

An asymptotic theory for recurrence relations based on minimization and maximization

HSIEN-KUEI HWANG AND TSUNG-HSI TSAI

Institute of Statistical Science

Academia Sinica

Taipei 115

Taiwan

November 20, 2001

Abstract

We derive asymptotic approximations for the sequence $f(n)$ defined recursively by $f(n) = \min_{1 \leq j < n} \{f(j) + f(n-j)\} + g(n)$, when the asymptotic behavior of $g(n)$ is known. Our tools are general enough and applicable to another sequence $F(n) = \max_{1 \leq j < n} \{F(j) + F(n-j) + \min\{g(j), g(n-j)\}\}$, also frequently encountered in divide-and-conquer problems. Applications of our results to algorithms, group testing, dichotomous search, etc. are discussed.

Key words. Divide-and-conquer, recurrence relations, slowly varying functions, power-of-two rules, optimal algorithms.

AMS subject classifications. 68W40 68Q25 05A20 11B37.

*Although this may seem a paradox,
all exact science is dominated by
the idea of approximation.
— Bertrand Russell (1872–1970)*

1 Introduction

The following recurrence relation

$$f(n) = \min_{1 \leq j < n} \{f(j) + f(n-j)\} + g(n) \quad (n \geq 2), \quad (1)$$

with $f(1)$ given, appeared often in diverse problems: random trees (see Hammersley and Grimmett [22], Harding [23], Fill [12]), Huffman coding (see Glassey and Karp [18], Chang and Thomas [7]), binomial group testing (see Hwang et al. [30], O'Geran et al. [43]), dynamic programming (see Fredman and Knuth [16]), dichotomous search problems (see Wong [56], Morris [41], Carlitz [6], Gal [17], Greene and Knuth [20]), design of electrical circuits (see Pelling and Rogers [46, 45]), divide-and-conquer problems (see Chen et al. [8]), etc.

Hammersley and Grimmett [22] proved that

- if $g(n)$ is nonincreasing then $f(n) = \sum_{1 \leq k \leq n} g(k)$;

- if $g(n)$ is nondecreasing and convex, namely $g(n+1) - 2g(n) + g(n-1) \geq 0$ for $n \geq 3$, then $f(n) = \psi(n)$, where $\psi(1) := f(1)$ and

$$\psi(n) = \psi(\lfloor n/2 \rfloor) + \psi(\lceil n/2 \rceil) + g(n) \quad (n \geq 2); \quad (2)$$

- if $g(n)$ is nondecreasing and concave, namely $g(n+1) - 2g(n) + g(n-1) \leq 0$ for $n \geq 3$, then $f(n) = \phi(n)$, where $\phi(1) := f(1)$ and

$$\phi(n) = \phi(2^{\lfloor \log_2 2n/3 \rfloor}) + \phi(n - 2^{\lfloor \log_2 2n/3 \rfloor}) + g(n) \quad (n \geq 2); \quad (3)$$

see also Fredman and Knuth [16], Glassey and Karp [18], Snir [52], Walsh [54]. Note that $2^{\lfloor \log_2 2n/3 \rfloor}$ is the unique power-of-two lying between $n/3$ and $2n/3$. Also observe that the convexity and concavity are used for $n \geq 3$ because $f(n)$ can be determined without uncertainty for $n \leq 5$ and the sign of $g(4) - 2g(3) + g(2)$ is needed for $f(6)$.

This result is, although elegant, less useful in practice for two reasons: First, floor and ceiling operators, ubiquitous in computer algorithms, define always functions that are neither convex nor concave, like $\lfloor n/2 \rfloor$, $\lfloor \sqrt{n} \rfloor$, $\lceil \log_2 n \rceil$, etc; second, the “merge cost” $g(n)$ is usually not known in its complete precision so as to determine its convexity or concavity for all $n \geq 3$, for example, $g(n) \sim \sqrt{\pi} n^{3/2}$, where, *here and throughout this paper*, the symbol $g(n) \sim \gamma(n)$ means that $g(n)/\gamma(n) \rightarrow 1$ as $n \rightarrow \infty$.

In general, when $g(n)$ is known to be bounded above by some scale in terms of O or o , the sequence $f(n)$ is easily estimated by the usual master theorems (applying to either $\phi(n)$ or $\psi(n)$ defined above; see Cormen et al. [9] and Roura [51]). On the other hand, when the asymptotic behavior of $g(n)$ is known (like $g(n) \sim n/\sqrt{\log_2 n}$), the problem of determining the asymptotic behavior of $f(n)$ becomes more involved because we need then an asymptotically precise *lower bound*, which is in general harder than the corresponding upper bound. Such a study was first initiated by Fredman and Knuth [16], and then followed by Kapoor and Reingold [34, 35], for more general recurrences but for restricted $g(n)$; see also Pippenger [47].

The purpose of this paper is to develop a theory for asymptotics of $f(n)$ for more general $g(n)$. In particular, we show that if

$$g(n) \sim n^\alpha \ell(n),$$

where $\ell(n)$ is slowly varying in the sense that $\ell(\lfloor bn \rfloor)/\ell(n) \sim 1$ for any $b > 0$ as $n \rightarrow \infty$, then

$$\begin{cases} f(n) \sim \psi(n) \sim n^\alpha \ell(n)/(1 - 2^{1-\alpha}), & \text{if } \alpha > 1; \\ f(n) \sim \psi(n) \sim \phi(n) \sim n \sum_{1 \leq 2^k \leq n} \ell(2^k), & \text{if } \alpha = 1, \end{cases}$$

provided that the partial-sum on the right-hand side diverges with n ; see Section 3 for more general results. Common examples of slowly varying functions include any real powers of $\log n$ (and iterated logs), $e^{(\log n)^{b_1} (\log \log n)^{b_2} \dots (\log \dots \log n)^{b_k}}$, where $b_1 < 1$, etc. Such a consideration covers almost all practical cases of interest.

An algorithmic consequence of this result is that if $\sum_{2^k \leq n} g(2^k)/2^k \rightarrow \infty$, then the usual *half-half divide-and-conquer* rule $n \mapsto (\lfloor n/2 \rfloor, \lceil n/2 \rceil)$ is preferable.

On the other hand, when $g(n)$ grows faster than any power of n , then

$$f(n) \sim \psi(n) \sim g(n);$$

see Theorem 1 for a precise statement. The remaining case is when $f(n)$ is asymptotically linear. By a well-known subadditive theorem of Hammersley [21], if $g(n)$ is nondecreasing, then the convergence of the series $\sum_{n \geq 1} g(n)/n^2$ is both necessary and sufficient for $f(n)/n \sim c$, for some constant c ; see

also de Bruijn and Erdős [5], Devroye [11]. While this case may be regarded as solved, the explicit characterization of c for general $g(n)$ remains a very hard problem. For more subadditive examples, see Steele [53] and Yukich [57].

We will develop tools for deriving effective lower bounds for $f(n)$. The tools are also useful for another type of recurrence

$$F(n) = \max_{1 \leq j < n} \{F(j) + F(n-j) + \min\{g(j), g(n-j)\}\} \quad (n \geq 2), \quad (4)$$

with $F(1)$ specified. Such a recurrence, although defined in a rather different manner, enjoys several common properties with $f(n)$; see Alonso et al. [2], Batty and Rogers [3], Li and Reingold [39], Wang [55], and Section 4 for more details. The corresponding asymptotic theory will also be developed but with most of the details omitted due to analogy. For other recurrences of a similar nature, see, for example, [16, 34, 49, 50].

For this category of problems, there are typically two different approaches to deriving bounds for $f(n)$: *combinatorial (or tree) approach*, which is roughly based on an inductive additivity argument on the associated optimal trees, and *computational approach*, which operates directly on the objective function via induction and additivity. While most previous papers rely on the first approach, we will use throughout this paper the second approach, which in many cases significantly simplifies the proof.

Another feature of recurrences defined via minimization or maximization is that they are very sensible to small variations of the given sequence $g(n)$ as far as indices attaining the minimum value are concerned. For example, if $g(n) = \lfloor \log_2 n \rfloor$, then the indices at which $f(n)$ attain the minimum value do not seem to have a simple pattern and $f(n)/n \rightarrow c$, for some constant $c \approx 1.66$. But if $g(n) = \lceil \log_2 n \rceil$, then it can be proved that $f(n) = \phi(n) \sim 2n$. Curiously, both $g(n)$ gives rise to the same asymptotic behavior for $F(n)$: $F(n) \sim n$; see Li and Reingold [39]. For another example, see Hammersley and Grimmett [22]. On the other hand, regularity of the behaviors of the recurrences in question emerges when we take an *asymptotic* viewpoint.

This paper is organized as follows. We first establish in the next section two lower bounds that are also of some interests *per se*; we then divide the analysis and discussions of the asymptotics of $f(n)$ in three classes according to the growth rate of $g(n)$: *rapidly growing*, *regularly varying* and *almost linear*. The parallel theory for $F(n)$ is given in Section 4. A new optimality property for the *balanced power-of-two* rule $n \mapsto (2^{\lfloor \log_2 2n/3 \rfloor}, n - 2^{\lfloor \log_2 2n/3 \rfloor})$ is derived, extending results given in Chen et al. [8]. Finally, we discuss briefly a few examples and indicate several lines of further investigation.

Notation. Throughout this paper, $f(n)$, $\psi(n)$ and $\phi(n)$, $F(n)$ are defined, respectively, by (1), (2), (3) and (4). For notational convenience, we sometimes write $f(n) = f[g](n)$ to indicate the dependence on the underlying “merge function” $g(n)$; the symbols $\psi[g](n)$, $\phi[g](n)$ and $F[g](n)$ are defined similarly. We implicitly assume that $g(n)$ is nonnegative for $n \geq 1$. We write $L_n = \lfloor \log_2 n \rfloor$ and $\rho(n) := \lfloor \log_2 2n/3 \rfloor$. The symbol $\{x\}$ denotes the fractional part of x .

2 Effective lower bounds for $f(n)$

Without loss of generality, we assume (here and throughout the paper) that $f(1) = g(0) = g(1) = \psi(1) = \phi(1) = F(1) = 0$.

We derive two lower bounds for $f(n)$ under different assumptions on $g(n)$.

2.1 $g(n)$ nondecreasing

Proposition 1. *If $g(n)$ is nondecreasing for $n \geq 1$ then*

$$f(n) \geq \sum_{k \geq 1} g(\lceil n/k \rceil) \quad (n \geq 1). \quad (5)$$

For methodological interests, we give two different proofs, each applicable to derive more general lower bounds. Write $A(n) = \sum_{k \geq 1} g(\lceil n/k \rceil)$. By induction, we need only to prove that

$$A(n) \leq A(j) + A(n-j) + g(n), \quad (6)$$

for $1 \leq j < n$. Note that $f(n)$ is the *maximal* solution of all sequences satisfying the subadditive inequality (6) (and $A(1) = 0$).

First proof of (6). Assume by symmetry $j \leq \lfloor n/2 \rfloor$. By collecting likewise terms, $A(n)$ can be written as

$$A(n) = \sum_{2 \leq k \leq n} \left(\left\lceil \frac{n}{k-1} \right\rceil - \left\lceil \frac{n}{k} \right\rceil \right) g(k).$$

Then

$$A(n) - A(j) - A(n-j) = \sum_{2 \leq k \leq n} (\delta_{n,j}(k-1) - \delta_{n,j}(k)) g(k),$$

where

$$\delta_{n,j}(k) := \left\lceil \frac{n}{k} \right\rceil - \left\lceil \frac{j}{k} \right\rceil - \left\lceil \frac{n-j}{k} \right\rceil.$$

Noting that $\delta_{n,j}(k)$ assumes either 0 or -1 for $1 \leq j < n$ and that $\delta_{n,j}(1) = 0$, $\delta_{n,j}(n) = -1$, $\delta_{n,j}(n-1) = 0$, we have

$$\begin{aligned} A(n) - A(j) - A(n-j) &= g(n) + \sum_{2 \leq k \leq n-2} \delta_{n,j}(k) (g(k+1) - g(k)) \\ &\leq g(n), \end{aligned}$$

as required. \blacksquare

In general, any sum of the form

$$X(n) := \sum_{1 \leq k \leq n} (x_{n,k-1} - x_{n,k}) g(k),$$

with $X(1) := 0$, is a lower bound of $f(n)$ if $\xi_{n,j}(k) := x_{n,k} - x_{j,k} - x_{n-j,k}$ satisfies

$$\xi_{n,j}(k) \begin{cases} \leq 0, & \text{if } 1 \leq k < n; \\ \geq -1, & \text{if } k = n. \end{cases}$$

Second proof of (6). Define $k_0 = 1$ and for $i \geq 1$

$$k_i = \min \left\{ k \geq 1 : \left\lceil \frac{j}{k} \right\rceil < \left\lceil \frac{n}{k+i} \right\rceil \right\}.$$

Define $k_i = n+1$ if $\lceil j/k \rceil \geq \lceil n/(k+i) \rceil$ for $1 \leq k \leq n$. Using the inequalities

$$\begin{aligned} \frac{j}{k_i} + 1 &\leq \left\lceil \frac{j}{k_i} \right\rceil + 1 \leq \left\lceil \frac{n}{k_i+i} \right\rceil, \\ \frac{n-1}{k_i+i} + 1 &\geq \left\lceil \frac{n}{k_i+i} \right\rceil, \end{aligned}$$

we have

$$n - j \geq (k_i + i) \left\lceil \frac{n}{k_i + i} \right\rceil - k_i - i + 1 - k_i \left\lceil \frac{n}{k_i + i} \right\rceil + k_i = i \left\lceil \frac{n}{k_i + i} \right\rceil - i + 1,$$

which implies that

$$\left\lceil \frac{n - j}{i} \right\rceil \geq \left\lceil \frac{n}{k_i + i} \right\rceil.$$

Thus

$$\begin{aligned} A(n) - g(n) &= \sum_{k \geq 1} g \left(\left\lceil \frac{n}{k + 1} \right\rceil \right) \\ &= \sum_{i \geq 1} g \left(\left\lceil \frac{n}{k_i + i} \right\rceil \right) + \sum_{i \geq 0} \sum_{k_i \leq k < k_{i+1}} g \left(\left\lceil \frac{n}{k + i + 1} \right\rceil \right) \\ &\leq \sum_{i \geq 1} g \left(\left\lceil \frac{n - j}{i} \right\rceil \right) + \sum_{i \geq 0} \sum_{k_i \leq k < k_{i+1}} g \left(\left\lceil \frac{j}{k} \right\rceil \right) \\ &= A(n - j) + \sum_{k \geq 1} g \left(\left\lceil \frac{j}{k} \right\rceil \right) \\ &= A(n - j) + A(j). \end{aligned}$$

This completes the proof of (6). \blacksquare

More generally, define $Y(1) := 0$ and

$$Y(n) := \sum_{1 \leq k \leq n} g(y_{n,k}) \quad (n \geq 2),$$

where $y_{n,k}$ is nonincreasing in k for fixed n and $y_{n,1} = n$. If $y_{n,k}$ satisfies the inequality

$$\max\{y_{j,k}, y_{n-j,i}\} \geq y_{n,k+i},$$

for $1 \leq k \leq j$, $1 \leq i \leq n - j$, and $k + i \leq n$, then $Y(n) \leq f(n)$. The proof follows the same line as above and is omitted.

The form of the lower bound (5) was inspired from the upper bound $\sum_{k \geq 2} g(\lfloor n/k \rfloor)$ for $F(n)$ derived in Alonso et al. [2] and Wang [55]. Both of our proofs apply *mutatis mutandis* to $F(n)$ and are much shorter; see Section 4.

Both sequences $\psi(n)$ and $\phi(n)$ can be expressed completely in terms of $g(n)$ as follows. For $n \geq 1$

$$\psi(n) = \sum_{0 \leq k \leq L_n} \left[2^k g \left(\left\lfloor \frac{n}{2^k} \right\rfloor \right) + 2^k \left\{ \frac{n}{2^k} \right\} \left(g \left(\left\lfloor \frac{n}{2^k} \right\rfloor + 1 \right) - g \left(\left\lfloor \frac{n}{2^k} \right\rfloor \right) \right) \right], \quad (7)$$

$$\phi(n) = \sum_{0 \leq k \leq L_n} \left[\left\lfloor \frac{n}{2^k} \right\rfloor \left(g(2^{k+1}) - g(2^k) \right) - g(2^{k+1}) + g \left(2^k \left(1 + \left\{ \frac{n}{2^k} \right\} \right) \right) \right]; \quad (8)$$

see Hammersley and Grimmett [22], Hwang [31], Hwang and Steyaert [32].

From the recursive definitions and these expressions, one easily derives the following bounds.

Corollary 1. *If $g(n)$ is nondecreasing, then*

$$\sum_{0 \leq k \leq L_n} 2^k g \left(\left\lfloor \frac{n}{2^k} \right\rfloor \right) \leq \psi(n) \leq \sum_{0 \leq k \leq L_n} 2^k g \left(\left\lceil \frac{n}{2^k} \right\rceil \right), \quad (9)$$

and

$$\sum_{0 \leq k \leq L_n} \left\lfloor \frac{n}{2^k} \right\rfloor g(2^k) \leq \phi(n) \leq g(n) + \sum_{0 \leq k \leq L_n} \left\lceil \frac{n}{2^k} \right\rceil g(2^k). \quad (10)$$

It is interesting to compare the corresponding terms in both bounds.

Proposition 2. *If $g(n)$ is nondecreasing then for all $n \geq 1$*

$$\frac{\psi(n) + g(n)}{2} \leq f(n) \leq \psi(n), \quad (11)$$

$$\frac{\phi(n) + g(n) - g(\tau(n))}{2} \leq f(n) \leq \phi(n), \quad (12)$$

where

$$\tau(n) = \begin{cases} \max\{2^{\rho(n)}, n - 2^{\rho(n)}\}, & \text{if } |n - 2^{\rho(n)+1}| > 1; \\ 0, & \text{if } |n - 2^{\rho(n)+1}| \leq 1. \end{cases}$$

Furthermore, if $g(2) > 0$, then for $n \geq 3$

$$f(n) \geq \frac{\psi(n) + g(n) + \lfloor n/2 \rfloor g(2)}{2}. \quad (13)$$

Proof. The first inequalities (11) follow from (9) and the relations

$$2f(n) \geq 2A(n) \geq 2g(n) + 2g(\lceil n/2 \rceil) + 4g(\lceil n/4 \rceil) + \dots.$$

To prove the better bound (13), observe that for $n \geq 2$

$$\psi(n) \leq \sum_{0 \leq k \leq L_n - 1} 2^k g\left(\left\lceil \frac{n}{2^k} \right\rceil\right) + \left\lfloor \frac{n}{2} \right\rfloor g(2),$$

since the number of terms contributing to $g(2)$ in (7) is at most $\lfloor n/2 \rfloor$. Now

$$\begin{aligned} 2f(n) &\geq 2A(n) \\ &\geq g(n) + \sum_{0 \leq k \leq L_n - 1} 2^k g\left(\left\lceil \frac{n}{2^k} \right\rceil\right) + 2(n - 1 - 2^{L_n - 1})g(2). \end{aligned}$$

Since

$$n - 1 - 2^{L_n - 1} \geq \left\lfloor \frac{n}{2} \right\rfloor,$$

for $2^{L_n} < n < 2^{L_n+1}$, (13) follows when $n \neq 2^{L_n}$. But (13) is easily checked when $n = 2^{L_n}$. This proves (13).

For the inequalities (12), we first prove the inequality

$$\sum_{k \geq 2} g\left(\left\lceil \frac{j}{k} \right\rceil\right) + \sum_{k \geq 2} g\left(\left\lceil \frac{n-j}{k} \right\rceil\right) \leq \sum_{k \geq 3} g\left(\left\lceil \frac{n}{k} \right\rceil\right), \quad (14)$$

for $n/2 \leq j \leq 2n/3$. [Compare this inequality with (6).] We sketch the proof since it is similar to the second proof of (6). Observe first that

$$\min \left\{ \frac{j}{k}, \frac{n-j}{i} \right\} \leq \frac{n}{i+k} \quad (i, k \geq 2).$$

Define for $i \geq 1$

$$k_i := \min \left\{ k \geq 2 : \frac{j}{k} < \frac{n-j}{i+1} \right\}.$$

Then

$$\frac{n-j}{i+1} \leq \frac{n}{k_i+i}, \quad \text{and} \quad \frac{j}{k} \leq \frac{n}{k+i+1},$$

for $k \geq k_i$ and $i \geq 1$. The remaining proof uses the same last step of the second proof of (6) (with the inequality reversed).

We now apply induction and (14) to prove that

$$\phi(n) \leq 2A(n) - g(n) + g(\tau(n)) \quad (n \geq 1).$$

By induction,

$$\begin{aligned} \phi(2^{\rho(n)}) + \phi(n - 2^{\rho(n)}) &\leq 2 \sum_{k \geq 2} \left(g \left(\left\lceil \frac{2^{\rho(n)}}{k} \right\rceil \right) + g \left(\left\lceil \frac{n - 2^{\rho(n)}}{k} \right\rceil \right) \right) \\ &\quad + g(2^{\rho(n)}) + g(n - 2^{\rho(n)}) + g(\tau(n - 2^{\rho(n)})), \end{aligned} \quad (15)$$

since $\tau(2^{\rho(n)}) = 0$. If $\tau(n) = 0$ then

$$g(2^{\rho(n)}) + g(n - 2^{\rho(n)}) \leq 2g(\lceil n/2 \rceil),$$

and $\tau(n - 2^{\rho(n)}) = 0$. If $\tau(n) > 0$ then

$$g(2^{\rho(n)}) + g(n - 2^{\rho(n)}) \leq g(\lceil n/2 \rceil) + g(\tau(n)),$$

and

$$g(\tau(n - 2^{\rho(n)})) \leq g(\lceil n/2 \rceil).$$

It follows, by (14) and (15), that

$$\begin{aligned} \phi(n) &= \phi(2^{\rho(n)}) + \phi(n - 2^{\rho(n)}) + g(n) \\ &\leq 2 \sum_{k \geq 3} g(\lceil n/k \rceil) + 2g(\lceil n/2 \rceil) + g(\tau(n)) + g(n) \\ &= 2A(n) - g(n) + g(\tau(n)). \end{aligned}$$

This proves the lower bound of (12). \blacksquare

Corollary 2. *Any divide-and-conquer algorithms using either the half-half rule or the balanced power-of-two rule have their costs at most twice the optimum (minimum) value if the “merge cost” is nondecreasing.*

2.2 $g(n)/n$ nondecreasing

In case when $g(n)/n$ is nondecreasing, we have a stronger lower bound when $g(n)$ does not grow too fast. Indeed we prove a more general result covering the case of nondecreasing $g(n)/n$ as a special case.

Proposition 3. Let $\ell(n) := g(n)/n \geq 0$. Define

$$\lambda(k) := \min_{2^k \leq m < 2^{k+1}} \ell(m) \quad (k \geq 0).$$

If $\lambda(k)$ is nondecreasing then

$$B(n) := n \sum_{0 \leq k \leq L_n} \lambda(k) - 2^{L_n+1} \lambda(L_n) \leq f(n),$$

for $n \geq 1$.

Proof. Since $B(1) = -\ell(1) = 0$, it suffices to prove that

$$B(n) - B(j) - B(n-j) \leq g(n), \quad (16)$$

for $1 \leq j < n$ and $n \geq 2$. Assume $j \leq \lfloor n/2 \rfloor$. There are two cases: $L_{n-j} = L_n$ and $L_{n-j} = L_n - 1$.

If $L_{n-j} = L_n$, then the inequality (16) is equivalent to

$$j \sum_{L_j < k \leq L_n} \lambda(k) + 2^{L_j+1} \lambda(L_j) \leq n\ell(n).$$

By applying the inequalities $j < 2^{L_j+1}$ and $a \leq 2^a - 1$ for *integral* $a \geq 0$, we have

$$\begin{aligned} j \sum_{L_j < k \leq L_n} \lambda(k) + 2^{L_j+1} \lambda(L_j) &\leq j\ell(n) + 2^{L_j+1}(L_n - L_j - 1)\ell(n) + 2^{L_j+1}\ell(n) \\ &\leq j\ell(n) + 2^{L_j+1}(2^{L_n-L_j-1} - 1)\ell(n) + 2^{L_j+1}\ell(n) \\ &= (j + 2^{L_n})\ell(n) \\ &\leq (j + n - j)\ell(n) \\ &= n\ell(n), \end{aligned}$$

as desired.

On the other hand, if $L_{n-j} = L_n - 1$, then the inequality (16) becomes

$$j \sum_{L_j < k < L_n} \lambda(k) \leq n\ell(n) - n\lambda(L_n) + 2^{L_n+1}\lambda(L_n) - 2^{L_j+1}\lambda(L_j) - 2^{L_n}\lambda(L_n - 1).$$

So we need only to prove the inequality

$$j \sum_{L_j < k < L_n} \lambda(k) \leq 2^{L_n+1}\lambda(L_n) - 2^{L_j+1}\lambda(L_j) - 2^{L_n}\lambda(L_n - 1). \quad (17)$$

By the same argument as above, we have

$$j \sum_{L_j < k < L_n} \lambda(2^k) \leq (2^{L_n} - 2^{L_j+1}) \lambda(L_n - 1),$$

which is obviously not greater than the right-hand side of (17). This completes the proof. \blacksquare

Corollary 3. If $g(n)/n$ is nondecreasing, then

$$n \sum_{0 \leq k \leq L_n} g(2^k)/2^k - 2g(2^{L_n}) \leq f(n),$$

for $n \geq 1$.

3 Asymptotics of $f(n)$

We derive asymptotic approximations for $f(n)$ under different assumptions on the growth order of $g(n)$.

3.1 $g(n)$ rapidly growing

Intuitively, when $g(n)$ grows faster than any power of n , the total cost $f(n)$ is dominated by $g(n)$, further divide-and-conquer cost being asymptotically negligible.

Theorem 1. *If $g(n)$ satisfies*

$$\frac{g(n)}{\max\{g(\lceil n/2 \rceil), g(\lfloor n/2 \rfloor)\}} \rightarrow \infty, \quad (18)$$

as $n \rightarrow \infty$, then

$$f(n) \sim \psi(n) \sim g(n).$$

Proof. By assumption (18), for any $M > 0$, there exists $n_0 > 1$ such that

$$g(\lfloor n/2 \rfloor) \leq \frac{g(n)}{M}, \quad g(\lceil n/2 \rceil) \leq \frac{g(n)}{M},$$

for $n \geq n_0$. Then, by (7),

$$\begin{aligned} \psi(n) - g(n) &\leq \sum_{L_n - L_{n_0} + 1 \leq k \leq L_n} O(2^k) + \sum_{1 \leq k \leq L_n - L_{n_0}} 2^k \left(g\left(\left\lfloor \frac{n}{2^k} \right\rfloor\right) + g\left(\left\lceil \frac{n}{2^k} \right\rceil\right) \right) \\ &\leq O(n) + 2 \sum_{k \geq 1} \left(\frac{2}{M}\right)^k g(n) \\ &= o(g(n)), \end{aligned}$$

where we used the identities

$$\left\lfloor \frac{\lfloor n/2^k \rfloor}{2} \right\rfloor = \left\lfloor \frac{n}{2^{k+1}} \right\rfloor, \quad \left\lceil \frac{\lceil n/2^k \rceil}{2} \right\rceil = \left\lceil \frac{n}{2^{k+1}} \right\rceil \quad (0 \leq k \leq L_n). \quad (19)$$

Since $g(n) \leq f(n) \leq \psi(n)$, this proves the theorem. \blacksquare

The truncating argument used in the proof will be applied again later.

A typical example of functions $g(n)$ satisfying (18) is $g(n) = e^{cn^\alpha}$, where $c, \alpha > 0$.

Note that $\phi(n)$ is in general larger than $\psi(n)$ when g is rapidly growing. We need stronger conditions on g in order that $\phi(n) \sim g(n)$.

3.2 $g(n)$ regularly varying

We first prove an asymptotic version of Hammersley and Grimmett's [22] result.

Proposition 4. *If $g(n) \sim v(n)$, where $v(n)$ is nondecreasing and convex for $n \geq 3$ with $v(1) = 0$, then*

$$f(n) \sim \psi[g](n) \sim \psi[v](n); \quad (20)$$

on the other hand, if $v(n)$ is nondecreasing, concave for $n \geq 3$, $v(1) = 0$ and $\sum_{1 \leq k \leq L_n} v(2^k)/2^k$ diverges with n , then

$$f(n) \sim \phi[g](n) \sim \phi[v](n) \sim n \sum_{1 \leq k \leq L_n} v(2^k)/2^k. \quad (21)$$

Proof. By assumption, for any $\varepsilon > 0$, there exists $n_0 > 0$ such that

$$(1 - \varepsilon)v(n) \leq g(n) \leq (1 + \varepsilon)v(n) \quad (n \geq n_0).$$

Let

$$e(n) = \begin{cases} (1 - \varepsilon)v(n) - g(n), & \text{if } g(n) < (1 - \varepsilon)v(n); \\ 0, & \text{otherwise.} \end{cases}$$

Then $(1 - \varepsilon)v(n) - e(n) \leq g(n)$ for $n \geq 1$. From this we deduce that

$$\psi[(1 - \varepsilon)v](n) + f[-e](n) = f[(1 - \varepsilon)v](n) + f[-e](n) \leq f[(1 - \varepsilon)v - e](n) \leq f[g](n) \leq \psi[g](n),$$

where the “=” sign follows from applying Hammersley and Grimmett’s [22] result. Since $e(n) = 0$ for $n \geq n_0$, we have

$$f[-e](n) = O(n).$$

But $v(n)$ is nondecreasing, we have, by (9),

$$\begin{aligned} \psi[v](n) &\geq \sum_{0 \leq k \leq L_n} 2^k v(\lfloor n/2^k \rfloor) \\ &\geq \sum_{0 \leq k \leq L_n} 2^k v(2^{L_n-k}) \\ &\geq 2^{L_n} \sum_{0 \leq k \leq L_n} \frac{v(2^k)}{2^k} \\ &\geq \frac{n}{2} \sum_{0 \leq k \leq L_n} \frac{v(2^k)}{2^k}, \end{aligned}$$

which implies that

$$f[-e](n) = o(\psi[v](n)).$$

Thus

$$1 - \varepsilon + o(1) = \frac{\psi[(1 - \varepsilon)v](n) + f[-e](n)}{\psi[v](n)} \leq \frac{f[g](n)}{\psi[v](n)} \leq \frac{\psi[g](n)}{\psi[v](n)} \leq 1 + \varepsilon + o(1).$$

Since ε is arbitrary, it follows that

$$\frac{f[g](n)}{\psi[v](n)} \rightarrow 1 \quad \text{and} \quad \frac{\psi[g](n)}{\psi[v](n)} \rightarrow 1,$$

from which we conclude that

$$\frac{f[g](n)}{\psi[g](n)} \rightarrow 1.$$

This proves (20). The proof of (21) is similar and omitted. \blacksquare

Recall that a function $\ell(n)$ is *slowly varying* (see Bingham et al. [4]) if

$$\frac{\ell(\lfloor bn \rfloor)}{\ell(n)} \sim 1,$$

for all $b > 0$, as $n \rightarrow \infty$.

Theorem 2. *If $g(n) \sim n^\alpha \ell(n)$, where $\alpha > 1$ and $\ell(n)$ is slowly varying, then*

$$f(n) \sim \psi(n) \sim n^\alpha \ell(n) / (1 - 2^{1-\alpha}).$$

Proof. First observe that the sequence $g_1(n) := n^\alpha \ell(n)$ is asymptotically convex

$$g_1(n+1) - 2g_1(n) + g_1(n-1) \sim \alpha(\alpha-1)n^{\alpha-2}\ell(n) > 0,$$

for $\alpha > 0$. We now apply Proposition 4, using the same truncating argument as in the proof of Theorem 1 if needed, and we obtain

$$f(n) \sim \psi[g_1](n),$$

the error introduced being $O(n)$. By (9) and slow variation of $\ell(n)$,

$$\begin{aligned} \psi[g_1](n) &\sim \sum_{k \geq 1} 2^k \left\lfloor \frac{n}{2^k} \right\rfloor^\alpha \ell\left(\left\lfloor \frac{n}{2^k} \right\rfloor\right) \\ &= n^\alpha \ell(n) \sum_{k \geq 1} 2^{-k(\alpha-1)} \frac{\ell(\lfloor n/2^k \rfloor)}{\ell(n)}. \end{aligned}$$

It remains to show that

$$\sum_{k \geq 1} 2^{-k(\alpha-1)} \frac{\ell(\lfloor n/2^k \rfloor)}{\ell(n)} \sim \sum_{k \geq 1} 2^{-k(\alpha-1)} = (1 - 2^{1-\alpha})^{-1}. \quad (22)$$

Then, by slow variation and (19), there is a $c_0 > 0$ such that for $k \geq 0$

$$\begin{aligned} \frac{\ell(\lfloor n/2^k \rfloor)}{\ell(n)} &= \frac{\ell(\lfloor n/2^k \rfloor)}{\ell(\lfloor n/2^{k-1} \rfloor)} \cdot \frac{\ell(\lfloor n/2^{k-1} \rfloor)}{\ell(\lfloor n/2^{k-2} \rfloor)} \cdots \frac{\ell(\lfloor n/2^{L_n-L_{n_0}+1} \rfloor)}{\ell(\lfloor n/2^{L_n-L_{n_0}} \rfloor)} \cdot \frac{\ell(\lfloor n/2^{L_n-L_{n_0}} \rfloor)}{\ell(\lfloor n/2^{L_n-L_{n_0}-1} \rfloor)} \cdots \frac{\ell(\lfloor n/2 \rfloor)}{\ell(n)} \\ &\leq c_0 e^{k\delta}, \end{aligned}$$

since $\lfloor n/2^{L_n-L_{n_0}-1} \rfloor \geq n_0$. Thus

$$\sum_{k > k_0} 2^{-k(\alpha-1)} \frac{\ell(\lfloor n/2^k \rfloor)}{\ell(n)} \rightarrow 0,$$

as $k_0 \rightarrow \infty$. On the other hand, for fixed k_0 ,

$$\sum_{1 \leq k \leq k_0} 2^{-k(\alpha-1)} \frac{\ell(\lfloor n/2^k \rfloor)}{\ell(n)} \rightarrow \sum_{1 \leq k \leq k_0} 2^{-k(\alpha-1)}.$$

Thus for any $\varepsilon > 0$, there is an $M > 0$ such that for $k_0 > M$

$$\limsup_{n \rightarrow \infty} \frac{\sum_{k \geq 1} 2^{-k(\alpha-1)} \ell(\lfloor n/2^k \rfloor) / \ell(n)}{\sum_{k \geq 1} 2^{-k(\alpha-1)}} \leq \frac{\sum_{k \leq k_0} 2^{-k(\alpha-1)} + \varepsilon}{\sum_{k \geq 1} 2^{-k(\alpha-1)}},$$

and

$$\liminf_{n \rightarrow \infty} \frac{\sum_{k \geq 1} 2^{-k(\alpha-1)} \ell(\lfloor n/2^k \rfloor) / \ell(n)}{\sum_{k \geq 1} 2^{-k(\alpha-1)}} \geq \frac{\sum_{k \leq k_0} 2^{-k(\alpha-1)}}{\sum_{k \geq 1} 2^{-k(\alpha-1)}}.$$

This proves (22) and the proposition. \blacksquare

3.3 $g(n)$ almost linear

We consider in this section the case $g(n) \sim n\ell(n)$, where $\ell(n)$ is slowly varying. This case is more involved than previous cases because we have no *a priori* information on convexity or concavity (even asymptotically).

Theorem 3. *If $g(n) \sim n\ell(n)$, where $\ell(n)$ is slowly varying, then*

$$f(n) \sim \psi(n) \sim \phi(n) \sim n \sum_{0 \leq k \leq L_n} \ell(2^k),$$

provided that $\sum_{0 \leq k \leq L_n} \ell(2^k) \rightarrow \infty$.

In the special case when $\ell(n)$ is nondecreasing, the theorem follows directly from Corollary 3, (9), and (10); also if $n\ell(n)$ is asymptotically concave, the theorem follows from Proposition 4. Note that the lower bound $A(n)$ satisfies ($g(n)$ being eventually nondecreasing)

$$A(n) \sim n \sum_{1 \leq j \leq n} \ell(j)/j \sim (\log 2)n \sum_{0 \leq k \leq L_n} \ell(2^k) \sim (\log 2)f(n),$$

by slow variation, which is insufficient for asymptotic equivalence of $f(n)$.

Proof. Obviously, $f(n) \leq \min\{\psi(n), \phi(n)\}$, and by (9) and (10),

$$\psi(n) \sim \phi(n) \sim n \sum_{0 \leq k \leq L_n} \ell(2^k).$$

Thus, in particular,

$$\limsup_{n \rightarrow \infty} \frac{\psi(n)}{n \sum_{0 \leq k \leq L_n} \ell(2^k)} \leq 1, \quad \limsup_{n \rightarrow \infty} \frac{\phi(n)}{n \sum_{0 \leq k \leq L_n} \ell(2^k)} \leq 1.$$

It remains to prove that

$$\liminf_{n \rightarrow \infty} \frac{f(n)}{n \sum_{0 \leq k \leq L_n} \ell(2^k)} \geq 1. \tag{23}$$

To that purpose, we need the *uniform convergence theorem* for slowly varying functions, which states that the convergence $\ell(\lfloor bn \rfloor)/\ell(n) \rightarrow 1$ is uniform on any finite interval of the positive real line; see Section 1.2 of Bingham et al. [4].

Then, for fixed $0 < \varepsilon < 1$, there exists $n_1 > 0$ such that for all $n \geq n_1$

$$\frac{\ell(\lfloor bn \rfloor)}{\ell(n)} \leq 1 + \frac{\varepsilon}{4}, \tag{24}$$

uniformly for $1/4 \leq b \leq 1$. On the other hand, by assumption, there exists $n_2 > 0$ such that for all $n \geq n_2$

$$\frac{n\ell(n)}{g(n)} \leq 1 + \varepsilon.$$

Take $n_0 := \max\{n_1, n_2\}$. Define

$$\ell_0(n) = \begin{cases} 0, & \text{if } n < n_0; \\ \ell(n), & \text{if } n \geq n_0, \end{cases}$$

and for $n \geq 1$

$$C(n) := \frac{n \sum_{0 \leq k \leq L_n} \ell_0(2^k) - 2^{L_n+1} \ell_0(2^{L_n})}{(1 + \varepsilon)^2}.$$

We show that $C(n)$ satisfies the subadditivity

$$C(n) \leq C(j) + C(n - j) + g(n), \quad (25)$$

for $n \geq 2$ and $1 \leq j < n$; this will imply that $C(n) \leq f(n)$ for $n \geq 1$, and that

$$\liminf_{n \rightarrow \infty} \frac{f(n)}{n \sum_{0 \leq k \leq L_n} \ell_0(2^k) - 2^{L_n+1} \ell_0(2^{L_n})} \geq \frac{1}{(1 + \varepsilon)^2}.$$

Since ε is arbitrary, we then obtain (23) by noting that

$$n \sum_{0 \leq k \leq L_n} \ell(2^k) = n \sum_{0 \leq k \leq L_n} \ell_0(2^k) + O(n),$$

and

$$\ell(2^{L_n}) = o\left(\sum_{0 \leq k \leq L_n} \ell(2^k)\right).$$

To prove (25), we assume $j \leq \lfloor n/2 \rfloor$. Then $L_j \leq L_n - 1$ and $L_{n-j} = L_n$ or $L_{n-j} = L_n - 1$. If $L_{n-j} = L_n$, then

$$\begin{aligned} (1 + \varepsilon)^2 (C(n) - C(j) - C(n - j)) &= j \ell_0(2^{L_n}) + j \sum_{L_j < k < L_n} \ell_0(2^k) + 2^{L_j+1} \ell_0(2^{L_j}) \\ &=: \Delta_1; \end{aligned}$$

if $L_{n-j} = L_n - 1$, then

$$\begin{aligned} (1 + \varepsilon)^2 (C(n) - C(j) - C(n - j)) &= n \ell_0(2^{L_n}) + j \sum_{L_j < k < L_n} \ell_0(2^k) + 2^{L_j+1} \ell_0(2^{L_j}) + 2^{L_n} \ell_0(2^{L_n-1}) - 2^{L_n+1} \ell_0(2^{L_n}) \\ &=: \Delta_2. \end{aligned}$$

We prove that

$$\max\{\Delta_1, \Delta_2\} \leq (1 + \varepsilon)n\ell(n),$$

which implies that

$$C(n) - C(j) - C(n - j) \leq \frac{n\ell_0(n)}{1 + \varepsilon} \leq g(n),$$

proving (25).

First we have, by (24),

$$\begin{aligned} j \sum_{L_j < k < L_n} \ell_0(2^k) + 2^{L_j+1} \ell_0(2^{L_j}) &\leq 2^{L_j+1} \sum_{1 \leq k \leq L_n - L_j} \ell_0(2^{L_n - k}) \\ &\leq 2^{L_j+1} \sum_{1 \leq k \leq L_n - L_j} (1 + \varepsilon/4)^{\lceil k/2 \rceil} \ell_0(2^{L_n}) \\ &\leq (1 + \varepsilon/4) 2^{L_j+1} \sum_{1 \leq k \leq L_n - L_j} 2^{\lceil k/2 \rceil - 1} \ell_0(2^{L_n}) \\ &\leq (1 + \varepsilon/4) 2^{L_n} \ell_0(2^{L_n}). \end{aligned}$$

Thus, we have

$$\begin{aligned}\Delta_1 &< j\ell_0(2^{L_n}) + (1 + \varepsilon/4)2^{L_n}\ell_0(2^{L_n}) \\ &\leq (1 + \varepsilon/4)^2 (j + 2^{L_n}) \ell_0(n) \\ &\leq (1 + \varepsilon)n\ell_0(n),\end{aligned}$$

where we used the inequality $2^{L_n-j} = 2^{L_n} \leq n - j$.

Similarly,

$$\begin{aligned}\Delta_2 &< n\ell_0(2^{L_n}) + (1 + \varepsilon/4)2^{L_n}\ell_0(n) + 2^{L_n}\ell_0(2^{L_n-1}) - 2^{L_n+1}\ell_0(2^{L_n}) \\ &\leq (n + (1 + \varepsilon/4)2^{L_n} + (1 + \varepsilon/4)2^{L_n} - 2^{L_n+1}) \ell_0(2^{L_n}) \\ &= (n + \varepsilon 2^{L_n-1}) (1 + \varepsilon/4)\ell_0(n) \\ &\leq (1 + \varepsilon)n\ell_0(n).\end{aligned}$$

This completes the proof. \blacksquare

Corollary 4. *If $g(n) \sim n(\log_2 n)^\alpha$, where $\alpha > -1$, then*

$$f(n) \sim \frac{n(\log_2 n)^{\alpha+1}}{\alpha + 1};$$

if $g(n) \sim n/\log_2 n$, then

$$f(n) \sim n \log_2 \log_2 n.$$

4 Asymptotics of $F(n)$

The previous results and methods of proof for $f(n)$ also apply to $F(n)$. We list the corresponding results with a sketch of proof if necessary.

4.1 Batty and Rogers's result

Batty and Rogers [3] derived an analogous version for $F(n)$ of Hammersley and Grimmett's results mentioned in Introduction. That is,

- if $g(n)$ is nonincreasing then $F(n) = \sum_{1 \leq k < n} g(k)$;
- if $g(n)$ is nondecreasing and convex, namely $g(n+1) - 2g(n) + g(n-1) \geq 0$ for $n \geq 2$, then $F(n) = \psi[g^*](n)$, where $g^*(n) := g(\lfloor n/2 \rfloor)$.
- if $g(n)$ is nondecreasing and concave, namely $g(n+1) - 2g(n) + g(n-1) \leq 0$ for $n \geq 2$, then $F(n) = \phi[g^\dagger](n)$, where $g^\dagger(n) := g(\min\{2^{\rho(n)}, n - 2^{\rho(n)}\})$; also $F(n) = \varphi(n)$, where

$$\varphi(n) = \varphi(2^{\lceil \log_2 n/2 \rceil}) + \varphi(n - 2^{\lceil \log_2 n/2 \rceil}) + g(\min\{2^{\lceil \log_2 n/2 \rceil}, n - 2^{\lceil \log_2 n/2 \rceil}\}) \quad (n \geq 2),$$

with $\varphi(1) = F(1)$.

The last recurrence appeared in the analysis of bottom-up mergesort; see Panny and Prodinger [44]. Note that $2^{\lceil \log_2 n/2 \rceil}$ is the largest power of two less than n .

For an extension of the above result, see Li and Reingold [39].

4.2 Upper bounds

The following result is due to Alonso et al. [2] and Wang [55].

Proposition 5. *If $g(n)$ is nondecreasing, then for $n \geq 1$,*

$$F(n) \leq \sum_{k \geq 2} g(\lfloor n/k \rfloor).$$

Proof. Both of our proofs for (6) applies. The first proof runs as follows. Let

$$Z(n) := \sum_{k \geq 2} g(\lfloor n/k \rfloor) = -g(n) + \sum_{k \geq 2} \left(\left\lfloor \frac{n}{k} \right\rfloor - \left\lfloor \frac{n}{k+1} \right\rfloor \right) g(k).$$

Assume $1 \leq j \leq \lfloor n/2 \rfloor$. Then

$$Z(n) - Z(j) - Z(n-j) = -g(n) + g(j) + g(n-j) + \sum_{k \geq 2} (\eta_{n,j}(k) - \eta_{n,j}(k+1)) g(k),$$

where $\eta_{n,j}(k) := \lfloor n/k \rfloor - \lfloor j/k \rfloor - \lfloor (n-j)/k \rfloor$ assumes either 0 or 1. Since $\eta_{n,j}(k) = 1$ for $n-j < k \leq n$, we have

$$\begin{aligned} Z(n) - Z(j) - Z(n-j) &\geq -g(n) + g(j) + g(n-j) + \sum_{n-j < k \leq n} \eta_{n,j}(k) (g(k) - g(k-1)) \\ &= g(j), \end{aligned}$$

proving, by induction and superadditivity, that $F(n) \leq Z(n)$. The second proof is omitted. \blacksquare

Corollary 5. *For nondecreasing $g(n)$ and $n \geq 2$*

$$\psi[g^*](n) \leq F(n) \leq 2\psi[g^*](n) - g(\lfloor n/2 \rfloor) + g(\lfloor n/3 \rfloor);$$

and

$$\phi[g^\dagger](n) \leq F(n) \leq 2\phi[g^\dagger](n) + g(\lfloor n/2 \rfloor).$$

Wang [55] proved that $F(n) \leq 2\psi[g^*](n)$.

By (9) and (10), we have that if $g(n)$ is nondecreasing, then

$$\sum_{1 \leq k \leq L_n} 2^{k-1} g\left(\left\lfloor \frac{n}{2^k} \right\rfloor\right) \leq \psi[g^*](n) \leq \sum_{1 \leq k \leq L_n} 2^{k-1} g\left(\left\lceil \frac{n}{2^k} \right\rceil\right);$$

and

$$\sum_{1 \leq k \leq L_n} \left\lfloor \frac{n}{2^k} \right\rfloor g(2^{k-1}) \leq \phi[g^\dagger](n) \leq g^\dagger(n) + \sum_{1 \leq k \leq L_n} \left\lceil \frac{n}{2^k} \right\rceil g(2^{k-1}).$$

Proposition 6. *Let*

$$\lambda(k) := \max_{2^k \leq m < 2^{k+1}} g(m)/m \quad (k \geq 0).$$

If $\lambda(k) = 0$ for $0 \leq k < k_0$ and is nonincreasing for $k \geq k_0$, where $k_0 \geq 1$, then for $n \geq 2^{k_0}$

$$F(n) \leq \frac{n}{2} \sum_{0 \leq k < L_n} \lambda(k) - 2^{L_n} \lambda(L_n) + n \lambda(k_0).$$

This differs from Proposition 3 where $\lambda(k)$ is *nondecreasing!* The proof is similar.

4.3 Asymptotic approximations

Theorem 4. If $g(n)/\max\{g(\lfloor n/2 \rfloor), g(\lceil n/2 \rceil)\} \rightarrow \infty$, then

$$F(n) \sim \psi[g^*](n) \sim g(\lfloor n/2 \rfloor).$$

Proposition 7. If $g(n) \sim v(n)$, where $v(n)$ is nondecreasing and convex for $n \geq 2$, then ($v(1) := 0$)

$$F(n) \sim \psi[v^*](n),$$

where $v^*(n) := v(\lfloor n/2 \rfloor)$; if $v(n)$ is nondecreasing, concave for $n \geq 2$ and the series $\sum_k v(2^k)/2^k$ diverges, then ($v(1) := 0$)

$$F(n) \sim \phi[v^\dagger](n) \sim \frac{n}{2} \sum_{1 \leq k < L_n} \frac{v(2^k)}{2^k},$$

where $v^\dagger(n) := v(\min\{2^{\rho(n)}, n - 2^{\rho(n)}\})$.

Theorem 5. If $g(n) \sim v(n) := n^\alpha \ell(n)$, where $\ell(n)$ is slowly varying, then for $\alpha > 1$

$$F(n) \sim \psi[v^*](n) \sim \frac{n^\alpha \ell(n)}{2(1 - 2^{1-\alpha})},$$

and for $\alpha = 1$

$$F(n) \sim \psi[v^*](n) \sim \phi[v^\dagger](n) \sim \frac{n}{2} \sum_{1 \leq k < L_n} \ell(2^k), \quad (26)$$

when $\sum_{1 \leq k < L_n} \ell(2^k) \rightarrow \infty$.

We sketch the proof of (26), which proceeds along the same line of the proof of Theorem 3. For $0 < \varepsilon < 1$, choose $n_0 > 0$ so large that

$$\frac{\ell(\lfloor bn \rfloor)}{\ell(n)} \geq 1 - \varepsilon \quad (n \geq n_0),$$

for $1/2 \leq b \leq 1$ and that

$$\frac{n\ell(n)}{g(n)} \geq 1 - \varepsilon \quad (n \geq n_0).$$

Instead of $C(n)$, we now take

$$D(n) = \frac{(1/2 + 2\varepsilon)n \sum_{1 \leq k < L_n} \ell_0(2^k) - 2^{L_n} \ell_0(2^{L_n}) + n\ell(n_0)}{(1 - \varepsilon)^2},$$

where

$$\ell_0(n) = \begin{cases} \ell(n_0), & \text{if } n \leq n_0; \\ \ell(n), & \text{if } n > n_0. \end{cases}$$

Then in a similar manner we can prove that

$$D(n) \geq D(j) + D(n - j) + g(j),$$

for $1 \leq j \leq \lfloor n/2 \rfloor$ and $n \geq 2$. From this inequality and the same liminf and limsup arguments as the proof of Theorem 3, we obtain (26).

5 Optimal power-of-two rule

While the usual half-half rule is almost the synonym of divide-and-conquer in diverse problems, there are many problems for which the balanced power-of-two rule ($n \mapsto (2^{\lceil \log_2 2n/3 \rceil}, n - 2^{\lceil \log_2 2n/3 \rceil})$) are definitely better, notably problems related to Huffman coding; see for instance [7, 8, 12, 18, 25, 22, 30]. We prove in this section that *if one has to divide one of the two subproblems in size that is a power-of-two, then the best strategy is the balanced power-of-two rule, provided that the “merge cost” is nondecreasing.*

In Chen et al. [8], it was proved (implicitly) that among the class of sequences defined by

$$T(n; \theta) = \min_{1 \leq j < n} \left\{ T(2^{\lceil \log_2 \theta n \rceil}; \theta) + T(n - 2^{\lceil \log_2 \theta n \rceil}; \theta) \right\} + g(n) \quad (n \geq 2),$$

with $T(1; \theta) := 0$, where $1/2 \leq \theta < 1$, the sequence $T(n; 2/3)$ satisfies $T(n; 2/3) \leq T(n; \theta)$ when $g(n)$ is nondecreasing. This proves the optimality of the balanced power-of-two rule in this class of divide-and-conquer algorithms when the “merge cost” is nondecreasing.

Theorem 6. *Define the sequence*

$$h(n) = \min_{1 \leq 2^j < n} \{ h(2^j) + h(n - 2^j) \} + g(n) \quad (n \geq 2),$$

with $h(1) := 0$. If $g(n)$ is nondecreasing, then $h(n) = \phi(n) = T(n; 2/3)$, namely, the balanced power-of-two rule is the best power-of-two rule.

Proof. We need only to prove that

$$\phi(2^j) + \phi(n - 2^j) \geq \phi(n) - g(n),$$

for $1 \leq 2^j < n$. For simplicity, write $k = \rho(n)$. Thus $3 \cdot 2^{k-1} \leq n < 3 \cdot 2^k$. We use induction.

First, if $j = k + 1$, then $2^j \geq n - 2^{j-1}$. We have, by induction hypothesis,

$$\begin{aligned} \phi(2^{k+1}) + \phi(n - 2^{k+1}) &= \phi(2^k) + \phi(2^k) + g(2^{k+1}) + \phi(n - 2^{k+1}) \\ &= \phi(2^k) + \left(\phi(2^k) + \phi(n - 2^{k+1}) \right) + g(2^{k+1}) \\ &\geq \phi(2^k) + \left(\phi(n - 2^k) - g(n - 2^k) \right) + g(2^{k+1}) \\ &\geq \phi(2^k) + \phi(n - 2^k). \end{aligned}$$

Second, if $j \leq k - 1$ and $\rho(n - 2^j) = k$, then, similarly as above,

$$\begin{aligned} \phi(2^j) + \phi(n - 2^j) &= \phi(2^j) + \phi(2^k) + \phi(n - 2^j - 2^k) + g(n - 2^j) \\ &= \left(\phi(2^j) + \phi(n - 2^j - 2^k) \right) + \phi(2^k) + g(n - 2^j) \\ &\geq \phi(n - 2^k) - g(n - 2^k) + \phi(2^k) + g(n - 2^j) \\ &\geq \phi(2^k) + \phi(n - 2^k), \end{aligned}$$

since $g(n - 2^j) \geq g(n - 2^k)$.

Finally, if $j \leq k - 1$ and $\rho(n - 2^j) = k - 1$, then $\rho(n - 2^{k-1}) = k - 1$, and we have

$$\begin{aligned} \phi(2^j) + \phi(n - 2^j) &= \phi(2^j) + \phi(2^{k-1}) + \phi(n - 2^j - 2^{k-1}) + g(n - 2^j) \\ &= \left(\phi(2^j) + \phi(n - 2^j - 2^{k-1}) \right) + \phi(2^{k-1}) + g(n - 2^j) \\ &\geq \phi(n - 2^{k-1}) - g(n - 2^{k-1}) + \phi(2^{k-1}) + g(n - 2^j) \\ &= \phi(2^{k-1}) + \phi(n - 2^k) + \phi(2^{k-1}) + g(n - 2^j) \\ &= \phi(2^{k-1}) + \phi(2^{k-1}) + g(2^k) - g(2^k) + g(n - 2^j) + \phi(n - 2^k) \\ &\geq \phi(2^k) + \phi(n - 2^k), \end{aligned}$$

since $g(n - 2^j) \geq g(2^k)$. The proof is complete. \blacksquare

In a parallel way, we also have the following result.

Theorem 7. *Define*

$$H(n) = \max_{1 \leq 2^j < n} \{H(2^j) + H(n - 2^j) + \min\{g(2^j), g(n - 2^j)\}\} \quad (n \geq 2),$$

with $H(1) := 0$. If $g(n)$ is nondecreasing, then $H(n)$ attains the maximum value with the choice $j = \rho(n)$.

6 Applications

Finding min and max. Consider first the problem of finding the minimum and the maximum in a sequence of n numbers. The least number of comparisons $W(n)$ in the worst case (among all algorithms using only comparisons) is known to be $W(n) = \lceil 3n/2 \rceil - 2$ for $n \geq 1$; see Pohl [48]. This problem is used as one of the standard divide-and-conquer examples in most textbooks (see Aho et al. [1]). The associated recurrence for the cost (of the half-half algorithm) is given by (2) with $g(1) = 0$, $g(2) = 1$ and $g(n) = 2$ for $n \geq 3$. The solution satisfies, by (7),

$$\psi(n) = \begin{cases} 2n - 2^{L_n-1} - 2, & \text{if } \lfloor n/2^{L_n-1} \rfloor = 2; \\ n + 2^{L_n} - 2, & \text{if } \lfloor n/2^{L_n-1} \rfloor = 3, \end{cases}$$

which can be written as $\psi(n) = P(\log_2 n)n - 2$, where $P(t) = P(\{t\})$ is a continuous periodic function of period 1 defined by

$$P(t) = \begin{cases} 2 - 2^{-1-t}, & \text{if } 0 \leq t < \log_2 3 - 1; \\ 1 + 2^{-t}, & \text{if } \log_2 3 - 1 \leq t < 1. \end{cases}$$

Note that the maximum value of $P(t)$ is equal to $5/3$ and that $\psi(n) = \lceil 3n/2 \rceil - 2$ for $n = 2^k - 1, 2^k, 2^k + 1$, and these n 's only! The algorithm (using the half-half rule) is thus not optimal for other values of n . On the other hand, since $g(n)$ is obviously concave for $n \geq 3$, we have, by (8),

$$f(n) = \phi(n) = \lceil 3n/2 \rceil - 2.$$

The balanced power-of-two rule is optimal for this problem for all $n \geq 1$. Indeed the minimum value of $f(n)$ is achieved for all j such that $(j \bmod 2) + (n - j \bmod 2) \leq (n \bmod 2)$. In other words, if n is odd, then any $1 \leq j < n$ is optimal; if n is even, then j must be even. This means that the reason why the half-half divide-and-conquer rule is not optimal is that it divides the problem of even size into two subproblems of odd sizes.

Binomial group testing. The recurrence (1) with $g(n) = 1 - q^n$, $0 \leq q < 1$, appeared in the analysis of a modified binomial group testing (see Hwang et al. [30] and O'Geran et al. [43]). From (8), it follows that

$$\frac{f(n)}{n} \sim 1 - \sum_{k \geq 1} 2^{-k} q^{2^k}.$$

Also tight bounds are provided by (10).

Linear function: mergesort, in-situ permutations, etc. The most frequently encountered case is when $g(n) = n$ (or $g(n) = n - 1$). In this case both $f(n)$ and $F(n)$ were extensively studied. The solution of $f(n)$ is given by (with $g(n) = n$)

$$f(n) = nL_n + 2n - 2^{L_n+1} \quad (n \geq 1).$$

This function has the property that $f(n) = f(j) + f(n - j) + g(n)$ for

$$\min\{2^{\rho(n)}, n - 2^{\rho(n)}\} \leq j \leq \max\{2^{\rho(n)}, n - 2^{\rho(n)}\},$$

which can be easily proved by using the expression $f(n) = \sum_{2 \leq k \leq n} \lceil \log_2 n \rceil + n - 1$; see Wong [56], Batty and Rogers [3]. From an algorithmic point of view, this means that the divide-and-conquer in each step can be more flexible if the “merge cost” is linear.

Some concrete problems involving essentially $f(n)$ are listed below.

- Optimal search in a sorted array: see Wong [56], Morris [41], Carlitz [6], Gal [17], Greene and Knuth [20];
- Best-case cost of quicksort and worst-case cost of mergesort: see Flajolet and Golin [13], Golin and Sedgewick [19], Chen et al. [8].

The function $F(n)$ satisfies (with $g(n) = n$)

$$\begin{aligned} F(n) &= F(\lfloor n/2 \rfloor) + F(\lceil n/2 \rceil) + \lfloor n/2 \rfloor \\ &= F(2^{\rho(n)}) + F(n - 2^{\rho(n)}) + \min\{2^{\rho(n)}, n - 2^{\rho(n)}\} \\ &= F(2^{\lceil \log_2 n/2 \rceil}) + F(n - 2^{\lceil \log_2 n/2 \rceil}) + n - 2^{\lceil \log_2 n/2 \rceil} \\ &= \sum_{1 \leq k < n} \nu(k), \end{aligned}$$

where $\nu(n)$ denotes the number of 1’s in the dyadic representation of n . Such an $F(n)$ appeared in many different problems; a brief list is as follows.

- Algorithms for in-situ permutations: see Knuth [36] or Knuth and Greene [20];
- A set merging process: see McIlroy [40];
- Connecting edges in n -cubes: see Hart [24];
- Merging networks: see Hong and Sedgewick [29];
- Maximum external left length of binary trees: see Li [38];
- Boustradophedonic random generation of labelled combinatorial structures: see Flajolet et al. [15];
- Best-case cost of mergesorts: see Flajolet and Golin [13], Panny and Prodinger [44], Chen et al. [8].

It is well known that $F(n)/n = \frac{1}{2} \log_2 n + Q(\log_2 n)$, where $Q(u)$ is continuous, periodic, and nowhere differentiable; see Delange [10] and Flajolet et al. [14]. A complete characterization of indices for which $F(n) = F(j) + F(n - j) + j$ is given by McIlroy [40]. This characterization may have further algorithmic applications.

On the other hand, take $g(n) = \lfloor n/2 \rfloor$. Although the indices j at which $f(n)$ attain the minimum value have no obvious pattern, our result readily gives $f(n) \sim \frac{1}{2}n \log_2 n$.

Cases when $g = g(j, n - j)$ are dependent on both j and n . Our results can also be applied to give effective bounds when g depends not only on n but also on j . The general pattern of the recurrence is

$$f(n) = \min_{1 \leq j < n} \{f(j) + f(n - j) + g(j, n - j)\} \quad (n \geq 2),$$

with $f(1)$ given. For example, $g(x, y) = ax + by$ and $g(x, y) = ax(x + y) + b(x + y)$ appeared in dichotomous search problems (see Murakami [42], Batty and Rogers [3], Hassin and Rotovely, [26], Hinderer [27]) and in matching heuristics (see Jünger and Pulleyblank [33]), and

$$\begin{aligned} g(x, y) &= x + y - x/(y + 1) - y/(x + 1), \\ g(x, y) &= \frac{x(2x + y)}{(y + 1)(y + 2)} + \frac{y(2y + x)}{(x + 1)(x + 2)} - \left(\frac{x}{y + 1} + \frac{y}{x + 1} \right)^2, \end{aligned} \quad (27)$$

appeared in mergesort (see [8]). Since the minimum and maximum values of $g(j, n - j)$ for $1 \leq j < n$ depend only on n (often in a simple way), we can apply previous results to obtain useful bounds on $f(n)$. For example, if $g(j, n - j) = aj + bn$, then

$$(b + o(1))n \log_2 n \leq f(n) \leq (a + b + o(1))n \log_2 n,$$

provided that $a, b > 0$; the other cases when one or both of them is negative can be considered similarly. For a unified treatment of search problems, see Hinderer and Stieglitz [28]. When g is given by (27), one can obtain more precise asymptotic approximation for $f(n)$

$$f(n) = cn + O(\log_2 n);$$

see [8] for more details and the expression of c .

7 Extensions

Many problems arise from our investigation; we mention some of these as follows.

Range of optimal indices. Based on simulations, we conjecture that if $g(n)$ is nondecreasing, then the largest index $j^* = j^*(n) \leq \lfloor n/2 \rfloor$ for which

$$f(n) = f(j^*) + f(n - j^*) + g(n),$$

satisfies

$$\liminf_{n \rightarrow \infty} \frac{j^*(n)}{n} \geq 1/3.$$

The result, if correct, would reduce considerably the computational complexity of $f(n)$.

Deeper connections between $f(n)$ and $F(n)$? Why there are so many similarities between the two sequences $f(n)$ and $F(n)$? Is there a more unified framework based on, say the theory of dynamic programming, to provide more “structural interpretations?” Our results imply that for nondecreasing $g(n)$,

$$F(n) + g(n) \leq \sum_{j \geq 1} g(\lfloor n/j \rfloor) \leq \sum_{j \geq 1} g(\lceil n/j \rceil) \leq f(n);$$

and that when $g(n) \sim n^\alpha \ell(n)$, $\alpha \geq 1$,

$$F(n) \sim \frac{f(n)}{2}.$$

[When $\alpha = 1$, we need $\sum_{2^j \leq n} \ell(2^j) \rightarrow \infty$.] How universal is the last estimate? and is there an intuitive interpretation of the constant $1/2$? [The above estimate holds also for $g(n) = \lceil \log_2 n \rceil$.]

Limit constants of subadditive sequences. As mentioned in Introduction, the determination of the limit $f(n)/n$, when $f(n)$ is linear, is in general a very challenging problem. A problem closer to the main theme of this paper is the constant $\lim_{n \rightarrow \infty} f(n)/n$ when $g(n) \sim n^\alpha \ell(n)$, where $0 < \alpha < 1$. Our estimates (5), (9) and (10) provide bounds for the limit constant when $g(n)$ is nondecreasing, but they are in general not tight enough.

Error terms. What is the magnitude of growth of the second-order term in the asymptotic expansion of $f(n)$ if more information of $g(n)$ is known? Our tools are in most cases too weak for this problem.

Half-half or balanced power-of-two? A rough conclusion of our results would be to prefer half-half rule when the “merge cost” $g(n)$ is at least linear, and prefer balanced power-of-two rule otherwise. Asymptotically, this would be sufficient for most practical uses. The situations are, however, more complicated in general, notably the case of linear $f(n)$. For example, when $g(n) = \lfloor \log_2 n \rfloor$, we have $\psi[g](n) \leq \phi[g](n)$ for $n \geq 1$. It would be interesting to find practical conditions (more general than convexity and concavity) under which one can prefer the use of balanced power-of-two rule to half-half rule and vice versa.

Lower bounds for monotone $g(n)/n$. We derived precise lower bounds for $f(n)$ when $g(n)/n$ is nondecreasing and for $F(n)$ when $g(n)/n$ is nonincreasing. What are the corresponding bounds for the remaining cases?

Limit constants of superadditive sequences. If $g(n)$ is nondecreasing, is it true that $F(n) \sim cn$ iff $\sum_{n \geq 1} g(n)/n^2 < \infty$? This requires an extension of Hammersley’s subadditive result (see [21]) to superadditive sequences.

8 Conclusions

From an algorithmic point of view, our results indicate that the usual half-half divide-and-conquer rule is good (asymptotically optimal under suitable conditions) if its cost $\psi(n)$ is larger than linear, namely, $\psi(n)/n \rightarrow \infty$. The case when its cost is linear is more delicate and reveals the intricacy of general subadditive problems, although the balanced power-of-two rule outperforms other rules in certain cases.

The following dependence of the asymptotic behavior of $f(n)$ on $g(n)$ is obvious from our results

$$f(n) \propto \begin{cases} \{g(1), g(2), g(3), g(4), \dots, g(n)\}, & \text{if } f(n) \text{ is linear;} \\ \{g(1), g(2), g(4), g(8), \dots, g(2^{L_n})\}, & \text{if } \log f(n)/\log n \sim 1, f(n)/n \rightarrow \infty; \\ g(n), & \text{if } \lim_{n \rightarrow \infty} \log f(n)/\log n > 1. \end{cases}$$

But it should be pointed out that in the intermediate (second) case, the dependence of $f(n)$ on $\{g(2^k)\}$ may be misleading, because this result reflects mainly our assumption of slow variation rather than the inherent structure of the underlying cost function. Thus one cannot conclude from it that in order to improve the efficiency of the original algorithm it suffices to improve subproblems of sizes a power of two.

Acknowledgements

We thank Philippe Flajolet, Edward Reingold and the two referees for useful suggestions.

References

- [1] A. V. Aho, J. E. Hopcroft and J. D. Ullman, “The Design and Analysis of Computer Algorithms,” Second Printing, Addison-Wesley, Reading, MA, 1975.
- [2] L. Alonso, E. M. Reingold and R. Schott, Multidimensional divide-and-conquer maximin recurrences, *SIAM Journal on Discrete Mathematics*, **8** (1995), 428–447.
- [3] C. J. K. Batty and D. G. Rogers, Some maximal solutions of the generalized subadditive inequality, *SIAM Journal on Algebraic and Discrete Methods*, **3** (1982), 369–378.
- [4] N. H. Bingham, C. M. Goldie and J. L. Teugels, “Regular Variation,” Cambridge University Press, Cambridge-New York, 1989.
- [5] N. G. de Bruijn and P. Erdős, Some linear and some quadratic recursion formulas. II, *Indagationes Mathematicae*, **14** (1952), 152–163.
- [6] L. Carlitz, A sorting function, *Duke Mathematical Journal*, **38** (1971), 561–568.
- [7] C.-S. Chang and J. A. Thomas, Huffman algebras for independent random variables, *Discrete Events Dynamic Systems: Theory and Applications*, **4** (1994), 23–40.
- [8] W.-M. Chen, H.-K. Hwang and G.-H. Chen, The cost distribution of queue-mergesort, optimal mergesorts, and power-of-2 rules, *Journal of Algorithms*, **30** (1999), 423–448.
- [9] T. H. Cormen, C. E. Leiserson and R. L. Rivest, “Introduction to Algorithms,” MIT Press, Cambridge, MA, 1990.
- [10] H. Delange, Sur la fonction sommatoire de la fonction “somme des chiffres”, *Enseignement Mathématique*, **21** (1975), 31–47.
- [11] L. Devroye, Moment inequalities for random variables in computational geometry, *Computing*, **30** (1983), 111–119.
- [12] J. A. Fill, On the distribution of binary search trees under the random permutation model, *Random Structures and Algorithms*, **8** (1996), 1–25.
- [13] P. Flajolet and M. Golin, Mellin transforms and asymptotics: the mergesort recurrence, *Acta Informatica*, **31** (1994), 673–696.
- [14] P. Flajolet, P. Grabner, P. Kirschenhofer, H. Prodinger and R. F. Tichy, Mellin transforms and asymptotics: digital sums, *Theoretical Computer Science*, **123** (1994), pp. 291–314.
- [15] P. Flajolet, P. Zimmerman, and B. Van Cutsem, A calculus for the random generation of labelled combinatorial structures, *Theoretical Computer Science*, **132** (1994), 1–35.
- [16] M. L. Fredman and D. E. Knuth, Recurrence relations based on minimization, *Journal of Mathematical Analysis and Applications*, **48** (1974), 534–559.
- [17] S. Gal, A discrete search game, *SIAM Journal on Applied Mathematics*, **27** (1974), 641–648.
- [18] C. R. Glassey and R. M. Karp, On the optimality of Huffman trees, *SIAM Journal on Applied Mathematics*, **31** (1976), 368–378.

- [19] M. J. Golin and R. Sedgewick, Queue-mergesort, *Information Processing Letters*, **48** (1993), 253–259.
- [20] D. H. Greene and D. E. Knuth, “Mathematics for the Analysis of Algorithms,” Third Edition, Birkhäuser, Boston, 1990.
- [21] J. M. Hammersley, Generalization of the fundamental theorem on sub-additive functions, *Proceedings of the Cambridge Philosophical Society*, **58** (1962), 235–238.
- [22] J. M. Hammersley and G. R. Grimmett, Maximal solutions of the generalized subadditive inequality, in “Stochastic Geometry,” edited by E. F. Harding and D. G. Kendall, John Wiley and Sons, London, pp. 270–284, 1974.
- [23] E. F. Harding, The probabilities of rooted tree-shapes generated by random bifurcation, *Advances in Applied Probability*, **3** (1971), 44–77.
- [24] S. Hart, A note on the edges of the n -cube, *Discrete Mathematics*, **14** (1976), 157–163.
- [25] R. Hassin and M. Henig, Dichotomous search for random objects on an interval, *Mathematics of Operations Research*, **9** (1984), 301–308.
- [26] R. Hassin and R. Rotovely, Asymptotic analysis of dichotomous search with search and travel costs, *European Journal of Operations Research*, **58** (1992), 78–89.
- [27] K. Hinderer, On dichotomous search with direction-dependent costs for a uniformly hidden object, *Optimization*, **21** (1990), 215–229.
- [28] K. Hinderer and M. Stieglitz, On polychotomous search problem, *European Journal of Operations Research*, **73** (1994), 279–294.
- [29] Z. Hong and R. Sedgewick, Notes on merging networks, in ACM Symposium on Theory of Computing (STOC’82), pp. 296–302, ACM Press, 1982.
- [30] F. K. Hwang, C. G. Pfeifer and P. Enis, An optimal hierarchical procedure for a modified binomial group-testing problem, *Journal of the American Statistical Association*, **76** (1981), 947–949.
- [31] H.-K. Hwang, Limit theorems for mergesort, *Random Structures and Algorithms*, **8** (1996), 319–336.
- [32] H.-K. Hwang and J.-M. Steyaert, On the number of heaps and the cost of heap construction, *Colloquium on Mathematics and Computer Science*, Versailles, September, 2002, accepted; available via the link <http://algo.stat.sinica.edu.tw>.
- [33] M. Jünger and W. Pulleyblank, New primal and dual matching heuristics, *Algorithmica*, **13** (1995), 357–380.
- [34] S. Kapoor and E. M. Reingold, Recurrence relations based on minimization and maximization, *Journal of Mathematical Analysis and Applications*, **109** (1985), 591–604.
- [35] S. Kapoor and E. M. Reingold, Optimum lopsided binary trees, *Journal of the ACM*, **36** (1989), 573–590.

- [36] D. E. Knuth, Mathematical analysis of algorithms, in “Information Processing’71,” (Proc. IFIP Congress, Ljubljana, 1971), Vol. 1: Foundations and Systems, pp. 19–27, North-Holland, Amsterdam, 1972.
- [37] D. E. Knuth, “The Art of Computer Programming, Volume III—Sorting and Searching,” Second Edition, Addison-Wesley, Reading, MA, 1998.
- [38] S. Y. R. Li, Binary trees and uniform distribution of traffic cutback, *Journal of Computer and System Sciences*, **32** (1986), 1–14.
- [39] Z. Li and E. M. Reingold, Solution of a divide-and-conquer maximin recurrence, *SIAM Journal on Computing*, **18** (1989), 1188–1200.
- [40] M. D. McIlroy, The number of 1’s in binary integers: bounds and extremal properties, *SIAM Journal on Computing*, **3** (1974), 255–261.
- [41] R. Morris, Some theorems on sorting, *SIAM Journal on Applied Mathematics*, **17** (1969), 1–6.
- [42] S. Murakami, A dichotomous search, *Journal of the Operations Research Society of Japan*, **14** (1971), pp. 127–142.
- [43] J. H. O’Geran, H. P. Wynn and A. A. Zhiglyavsky, Search, *Acta Applicandae Mathematicae*, **25** (1991), 241–276.
- [44] W. Panny and H. Prodinger, Bottom-up mergesort—a detailed analysis, *Algorithmica*, **14** (1995), pp. 340–354.
- [45] M. J. Pelling and D. G. Rogers, A combinatorial problem in the design of electrical circuits, *Southeast Asian Bulletin of Mathematics*, **4** (1980), 29–33.
- [46] M. J. Pelling and D. G. Rogers, A problem in the design of electrical circuits, a generalized subadditive inequality and the recurrence relation $j(n, m) = j(\lfloor n/2 \rfloor, m) + j(\lfloor (n+1)/2 \rfloor, m) + j(n, m-1)$, *Lecture Notes in Mathematics*, 622, pp. 153–169, Springer, Berlin-New York, 1977.
- [47] N. Pippenger, An elementary approach to some analytic asymptotics, *SIAM Journal on Mathematical Analysis*, **24** (1993), 1361–1377.
- [48] I. Pohl, A sorting problem and its complexity, *Communications of the ACM*, **15** (1972), 462–464.
- [49] E. M. Reingold and R. E. Tarjan, On a greedy heuristic for complete matching, *SIAM Journal on Computing*, **10** (1981), 676–681.
- [50] E. M. Reingold and J. S. Tilford, Tidier Drawings of Trees, *IEEE Transactions on Software Engineering*, **7** (1981), 223–228.
- [51] S. Roura, Improved master theorems for divide-and-conquer recurrences, *Journal of the ACM*, **48** (2001), 170–205.
- [52] M. Snir, Exact balancing is not always good, *Information Processing Letters*, **22** (1986), 97–102.
- [53] J. M. Steele, “Probability Theory and Combinatorial Optimization,” CBMS-NSF Regional Conference Series in Applied Mathematics, 69, SIAM, PA, 1997.
- [54] T. R. Walsh, How evenly should one divide to conquer quickly? *Information Processing Letters*, **19** (1984), pp. 203–208.

- [55] B.-F. Wang, Tighter bounds on the solution of a divide-and-conquer maximin recurrence, *Journal of Algorithms*, **23** (1997), 329–344.
- [56] E. Wong, A linear search problem, *SIAM Review*, **6** (1964), 168–174.
- [57] J. E. Yukich, “Probability Theory of Classical Euclidean Optimization Problems,” Lecture Notes in Mathematics, 1675, Springer-Verlag, Berlin, 1998.