n + q_n - q_{n-1} = \delta_n + 2(H_n - 1)$$

Thus using our previous expressions for δ_n, we have, for n odd,

$$\begin{aligned}\frac{\mu_n}{2} &= nH_{n-1} - (n+1)H_{(n-1)/2} + H_n - 1 \\ &= (n+1)(H_{n-1} - H_{(n-1)/2}) + \frac{1}{n} - 1\,,\end{aligned}$$

and for n even and at least two,

$$\begin{aligned}\frac{\mu_n}{2} &= nH_{n-1} - (n+1)H_{n/2} + 1 + H_n - 1 \\ &= (n+1)(H_{n-1} - H_{(n-1)/2}) + \frac{1}{n}\,.\end{aligned}$$

Thus to show that $\mu_n \leq 2n-2$, we need to show that

$$H_{n-1} - H_{(n-1)/2} \leq \frac{n - \frac{1}{n}}{n+1} \text{ for } n \text{ odd, and}$$

$$H_{n-1} - H_{n/2} \leq \frac{n - (1 + \frac{1}{n})}{n+1} \text{ for } n \text{ even and at least two.}$$

This can be checked directly for small n; for larger values of n, it follows from observing that $\ln 2$ can be squeezed between the left and right hand sides of the above inequalities. This completes the proof of lemma 16.

5 Appendix 2: Computations, Simulations and Figures

In this section we present some figures relevant to the theoretical results we have obtained. Specifically, for both the basic and median-of-3 variants of quicksort, we present three figures: the *exact* key comparison frequency distribution for $n = 100$, a (pseudo-random number generated) simulation of the distribution for $n = 120{,}000$, and a comparison of the theoretical (as from our theorems) and simulated (as from the simulations) behaviour of $\Pr\left[\left|\frac{Q_n}{q_n} - 1\right| > \varepsilon\right]$.

Let $f(n,k)$ be the probability that a basic quicksort of n keys takes k key comparisons, assuming all possible input permutations are equally likely. The exact key comparison frequency distribution figures were created by picking a suitably small n (small enough to allow the computations to finish in a reasonable time), and computing $f(n,k)$ for all possible values of k. This complete distribution for n keys can be computed in time $O(n^6)$ and space $O(n^3)$ as follows:

Set $f(0,k) = \delta_{0,k}$ for all k (where $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise), and for $n \geq 1$

$$f(n,k) = \sum_{r=1}^{n} \frac{1}{n} \sum_{x,y} \{ f(r-1,x) f(n-r,y) \; : \; x + y = k - n + 1 \} \,.$$

Figures 1a and 1b show the exact quicksort key comparison frequency distribution for $n = 100$. In order to show better the behaviour in the tail, the logarithm of the distribution is also shown.

A similar computation works for the median-of-3 case, although we must be careful to specify the number of key comparisons in the basis cases. Also, we assume that for $n \geq 3$ keys, 3 comparisons are performed to find the median, and so partitioning takes a total of $3 + n - 3 = n$ comparisons.

Let $g(n, k)$ be the probability that a median-of-3 quicksort of n keys takes k key comparisons, assuming all input permutations to be equally likely. For $k \geq 0$, set

$$g(0,k) = g(1,k) = \delta_{0,k}, \quad \text{and} \quad g(2,k) = \delta_{1,k} \ .$$

Then for $n \geq 3$ and $k \geq 0$,

$$g(n,k) \ = \sum_{r=1}^{n} \frac{(r-1)(n-r)}{\binom{n}{3}} \sum_{x,y} \{ g(r-1,x) g(n-r,y) \ : \ x + y = k - n \} \ .$$

The simulation figures 2a and 2b were produced by simulating 10,000 trials of quicksort. The minimum and maximum values on the horizontal scale are the respective minimum and maximum observed during the trials; once these values were determined, the range of observed values was partitioned into a number of buckets, and the density of each bucket (as shown by the height of the corresponding bar) was determined. Note that the maximum observed number of key comparisons is far less than the maximum possible number of key comparisons.

The final two figures show how the bounds established in our two theorems compare with simulations of quicksort, for a fixed value of ϵ, namely $\varepsilon = 0.05$. For each of 60 values of n from 2,000 through 120,000 (incrementing by 2,000), 1,000 trials of quicksort were performed. For each set of 1000 trials, the empirical value $\Pr\left[\left| \frac{Q_n}{q_n} - 1 \right| > \varepsilon \right]$ (that is, the proportion of the trials in which $| \ Q_n / q_n - 1 \ | > \varepsilon$) and the quantity $n^{-2\varepsilon(\ln \ln n - c)}$ are shown, for a certain value of c. (The value of c was chosen to least-squares fit the larger half of the sample points for n.)

All simulations were performed on a cluster of Suns and Sparcs. Mike Hallett, Fahir Ergincan and Doug Goodman helped with the running of the quicksort simulations.

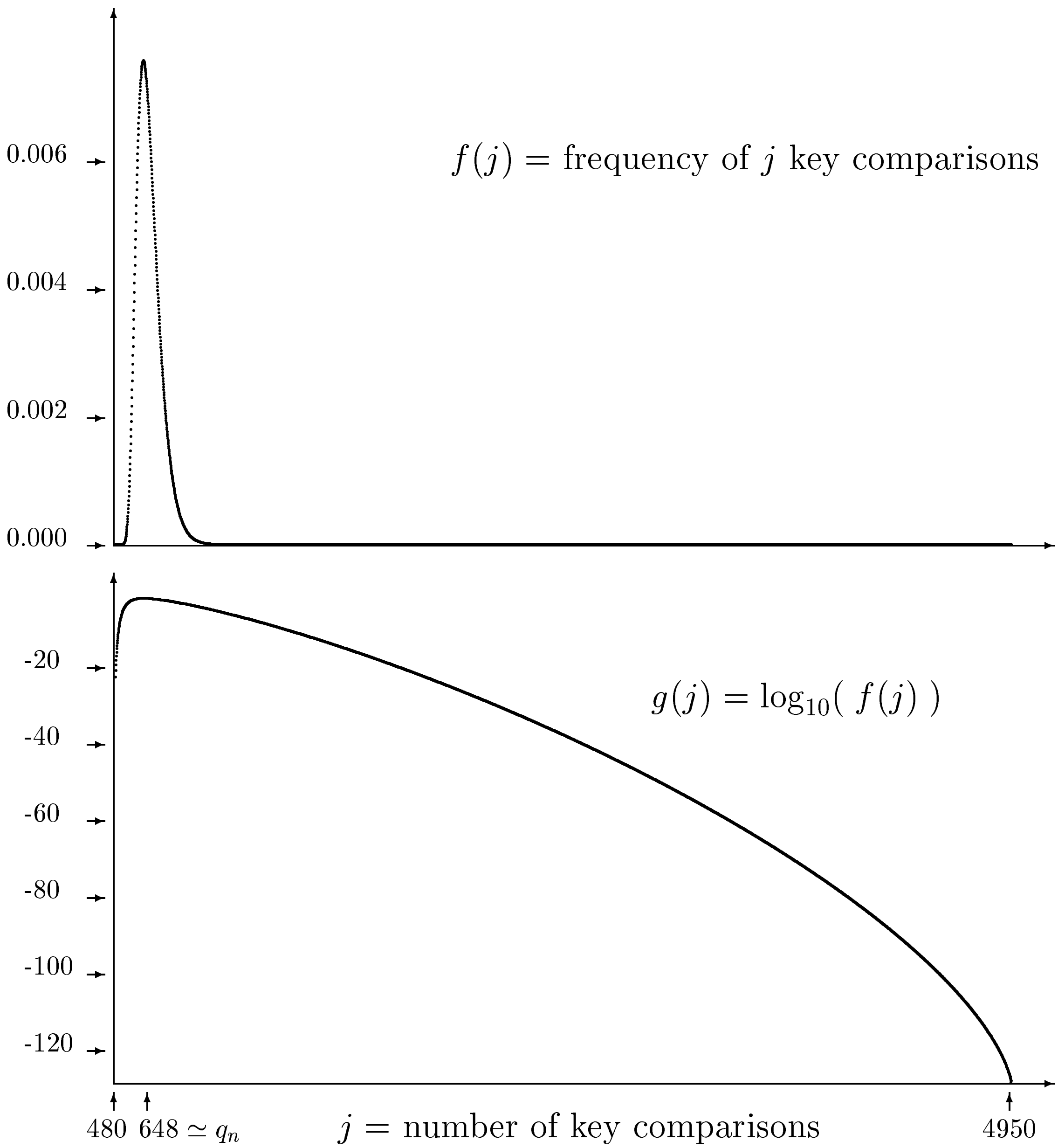

Figure 1.a
Basic Quicksort Key Comparison Frequency
computed exactly
$n = 100$ keys

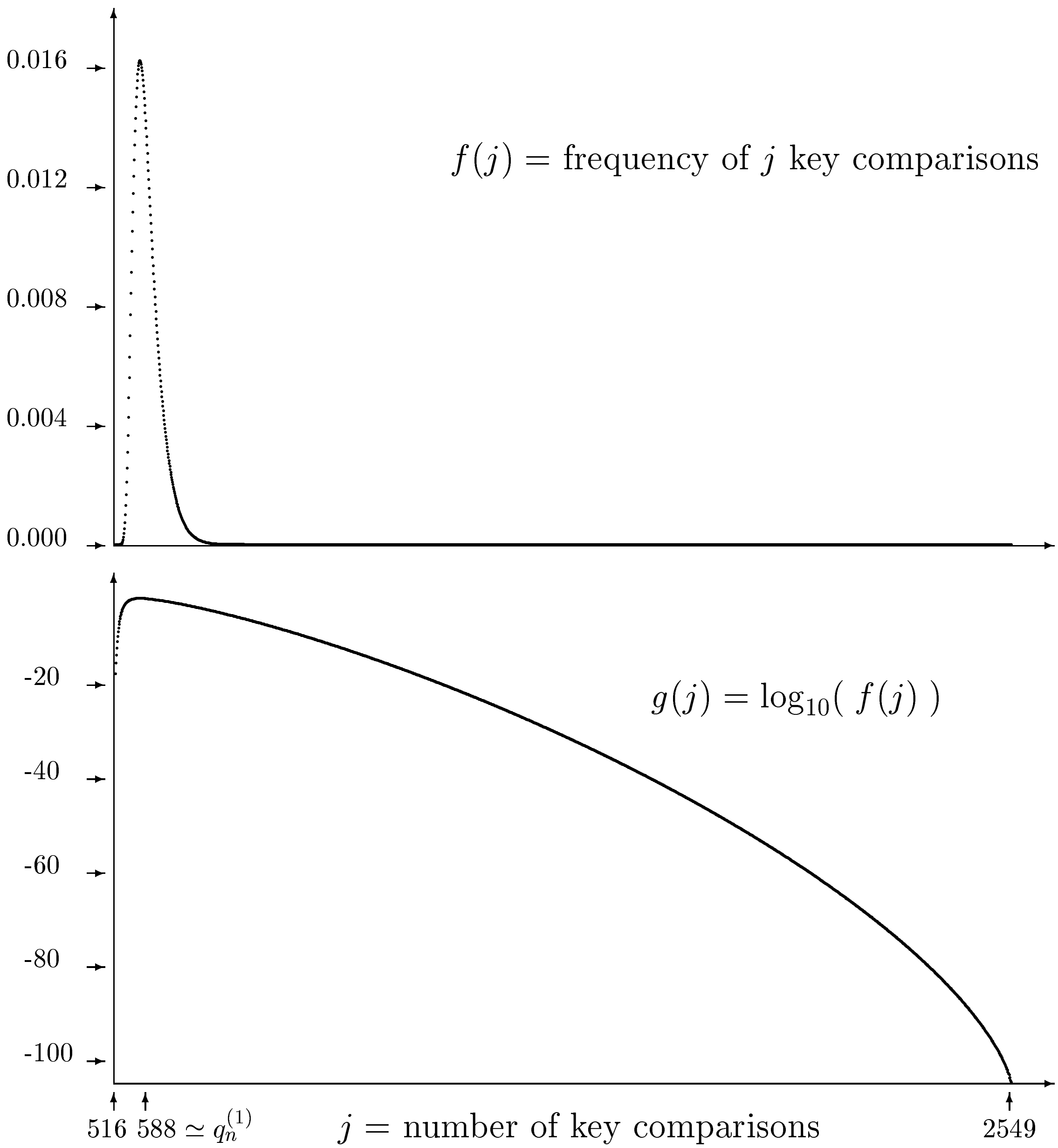

Figure 1.b
Median-of-3 Quicksort Key Comparison Frequency
computed exactly
$n = 100$ keys

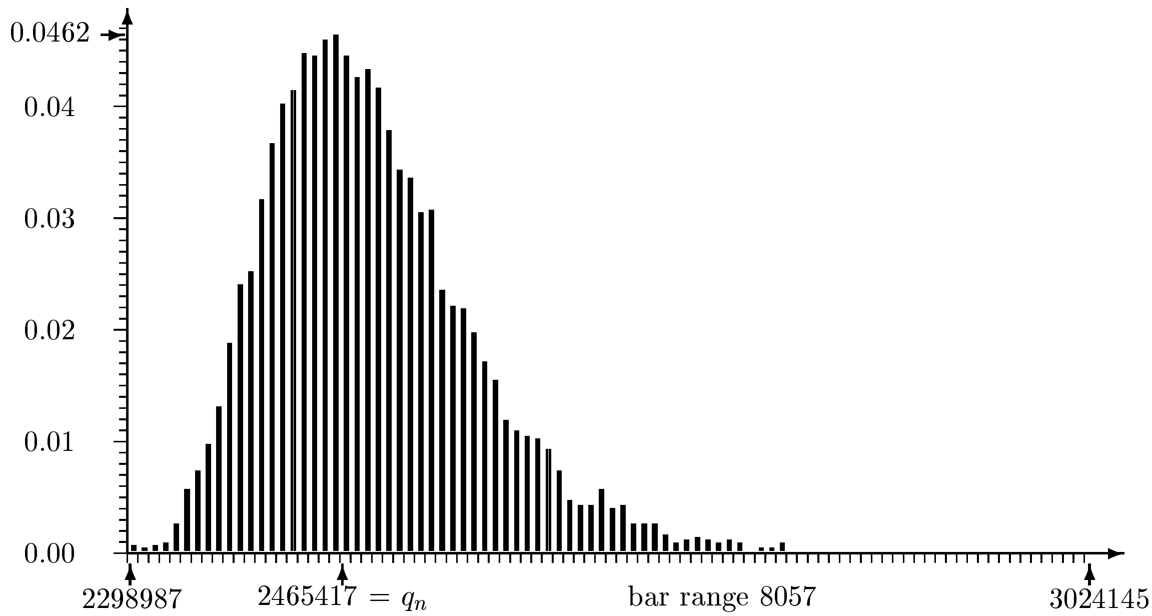

Figure 2.a
Basic Quicksort Key Comparison Frequency
from simulation: 10,000 trials
$n = 120,000$ keys

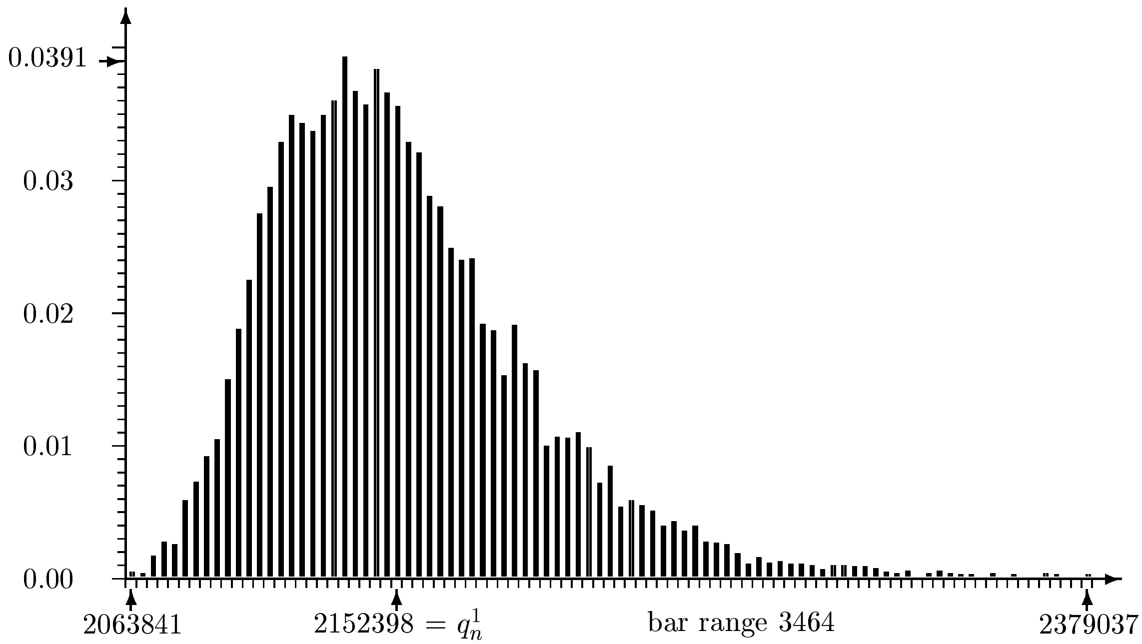

Figure 2.b
Median-of-3 Quicksort Key Comparison Frequency
from simulation: 10,000 trials
$n = 120,000$ keys

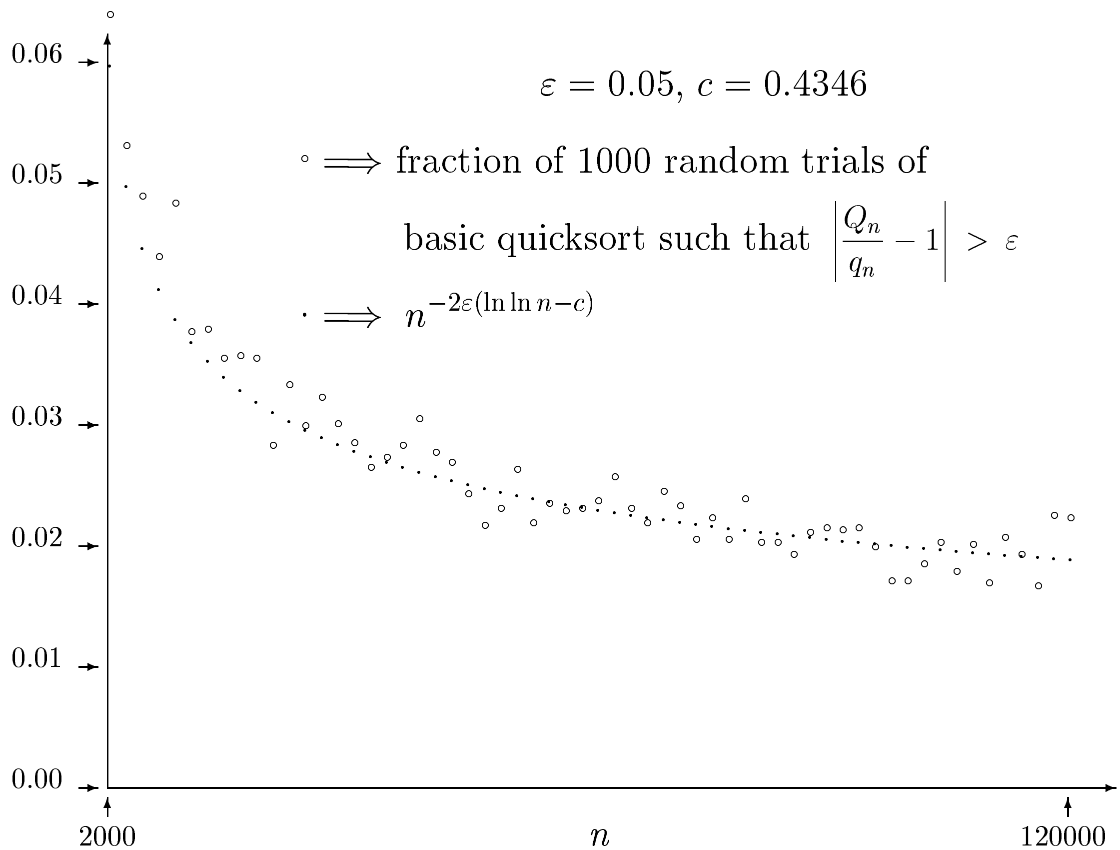

Figure 3.a
Basic Quicksort
Theoretical vs. Observed Deviations

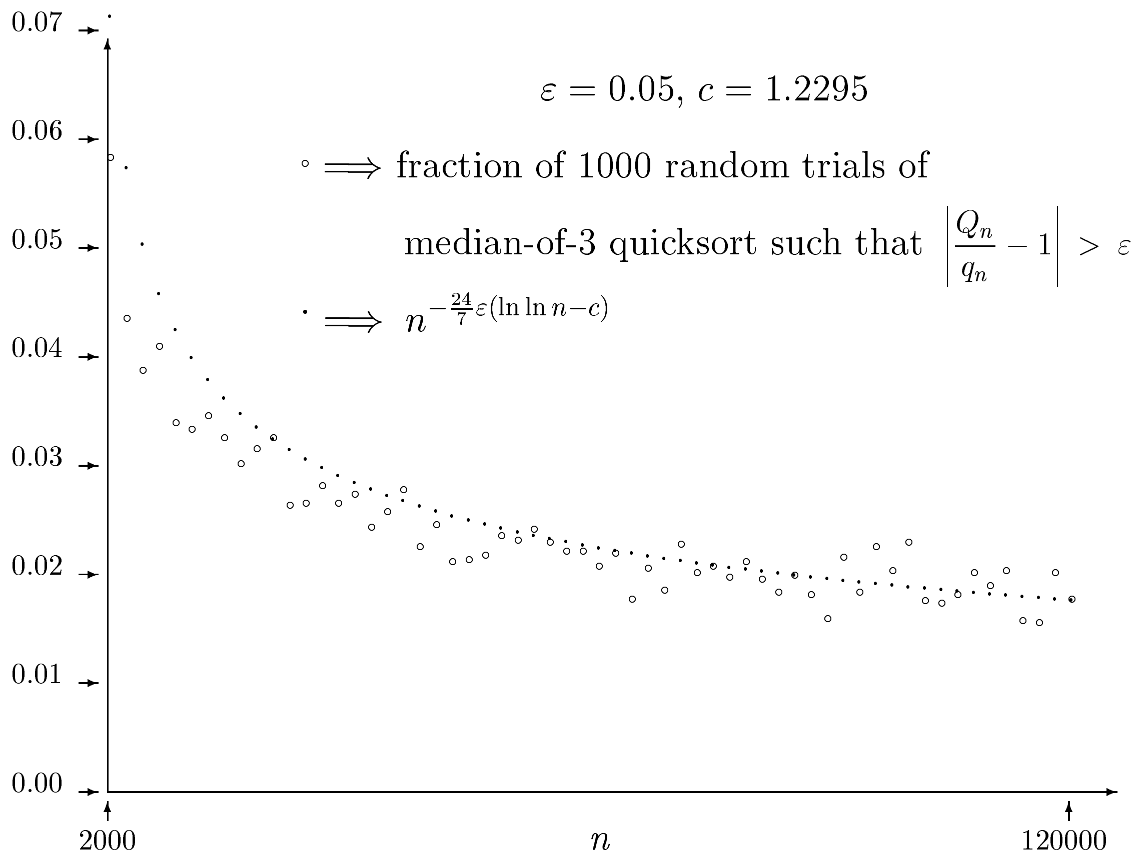

Figure 3.b
Median-of-3 Quicksort
Theoretical vs. Observed Deviations

Acknowledgements. The first steps in this investigation were taken in the ideal working environment of the Workshop on Random Graphs and Probabilistic Algorithms, organized by Bruce Reed, and held at Bellairs Research Institute of McGill in February 1988.

We had interesting discussions with Harry Kesten and Joel Spencer concerning the proof of Lemma 7. A question of Richard Karp prompted us to strengthen earlier versions of our two theorems, so as to handle the case for epsilon tending to zero.

References

[Dev86] L. Devroye. A note on the height of binary search trees. *Journal of the ACM*, 33(3):489–498, July 1986.

[Hen89] P. Hennequin. Combinatorial analysis of quicksort algorithm. *Theoretical Informatics and Applications*, 23(3):317–333, 1989.

[Hoa61] C.A.R. Hoare. Partition (algorithm 63), quicksort (algorithm 64), and find (algorithm 65). *J. ACM*, 7:321–322, 1961.

[Hoa62] C.A.R. Hoare. Quicksort. *Computer J.*, 5:10–15, 1962.

[Knu73] D. Knuth. *The Art of Computer Programming*. Addison-Wesley, first edition, 1973.

[Mah92] H.M. Mahmoud. *Evolution of Random Search Trees*. John Wiley and Sons, New York, 1992.

[McD89] C.J.H. McDiarmid. On the method of bounded differences. In J Siemons, editor, *Surveys in Combinatorics*. London Math Society, 1989.

[MH92] C.J.H. McDiarmid and R.B. Hayward. Strong concentration for quicksort. In *Proc. of the 3rd ann. ACM-SIAM Symposium on Discrete Algorithms*, pages 414–421, Orlando, Florida, January 1992.

[Nev75] J. Neveu. *Discrete-Parameter Martingales*. North-Holland/American Elsevier, 1975.

[Rég89] M. Régnier. A limiting distribution for quicksort. *Theoretical Informatics and Applications*, 23(3):335–343, 1989.

[Rös91] U. Rösler. A limit theorem for quicksort. *Theoretical Informatics and Applications*, 25(1):85–100, 1991.

[Sed80] R. Sedgewick. *Quicksort*. Garland Publishing Inc., first edition, 1980.

[Van70] M.H. VanEmden. Increasing the efficiency of quicksort. *Communications of the ACM*, 13(9):563–567, September 1970.

[Wil91] D. Williams. *Probability with Martingales*. Cambridge University Press, 1991.