1989

(Probably) Optimal Solution to Some Problems Not Only on Graphs

Wojciech Szpankowski
Purdue University, spa@cs.purdue.edu

Report Number:
89-872

http://docs.lib.purdue.edu/cstech/741

This document has been made available through Purdue e-Pubs, a service of the Purdue University Libraries. Please contact epubs@purdue.edu for additional information.
(PROBABLY) OPTIMAL SOLUTION TO SOME PROBLEMS NOT ONLY ON GRAPHS

Wojciech Szpankowski

CSD-TR-872
March 1989
A mathematical abstraction of the model studied in this paper can be formulated as follows: find the optimal value of $Z_{\text{max}} = \max \{ \Xi w_i(\alpha) \}$ (respectively $Z_{\text{min}}$), where $\Xi$ is an operator (e.g., $\Xi = \min$ for bottleneck problems, $\Xi = \Sigma$ for separable objective functions, and so on), $n$ is an integer, $B_n$ is the set of all feasible solutions, $S_n(\alpha)$ is the set of objects belonging to the $\alpha$-th feasible solution, and $w_i(\alpha)$ is the weight assigned to the $i$-th object.

Our interest lies in finding an asymptotic solution to this (optimization) problem in a probabilistic framework, that is, under assumption that the weights are random variables drawn independently and having identical distribution function. Such a general formulation of the problem, and the proposed methodology through a novel application of order statistics, allow us to study in a uniform manner a large class of problems investigated vigorously in computer science over the last two decades. Among others we mention here: the assignment problem or perfect matching in bipartite graphs, the traveling salesman problem, the minimum spanning tree, the minimum weighted $k$-clique, the bottleneck assignment and traveling salesman problems, location problems on graphs, and finally such unrelated problems as the height of digital trees and the maximum queue length in a queue. Finally, we discuss some consequences of our investigations. In particular, we present some sufficient conditions under which the asymptotic performance of a greedy algorithm and the optimal one is the same in a probabilistic sense.

1. MOTIVATIONS

Most algorithm designs are finalized to the optimization of the asymptotic worst-case performance [AHU74]. Insightful, elegant and generally useful constructions have been set up in this endeavor. Along these lines, however, the design of an algorithm is usually targeted at coping efficiently with unrealistic, even pathological inputs and the possibility is neglected that a simpler algorithm that works fast "on average" might perform just as well, or even better in practice. This alternative solution called also a probabilistic approach became an important issue a decade ago when it became clear that the prospects for showing the existence of
polynomial time algorithms for NP-hard problems, were very dim. (Students of operations research, as opposed to those studying computer science, are convinced in probabilistic heuristics, since in the very early years of their study, they became familiar with the simplex method of the linear programming which has exponential worst case behavior, but acceptable average case, i.e., practical complexity.) This fact, and apparently high success rate of heuristic approaches to solving certain difficult problems, led Richard Karp [KAR76] to undertake a more serious investigation of probabilistic approximation algorithms. But one must realize that there are problems which are also "hard" "on average" [LEV86]. The last few years witnessed an increasing interest in the probabilistic approach to the NP-hard problems [BOL85, GAJ79, KAR76, KAR77, KNU773, LOU87, LUK81, PAL85, PAP81, WEI80]. Setting aside the realm of approximation algorithms for NP-hard problems, then achieving a good average case performance is rarely the primary objective of algorithm design. This may seem surprising, since algorithms that achieve this objective are also likely to be practically efficient. In assessing algorithmic performance, the average case analysis might be often an interesting and more fruitful approach.

Enlightened by these motivations, we undertake in this paper a study of a class of problems in a probabilistic framework. A general mathematical model of these problems can be formulated as follows. For every integer \( n \), find the optimal value of \( Z_{\max} = \max_{a \in B_n} \{ \sum_{i \in S_n(a)} w_i(a) \} \)
(\( Z_{\min} \) respectively), where \( \Xi \) is an operator (e.g., \( \Xi = \Sigma \) or \( \Xi = \min \), etc.), \( B_n \) is the set of all feasible solutions, \( S_n(a) \) is the set of all objects belonging to the \( a \)-th feasible solution, and \( w_i(a) \) is the weight assigned to the \( i \)-th object. For example, in the traveling salesman problem [BOR62, GAJ79, KAR76, KAR77, ANV79, WEI80, LLK85], the operator \( \Xi \) becomes a sum \( \Sigma \) operator, \( B_n \) represents the set of all Hamiltonian paths in a graph with \( n \) vertices, \( S_n(a) \) is the set of edges which fall into the \( a \)-th Hamiltonian path, and \( w_i(a) \) is the length of the \( i \)-th edge; for the bottleneck traveling salesman problem [GAJ78, WEI80, SZP88c] the operator \( \Xi \)
becomes "min" operator. Some other examples of our general optimization problem include the assignment problem [BOR62, FHR87, WAL79, WEI80, LLK85], the minimum spanning tree [BER73, BOL85, G0J83, KNU73], the minimum weighted k-clique problem [LUK81, BOL85, BER73], the bottleneck and capacity assignment problems [GaG78, WEI80, SZP88c], geometric location problems [PAP81], and some others not directly related to optimization such as the height of digital trees [ApS88, FlO82, KNU73, SZP86], the maximum queue length [IGL72, SZP88b] and hashing with lazy deletion [MSW87, SZP88b]. In our probabilistic framework, we assume that the weights \( w_i(\alpha) \) are random variables drawn independently with a common distribution function \( F(\cdot) \). Our interest lies in finding an asymptotic solution to \( Z_{\text{max}} \) and \( Z_{\text{min}} \) in probability, in mean and almost surely (with probability one) sense for a large class of distribution functions \( F(\cdot) \), and apply these findings to design heuristic algorithms.

It is our understanding that designing of algorithms "on average" and the average case analysis have experienced some setbacks in the past due to two reasons. The first one is related to some \textit{a priori} assumptions regarding distribution of inputs, and this conveys a flavor of arbitrariness. The second reason lies in the fact that the average case analysis is usually more intricate, and there are no widely accepted tools to analyze and design algorithms that work well for a typical input. We address both issues in this paper, and we intend to shed some light on them for the class of problems discussed above. First of all, one might argue that basing the design of algorithms on the worst case approach is just as arbitrary, in that it subtends the rather strong assumption that pathological inputs are somehow more likely than others. Our analysis \textit{does not} assume any \textit{a priori} information about distribution of inputs, and we derive our results for arbitrary distribution function of the weights. We, however, identify two types of weight distributions that lead to very tight asymptotic expansions of the optimal values \( Z_{\text{max}} \) and \( Z_{\text{min}} \). In addition, we present two different techniques for bounding the optimal value from below, and three approaches for obtaining upper bounds on the optimal objective function (see our Proposition in
Section 3). These results are consequences of our novel applications of order statistics [GAL87] (see Main Lemma). Knowing asymptotic expansions for the optimal values $Z_{\text{max}}$ and $Z_{\text{min}}$ one may design and tune up any heuristic by comparing the value $Z_{\text{appr}}$ of the approximate solution with the asymptotics of the optimal solution. If the relative error between these two solutions becomes smaller and smaller for large inputs, then the heuristic is a near optimal solution [KAR76, WEI80]. In particular, we compare performance of the optimal algorithm and a greedy heuristic, and provide sufficient conditions which assure that a greedy algorithm achieves asymptotically the same performance as the optimal one (see Theorem 1 in Section 3). Finally, for bottleneck problems (i.e., $\Xi=\min$) and for separable objective functions (i.e., $\Xi=\Sigma$) we present even stronger results that give an ultimate answer to the optimization problem keeping modeling assumptions as minimum as possible (see Theorem 2 and 3 in Section 3).

To our best knowledge, the literature is very scarce in the results of our type, although the literature on the optimization problems is very huge (cf. [LLK85]). Some of the problems discussed here have been investigated in the past [ApS88, BOR62, FHR87, KAR77, LOU87, LUK81, SZP86, WAL79, WEI80], however, the approach adopted in this paper is a little similar only to the work of Weide [WEI80], partially to Luker [LUK81], and it has something in common with the work of Frenk et al [FHR87]. Nevertheless, Weide in his work has rather concentrated on (random) graphs, while we do not. We solve also, the open problem suggested by Weide, that is, we obtain asymptotically exact solutions in the cases the author of [WEI80] provides only upper bounds. In addition, our techniques are completely different. Weide, as well as Luker [LUK81] and others, in order to obtain their estimates, need to know the solution of the problem for unweighted random graphs, which might be a serious problem by itself. Finally, Frenk et al [FHR87] have obtained some results for a class of distribution functions of inputs, but they have focused only on the linear assignment problem, while we do not.

This paper is organized as follows. In the next section, we rigorously formulate our gen-
eral problem and illustrate it by some relevant examples taken from such diverse areas as optimization on graphs, analysis of digital trees, queueing theory and hashing. Section 3 is the heart of our paper and it contains all the main results. In this section we also provide some of the derivations, but most of the cumbersome ones are delayed to Section 5. Finally, Section 4 presents solutions to problems discussed in Section 2, using our main results form Section 3.

2. PROBLEM STATEMENT

We start this section with a particular representation of a more general problem which is formulated at the end of this section. Let \( n \) be an integer (e.g., number of vertices in a graph, number of keys in a digital tree, etc.), and \( S \) a set of objects (e.g., set of vertices, keys, etc). We shall investigate the optimal values \( Z_{\text{max}} \) and \( Z_{\text{min}} \) defined as follows

\[
Z_{\text{max}} = \max_{\alpha \in B_n} \sum_{i \in S_n(\alpha)} w_i(\alpha) \quad Z_{\text{min}} = \min_{\alpha \in B_n} \sum_{i \in S_n(\alpha)} w_i(\alpha),
\]

where \( B_n \) is a set of all feasible solutions, \( S_n(\alpha) \) is a countable set of objects from \( S \) belonging to the \( \alpha \)-th feasible solution, and \( w_i(\alpha) \) is the weight assigned to the \( i \)-th object in the \( \alpha \)-th feasible solution (in addition, by \( w_{ij} \) we denote a weight assigned to a pair of objects \((i,j)\) in \( S \)). Throughout this paper, we adopt the following assumptions:

(A) The cardinality \( |B_n| \) of \( B_n \) is fixed and equal to \( m \). The cardinality \( |S_n(\alpha)| \) of the set \( S_n(\alpha) \) does not depend on \( \alpha \in B_n \), and for all \( \alpha \) it is equal to \( N \), i.e., \( |S_n(\alpha)| = N \).

(B) For all \( \alpha \in B_n \) and \( i \in S_n(\alpha) \) the weights \( w_i(\alpha) \) (i.e., the weights \( w_{ij} \)) are identically and independently distributed (i.i.d) random variables with common distribution function \( F(.) \), and the mean value \( \mu \) and the variance \( \sigma^2 \).

The assumption (B) defines a probabilistic model of our problem (2.1), and therefore, the objective functions \( Z_{\text{max}} \) and \( Z_{\text{min}} \) are random variables. We shall explore the asymptotic behaviors of all moments of \( Z_{\text{max}} \) and \( Z_{\text{min}} \) as \( n \) becomes large. In addition, we investigate
asymptotic behavior of $Z_{\text{max}}$ and $Z_{\text{min}}$ that hold either in probability or almost surely, i.e., with probability one (see [FEL71, GAL87, REN70, LUE81] for definitions). Before we plug into the analysis, we discuss some important examples of our problem.

Example 2.1. Linear assignment problem or perfect matching in bipartite graph

Given an $n \times n$ matrix $\{a_{ij}\}_{i,j=1}^{n}$, the problem is to find a permutation $\alpha: \{1, 2, \ldots, n\} \rightarrow \{1, 2, \ldots, n\}$ that maximizes or minimizes $\sum_{i=1}^{n} a_{\alpha(i)}$. In our notations, $B_n$ is a set of all permutations of $\{1, 2, \ldots, n\}$, $S_n(\alpha) = \{1, 2, \ldots, n\}$, and

$$Z_{\text{max}} = \max_{\alpha \in B_n} \sum_{i=1}^{n} a_{\alpha(i)}$$

(2.2)

Note that $|B_n| = n!$, $|S_n(\alpha)| = n$ and the weights $w_i(\alpha) = a_{\alpha(i)}$. This problem is equivalent to the perfect matching in a bipartite graph [BER73, BOL85, WAL79].

Example 2.2. Traveling salesman problem

Let $G_n$ be a graph with $n$ vertices. We assign a (random) weight $w_{ij}$ for every edge $(i,j)$, $i,j = 1, 2, \ldots, n$ belonging to the graph $G_n$. The traveling salesman problem is to find a path through all vertices with the minimum total weight. Of course, this can be formulated as our problem (2.1) with $B_n$ being the set of all Hamiltonian paths and $S_n(\alpha)$ is a set of $n - 1$ edges in the $\alpha$-th Hamiltonian path, that is, $N = |S_n(\alpha)| = n - 1$. The cardinality of $B_n$ depends on the structure of the graph, and general formula for $m = |B_n|$ can be found in [GoJ83]. For example, if $G_n$ is a complete graph, then $|B_n| = (n - 1)!$

Example 2.3. The minimum spanning tree

As in the previous example, the graph $G_n$ with $n$ vertices is given. We optimize $Z_{\text{min}}$ with $B_n$ being the set of all spanning trees, and $S_n(\alpha)$ of cardinality $N = n - 2$, being the set of edges belonging to the $\alpha$-th spanning tree. The cardinality of $B_n$ depends on the structure of $G_n$ and a
general formula on $|B_n|$ can be found in [GoJ83]. For example, in [GoJ83] we find that in a complete graph, there is $|B_n| = m = n^{n-2}$ rooted labeled trees.

Example 2.4. The minimum weighted $k$-clique

In a graph $G_n$ with $n$ vertices, we call a subgraph a $k$-clique if it is spanned over $k$ adjacent vertices [BER73, BOL85, LUE81]. In addition, it is assumed that a weight $w_{ij}$ is assigned to each edge $(i,j)$ in $G_n$, $i,j = 1, 2, \ldots, n$. The objective function $Z_{\text{min}}$ has the form of (2.1) with $B_n$ being the set of all $k$-cliques and $S_n(\alpha)$ the set of edges belonging to the $\alpha$-th $k$-clique. The cardinalities of $|B_n| = m$ and $|S_n(\alpha)| = N$, in general, depend on the structure of $G_n$, but for instance in the complete graph $G_n$ one immediately finds $m = \binom{n}{k}$ and $N = k\binom{k}{2}$. Throughout the paper, we shall use the notation $C_n^k$ for the Newton coefficient $\binom{n}{k}$ to simplify some of our formulas.

In the next examples, either the weight function is not given explicitly or/and the distribution function $F(\cdot)$ of the weight must be computed from the model description. In addition, the next two examples are not explicitly optimization problems. Nevertheless, they can be represented, as we shall see, in the form of (2.1), and more importantly they can be attacked by the same methodology as the general optimization problem (2.1).

Example 2.5. The height of digital trees (trie)

In this example we deal with a digital data structure called a trie [AHU74, KNU73], and our interest lies in computing the height of this tree. We shall show that we can pose the problem in terms of our form (2.1). Let $X_1, X_2, \ldots, X_n$ be $n$ strings of (possible) unbounded lengths formed by symbols from a binary alphabet $\Sigma = \{0, 1\}$ (generalization to a finite $V$-ary alphabet is trivial). It is assumed that symbols 0 and 1 from the alphabet occur independently and with probabilities $p$ and $q = 1-p$ respectively. It turns out (see [SZP86, ApS88]) that the
height of these trees can be easily evaluated through the so called *alignment* or *common* operator. The common (alignment) operator \( C_{ij} = \text{com}(X_i, X_j) \) is defined as the length of the longest string, that is, prefix of \( X_i \) and \( X_j \). Thus, \( C_{ij} = k \) iff \( X_i \) and \( X_j \) agree exactly on their first \( k \) positions, but differ on their \((k+1)\)-st. It is easy to show that the height \( H_n \) of a trie built from \( X_1, X_2, \ldots, X_n \) is given by

\[
H_n = \max_{1 \leq i < j \leq n} \{ C_{ij} \} + 1
\]

(2.2)

Therefore, the problem is reduced to our original formulation (2.1), if one defines \( Z_{\text{max}} = H_n \) and the weights as \( C_{ij} = \text{com}(X_i, X_j) \) with the cardinality of the set \( B_n \) equals to \( m = n(n - 1)/2 \) and \( I_{\text{S}}(\alpha) = N = 1 \). Note however, that the distribution of the weights \( C_{ij} \) is not given in an explicit form. But under independence assumptions we have postulated, it is an easy task to show that

\[
Pr(C_{ij} = k) = P^k(1-P)^{k-1} \quad k = 0, 1, \ldots.
\]

(2.3)

where \( P = p^2 + q^2 \). Hence, the distribution function for the weights \( C_{ij} \) is \( F(l) = Pr(C_{ij} \leq l) = 1 - P^l \), and this completes the description of the model in terms of our original problem. □

**Example 2.6** *Buffer occupancy problem, that is, maximum queue length*

Let us consider a dynamic data structure called queue [AHU74] with random insertions and deletions of items. Any item, which we further call a customer, may arrive at a random moment of time and stay in the queue for a random time (waiting time) until it is taken to a service for a random service time, and then deleted. If interarrival times between insertions and service times are independent random variables with general distribution functions, then in queueing theory terminology the data structure is called GI|GI|1 queue [KLE76]. The quantities of interest are queue length \( Q_k \) and waiting time \( W_k \) at the moment of the \( k \)-th arrival of a customer. In fact, the maximum size of the queue is a fundamental quantity, which is directly related to many problems of resource allocations, design of nodes in a computer network (occupancy
problem for a buffer in the node), etc. Therefore, our purpose is to study the following quantities

\[ Q_{\text{max}} = \max_{1 \leq k \leq n} Q_k \quad W_{\text{max}} = \max_{1 \leq k \leq n} W_k \]  

(2.3)
as \( n \) tends to infinity. Naturally (2.3) is another form of our general problem (2.1) with the cardinality of \( B_n \) equal to \( n \) and \( |S_n| = 1 \). The problem is interesting since the distributions of weights, that is, the queue and the waiting times seen by the arrival of the \( k \) customer, are not known for the general GI/GI queue. Nevertheless, the tail of the distributions \( F_Q(m) = \Pr \{ Q_k \leq m \} \) and \( F_W(x) = \Pr \{ W_k < x \} \) in a stationary case is possible to evaluate. Indeed, Feller [FEL71] has shown that

\[ 1 - F_W(x) = c_1 e^{-\theta}(1 + o(1)) \quad 1 - F_Q(m) = c_2 \omega^m(1 + o(1)) \]  

(2.4)
for \( x \to \infty \) and \( m \to \infty \), where \( c_1 \) and \( c_2 \) are constants, and \( \theta \) and \( \omega \) are parameters of the GI/GI queue. The latter parameters can be computed as follows. Let \( A^*(\theta) \) and \( B^*(\theta) \) stand for the Laplace-Stieltjes transforms [AbS64, KLE76] for the interarrival times and service times respectively. Then, \( \theta \) is defined as a unique solution of the following complex equation

\[ A^*(\theta)B^*(-\theta) = 1 \]  

(2.5)
and \( \omega \) is given by \( \omega = A^*(\theta) [TAK81] \). Several generalizations and applications of this problem are possible. First of all, due to a result of Takahashi [TAK81], we know that (2.4) holds also for some GI/GI queues where \( c \) stands for the number of servers in a queueing system. Moreover, Szpankowski in [SZP88b], using some results of Morrison et al [MSW87], applied this approach to M/M/\infty queue to study the occupancy problem for hashing with lazy deletion.

So far, we have restricted our attention to problems which can be represented as (2.1), that is, the objective function is a sum of weights (a separable function). In practice some other objective functions play an important role. For example, in a class of bottleneck and capacity problems [GaG78] '\( \Sigma \) ' in (2.1) is replaced by 'max' and 'min', respectively; in Example 2.5 on
digital trees, the objective function (2.2) can be represented as

\[ H_n = \max_{1 \leq i < j \leq n} \text{com}(X_i, X_j) + 1, \]

hence "common" can be interpreted as an operator. In general, our formulation (2.1) needs extension to include most of the interesting problems, namely we define

\[ Z_{\text{max}} = \max_{\alpha \in B_n} \Xi_{\in S_n(\alpha)} w_i(\alpha) \quad Z_{\text{min}} = \min_{\alpha \in B_n} \Xi_{\in S_n(\alpha)} w_i(\alpha) \quad (2.6) \]

where \( \Xi \) is an operator applied to a set \( \{w_i(\alpha), i \in S_n(\alpha)\} \), e.g., in (2.1) \( \Xi = \Sigma \). Below we present some more relevant examples.

**Example 2.7 Bottleneck and capacity assignment problems**

Let \( A = \{a_{ij}\}_{j=1}^n \) be an \( n \times n \) matrix of real numbers (weights) and by \( \sigma(\cdot) \) we denote a permutation of the set of indices \( \{1,2,\ldots,n\} \). The set of all permutations of \( \{1,2,\ldots,n\} \) is denoted by \( B_n \), and naturally the cardinality of \( B_n \) is \( n! \). The bottleneck assignment problem (BAP) seeks such a permutation \( \sigma \) that minimizes \( \max_{1 \leq i \leq n} a_{i, \sigma(i)} \). That is, the objective function \( Z_{\text{min}} \) for BAP is

\[ Z_{\text{min}} = \min_{\alpha \in B_n} \left( \max_{1 \leq i \leq n} a_{i, \sigma(i)} \right) \quad (2.7a) \]

On the other hand, the objective function \( Z_{\text{max}} \) for the capacity assignment problem (CAP) is a reverse to (2.7a), that is,

\[ Z_{\text{max}} = \max_{\alpha \in B_n} \left( \min_{1 \leq i \leq n} a_{i, \sigma(i)} \right) \quad (2.7b) \]

These two problems fall into (2.6) with the operator \( \Xi \) being 'max' and 'min' respectively.

**Example 2.8 Bottleneck and capacity traveling salesman problem**

With the notation as in Example (2.2), the objective functions for bottleneck and capacity traveling salesman problems are

\[ Z_{\text{min}} = \min_{\alpha \in B_n} \left( \max_{i \in S_n(\alpha)} w_i(\alpha) \right) \quad \text{and} \quad Z_{\text{max}} = \max_{\alpha \in B_n} \left( \min_{i \in S_n(\alpha)} w_i(\alpha) \right) \]

respectively, where the set of feasible solution \( B_n \) represents all Hamiltonian circuits and \( S_n(\alpha) \) is a set of all vertices belonging to the \( \alpha \)-th Hamiltonian circuit. We should also notice that
these problems can be represented in terms of an appropriate assignment problem with an additional requirement that the choice of elements for the matrix $A$ must form a tour. □

Finally in some situations, our basic assumptions (A) and (B) are too restrictive. Therefore, we also consider an additional generalization of problem (2.6) by extending our basic assumption (A) and (B), namely

(A') The cardinality of $S_n(\alpha)$ depends on $\alpha \in B_n$, that is $|S_n(\alpha)| = N(\alpha)$.

(B') The weights $w_i(\alpha)$ are dependent random variables with different distribution functions.

The last four examples illustrate problem (2.6) with assumptions (A') and (B').

Example 2.9 Symmetric assignment problem

The setting of the problem is the same as in Example 2.1, except that the matrix $A = [a_{ij}]_{i,j=1}^n$ is symmetric. In terms of the perfect bipartite matching, we assume that the graph is undirected. Then assumption (B) does not hold any longer, since some of $a_{i(a(i)}$ might be dependent. For example, let $n = 2$ and $a(1) = 2$ and $a(2) = 1$, hence $Z_{\text{max}} = a_{12} + a_{21} = 2a_{12}$, since $a_{12} = a_{21}$ by symmetry.

Example 2.10 Suffix tree

A suffix tree is a digital tree (i.e., trie), as the one we discussed in Example 2.5, but the keys are very dependent. More precisely, let $X = x_1x_2x_3 \cdots$ be a string of (possible) unbounded length, and let $S_i = x_i x_{i+1} \cdots$ be the $i$th suffix of $X$, $i = 1, 2, \ldots, n$. We store the first $n$ suffixes of $X$ in a trie in the same manner as discussed in Example 2.5. Such a digital tree is called suffix tree or position tree [ApS88, AHU74]. As in Example 2.5 we can argue that the height of a suffix tree can be computed through the knowledge of the so-called self-alignments $C_{ij}$ i.e., $C_{ij} = k$ iff $S_i$ and $S_j$ agree exactly on $k$ symbols, but differ on their $(k+1)$st. Then, the problem falls into our formulation (2.6), but this time neither the assumption (A)
nor (B) hold, and one must consider (A') and (B'). In fact, the self-alignments \( C_{ij} \) depends on \( i \) and \( j \), but fortunately in such a manner that the distribution of \( C_{ij} \) depends only on the difference \( d = |j - i| \), so we denote it by \( C_d \). In [ApS88] we have estimated the complement of distribution function of \( C_d \), namely

\[
R_d(k) = Pr\{C_d \geq k\} = (p^{i+2} + q^{i+2}y(p^{i+1} + q^{i+1})^{d-r}
\]

where \( k = dl + r \) with \( l = 0, 1, 2, \ldots \) and \( r = 0, 1, \ldots, d - 1 \).

Example 2.11 Minimum diameter spanning tree

We consider the same setting as in Example 2.3, and we are seeking a minimum diameter spanning tree. The diameter in a graph is defined to be the maximum of the weights between any pair of vertices. More precisely, the objective function is

\[
Z_{\text{min}} = \min_{a \in B_n} \{ \max_{i,j \in S(a)} \sum_{k \in P(i,j)} w_k(\alpha) \}
\]

where \( B_n \) is a set of spanning trees, \( S(a) \), a set of vertices belonging to \( \alpha \), and \( P(i,j) \) is a set of edges in the shortest path between the \( i \)-th and the \( j \)-th vertex. Definitely the problem falls into (2.6) with \( \Xi = \max \Sigma \). We note that for a given \( i \) and \( j \), the weights \( w_k(\alpha) \) under the sum in (2.9) are dependent causing the problem to be difficult. This problem was posed by D.T. Lee.

Example 2.12 Location problems

A general location problem can be formulated as follows. Let \( x_1, x_2, \ldots, x_N \) be a given set of points. An \( L \) median problem selects \( L \) points \( c_1, c_2, \ldots, c_L \) so as to minimize (maximize) the distance between these points and the points \( x_1, x_2, \ldots, x_N \). To formulate the problem in terms of our setting (2.6), we introduce a distance function (random variable) \( d(x_i, x_j) \) which represents weights for a pair \( (x_i, x_j) \). As a feasible solution \( \alpha = (c_1, \ldots, c_L) \), we accept any choice of \( L \) points out of \( n \), so the cardinality of \( |B_n| = C_n^L \). Then, we have

\[
Z_{\text{min}} = \min_{a \in B_n} \{ \min_{1 \leq i \leq L} \{ \sum_{i=1}^{n-L} \min_{d(x_i, c_j)} \} \} = C_n^L
\]
Naturally, it falls into our general problem (2.6) with assumptions (A') and (B'). The problem is quite difficult. To illustrate it let us concentrate on a linear location problem. In this formulation, \( n \) points are randomly distributed on a line and a distance between any two consecutive points \( d(x_i, x_{i+1}) \) is distributed according to \( F(\cdot) \). Then a weight \( w_{ij}(\alpha) \) can be defined as

\[
w_{ij}(\alpha) = \min d(x_i, c_j) = d(x_i, x_{i+1}) + d(x_{i+1}, x_{i+2}) + \ldots + d(x_i, c_j)
\]

and eventually formulation (2.10) can be reduced to our initial problem (2.1) with \( \Xi = \Sigma \). Note, however, that the set \( S_n(\alpha) \) is no longer of the same cardinality for every \( \alpha \). Moreover, the weights under the sum are dependent (for geometric location problems see [PAP 81]). Finally, we mention that some simplification of the problem can be achieved if one considers the location problem on a (complete directed) graph. Indeed, let \( w_{ij} \) be a weight assigned to the \((i,j)\)-edge and distributed according to \( F(\cdot) \). By a feasible solution, we understand a subset \( \alpha \subset M = \{1, 2, \ldots, n\} \) of cardinality \( L \) of vertices in a graph \( G_n \). Then we can define the \( L \) median problem as follows

\[
Z_{\max} = \max_{\alpha \in B_n} \sum_{i \in M-\alpha} \max_{j \in \alpha} w_{ij}
\]

(2.11)

where \( |B_n| = C_n^L \). We note that for complete directed graphs, the weights in (2.11) are independent, which makes the problem a little easier.

3. MAIN RESULTS

In this section we present our main results and discuss some consequences of these findings. The plan for this section is as follows. We first argue that our general problem (2.6) can be reduced to an appropriate problem in order statistics [GAL87, FEL71, REN70]. Then, by identifying three types of distribution functions, we study some asymptotic properties of the order statistics. After presenting our Main Lemma on asymptotic behaviors of order statistics, we focus again on our principal problem (2.6). Using the Main Lemma, we establish lower and upper bounds (in a probability sense) on \( Z_{\max} \) and \( Z_{\min} \), and we present asymptotic results for
Finally, we discuss some consequences of our main findings. In particular, we present sufficient conditions under which the performance of a greedy algorithm is asymptotically the same as the optimal solution. This generalizes previous studies on matroids and greedoids [KoL83] in the sense that a greedy algorithm is asymptotically equivalent in a probability sense to the optimal solution. Then, we discuss solutions for some particular operators $\Xi$, and present asymptotic analysis. Surprisingly enough, in some cases we will be able to present very general results which hold for a large class of distribution functions. Two cases are worth mentioning here. For bottleneck and capacity problems we shall prove for a wide class of instances that $Z_{\text{max}} \sim F^{-1}(1 - \log n/n)$ and $Z_{\text{min}} \sim F^{-1}(\log n/n)$ in probability respectively.

Next, for separable objective functions, that is, for problem (2.1) (in this case $\Xi$ becomes $\Sigma$), we shall prove a rather interesting result. Namely, for all possible instances of the problem with $m = \Theta(N^p)$, (i.e., the cardinality of the set of feasible solutions is in a polynomial relationship with the cardinality of a set of all objects belonging to a feasible solution) asymptotically $\mathbb{E}Z_{\text{max}} - \mathbb{E}Z_{\text{min}} = N \mu$, where $\mu \neq 0$ is the average weight. We also present some general results in the case when $m = \Theta(N^k)$. In this section we omit most of the proofs, and delay them until the last section of the paper.

To formulate our general problem in terms of order statistics, we define a random variable $X_\alpha$ as $X_\alpha = \frac{1}{\Xi S_n(\alpha)} \sum_{i \in S_n(\alpha)} w_i(\alpha)$ where $\alpha$ is a feasible solution, and without loss of generality, we can assume that $\alpha \in \{1, 2, \ldots, m = \lfloor B_n \rfloor\}$. Then our problem reduces to the evaluation of the following two order statistics $Z_{\text{max}} = \max_{1 \leq \alpha \leq n} \{X_\alpha\}$ and $Z_{\text{min}} = \min_{1 \leq \alpha \leq n} \{X_\alpha\}$. We note, however, that the random variables $X_\alpha$ are dependent and this causes some analytical difficulties (see Main Lemma below). The distribution function of $X_\alpha$ depends on the operator $\Xi$ and, under assumptions (A) and (B), on the size $N$ of $S_n(\alpha)$. For simplicity of notation we denote this distribution by $F_N(x) = P[X_\alpha < x]$. In some cases there exists explicit relationship between the distribution function $F_N(\cdot)$ and the original distribution $F(\cdot)$ of weights. This follows, in
particular, from assumptions (A) and (B), which are enforced throughout this section, if not stated otherwise. By assumption (B) the weights are i.i.d. random variables. This suggests the following formula on $F_N(x)$ [FEL71]:

- If $\Xi = \Sigma$ (problem (2.1)) then $F_N(x) = F(x)^*F(x)^* \cdots F(x)^* = F^{*N}(x)$, where $*$ is the convolution operator,
- If $\Xi = \text{max}$ (e.g., Examples 2.7 and 2.8), then $F_N(x) = F^N(x),$
- If $\Xi = \text{min}$ (e.g., Examples 2.7 and 2.8), then $F_N(x) = 1 - (1 - F(x))^N,$
- If $\Xi = \text{com}$ (e.g., Example 2.5 and 2.10), then $F_N(x)$ is geometrically distributed as stated in (2.3) and (2.8).

In the next considerations, we shall assume that the distribution function $F_N(x)$ is given or can be computed from the description of the problem.

From the previous arguments, it should be clear that a successful solution of our problem, that is, obtaining asymptotics for $Z_{\text{max}}$ and $Z_{\text{min}},$ depends upon establishing asymptotic relationships for some order statistics; in particular, for maximum and minimum of dependent random variables. So, let us consider the following abstract problem: given $n$ random variables $X_1, X_2, \ldots, X_n$ with distribution functions $G_1(x), \ldots, G_n(x)$ respectively, evaluate for large $n$ the behavior of $Z_{\text{max}} = \max_{1 \leq k \leq n} \{X_k \}$ and $Z_{\text{min}} = \min_{1 \leq k \leq n} \{X_k \}.$ It turns out that the solution to such a problem depends on the shape of the distribution function, in particular, its behavior at infinity. We shall consider three types of distribution functions described in the next definition.

Definition 1. (i) A general distribution function $G()$ is called Type I distribution.

(ii) A distribution function satisfying the following two conditions:

\[
G(x) < 1 \text{ for } x < \infty \tag{3.1a}
\]

\[
\lim_{x \to \infty} \frac{1 - G(xc)}{1 - G(x)} = 0 \text{ for all } c > 1 \tag{3.1b}
\]
is called Type II distribution. If conditions (3.1) are replaced by

\[ G(x) > 0 \text{ for } x > -\infty \]  
\[ \lim_{x \to -\infty} \frac{G(x+c)}{G(x)} = 0 \text{ for all } c < 1 \]  

then we have Type II' distributions.

(iii) Assuming (3.1a) holds and

\[ \lim_{x \to -\infty} \frac{1-G(x+c)}{1-G(x)} = 0 \text{ for all } c > 0 \]  

we obtain Type III distributions. Finally, we require condition (3.2a) and

\[ \lim_{x \to -\infty} \frac{G(x+c)}{G(x)} = 0 \text{ for all } c < 0 \]  

for Type III' distributions. \(\Box\)

In the further investigations we shall often use the following representatives of the above types of distribution functions:

* uniform distribution \(U(0,1)\) over interval \([0,1]\), hence

\[ G(x) = x \quad 0 \leq x \leq 1. \]  

It belongs to type I distributions. In general, a composition of several uniform distributions in the form \(G(x) = x^a(1 + o(1))\) [FEL71], where \(a \leq x \leq b\), belongs to the same class of distribution functions.

* gamma distribution \(\text{gamma}(\beta, \lambda)\) with the density function \(g(x) = G'(x)\) given by [FEL71, REN70]

\[ g(x) = \frac{\lambda^\beta}{\Gamma(\beta)} x^{\beta-1} e^{-\lambda x} \quad x \geq 0 \]  

belongs to type II distribution functions. This simple fact can be easily proved if one notes that condition (3.1b) can be equivalently expressed in terms of density functions as

\[ \lim_{x \to \infty} g(x+c)/g(x) = 0 \text{ for all } c > 1. \]  

If the condition \(x > 0\) in (3.4b) is replaced by \(x < 0\), then
one obtains negative gamma distributions which belong to type II', as it is easy to verify. 

- normal distribution $N(\mu, \sigma)$ with the density functions as below \cite{FEL71}

$$g(x) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{(x-\mu)^2}{2\sigma^2} \right] \quad -\infty < x < \infty$$

belongs to type II and II' distribution functions.

Finally, to finish with definitions, for non-identically distributed random variables we need slightly more restrictive types than those discussed in Definition 1, namely we need to assume that the types hold uniformly in the class of distribution functions. More precisely,

**Definition 2.** A sequence of distribution functions $G_1(x), \ldots, G_n(x)$ is *uniformly* Type II and III (resp. II' and III') if conditions (3.1b) and (3.3a) (resp. (3.2b) and (3.3b)) hold uniformly in $n$. For example, if the following holds

$$\lim_{x \to \infty} \sup_n \frac{1 - G_n(xc)}{1 - G_n(x)} = 0 \quad \text{for all} \quad c > 1$$

then the sequence of distributions is uniformly Type II. □

Now we are almost ready to formulate our Main Lemma on the order statistics. In the lemma we often use a sequence $a_n$ defined as the smallest solution of the following equation

$$\sum_{k=1}^{n} [1 - G_k(a_n)] = 1,$$

and another sequence $b_n$ as the largest solution of

$$\sum_{k=1}^{n} G_k(b_n) = 1$$

Then, in the last section of this paper, we prove the following important lemma.

**Main Lemma.** (i) Let $G_1(\cdot), \ldots, G_n(\cdot)$ belong to Type I distributions. Then the following bounds for the $r$-th moments $EZ_{\max}^r$ and $EZ_{\min}^r$ of $Z_{\max} = \max_{1 \leq k \leq n} \{X_k\}$ and $Z_{\min} = \min_{1 \leq k \leq n} \{X_k\}$
hold

\[ E Z_{\text{max}}^r \leq \bar{a}_n + \sum_{k=1}^{n} \int_{1}^{\infty} [1 - G_k(x^{1/r})] dx \]  (3.7)

and

\[ E Z_{\text{min}}^r \geq \bar{b}_n - \sum_{k=1}^{n} \int_{1}^{1/r} G_k(x^{1/r}) dx \]  (3.8)

where \( \bar{a}_n \) and \( \bar{b}_n \) are the smallest and the largest solutions of the following

\[ \sum_{k=1}^{n} [1 - G_k(\bar{a}_n^{1/r})] = 1; \quad \sum_{k=1}^{n} G_k(\bar{b}_n^{1/r}) = 1 \]  (3.9)

respectively. Note that for \( r = 1 \), \( \bar{a}_n \) and \( \bar{b}_n \) coincide with \( a_n \) and \( b_n \) defined in (3.5) and (3.6) respectively.

(ii) For uniformly Type II and II' distributions, the following holds

\[ \lim_{n \to \infty} \frac{Z_{\text{max}}}{a_n} \leq 1 \quad \text{in probability (in short: pr.)} \]  (3.10a)

\[ \lim_{n \to \infty} \frac{Z_{\text{min}}}{b_n} \geq 1 \quad \text{in probability (pr.)} \]  (3.10b)

respectively. If, in addition, either the following inequality

\[ P \{X_1 < x_1, X_2 < x_2, \ldots, X_n < x_n\} \leq \delta \cdot G_1(x_1) \cdot G_2(x_2) \cdots G_n(x_n) \]  (3.11a)

or the next inequality holds

\[ P \{X_1 > x_1, \ldots, X_n > x_n\} \leq \delta \cdot [1 - G_1(x_1)] \cdots [1 - G_n(x_n)] \]  (3.11b)

for some \( \delta = \mathcal{O}(1) \), then respectively

\[ \lim_{n \to \infty} \frac{Z_{\text{max}}}{a_n} = \lim_{n \to \infty} \frac{Z_{\text{min}}}{b_n} = 1 \quad \text{in probability} \]  (3.12)

which equivalently can be written as \( Z_{\text{max}}/a_n \to 1, Z_{\text{min}}/b_n \to 1 \), or \( Z_{\text{max}} = a_n(1 + o(1)) \) and \( Z_{\text{min}} = b_n(1 + o(1)) \) (pr.). Finally, assuming the \( r \)-th moments of \( X_1, X_2, \ldots, X_n \) exist, that is, denoting by \( g_k(x) \) the density function of \( X_k \) we postulate the following for some \( r \)
\[
\int_{-\infty}^{\infty} x^r g_k(x) \, dx < \infty, \quad (3.13)
\]

and then, the convergence in probability can be replaced by convergence in mean, namely

\[
\lim_{n \to \infty} E|Z_{\text{max}}/a_n - 1|^r = 0, \quad (3.14)
\]
hence, \(EZ'_{\text{max}} = a_n^r(1 + o(1)) - a_n^r\).

(iii) For uniformly Type III and III' distribution stronger asymptotics can be established, namely

\[
\lim_{n \to \infty} (Z_{\text{max}} - a_n) \leq 0 \quad (pr.) \quad (3.15a)
\]
\[
\lim_{n \to \infty} (Z_{\text{min}} - b_n) \geq 0 \quad (pr.) \quad (3.15b)
\]
respectively. If, in addition, inequalities (3.11) hold respectively, then

\[
Z_{\text{max}} - a_n \overset{p}{\to} 0 \quad \text{and} \quad Z_{\text{min}} - b_n \overset{p}{\to} 0 \quad (3.16)
\]

Finally, for distributions satisfying (3.14) the convergence in mean holds, that is,

\[
\lim_{n \to \infty} E|Z_{\text{max}} - a_n|^r = \lim_{n \to \infty} E|Z_{\text{min}} - b_n|^r = 0 \quad (3.17)
\]

(iv) If \(X_1, X_2, \ldots, X_n\) are identically independently distributed (i.i.d.) random variables satisfying either (3.1a) or (3.2a), then

\[
P \left[ \lim_{n \to \infty} \frac{Z_{\text{max}}}{a_n} = 1 \right] = P \left[ \lim_{n \to \infty} \frac{Z_{\text{min}}}{b_n} = 1 \right] = 1, \quad (3.18a)
\]
that is, \(Z_{\text{max}}/a_n \to 1\) and \(Z_{\text{min}}/b_n \to 1\) almost surely, if and only if the following holds for all \(\beta > 1\)

\[
\sum_{n=1}^{\infty} [1 - G(\beta a_n)] < \infty \quad \text{and} \quad \sum_{n=1}^{\infty} G(\beta b_n) < \infty \quad (3.18b)
\]
respectively. \(\Box\)

Investigating further we shall see that the almost sure convergence discussed in (iv) above can be extended to some dependent random variables. This is achieved by applying a similar
approach to the one presented in (ii) and (iii) (see Section 5 for the proof, and more detailed discussion) and referring to the Borel-Cantelli Lemma [BIL86, REN70, FEL71].

The main lemma plays a crucial role in establishing our results on asymptotic behaviors of $Z_{\text{max}}$ and $Z_{\text{min}}$. It turns out that very general results can be formulated for two important classes of operators $\Xi$, namely, the so called nondecreasing and ranking-dependent operators.

An operator $\Xi$ is nondecreasing if the following holds

$$w_i(\alpha) \leq w'_i(\alpha) \implies \Xi_{\alpha} w_i(\alpha) \leq \Xi_{\alpha} w'_i(\alpha)$$

(3.19)

for every $\alpha$ and $i$. For ranking-independent operator $\Xi$, we require that the optimal solution $\alpha_{\text{opt}}$ depends only on the ranks of the weights $w_i(\alpha)$ and not on the concrete values of the weights, that is, for any increasing function $f(\cdot)$ the following holds

$$f(Z_{\text{max}}) = \max_{\alpha \in \mathcal{B}, \alpha \in S(\alpha)} \Xi f(w_i(\alpha))$$

(3.20)

For example, the operator $\Sigma$, "max" and "min" are nondecreasing operators. In addition, "max" and "min" are also ranking-dependent operators, that is, the bottleneck and capacity problems defined in Example 2.7 and 2.8 belong to that class.

Now we are prepared to present our main results. We assume that the operator $\Xi$ is either nondecreasing or ranking-dependent. For nondecreasing operators, we implicitly assume that distribution functions of the weights are either of type II or II' and III or II'. Then, we write $Z \sim a_n$ and understand this notation as $Z/a_n \rightarrow 1$ for type II or II' distributions (see Main Lemma (i)), and $Z \sim a_n \rightarrow 0$ for type III or III' distributions (see Main Lemma (iii)). For ranking-dependent operators, it is assumed that distribution functions are strictly continuous (i.e., increasing functions). Finally, we shall derive results only for $Z_{\text{max}}$, since by analogy one easily obtains similar conclusions for $Z_{\text{min}}$.

**Upper (Lower) Bounds on $Z_{\text{max}}$ ($Z_{\text{min}}$)**
In this subsection, we establish two upper bounds on $Z_{\text{max}}$. The first upper bound is valid for all possible operators $\Xi$, while the second bound is restricted to a class of nondecreasing operators. We recall that $F_N(\cdot)$ denotes the distribution function of $X_{\alpha} = \Xi w_i(\alpha)$ for every $\alpha$ (cf. assumption (B)). Let us also define the reliability function $R_N(x)$ as $R_N(x) = 1 - F_N(x)$.

We assume that $F_N(\cdot)$ is either type II or type III distribution function. In order to establish the first upper bound for $Z_{\text{max}} = \max_{\alpha \in \mathcal{B}_n} \{X_{\alpha}\}$ we apply our Main Lemma (ii) and (iii) to show that $Z_{\text{max}} \leq \overline{Z} - a_n$ (pr.), where $a_n$ is the smallest solution of (3.5), that is, in our case $mR_N(a_n) = 1$. In the above, $\leq$ means stochastically smaller in the sense of Stoyan [ST083], and the existence of the random variable $\overline{Z}$ is proved in [LaR77] ($\overline{Z}$ is called maximally dependent random variable). We adopt the following short notations. If $Z_{\text{max}} \leq \overline{Z} - a_n$ (pr.) for some random variable $\overline{Z}$ and $a_n = R_N^{-1}(1/m)$, then we write $Z_{\text{max}} \leq R_N^{-1}(1/m)$ (pr.). For further simplification, we write "$\leq$" instead of "$\leq$". So, in summary we obtain

$$Z_{\text{max}} \leq R_N^{-1}(1/m) \quad \text{(pr.)}$$

To establish second bound on $Z_{\text{max}}$ we assume that $\Xi$ is nondecreasing. Since $w_k \leq \max_{i \in S} w_i$, for every $k \in S$ where $S$ is a set of all objects, we immediately obtain (see also [ST083] for stochastic inequalities)

$$Z_{\text{max}} = \max_{\alpha \in \mathcal{B}_n} \Xi w_i(\alpha) \leq \Xi \max_{i \in S_n} \max_{\alpha \in \mathcal{B}_n} \{w_i(\alpha)\}$$

and in the above we simplify the notation $S_n(\alpha)$ to $S_n$. To modify the RHS of (3.22) we note that the sequence $\{w_i(\alpha)\}_{\alpha \in \mathcal{B}_n}$ contains many identical weights which, if deleted, do not change the value of $\max_{\alpha \in \mathcal{B}_n} \{w_i(\alpha)\}$. Let

$$\max_{\alpha \in \mathcal{B}_n} \{w_i(\alpha)\} = \max_{j \in O(i)} \{w_{ij}\}$$

where $O(i)$ enumerates all distinct weights $w_i(\alpha)$ over all feasible solutions $\alpha \in \mathcal{B}_n$, and we
denote these distinct weights by $w_{ij}$. In fact, the set $O(i)$ can be interpreted as a set of "direct neighbors" of the $i$-th object. For example, in the assignment problem (or any combinatorial problem on matrices), $O(i)$ is the set of all elements in the $i$-th row or column; in graph problems $O(i)$ represents vertices directly connected to the $i$-th vertex, etc. In accordance to our assumption (B), the weights $w_{ij}$ are i.i.d. and assuming the operation in (3.23) does not introduce dependency, we immediately show, by quoting Main Lemma (ii)-(iii), that

$$\max_{j \in O(i)} \{ w_{ij} \} \sim R^{-1}(1/K) \quad (pr.),$$

provided $|O(i)| = K$ tends to infinity with $n$, i.e., and $K \to \infty$ with $n \to \infty$. The assumption $|O(i)| = K$ restricts the class of random structures considered in this paper to the regular ones [BOL85]. Now, to establish an upper bound we try to incorporate (3.23) into (3.22), but this needs some care. In particular, we must know whether the operator $\Xi$ is convergence preserving. To simplify further considerations and to present more specific results, we consider two important operators, namely $\Xi = \Sigma$ and $\Xi = \min, \max$.

We shall first talk about the sum operator, i.e., $\Xi = \Sigma$. Quoting the Strong Law of Large Numbers [BIL86, FEL71], we may establish the following implications

$$\sum_{N} \max_{j \in O(i)} w_{ij} \overset{a.s.}{\to} E \max_{j \in O(i)} w_{ij} \overset{P}{\to} R^{-1}(1/K)$$

(3.24)

provided $N, K \to \infty$ for $n \to \infty$. The latter convergence in probability can be replaced by almost sure convergence for those distributions that satisfy condition (3.18b) in Main Lemma (iv). In particular, this is the case for the gamma distribution (3.4b) and the normal distribution (3.4c). Finally, (3.24) and (3.22) imply

$$Z_{\max} \leq NR^{-1}(1/K) \quad (a.s.) \quad \text{or} \quad (pr.) \quad Z_{\min} \geq NF^{-1}(1/K)$$

(3.25)

provided $K, N \to \infty$ with $n \to \infty$.

It turns out (see Theorem 2 of this section, and also Section 4) that this bound is even more important for max-min and min-max problems. We assume now $\Xi = \min$, so (3.22) is read as
\[
Z_{\text{max}} \leq \min_{i \in S_n} \max_{j \in O(i)} \{w_{ij}\}
\]

(3.26)

where \(1_{S_n}1 = N\) and \(1_{O(i)}1 = K\). Note that due to assumption (B), the distribution function of \(\max_{j \in O(i)} \{w_{ij}\}\) is equal to \(P\{\max_{j \in O(i)} \{w_{ij}\} < x\} = F^K(x)\), where \(F(\cdot)\) is the distribution of the weights \(w_{ij}\). Then for \(K, N \to \infty\), Main Lemma (ii), (iii) leads to the following conclusions

\[
Z_{\text{max}} \leq F^{-1}(N^{-1/K}) \quad (\text{pr.}) \quad Z_{\text{min}} \geq R^{-1}(N^{-1/K})
\]

(3.27)

provided \(N, K \to \infty\) as \(n \to \infty\). The latter assumption regarding \(N\) and \(K\) is important, as one can assess it considering the clique problem (see Section 4).

**Lower (Upper) Bounds for \(Z_{\text{max}} (Z_{\text{min}})\)**

The lower bound on \(Z_{\text{max}}\) is either established by using inequality (3.11) or by the fact that for any solution \(\alpha\) the value \(Z(\alpha)\) of the objective function must be smaller than \(Z_{\text{max}}\), i.e., \(Z(\alpha) \leq Z_{\text{max}}\). A proper choice of \(\alpha\) satisfying the above is discussed in this subsection.

We start our discussion with the inequality approach mentioned above. Let

\[
X_{\alpha} = \sum_{i \in S_{\alpha}(\alpha)} w_i(\alpha).
\]

Then, assuming

\[
P\{X_{\alpha} < x_{\alpha}, \alpha \in B_n\} \leq \delta \prod_{\alpha \in B_n} P\{X_{\alpha} < x_{\alpha}\}
\]

(3.28)

with \(\delta = O(1)\), one immediately proves, using inequalities (3.11) of Main Lemma (ii), (iii), that

\[
Z_{\text{max}} \geq R_N^{-1}(1/m) \quad (\text{pr.}) \quad Z_{\text{min}} \leq R_N^{-1}(1/m)
\]

(3.29)

The inequalities (3.28) can be called *mixing conditions*, since they resemble the strong-mixing conditions already widely used for the weak convergence of (weakly) dependent random variables [BIL86, FEL71]. We note that (3.28) holds if a weaker inequality is satisfied, namely

\[
P\{X_{\alpha_1} < x_{\alpha_1}, X_{\alpha_2} < x_{\alpha_2}\} \leq P\{X_{\alpha_1} < x_{\alpha_1}\} P\{X_{\alpha_2} < x_{\alpha_2}\}.
\]

In some circumstances the bound in probability can be replaced by an almost sure bound (see strong mixing conditions in [BIL86, FEL71, LaR76]). Finally, we note that for (3.29) we do not require monotonicity of the operator \(\Xi\).
Most lower bounds for the optimization problem are based on the simple observation that for any $\alpha \in B_n$ the solution $Z(\alpha)$ must be smaller than $Z_{\text{max}}$, so $Z(\alpha) \leq Z_{\text{max}}$. The open question is how to choose $\alpha$. We present two methods. One is based on some known results for unweighted random structures (e.g., graphs), and the other method assumes a greedy solution $\alpha_{\text{greedy}}$. In the first method, which is a slight generalization of Weide's approach [WEI80] (however we do not restrict ourselves to graphs), we consider a random structure which mimics our structure under considerations, except that objects do not have weights. This random structure grows randomly like random graphs [BOL85], that is, with probability $p_n$, an object is selected to the structure, and with probability $1 - p_n$ is not included in the structure. We assume that such a random structure is easier to analyze, and there are some already known results regarding properties of interest to us. For example, a great deal of results (e.g., existence of Hamiltonian circuits, connected components, spanning trees, etc.) is known about random (unweighted) graphs [BOL85]. We assume, in particular, that the following is known.

A feasible set $B_n$ is nonempty (a.s.) or (pr.) if objects of a random (unweighted) structure are selected with at least probability $p_n$. 

(3.30)

If (3.30) is satisfied, then we build a random weighted structure by choosing weights $w_i \geq R^{-1}(p_n)$ where $F(x) = 1 - R(x)$ is the distribution function of the desired weights. But, such a structure has in fact weights distribution given by $P\{w_i \geq R^{-1}(p_n)\} = R\{R^{-1}(p_n)\} = p_n$. Therefore, $\alpha = \{i \in G_n: w_i \geq R^{-1}(p_n)\}$, where $G_n$ is the random structure, is a particular solution to our problem. By (3.30) such a solution almost surely or in probability exists, hence the following holds

$$Z_{\text{max}} \geq \sum_{i \in S_n} R^{-1}(p_n) \quad \text{(a.s.) or (pr.)} \quad Z_{\text{min}} \leq \sum_{i \in S_n} F^{-1}(p_n) \quad (3.31)$$

where $|S_n| = N$. We must note however, that a successful application of (3.31) depends on a priori knowledge about the critical probability $p_n$. For example, in the random graph theory, it is known [BOL85] that $p_n \geq c \frac{\log n}{n}$ for some $c > 1$ implies existence of a Hamiltonian circuit.
in the graph, etc.

At last, we turn our attention to the greedy algorithm approach, which seems to be the most promising from the algorithmic viewpoint. In order to present the bound in a compact form, we need a general definition of a greedy algorithm. We follow the approach proposed in [KoL86]. We note that for any solution \( \alpha \in B_n \), we can write \( \alpha = x_1 x_2 \cdots x_N \) where \( x_i \in S \) (\( S \) is the set of objects), that is, a solution consists of a sequence of \( N \) objects. A greedy algorithm is a sequential procedure that selects in the \( k \)-th step an object \( x_k \) which locally optimizes the objective function. To be more formal, we extend the set of feasible solution \( B_n \) to a set \( \tilde{B}_n \) (so-called hereditary language [KoL 86]) possessing the following property: if \( \beta = x_1 x_2 \cdots x_k \subseteq \alpha \), then \( \beta \in \tilde{B}_n \) provided \( \alpha \in \tilde{B}_n \). In other words, \( \tilde{B}_n \) is a set of all partial solutions to our optimization problem. In particular, if \( \beta \in \tilde{B}_n \), then selecting an object \( x_k \) and adding it to \( \beta \) we might form another feasible solution, that is, if \( \beta \in \tilde{B}_n \) and \( x_k \in S \) then \( \beta x_k \in \tilde{B}_n \) if \( \beta x_k \) is a partial feasible solution. Now, we are in a position to present a general description of a greedy algorithm.

**Greedy Algorithm:**

For \( k = 1, 2, \ldots, N \) select an element \( x_k \in S \) such that \( \beta \in \tilde{B}_n \) implies \( \beta x_k \in \tilde{B}_n \) and

\[
\sum_{i \in S, (\beta x_k)} w_i(\beta x_k) = \max_{y \in S} \sum_{i \in S, (\beta y)} w_i(\beta y)
\]  

(3.32)

Stop if no such \( x_k \) exists.

The application of the greedy procedure (3.32) depends on the random structure under consideration and the operator \( \Xi \). Nevertheless, a pretty general class of greedy algorithms can be presented in a relatively uniform way, as we show below. The main problem in presenting a general structure for the greedy algorithm is to assure that in each step of the algorithm, we select a feasible sub-solution. To avoid this type of problem, we focus on complete graphs, and first assume \( \Xi = \Sigma \). Let \( \beta_i = x_1 x_2 \cdots x_i \in \tilde{B}_n \) be a partial solution found up to the \( i \)-th step in a
greedy procedure. Then, the \( i+1 \)-st object \( x_{i+1} \) is selected from the (presumably nonempty) neighborhood \( A(i) \) of the \( i \)-th object such that \( \beta_i x_{i+1} \) is a feasible subsolution. Naturally, the cardinality of \( A(i) \) depends on the step \( i \), and \( A(i) \subseteq O(i) \) where \( O(i) \) defined in (3.23) is a set of all direct neighbors of \( x_i \). With this notation in mind, the greedy algorithm (3.32) for \( \Xi = \Sigma \) might look as follows

\[
Z(\beta_{i+1}) = Z(\beta_i) + \max_{j \in A(i)} w_{ij} \tag{3.33}
\]

where \( Z(\beta) \) for \( \beta \in B_n \) is the value of the objective function \( Z \) for the partial feasible solution \( \beta \), and \( w_{ij} \) is a weight assigned to the pair \( (i, j) \), that is, to the \( (i, j) \) edge in a graph. We note, however, that (3.33) is not the only possible greedy approach. For example, the following one presents an alternative algorithm

\[
Z(\beta_{i+1}) = Z(\beta_i) + \max_{j \in A(i)} \sum_{k=1}^{M_i} w_{kj} \tag{3.34}
\]

where in the \( i \)-th step, the algorithm select \( M_i \) objects (see for example greedy algorithm for \( k \)-clique problem in the next section). Then, solving the recurrences (3.33) and (3.34) one finds out that the ultimate value of the objective function \( Z_{\text{grad}} \) is equal to

\[
Z_{\text{grad}} = \sum_{i \in S_n} \max_{j \in A(i)} w_{ij} + \sum_{i \in S_n - S'_n} w_i(\alpha_{\text{grad}}) \tag{3.35a}
\]

\[
Z_{\text{grad}} = \sum_{i \in S'_n} \max_{j \in A(i)} \sum_{k=1}^{M_i} w_{kj} + \sum_{i \in S_n - S'_n} w_i(\alpha_{\text{grad}}) \tag{3.35b}
\]

where \( N' = |S'_n| \leq S_n(\alpha_{\text{grad}}) = N \). In the above \( S'_n \) is the set of objects found by the greedy algorithm (before the stopping rule suggested in (3.32) is applied) that constitutes of a partial solution to the problem, and that assures a feasible solution can be constructed from this partial solution (see the second term of (3.35) ). For example, a greedy algorithm for a \( k \)-clique problem may select in the first step \( k \) smallest weights of outgoing edges of a vertex, and the rest of the edges (i.e., \( C_k^2 - k \) edges needed to fill up the \( k \)-clique) are completely determined by this first choice. That is, in this case \( N' = |S'_n| = k \) and \( |S_n| = C_k^2 = N \).
To analyze (3.35a) we assume that the weights in the set $A(i)$ are i.i.d., hence for type II and III distribution functions of the weights one finds $\max_{j \in A(i)} w_{ij} \sim R^{-1}(1A(i)\sim^{-1})$ provided \(|A(i)| \to \infty\) as \(n\) tends to infinity (see Section 5 for some more comments on this). In Section 5 we shall argue that the above implies the following

$$Z_{\text{grd}} = \sum_{i \in S_n} R^{-1}(1A(i)\sim^{-1}) + (N - N')\mu \quad (pr.)$$  \hspace{1cm} (3.36a)$$

where $\mu = EW_{ij}$ denotes the average weight. For (3.35b) the asymptotic formula is as follows

$$Z_{\text{grd}} \sim \sum_{i \in S_n} R_{M_i}(1A(i)\sim^{-1}) + (N - N')\mu \hspace{1cm} (3.36b)$$

where $F_{M_i}(x) = 1 - R_{M_i}(\cdot)$ denotes convolution of $M_i$ distribution functions of the weights.

The above is not restricted to the sum operator $\Sigma$. For example, the following is an obvious greedy algorithm for $\Xi = \min$ with an additional assumption that $S_n = S'_n$

$$Z_{\text{grd}} = \min_{i \in S_n} \max_{j \in A(i)} w_{ij} \quad (3.37)$$

with the same notation as above. In particular, if $N \to \infty$ as $n \to \infty$ and $F^{1A(i)}(\cdot)$ denotes the $|A(i)|$-th power of the distribution function $F(\cdot)$, then $Z_{\text{grd}} \sim b_n$ where $b_n$ is the largest root of the following equation

$$\sum_{i \in S_n} F^{1A(i)}(b_n) = 1 \hspace{1cm} (3.38)$$

In general, for nondecreasing operator $\Xi$ the greedy algorithm discussed above can be represented as

$$Z_{\text{grd}} = \Xi_{i \in S_n} \max_{j \in A(i)} w_{ij} \hspace{1cm} (3.39)$$

assuming $S_n = S'_n$. In Section 4, we illustrate some applications of these studies, and we shall see that for some problems a determination of the set $S'_n$ is the most cumbersome.
Summary of Main Results

In this subsection, we summarize our main results in the form of a proposition, and present some consequences of our findings. In particular, we shall discuss some sufficient conditions that guarantee the same asymptotic performance in a probability sense for the optimal solution and a greedy algorithm. This result can be considered as a first step towards probabilistic generalization of matroids and greedoids [KoL83]. In addition, we shall show some more results on bottleneck problems ($\Xi = \min$) and separable objective function problems ($\Xi = \Sigma$).

In the previous two subsections, we have derived our main results based on the Main Lemma. They can be summarized as follows.

**PROPOSITION.** Throughout this proposition we assume the following: (1) assumptions (A) and (B) hold; (2) $\Xi$ is either a nondecreasing operator or ranking-dependent operator; (3) for $\Xi$ nondecreasing, the weights are either distributed according to type II (II') or type III (III') (in the case of type I distributions the asymptotic results presented below should be replaced by an appropriate inequality from Main Lemma (i)); (4) for $\Xi$ ranking-dependent, the distribution function of weights can be any type as long as it is strictly increasing function;

(i) If $m \to \infty$ as $n \to \infty$, then (3.21) holds in probability, that is,

$$Z_{\text{max}} \leq R_N^{-1}(1/m) \quad (\text{pr.}) \quad Z_{\text{min}} \geq F_N^{-1}(1/m)$$

(3.40)

and (3.40) holds also in mean if (3.13) of the Main Lemma can be verified.

(ii) Let $K = 10(i)l \to \infty$ and $N \to \infty$ as $n \to \infty$, where $O(i)$ is defined in (3.23). Then, for $\Xi = \Sigma$

$$Z_{\text{max}} \leq NR^{-1}(1/K) \quad (a.s.) \quad Z_{\text{min}} \geq NF^{-1}(1/K)$$

(3.41)

if the distribution function $F(\cdot)$ satisfies (3.18b), otherwise (3.41) holds in probability, and in
For $\Xi = \min$ in the case of $Z_{\max}$, and $\Xi = \max$ for $Z_{\min}$

$$Z_{\max} \leq F^{-1}(N^{1/K}) \quad (pr.) \quad Z_{\min} \geq R^{-1}(N^{-1/K}) \quad (3.42)$$

(iii) If mixing conditions (3.28) hold, then for $N \to \infty$ with $n \to \infty$

$$Z_{\max} \geq R^{-1}_N(1/m) \quad (pr.) \quad Z_{\min} \leq F^{-1}_N(1/m) \quad (3.43)$$

and the above holds also in mean if, in addition, (3.13) is assumed.

(iv) Let $p_n$ be the critical probability defined in (3.30) that assures nonemptiness of $B_n \ a.s.$ (pr.). Then,

$$Z_{\max} \geq \Xi \sum_{i \in S_n} R^{-1}(p_n) \quad (a.s.) (pr.) \quad Z_{\min} \leq \Xi \sum_{i \in S_n} F^{-1}(p_n) \quad (3.44)$$

(v) For $\Xi = \Sigma$, with notations and assumptions already discussed above, the following holds

$$Z_{\max} \geq \sum_{i \in S_n} R_M^{-1}(1\{A(i)\}^{-1}) + (N - N')\mu \quad (pr.) \quad (3.45a)$$

$$Z_{\min} \leq \sum_{i \in S_n} F_M^{-1}(1\{A(i)\}^{-1}) + (N - N')\mu \quad (pr.) \quad (3.45b)$$

provided the objects in the set $A(i)$ have i.i.d. weights, and $S'_n$ represents a maximal set built by the greedy algorithm that also assures the existence of a feasible solution constructed from this set.

For $\Xi = \min$ in the case of $Z_{\max}$ and $\Xi = \max$ in the case of $Z_{\min}$, and $N' = N$ the following is true

$$Z_{\max} \geq b_n \quad (pr.) \quad Z_{\min} \leq a_n \quad (3.46)$$

where $b_n$ and $a_n$ are roots of the next two equations

$$\sum_{i \in S_n} F^{1\{A(i)\}^{-1}}(b_n) = 1 \quad \sum_{i \in S_n} R^{1\{A(i)\}^{-1}}(a_n) = 1 \quad (3.47)$$

where $F^{1\{A(i)\}^{-1}}(x) = \prod_{k=1}^{1\{A(i)\}} F(x)$ and $R^{1\{A(i)\}^{-1}}(x) = \prod_{k=1}^{1\{A(i)\}} R(x)$. If the distribution function of a
weight satisfies (3.18a) or (3.13), then the above bounds hold *almost surely* or *in mean*.

**Proof.** Proposition parts (i)-(iv) follow directly from the above discussion. Section 5.1 presents some more details about proving Proposition (v). □

The Proposition provides tools to investigate asymptotic behavior of the optimal value $Z_{opt}$ of the objective function. In many cases the asymptotic bounds provide the exact value of the first leading term in an asymptotic expansion. This has its own implications, and finds many applications in the design of heuristic (approximate) algorithms for some problems. To be more specific, we must agree on what a "good" heuristic means. In the spirit of Karp's idea [KAR76], we define a relative error $e_n$ of a solution $\alpha \in B_n$ in the following way

$$e_n = \frac{|Z_{opt} - Z(\alpha)|}{Z_{opt}}$$ (3.48)

Then, we agree to call a solution $\alpha$ as near-optimal (a good heuristic) if $e_n \to 0$ as $n \to \infty$ *almost surely*, or *in probability* or *in mean*. This condition implies, as it is easy to see and it was explicitly stated by Weide [WEI80], that $\alpha$ is near-optimal if and only if

$$\lim_{n \to \infty} \frac{Z(\alpha)}{Z_{opt}} = 1$$ in an appropriate probability sense, that is, the leading components in both $Z(\alpha)$ and $Z_{opt}$ agree. Our Proposition gives sufficient conditions for $\alpha$ to be near-optimal. In particular, the Proposition can predict when a greedy algorithm achieves near-optimality. We formulate these conclusions in the form of the following theorem.

**Theorem 1.** (i) Let $\alpha$ be a feasible solution $\alpha \in B_n$ such that

$$\lim_{n \to \infty} \frac{R^{-1}_N(1/m)}{Z(\alpha)} = 1 \quad (pr.) \quad \lim_{n \to \infty} \frac{F^{-1}_N(1/m)}{Z(\alpha)} = 1$$ (3.49)

then $\alpha$ is near-optimal, that is, $e_n \to 0$ (pr.) provided $m \to \infty$ as $n \to \infty$ and $F(\cdot)$ is type II and III (II' or III') distribution function. In addition, for $\Xi = \Sigma$ this remains true if (3.49) is replaced by
\[
\lim_{n \to \infty} \frac{NR^{-1}(1/K)}{Z(\alpha)} = \lim_{n \to \infty} \frac{NF^{-1}(1/K)}{Z(\alpha)} = 1 \quad (a.s.)
\]
(3.50a)

and
\[
\lim_{n \to \infty} \frac{F^{-1}(N^{-1}/K)}{Z(\alpha)} = \lim_{n \to \infty} \frac{R^{-1}(N^{-1}/K)}{Z(\alpha)} \quad (pr.)
\]
(3.50b)

respectively, provided \( K, N \to \infty \) as \( n \to \infty \) and \( F(\cdot) \) is type II (II') or III (III') distribution functions.

(ii) When is greedy near optimal? Let \( \Xi = \Sigma \). Then, under the same hypotheses as in (i), and with notation adopted from Proposition (v), the greedy algorithm \( \alpha_{grd} \) defined in (3.35) is near optimal (i.e., \( Z_{grd} \to Z_{\max} \)) if
\[
\lim_{n \to \infty} \frac{R^{-1}(1/m)}{\sum_{i \in S_n} R^{-1}(1A(i)^{-1}) + (N-N')\mu} = 1 \quad \text{or} \quad \lim_{n \to \infty} \frac{R^{-1}(1/K)}{\sum_{i \in S_n} R^{-1}(1A(i)^{-1}) + (N-N')\mu} = 1,
\]
(3.51)

and for \( \Xi = \min \) with \( n = N' \) the greedy becomes the near optimal one if
\[
\lim_{n \to \infty} \frac{R^{-1}(1/m)}{b_n} = 1 \quad \text{or} \quad \lim_{n \to \infty} \frac{NR^{-1}(1/K)}{b_n} = 1
\]
(3.52)

where \( b_n \) is a solution of the equation defined in (3.47).

\textbf{Proof}: It is a simple consequence of our Proposition. Details are left to the reader. \( \blacksquare \)

In Section 4, we apply these criteria to some of the problems presented in Section 2 and we present some near optimal algorithms.

\textbf{Some More Results for } \Xi = \max \text{ and } \Xi = \min

Now, we present one general result for \( \Xi = \min \) (max) operator, that is, for ranking dependent problems. We adopt all notations and assumptions from our Proposition. In particular, we assume that \( F(\cdot) \) is a strictly continuous function. As stated before for ranking-dependent problems, we may concentrate on one distribution and then translate the results to all other distributions by noting that \( X = F^{-1}(U) \) where \( X \) is \( F(\cdot) \)-distributed random variables and
\( U \) is uniformly distributed in the interval \([0,1]\) [FEL71]. It is convenient to assume that weights are exponentially distributed, that is, \( F(x) = 1 - e^{-x} \). Let \( K, N \to \infty \) as \( n \to \infty \). Then by Proposition (ii), formula (3.41) \( Z_{\text{max}} \leq \frac{1}{K} \log N \) (pr.). For a lower bound, we apply the greedy approach from Proposition (v) formula (3.46). We assume, in addition, that \( |A(i)| = K - i \), which is satisfied for many (complete) graphs and matrix problems. With these assumptions \( Z_{\text{max}} \geq b_n \) where \( b_n \) is the solution of the following equation

\[
\sum_{i=1}^{N} e^{-\frac{(K - i)K}{b_n}} = 1 \tag{3.53}
\]

It is not difficult to show that asymptotically \( b_n = \frac{\log N}{K} \) (for details see [SZP88c]). Therefore, we can formulate the following theorem.

**Theorem 2.** Let \( \Xi = \min (\max) \) and \( K, N \to \infty \) with \( n \to \infty \). In addition, we assume that \( N = N' \) and \( |A(i)| = K - i \), and the distribution of weights is a strictly continuous function. Then the following holds

\[
Z_{\text{min}} \sim F^{-1} \left( \frac{\log N}{K} \right) \quad \text{(pr.)} \quad Z_{\text{max}} \sim F^{-1} \left( 1 - \frac{\log N}{K} \right) \tag{3.54}
\]

respectively.

**Proof.** It is enough to note that the leading factor of \( b_n \) for the uniform distribution is the same as for the exponential case. \( \square \)

**Some More Results for \( \Xi = \Sigma \)**

Optimization problem (2.1) (i.e., problem (2.6) with \( \Xi = \Sigma \)) seems to be the most popular in the class of combinatorial optimizations. Therefore, we present here some more results in this case. We first note that an application of our Proposition crucially depends on satisfactory solutions of the following two problems:
(1) Explicit formula for the \( N \)-th convolution of the distribution function \( F(\cdot) \), that is, \( F_N(\cdot) \).

(2) Asymptotic solution to the (nonlinear) equations (3.5) and (3.6). More specifically, we need to investigate solutions \( a_n \) and \( b_n \) of

\[
m \cdot R_N(a_n) = 1 \quad m \cdot F_N(b_n) = 1
\]

for \( n \to \infty \).

One way to get around these difficulties is to consider special classes of distributions for which \( F_N(x) \) can be computed. We investigate three kinds of distributions, namely \( \text{gamma}(\beta, \lambda) \), normal \( N(\mu, \sigma) \) and uniform \( U(0, 1) \) as representatives of type II, III and I distributions (see (3.4a)-(3.4c)). It is well known that the sum of \( N \) i.i.d. gamma distributions \( \text{gamma}(\beta, \lambda) \) and normal distributions \( N(\mu, \sigma) \) are \( \text{gamma}(N\beta, \lambda) \) and \( N(N\mu, \sigma\sqrt{N}) \) respectively [FEL71, REN70]. This implies that the distribution function \( F_N(x) \) is

- for the gamma distribution

\[
F_N(x) = \frac{\gamma(N\beta, \lambda x)}{\Gamma(N\beta)}
\]

where \( \gamma(a, x) \) is the incomplete gamma function, that is [AbS64]

\[
\gamma(a, x) = \int_0^x t^{a-1} e^{-t} \, dt
\]

- for the normal distribution

\[
F_N = \Phi \left[ \frac{x - N\mu}{\sigma\sqrt{N}} \right]
\]

where \( \Phi(x) \) is the error function defined as [AbS64]

\[
\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-y^2/2} \, dy
\]

- for the uniform distribution, see Feller [FEL71] for a proof of the following

\[
F_N(x) = \frac{1}{N!} \sum_{k=0}^N (-1)^k \binom{N}{k} (x - k)^N
\]
where \( x^+ = \max \{0, x\} \).

The next problem is to solve asymptotically equations (3.55), which by Proposition (i) constitute an upper bound for \( Z_{\max} \) and a lower bound on \( Z_{\min} \), that is, \( Z_{\max} \leq a_n = R_N^{-1}(1/m) \) and \( Z_{\min} \geq a_n = F_N^{-1}(1/m) \) \((pr.)\). The following Corollary, for which the proof is presented in Section 5.2, is useful in assessing asymptotics of \( a_n \) and \( b_n \).

Corollary. The solutions \( a_n \) and \( b_n \) of equations (3.55) possess the following asymptotic expansions provided \( N \to \infty \) as \( n \to \infty \)

(i) for the gamma \((\beta, \lambda)\) distribution (see (3.4b))

\[
\begin{align*}
a_n &= R_n^{-1}(1/m) = \begin{cases} 
\frac{N}{\lambda} \log N + \frac{N\beta}{\lambda} \log \log N + O(N) & m = O(N!)
\end{cases} \\
&= \frac{N\beta}{\lambda} (1 + o(1)) & m = O(N^d)
\end{align*}
\]

and

\[
\begin{align*}
b_n &= F_n^{-1}(1/m) = \begin{cases} 
\frac{\beta}{\lambda} \left( \frac{N}{\epsilon} \right)^{1-1/\beta} (1 + o(1)) & m = O(N!)
\end{cases} \\
&= \frac{N\beta}{\lambda} (1 + o(1)) & m = O(N^d)
\end{align*}
\]

(ii) for the normal distribution \( N(\mu, \sigma) \) (see 3.4b))

\[
\begin{align*}
a_n &= N\mu + \sigma \sqrt{N} \sqrt{2 \log m - \log \log m + O(1)} \\
b_n &= N\mu - \sigma \sqrt{N} \sqrt{2 \log m - \log \log m + O(1)}
\end{align*}
\]

(iii) for the uniform distribution \( U(0, 1) \)

\[
\begin{align*}
a_n &= \begin{cases} 
N - 1 & m = O(N!)
\end{cases} \\
&= \frac{N}{2} & m = O(N^d)
\end{align*}
\]

and

\[
\begin{align*}
b_n &= \begin{cases} 
1 & m = O(N!)
\end{cases} \\
&= \frac{N}{2} & m = O(N^d)
\end{align*}
\]
Equipped with our Corollary and Proposition, we are ready to present the main result of this subsection. The Corollary solves, in some sense, the two difficulties (1) and (2) mentioned at the beginning of this subsection. Nevertheless, in the Corollary we restrict our interest only to a special class of distributions for which \( F_N(x) \) is known. However, there is a possibility to obtain the leading factor in the asymptotics of \( a_n \) and \( b_n \) (i.e., upper and lower bounds on \( Z_{\text{max}} \) and \( Z_{\text{min}} \)) without knowing \( F_N(x) \). Indeed, let \( \mu \) and \( \sigma^2 \) be the average and the variance of the weight \( w_i(\alpha) \) that is, \( \mu = E w_i(\alpha) \) and \( \sigma^2 = \text{var} w_i(\alpha) \). Rewriting our combinatorial optimization problem (2.1) as

\[
Z_{\text{max}} = N \mu + \sigma \sqrt{N} \max_{\alpha \in B_n} \left\{ \sum_{i \in S_{\alpha}(\alpha)} \frac{w_i(\alpha) - N \mu}{\sigma \sqrt{N}} \right\} 
\]

one shows immediately that the expression in the square brackets tends to the normal distribution, because of the famous Central Limit Theorem [FEL71]. It is very tempting to draw quick conclusions regarding the solutions \( a_n \) and \( b_n \) of (3.55), but one needs to note that these solutions depend also on the rate (error) in which our expression in (3.64) tends to the normal distribution. Therefore, we may or may not use the central limit theorem to approximate the solutions of (3.55) (the error might be comparable with \( m^{-1} \) and the solution (3.55) might give an incorrect answer). Details are provided in Section 5, while here we present our main findings.

In Section 5.3 we shall prove that the Central Limit Theorem approach can be adopted to solution of (3.55) if \( m \) is polynomially related to \( N \), that is, for some \( d, m = O(N^d) \). Therefore, in this case for \( \mu \neq 0 \), Proposition (i), Corollary (ii) and (3.64) imply that \( Z_{\text{max}} \leq N \mu (1 + o(1)) \) and \( Z_{\text{min}} \geq N \mu (1 + o(1)) \) (pr.) On the other hand, the following lower (upper) bound for \( Z_{\text{max}} \) (\( Z_{\text{min}} \)) is easy to justify. Let \( X_\alpha = \sum_{i \in S_{\alpha}(\alpha)} w_i(\alpha) \) and note that \( X_k \leq \max_{\alpha \in B_n} \{X_\alpha\} \). Therefore, by Strong Law of Large Numbers [FEL71], we obtain
and finally $Z_{\text{max}} \geq N \mu \ (a.s.)$. In conclusions, we formulate our main result in the form of the following theorem.

Theorem 3. Let $m$ be polynomially related to $N$, that is, for some $d > 0$ we have $m = O(N^d)$, and $N \to \infty$ with $n \to \infty$. If the average weight $\mu = Ew \neq 0$, then for large $n$ the following holds

$$Z_{\text{min}} = N\mu(1 + o(1)) = Z_{\text{max}} \ (pr.)$$

(3.65)

In fact, the lower (upper) bound for $Z_{\text{max}}$ ($Z_{\text{min}}$) holds almost surely. ■

Note that for $\mu = 0$ the leading factor in $a_n$ and $b_n$ is $\sqrt{2N \cdot d \cdot \log N}$, and only bounds can be established. More precisely, under the hypothesis of Theorem 3 and $\mu = 0$, $Z_{\text{max}} \leq \sigma \sqrt{2N \cdot d \cdot \log N} (1 + o(1))$ and $Z_{\text{min}} \geq -\sigma \sqrt{2N \cdot d \cdot \log N} (1 + o(1)) \ (pr.)$. Some further consequences of Theorem 3 will be discussed in the next section, which presents applications of our results to some optimization problems discussed in Section 2.

4. APPLICATIONS

In this section we apply our main results from Section 3 to solve some of the problems discussed in Section 2. For simplicity and clarity of our further discussions, we derive our results only for $Z_{\text{max}}$, while $Z_{\text{min}}$ is obtained by analogy. Throughout this section we assume type II (II') and III (III') of distribution functions unless explicitly stated otherwise.

PROBLEM 1. Linear Assignment Problem

In this case $m = n!$ and $N = n$. We assume that the distribution function of weights $\{a_{ij}\}_{i,j=1}^n$ is either of type II or of type III. Then, Proposition (i) and (ii) imply $Z_{\text{max}} \leq R^{-1}(1/n!) \ (pr.)$ and $Z_{\text{max}} \leq nR^{-1}(1/n) \ (a.s.)$ respectively. A lower bound follows from
the greedy approach discussed in Proposition (v) (see also [BOR62]). The greedy algorithm first finds the maximum element in the first column, and then deletes all elements from this column and the row containing this element. Next, the second column is examined, that is, \( n - 1 \) elements in this column exempting the row found in the first step of the algorithm. The maximum element is found and the column as well as the row containing this element are disregarded. The process continues up to the last column. Therefore, referring to notation in Proposition (v), we have \( |A(k)| = n - k, N = N', \) and finally

\[
\sum_{k=1}^{n} R^{-1}(1/k) \leq Z_{\text{max}} \leq R^{-1}(1/n!) \quad (pr.)
\]

(4.1)

Recall, that \( F_1(x) = 1 - R_n(x) \) represents the \( n \)-th convolution of weight distribution functions. The RHS of (4.1) can be replaced by \( nR^{-1}(1/n) \) (a.s.), and for distribution functions satisfying condition (3.18b) of Main Lemma, the almost sure convergence also holds in the LHS of (4.1).

In particular, for the gamma distributions \( \text{gamma}(\beta, \lambda) \) and the normal distributions \( N(\mu, \sigma) \) the following almost sure convergence can be easily derived from the above and our Corollary

\[
Z_{\text{max}} = (1 + o(1)) \frac{n}{\lambda} \log n \quad (a.s.)
\]

(4.2a)

\[
Z_{\text{max}} = n\sigma^2 \log n + n\mu + o(n \log n) \quad (a.s.)
\]

(4.2b)

For type II' and III' distributions, similar results can be found. In particular, for the negative exponential distribution and the normal distribution, one proves \( Z_{\text{min}} = -n \log n \) and \( Z_{\text{min}} = -n\sqrt{2} \log n \) (a.s.) respectively. These asymptotics cannot be, in general, extended to distribution functions of type I. For example, using Main Lemma (i), we can only show that for the uniform distribution \( 1 \leq Z_{\text{min}} \leq H_n \) (pr.), where \( H_n \) is the \( n \)-th harmonic number [KNU73]. Using some other ingenious methods, Walkup [WAL79] and Karp [KAR8] have shown that \( Z_{\text{min}} \leq 3 \) and \( Z_{\text{min}} \leq 2 \) respectively. For some more results on the linear assignment problem similar to ours, see also [FHR87].
PROBLEM 2. *Traveling Salesman Problem*

Let $G_n$ be a complete directed graph built on $n$ vertices. Then $|B_n| = (n - 1)!$ and $N = n$. In particular, the upper bounds for the assignment problem hold also in this case. A lower bound can be either derived from some known results on random (unweighted) graphs (see Proposition (iv)) or from the greedy approach (see Proposition (v)). The former method is based on the fact that for $p \geq c \log n/n$, $c > 1$, every random graph almost surely contains a Hamiltonian circuit [BOL85]. Then, by Proposition (iv) $Z_{\max} \geq nR^{-1}(c \log n/n)$ (a.s.) for $c > 1$. This approach, however, is not constructive. In the greedy method, we show how to build a solution that approximates in a probability sense the optimal path. The greedy algorithm works as follows: we start with a vertex, and select an outgoing edge with the maximum weight. This leads to another vertex, and the algorithm again selects a maximum weighted edge that does not form a cycle with the previously chosen edges. This means that at this stage of the algorithm the maximum is taken over $n - 1$ outgoing edges. In the third vertex and the others, we repeat the algorithm and at the $k$-th step a maximum weighted edge, out of $n - k$ outgoing edges, is found provided a cycle is not created. Finally, at the last $n$-th step, we select an edge from the last vertex to the first one so that the path is closed. In terms of our notation in Proposition (v), we have $N' = 1$, $|A(k)| = n - k$, so $Z_{\max} \geq \sum_{k=1}^{n} R^{-1}(1/k) + \mu$ (pr.), where $\mu$ is the average weight, and the distribution function is either type II or type III. In particular, the above greedy algorithm builds the near-optimal solution in the case of the gamma and the normal distributions. The asymptotic optimal values of the objective function $Z_{\max}$ are the same as in the assignment problem, and given in (4.2). For type I distribution (e.g., uniform), a separate analysis is required. The reader is referred to Karp [KAR77] and Steele [STE78] for a solution of the Euclidean traveling salesman problem with the uniform distribution of points in a unit square (see also [LLK85]). Finally, we note that our bounds, except the greedy one, work not only for complete graphs. The complete graph assumption is necessary only for the greedy
approximation to assure the existence of a feasible solution. However, even in this case the complete graph assumption can be relaxed in many cases, and then a 'patching' algorithm suggested by Karp [KAR77] needs to be applied in order to construct a circuit. Finally, we note that for \( m = O(n^d) \) Theorem 3 implies that \( Z_{\text{max}} - Z_{\text{min}} = \mu n \).

PROBLEM 3. Spanning Trees

A complete directed graph is assumed again, so \( |\mathcal{B}| = n^{n-2} \) [GoJ83] and \( N = n - 2 \). In particular, all conclusions from the traveling salesman problem and the assignment problem can be easily transferred to this case. In particular (4.2) holds. However, there is a difference between this and the other problems. Namely, it is well known that Kruskal's optimal algorithm for finding minimum (maximum) spanning tree is a greedy algorithm [AHU74]. This fact is crucial in deriving the exact value of the objective function in the case of the uniform distribution weights. In 1985 Frieze proved that \( Z_{\text{min}} = \zeta(3) \approx 1.202 \) [BOL85] where \( \zeta(z) \) is the zeta function [AbS64].

PROBLEM 4. The Optimal Weighted k-Clique

This is an interesting problem since \( N = C_k^2 \) and for bounded \( k \) the size \( N \) of a feasible solution is also bounded. Therefore, most of our methods do not give precise estimates of the asymptotic behavior, and as we shall see, a greedy algorithm does not perform asymptotically as good as the optimal solution. Following Lueker [LUE81], we first note that a random (unweighted) graph possesses almost surely a \( k \)-clique if the probability of an edge \( p \) is not smaller than \( n^{-2/(k-1)} + \epsilon \), where \( \epsilon > 0 \). So Propositions (i) and (iv) imply

\[
(a.s.) \quad C_k^2 \cdot R^{-1}(n^{-2/(k-1)}) \leq Z_{\text{max}} \leq R_{C_k^2}^{-1}(1/C_k^2) \quad (pr.) \tag{4.3}
\]

We recall that \( R_{C_k^2}(\cdot) \) represents the \( C_k^2 \) convolutions of the weight distributions. A solution to the RHS of (4.3) is required for every particular distribution (our Corollary and Theorem 3 can-
not be applied since \( N \) is bounded as \( n \) grows to infinity). In particular, the following can be derived from (4.3):

- for the gamma distribution \( \text{gamma} (\beta, \lambda) \)
  
  \[ Z_{\text{max}} - k \log n \quad \text{(pr.)} \]  
  \[ Z_{\text{max}} = k \beta \log (\lambda n) \]  
  \[ Z_{\text{max}} = k \beta \log n \quad \text{(pr.)} \]  

- for the normal distribution \( N(\mu, \sigma) \)
  
  \[ Z_{\text{max}} - k \sigma \sqrt{(k - 1) \log n} \quad \text{(pr.)} \]  
  \[ Z_{\text{max}} = k \sigma \sqrt{(k - 1) \log n} \]  
  \[ Z_{\text{max}} = k \sigma \sqrt{(k - 1) \log n} \quad \text{(pr.)} \]  

- for the uniform distribution \( U(0, 1) \)
  
  \[ \left( \frac{k! c_k^2}{1/(C_k^2 + 1)} \right) n \leq E Z_{\text{min}} \leq C_k^2 n \left( \frac{2}{k - 1} \right)^{2(k - 1)} \]  
  \[ Z_{\text{min}} = \frac{k! c_k^2}{1/(C_k^2 + 1)} \]  

In the case of the uniform distribution \( Z_{\text{min}} \) is shown rather than \( Z_{\text{max}} \), but of course every result from the above can be easily translated from \( Z_{\text{max}} \) to \( Z_{\text{min}} \) and vice versa. We also point out that LHS of (4.4c) was directly derived from Main Lemma (i).

Now we investigate the greedy approach. One can invent at least three different greedy approaches, as shown below.

**Greedy 1.**

begin
  select any vertex
  take \( k \) edges with the largest (lowest) weights of the chosen vertex
  add \( C_k^2 - k \) remaining edges
end

An almost equivalent algorithm is the following one.

**Greedy 2.**

begin
  select any vertex
  do for \( i = 1 \) to \( k \)
    for the \( i \)-th vertex select maximum weighted edge not chosen in the previous steps
  end
  add \( C_k^2 - k \) remaining edges
end

The last greedy algorithm is the most sophisticated and gives the best approximation.
Greedy 3.
begin
select an edge with the maximum weight
for \( i = 2 \) to \( k \)
select \( i \) edges of the maximum total weight and add them to subgraph formed in the first \( i-1 \)st steps
end
end

The first and the second greedy algorithms give the same asymptotic performance. In particular, for the second algorithm \( Z_{\text{grad}} \sim \sum_{i=0}^{k} R^{-1}(1/(n-i)) + (N-k)\mu \). For the gamma distribution one finds \( Z_{\text{grad}} \sim k \log n - Z_{\text{max}} \), so the greedy gives the same performance as the optimal solution. This is not, however, any longer true for the normal distribution. The greedy algorithm gives in this case \( Z_{\text{grad}} \sim k \sqrt{2} \log n \). An improvement can be achieved if one applies the third greedy algorithm. To derive an appropriate asymptotic formula, we first note that the value \( Z_{\text{grad}} \) of the objective function is equal in this case to

\[
Z_{\text{grad}} = \sum_{i=1}^{k} \max_{1 \leq j \leq n-i} \sum_{l=1}^{i} w_{lj}
\]

(4.5)

Let us assume that the weights are normally distributed \( N(0,\sigma^2) \) with zero mean. Then, \( S_i = \sum_{l=1}^{i} w_{lj} \) is distributed as \( N(0,i\sigma^2) \), and therefore by our previous results,

\[
\max_{1 \leq j \leq n-i} S_i - \sigma \sqrt{i} \sqrt{2} \log n.
\]

So, after \( k \) steps, one proves

\[
Z_{\text{grad}} \sim (2 + \sum_{i=2}^{k-1} \sqrt{2i} \sqrt{i} \log n)
\]

(4.6)
which agrees with Lueker’s result [LUE81].

Finally for the uniform distribution, we apply Greedy 3 for finding the lightest \( k \)-clique. In the first step, we find the lightest weight out of \( n^2 \) weights, therefore its value is \( 1/n^2 \) [GAL87]. Next, we note that \( S_i = \sum_{l=1}^{i} w_{lk} \) is distributed as \( i \) convolutions of uniform distributions, hence the distribution function \( F_i(x) \) of \( S_i \) is given by (3.58). Applying Main Lemma (i) we find out that \( b_n \), which is a solution of \( (n-i)F_i(b_n) = 1 \) is equal to \( b_n \sim (i/n)^{1/i} \). After
some further algebra we come to the following solution

\[
EZ_{\text{avg}} - \frac{1}{n^2} + \sum_{l=1}^{k-1} \frac{l^\frac{1}{n}}{l+1} \left[ 1 - \frac{1}{l+1} \right]
\]

(4.7)

which should be compared with (4.4c).

PROBLEM 5. **Height of a Trie With Independent Keys**

In Section 2 we have argued that the height of a digital tree (i.e., trie) with \( n \) independent keys, is maximum over \( m = n(n-1)/2 \) (dependent) alignments \( C_{ij} \) (cf. (2.2)) each being geometrically distributed as reported in (2.3). Therefore, straightforward implementation of Proposition (i) implies that \( H_n \leq -2 \log P \) (pr.) where \( P = p^2 + q^2 \). It can be proved that this bound holds also almost surely [SZP86], since \( H_n \) is a monotonically increasing sequence. To prove the lower bound, namely \( H_n \geq -2 \log P \), we need a little more intricate approach, the one established in Proposition (iii). The required inequality (i.e., mixing condition) derives from [PTT85, SZP86], and it becomes for binary trees

\[
P \{H_n \leq r\} \leq n \left[ \frac{2^r}{n} \right]^n \leq 2^{n-r} [1 - F(r)]^n
\]

(4.8)

where \( F(r) = Pr \{C_{ij} \leq r\} = 1 - P^{r+1} \), so \( R(r) = F(r) = P^{r+1} \). If \( a_n = R^{-1}(1/n^2) \), then the following estimates is a consequence of the above and (4.8)

\[
Pr \{H_n \leq (1 - \varepsilon) a_n\} \leq 2^{2n(1 - \varepsilon) \log n} \cdot n^{-2n(1 - \varepsilon)}
\]

(4.9)

The last inequality implies \( H_n \geq - (1 + o(1)) \log P \) (a.s.) by appealing to Borel-Cantelli Lemma [FEL71]. This proves also that \( H_n \sim (2/\log P^{-1}) \cdot \log n \) as required (see also [SZP86]).

PROBLEM 6. **Maximum Queue Length**

This problem is discussed in Example 2.6. In particular, we have shown there that tails of the (asymptotic) distributions for the queue length \( Q_n \) and for the waiting time \( W_n \) are given by
This and Proposition (i) immediately prove that $Q_{\text{max}} \leq (1 + o(1)) \log \omega n^{-1}$ (pr.) and $W_{\text{max}} \leq (1 + o(1)) \log n^{1/9}$ (pr.) where $\omega$ and $0$ are defined in Example 2.6 (see (2.5)). For a lower bound, we need a different approach which resembles our idea from Proposition (iv). Let $L_n$ denote a random number of busy periods completed just prior to the arrival of the $n$-th customer, and by $U_k$, $1 \leq k \leq L_n$ we denote the queue length seen by a random customer arriving in the $k$-th busy period. Naturally, $Q_{\text{max}} \geq \max_{1 \leq k \leq L_n} U_k$. But, $\{U_k\}_{k=1}^{L_n}$ are independently identically distributed (i.i.d.) random variables with the tail distribution satisfying (2.4). Indeed, we note that $\Pr\{U_k = t\} = \Pr\{Q_k = t, Q_k > 0\} = \Pr\{Q_k = t\}/\Pr\{Q_k > 0\}$, so (2.4) holds. Some difficulties arise because the maximum in $\max_{1 \leq k \leq L_n} U_k$ is taken over a random number $L_n$ of periods. But, it can be easily proved that $L_n \sim \alpha n$ (a.s.) [IGL85, SZP88b] where $\alpha$ is a constant, and therefore, it can be further shown that one may apply our Main Lemma (ii) (3.12) replacing $n$ by $\alpha n$ (see for details [GAL87, SZP88b]). This implies $Q_n \sim \log \omega (n \alpha)^{-1}$ (pr.), and hence with the upper bound discussed above, we prove that $Q_n \sim \log n^{-1}$ (pr.) and $W_n \sim \log n^{1/9}$ (pr.).

PROBLEM 7. Bottleneck and Capacity Assignment Problems

The bottleneck and capacity assignment problems are discussed in Example 2.7 and a general solution to these problems is presented in Theorem 2. In particular, Theorem 2 implies that for the bottleneck assignment problem $Z_{\text{min}} \sim F^{-1}(\log n/n)$ (pr.) and for the capacity assignment problem $Z_{\text{max}} \sim F^{-1}(1 - \log n/n)$ provided that the distribution weight function $F(.)$ is strictly increasing. The same conclusions hold for the bottleneck and capacity traveling salesman problems on complete directed graphs.

PROBLEM 8. Suffix Tree

The suffix tree problem is discussed in Example 2.10. Our purpose is to find an upper bound on the average height $Eh_n$ of a suffix tree built from $n$ suffixes of a random string. This
problem is significantly more difficult than the others, since the self-alignments are not only dependent but also nonidentically distributed. In addition, the keys are strongly dependent, and therefore an approach taken in digital trees with independent keys (cf. Problem 5), cannot be easily extended to formulate a reasonable lower bound. In short, we restrict ourselves to finding an upper bound on the average $E H_n$ of the height in a suffix tree.

Applying Main Lemma (i) (cf. (3.7)) one shows that

$$E H_n \leq a_n + \sum_{j=a_n}^{\infty} \sum_{d=1}^{n} (n-d)R_d(j)$$

where $a_n$ is a solution of the following equation

$$\sum_{d=1}^{n} (n-d)R_d(a_n) = 1$$

(4.11)

and the distribution function $F_d(k) = 1 - R_d(k)$ is given by (2.8). Formula (4.11) is a consequence of the fact that the set of self-alignments $\{C_d\}_{d=1}^{n-1}$ can be grouped into $(n-1)$ alignments of type $C_1$, $(n-2)$ of type $C_2$, ..., and one of type $C_{n-1}$. The solution to (4.11) can be upper bounded by a solution to a simpler equation given below (for details see [ApS88])

$$m(p \cdot a + q \cdot a) = 1$$

(4.12)

where $m = n^2$ and $a_n \leq \bar{a}_n$. The asymptotic solution to (4.12) can be easily obtained and one proves

$$\bar{a}_n = 2(1 + o(1)) \log_{p_{\text{max}}} n^{-1}$$

(4.13)

where $p_{\text{max}} = \max\{p, q\}$. Finally, to complete our analysis, we need to evaluate the second term in (4.10), that is, in our case $\sum_{j=a_n}^{\infty} \sum_{d=1}^{n} (n-d)R_d(j)$. Using (4.11) and the bound (see [ApS88] for details)

$$R_d(k) \leq (p^{f+1} + q^{f+1})^d$$

where $f = \lfloor \frac{k}{d} \rfloor$ and $\lfloor \cdot \rfloor$ is the floor operation, we prove
\[ \sum_{j=1}^{n} \sum_{d=1}^{n} (n-d)R_d(j) = \sum_{k=0}^{\infty} \sum_{d=1}^{n} (n-d)R_d(a_n + k) = O \left( \sum_{d=1}^{\infty} (n-d)R_d(n) \sum_{k=0}^{\infty} p_{d,k}^k \right) = O(1) \]

Hence, by the above and (4.13) we finally obtain

\[ EH_n \leq \frac{2}{\log P^{-1}} \log n + c \quad (4.14) \]

where \( c \) is a constant. Note that in the symmetric case, all self-alignments \( C_{ij} \) are identically distributed (e.g., set \( p = q \) in (2.8)), and then one may show that \( EH_n \sim 2 \log n \). Surprisingly enough, simulation results show that the average height of a suffix tree has the same leading factor in an asymptotic expansion as the average height of a digital trie with independent keys, that is, by the solution to Problem 5 we have \( EH_n \sim (2/\log P^{-1}) \log n \). It is, however, an open problem to prove this assessment.

The consequences of these results are discussed in detail in [ApS88]. Here we only point out that (4.14) suggests that a direct, natural construction of a suffix tree (that is, by consecutive insertion of suffixes) takes \( O(n \log n) \) time on the average, while the rather sophisticated method of Weiner [AHU74] takes \( O(n) \) time. However, the latter uses much more complicated data structure. On the other hand, using this direct construction, one can prove that computing the full statistics without overlap of all substrings of a word takes \( O(n \log n) \) expected time, while a more sophisticated method oriented on the worst case analysis, takes \( O(\log^2 n) \), and so on [ApS88].

**PROBLEM 9. Location Problem on Graphs**

We do restrict our investigation to a location problem on graphs. The objective function is expressed in (2.11), and for reader's convenience we repeat it here again

\[ Z_{\text{max}} = \max_{\alpha \in \mathcal{B}_n} \sum_{i \in M - \alpha} \max_{j \in \alpha} w_{ij} \quad (4.15) \]

where a feasible solution \( \alpha = (c_1, c_2, \ldots, c_L) \) consists of \( L \) selected vertices in the graph. Naturally, \( |\mathcal{B}_n| = \binom{n}{L} \sim n^L/L! \) for bounded \( L \). Let us define \( W(\alpha) = \max_{j \in \alpha} w_{ij} \), and we note
that \( W_i(\alpha) \) is i.i.d., random sequence. Since the weights \( w_{ij} \) are i.i.d. random variables with a distribution function \( F(\cdot) \), then \( W_i(\alpha) \) is easy to estimate. In particular, the distribution function \( F_W(x) \) of \( W_i(\alpha) \) is equal to \( F^L(x) \), since \( \alpha \) has cardinality \( L \). The average value \( EW \) of \( W_i(\alpha) \) is rather easy to evaluate in most of the interesting cases. For example, if the weights are exponentially distributed, then \( EW = H_L \) [GAL87] where \( H_L \) is the \( L \)-th harmonic number [KNU73]. Next, we note that \( m = |B_n| = n^L / L! \), so \( m \) is polynomially related to \( n \), and one may consider applying Theorem 3, since \( N = |M - \alpha| = n - L \rightarrow \infty \) as \( n \rightarrow \infty \). In addition, for most weight distribution functions \( EW \neq 0 \) even if the average weight \( \mu \) is equal to zero. In summary, by Theorem 3 we obtain

\[
Z_{\text{max}} - Z_{\text{min}} - (n - L)EW + \sigma W \sqrt{2nL \log n} - (n - L)EW
\]  

(4.16)

In particular

\[
Z_{\text{max}} - Z_{\text{min}} - \frac{(n - L)}{\lambda} H_L
\]  

(4.17)

for the exponential distribution of weights. □

This section has discussed only a selected number of possible applications of our main results. We can envision many more applications of our methodology. In particular, we are working towards estimating a measure of the homology between two or more (biological) sequences represented on a directed acyclic graph.

5. REMAINING PROOFS

In this section we present some of the remaining derivations. In particular, we prove our Main Lemma and discuss some aspects of the Proposition (v). Then, we turn attention to our Corollary, and finally we prove Theorem 3.

5.1 How to Prove Main Lemma and Proposition

The Main Lemma is the heart of all our derivations and we show now how it can be
proved. We restrict our considerations to $Z_{\text{max}}$ since $Z_{\text{min}} = - \max_{1 \leq k \leq n} \{-X_k\}$.

We start with proof of (i). Note that $Z_{\text{max}}^r = \max\{X_1^r, X_2^r, \ldots, X_n^r\}$, hence $Pr \{X_1^r < x\} = Pr \{X_i < x^{1/r}\} = G(x^{1/r})$. Then, as in Lai and Robbins [LaR76] one easily shows that

$$Z_{\text{max}}^r \leq a + \sum_{k=1}^{n} (X_k^r - a)^+$$  \hfill (5.1)

where $a$ is any number and $x^+ = \max(0, x)$. Taking the expectation of (5.1), we obtain

$$EZ_{\text{max}}^r \leq a + \sum_{k=1}^{n} \int_{a}^{\infty} [1 - G(y^{1/r})] dy$$  \hfill (5.2)

Finally, minimization of the RHS of (5.2) with respect to $a$ leads to $\tilde{a}_n$ as defined in (3.9), and thus proves Main Lemma part (i).

We now turn to a proof of Main Lemma part (ii) Let $R_k(x) = 1 - G_k(x) = Pr \{X_k > x\}$. Then one shows [FEL71]

$$Pr \{M_n > r\} = Pr \{X_1 > r \text{ or } X_2 > r \text{ or } \cdots \text{ or } X_n > r\} \leq \sum_{k=1}^{n} R_k(r).$$  \hfill (5.3)

Let $r = a_n(1 + \epsilon)$ with $\epsilon > 0$, and note that condition (3.1b) and (3.1c) for Type II distributions implies $R_k(a_n(1 + \epsilon)) = o(1)R_k(a_n)$ uniformly in $k$, so from (5.3) we obtain

$$Pr \{Z_{\text{max}} > (1 + \epsilon)a_n\} = o(1) \sum_{k=1}^{n} R_k(a_n) = o(1)$$

where the last equality follows from the definition of $a_n$ given in (3.5). This proves (3.10a). To show (3.12) we postulate, in addition, that inequality (3.11a) holds. Then

$$Pr \{Z_{\text{max}} \leq r\} = Pr \{X_1 \leq r, X_2 \leq r, \ldots, X_n \leq r\} \leq \delta G_1(r) \cdots G_n(r)$$  \hfill (5.4)

for $\delta = O(1)$. It is more convenient to deal with the logarithm of (5.4). We first note that for $\nu \to 1$ [GAL87] $\log \nu = \log[1 - (1 - \nu)] \leq \nu - 1$, hence from (5.4)

$$\log Pr \{Z_{\text{max}} < r\} \leq - \sum_{k=1}^{\infty} [1 - G_k(r)]$$  \hfill (5.5)
Let \( r = (1 - \varepsilon)a_n, \varepsilon > 0 \). Then to prove (3.12), we need to show that the RHS of (5.5) tends to \(-\infty\) when \( n \to \infty \). But substituting in (3.16) \( x = z/c \) for \( c > 1 \) one finds \( 1 - G_k(x) = o(1)(1 - G_k(z/c)) \). Let \( 1/c = 1 - \varepsilon < 1 \), then by the definition of \( a_n \) given in (3.5) and the above, one shows

\[
\sum_{k=1}^{n} [1 - G_k(1 - \varepsilon)a_n]] = \frac{1}{o(1)} \sum_{k=1}^{n} [1 - G_k(a_n)] = \frac{1}{o(1)} \tag{5.6}
\]

so this proves that \( \log P_r\{Z_{\text{max}} < (1 - \varepsilon)a_n\} \to -\infty \), and this implies \( \lim_{n \to \infty} Z_{\text{max}}/a_n \geq 1 \) (pr.), hence together with (3.10a) this completes the derivation of (3.12). It only remains to prove the convergence in mean (3.14). But, by the Mean Convergence Theorem [BIL86], to prove (3.14) it suffices to show that \( Z_{\text{max}}/a_n \) is uniformly integrable [BIL86]. Under hypothesis of our Lemma, this can be easily shown, and the reader is referred to Lai and Robbins [LaR77, pp. 103-106] for a detailed proof.

Finally, part (iii) is shown in the same manner as part (ii), and it is left to the reader. We only hint that the crucial observation follows from the definition (3.3a) of Type III distribution functions and Definition 2 (cf. 3.1c). Namely, for \( a_n \) defined in (3.5) and any \( c > 0 \), one notes that \( R_k(a_n + c) = o(1)R_k(a_n) \). The rest of the proof follows the same line of arguments as in part (ii). At last, part (iv) of Main Lemma is proved in [GAL87].

Now, few comments on our Proposition are in the sequel. In most cases the Proposition was "smoothly" derived in Section 3 from our Main Lemma. Only part (v), formula (3.45) needs some explanations. Recall that the value \( Z_{\text{grad}} \) of the objective function for the greedy algorithm is given in (3.35a), and we again repeat the formula below

\[
Z_{\text{grad}} = \sum_{i \in S_n} \max_{j \in A(i)} w_{ij} + \sum_{i \in S_n - S_n'} w_i(a_{\text{grad}}) \tag{5.7}
\]

where \( 1S_n' = N' \) and \( 1S_n = N \). The above two terms of (5.7) need to be evaluated separately. With respect to the first term, we note that the weights \( w_{ij} \) are i.i.d., so our Main Lemma can be applied provided the distribution function of weights is either of Type II or III. In particular, we
note that $\max_{j \in |A(i)|} w_{ij} \sim R^{-1}(|A(i)|^{-1})$ (pr.) assuming $|A(i)|$ is large. But for large $i$, the cardinality of $|A(i)|$ is small (e.g., in most cases $|A(i)| = n - i$). This, however, should not bother us, since under assumption (3.1a) of Definition 1 $R^{-1}(|A(i)|^{-1}) \to \infty$ as $|A(i)| \to \infty$. Therefore, for large $n$, the contribution of $R^{-1}(|A(i)|^{-1})$ when $|A(i)|$ is small, is negligible.

So far, we have shown that $\max w_{ij} \sim R^{-1}(|A(i)|^{-1})$ (pr.), but the first term of (5.7) needs to sum up all such asymptotics. Two problems arise. Firstly, note that $a_n - b_n$ does not apply that $\sum a_n - \sum b_n$ [BEN74], however, the last implication holds if $a_n - b_n$ holds uniformly in $n$ [BEN74]. Since the maximum operator in $\max \{w_{ij}\}$ is applied to random variables $w_{ij}$ distributed in the same fashion and $|A(i)|$ decreases monotonically to 1, hence the uniform convergence holds as needed. Secondly, we need to know whether for random variables $X_n \sim Y_n$ (pr.) implies $\sum X_n - \sum Y_n$ (pr.). The answer is affirmative [BIL86] and this finally settles the first term in (5.7). Naturally, the second term of (5.7) tends in probability to $(N - N')\mu$ by the Weak Law of Large Numbers [FEL71], so from our discussion, we see that Proposition (v) formula (3.45) holds.

Finally, we point out that in our Proposition the distribution function $F_N(\cdot)$ of $\sum_{\alpha \in A(\alpha)} w_\alpha(\alpha)$ must be of Type II (II') or III (III'). This might cause some problems, and a question arises whether a condition on the weight distribution function $F(\cdot)$ is not sufficient. In most cases, the answer is positive, and below we present one result of this sort for $\Xi = \Sigma$.

**Property:** If $F(\cdot)$ is of Type II, and $F_N(\cdot)$ represents convolution of $N$ distributions of $F(\cdot)$, then

$$\lim_{x \to \infty} \frac{1 - F_N(cx)}{1 - F(x)} = 0 \text{ for all } c > 1$$

(5.8)

that is, $F_N(\cdot)$ is of Type II, too.

**Proof.** To prove (5.8), note that

$$1 - F_N(cx) = Pr \{ Y_1 + Y_2 + \cdots + Y_N > cx \} \leq N[1 - F(cx)]$$
\[ 1 - F_N(x) = \Pr\{Y_1 + Y_2 + \cdots + Y_N > x\} \geq \Pr\{Y_1 > x\} = 1 - F(x) \]

therefore, (5.8) follows immediately from (3.1b) and the above. \(\blacksquare\)

5.2 The corollary is easy to prove

In this subsection, we prove our Corollary in the case of \(m = O(N!)\). The corollary in the case of \(m = O(N^d)\) follows directly from the normal approximation discussed in Theorem 3 (see also next subsection for a formal proof). We start with the gamma distribution with parameters \(\beta\) and \(\lambda\). Assuming \(|S_n(\alpha)| = N = n\), we note that the sum of \(n\) i.i.d. gamma distributions \(\text{gamma}(\beta, \gamma)\) is \(\text{gamma}(n\beta, \gamma)\). Then, simple arguments lead to the following formulas on the distribution function \(F_n(x)\) and the reliability function \(R_n(x)\) of the \(\text{gamma}(n\gamma, x)\) distribution:

\[
F_n(x) = \frac{\gamma(n\beta, \lambda x)}{\Gamma(n\beta)}, \quad R_n(x) = \frac{\Gamma(n\beta, \lambda x)}{\Gamma(n\beta)},
\]

where the incomplete gamma functions \(\gamma(a, x)\) and \(\Gamma(a, x)\) are defined as [AbS64]:

\[
\gamma(a, x) = \int_0^x e^{-t} t^{a-1} dt, \quad \Gamma(a, x) = \int_x^{\infty} e^{-t} t^{a-1} dt
\]

The purpose of our analysis is to derive asymptotic approximations for solutions \(a_n\) and \(b_n\) of equations (3.55), that is \(m R_n(a_n) = 1\) and \(m F_n(b_n) = 1\). These solutions strongly depend on the value of \(m\), and as stated above we restrict our analysis to \(m = n!\). Without loss of generality we assume \(\beta \beta = 1\). At first, we consider \(r_{\text{max}}\), that is, we search for a solution \(a_n\) of \(n! R_n(a_n) = 1\). It is known that for \(n > 1\) and \(x > n - 1\) [AbS64]

\[ e^{-\lambda x} (\lambda x)^{n-1} \leq \Gamma(n, \lambda x) \leq \frac{e^{-\lambda x} (\lambda x)^{n}}{(\lambda x - n + 1)} \]

This, and a rough estimation of \(a_n\) (i.e., \(a_n = O(n \log n)\)), suggest to approximate \(\Gamma(n, \lambda x)\) by the asymptotic formula [AbS64] \(\Gamma(n, \lambda x) \sim e^{-\lambda x} (\lambda x)^{n-1}\) which holds for \(x \to \infty\) and \(n = o(x)\) [AbS64]. Then, the problem lies in solving the following equation \(\lambda \beta = 1\)

\[ e^{-\alpha_n} (a_n)^{n-1} = 1 \]

or equivalently

\[ a_n - (n - 1) \log a_n - \log n = 0 \]  

(5.11)

for large \( n \). Let the LHS of (5.11) be denoted as \( f(a_n) \). We find such \( a_n \) and \( \overline{a}_n \) that \( a_n \leq a_n \leq \overline{a}_n \), that is, \( f(a_n) > 0 \) and \( f(\overline{a}_n) < 0 \). Let

\[ a_n = n \log n + n \log \log n \]

Then

\[ f(a_n) = n \log \log n - (n - 1) \log \log n \log n < 0 \]

for large \( n \). On the other hand, for any \( \varepsilon > 0 \) define

\[ \overline{a}_n = n \log n + \log \log n^{1+\varepsilon} \]

Note that

\[ f(\overline{a}_n) = n \log \log n^{1+\varepsilon} - (n - 1) \log \log n \log n^{1+\varepsilon} > 0 \]

for sufficiently large \( n \). Hence we prove that \( a_n = n \log n + n \log \log n + O(n) \), which establishes Corollary (i) formula (3.59). The proof for \( b_n \) goes the same way, except that \( \Gamma(n,x) \) is replaced by \( \lambda(n,x) \) and the following asymptotic approximation \( \lambda(n,x) \sim \frac{x^a e^{-x}}{n} \) is used for \( n \to \infty \) and \( x = o(n) \) [AbS64] since we know that \( b_n \) is bounded. Details are omitted and left to the reader.

In the proof of (3.61) in Corollary (ii), we use the following representation of our problem

\[ Z_{\text{max}} = N \mu + \sigma \sqrt{N} \left[ \sum_{i \in S_\alpha(x)} \frac{w_i(x) - N \mu}{\sigma \sqrt{N}} \right] \]  

(5.12)

Then, the expression in the parentheses is the standard normal distribution with distribution function \( \Phi(x) \) as defined in (3.57b). Using the following inequality [FEL71]

\[ \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \left[ \frac{1}{x} - \frac{1}{x^3} \right] \leq 1 - \Phi(x) \leq \frac{1}{\sqrt{2\pi}} e^{-x^2/2} \]  

(5.13)
or equivalently \( \Phi(x) = e^{-x^2/2}/(\sqrt{2\pi}) \) for \( x \to \infty \), and applying the same line of arguments as above, we prove Corollary (ii). The proof of the last part of the Corollary for the uniform distribution of the weights is rather simple. We just note that for \( b_n \) we need to solve \( x^2 = 1 \) so \( b_n = 1 \), while \( a_n = n \) solves (3.55).

5.3 Be careful with the normal approximation

We now prove Theorem 3 and for this we need to solve equations (3.55) for large values of \( n \), that is, \( mR_n(a_n) = 1 \) and \( mF_n(b_n) = 1 \) in the case \( \Xi = \Sigma \). It is tempting to apply the Central Limit Theorem and approximate the distribution function \( F_n(x) \) by the standard normal distribution \( N(0,1) \) (see representation (5.12)). However, one must be very careful not applying this approximation to (3.55) blindly. This subsection shows how to cope with the problem, that is, how to prove our theorem.

Let for the purpose of this subsection, \( F_N(x) \) denote the distribution function of

\[
\left[ \sum_{i \in S_n(n)} w_i(x) - N\mu \right] / \sqrt{N}
\]

where \( \mu \) and \( \sigma^2 \) are the average and the variance of the weights \( w_i(x) \) respectively. By the central limit theorem, we know that \( \lim_{N \to \infty} F_N(x) = \Phi(x) \) where \( \Phi(x) \) is the error function given in (3.57). So \( F_N(x) = \Phi(x) + e(x) \) where \( e(x) \) is the error function, and the value of \( e(x) \) is crucial for the solution of \( mR_N(x) = 1 \) and \( mF_N(x) = 1 \) and establishing \( a_n \) and \( b_n \). In other words, to obtain sound approximations of \( a_n \) and \( b_n \) using the central limit theorem approach, the following condition \( e(x) = o(m^{-1}) \) must hold. Indeed, the equation \( mR_N(x) = 1 \) leads to \( \Phi(x) + e(x) = \frac{1}{m} \), and the values of \( e(x) \) can be neglected only if \( e(x) < < m^{-1} \).

From Feller [FEL71] we know that

\[
F_N(x) = \Phi(x) + \phi(x) \sum_{k=3}^{r} N^{-k+1} P_k(x) + o(N^{-r+1})
\]

(5.14)

where \( \phi(x) = e^{-x^2/2}/(\sqrt{2\pi}) \) is the density function for the normal distribution \( N(0,1) \), and \( P_k(x) \) is
a polynomial of degree k, dependent only on the moments of the weights \( w_i(\alpha) \), but not on \( N \) and \( r \), where \( r \) is any integer (the larger \( r \) is the better the approximation is). For practical purposes, we can approximate \( 1 - \Phi(x) \sim e^{-x^2/2}/(x\sqrt{2\pi}) \) [FEL71]. Note also, that the polynomial \( P_k(x) \) does not change significantly the asymptotic solution of (3.55) (since exponential `swallow' polynomials). Therefore, selecting \( r \) such that \( N^{-h}r + 1 = o(m^{-1}) \) is sufficient to obtain valid asymptotics for \( a_n \) and \( b_n \). For instance, \( m = O(N^d) \), \( d \) is a constant, satisfies the condition, hence with the help of Corollary (ii), we prove our Theorem 3.

References


[SZP88b] Szpankowski, W., On the maximum queue length with applications to data structures: A simple but yet asymptotically exact approach, Purdue University, CSD TR-785, 1988.

[SZP88c] Szpankowski, W., On the bottleneck and capacity assignment problems, Purdue University, CSD TR-841, 1988.

