International Institute for Applied Systems Analysis
I I A S A www.iiasa.ac.at

An Epsilon-Approximation Scheme for Minimum Variance Combinatorial Problems

Katoh, N.

IIASA Working Paper

WP-87-117

November 1987

WORKING PAPER

AN ϵ-APPROXIMATION SCHEME FOR MINIMUM VARIANCE COMBINATORIAL PROBLEMS

N.Katoh

November 1987
WP-87-117

IIASA
International Institute for Applied Systems Analysis

An ϵ-approximation Scheme for Minimum Variance Combinatorial Problems

Naoki Katoh

November 1987
WP-87-117

INTERNATIONAL INSTITUTE FOR APPLIED SYSTEMS ANALYSIS
A-2361 Laxenburg, Austria

Foreword

Suppose that we are given a finite set E, a family of feasible subsets of E and an integer cost associated with each element in E. The author considers the problem of finding a feasible subset such that the variance among the costs of elements in the subset is minimized. The author shows that if one can solve the corresponding minimum cost problem in polynomial time, it is possible to construct a fully polynomial time approximation scheme for the above minimum variance problem.

Alexander B. Kurzhanski
Chairman
System and Decision Sciences Program

An ϵ-approximation Scheme for Minimum Variance Combinatorial Problems

Naoki Katoh

1. Introduction

A recent paper by Martello et al. [19] considered the following problem which they call a *"balanced optimization problem"*. Suppose that we are given a finite set $E = \{1,2,\ldots,|E|\}$, a family F of *"feasible subsets"* of E and an integer cost c_j associated with every $j \in E$. The balanced optimization problem is then described as follows:

$$\text{BALANCE}: \underset{S \in F}{\text{minimize}}\ d(S) \equiv \max\{c_j - c_{j'} \mid j, j' \in S\} \quad . \tag{1}$$

In other words, this problem tries to make the difference in value between the largest and smallest costs used as small as possible. [19] showed that if we can efficiently answer the feasibility question then we can efficiently solve Problem BALANCE. [19] also gave real-life examples in which balanced optimizations arise.

We may alternatively consider the variance for another measure of the balance among costs used. The *variance* among costs in $S(\in F)$ is defined by

$$var(S) \equiv \frac{1}{|S|} \sum_{j \in S} (c_j - \frac{1}{|S|} \sum_{j \in S} c_j)^2 \quad . \tag{2}$$

The *minimum variance problem* is then described as follows:

$$P: \underset{S \in F}{\text{minimize}}\ var(S) \quad . \tag{3}$$

We consider in what follows the class of P satisfying the following three assumptions.

(A1) $\qquad |S| = p \quad \text{for all} \quad S \in F \qquad (4)$

holds and that p depends only on $|E|$. We assume that p is a positive integer with $p \geq 2$, since the case of $p = 1$ is trivial.

(A2) F is given in concise form, i.e., all feasible subsets are not listed in advance but they are described through an oracle which says, within polynomial time in $|E|$, whether any given subset of E contains an $S \in F$ or not, and the input length needed for specifying this oracle is polynomial in $|E|$. We call this oracle the ***feasibility oracle*** and the time required to call the oracle (i.e., to test the feasibility of a given subset of E) is denoted by $f(|E|)$.

(A3) For any given subset E' of E and any given real numbers c'_j, $j \in E'$, we can produce $S' \in F$ with $S' \subseteq E'$ in polynomial time in $|E|$ such that S' is optimal to the following minimum-cost problem.

$$\text{minimize} \{ \sum_{j \in S} c'_j | S \in F, S \subseteq E' \} \quad . \tag{5}$$

If there is no $S \in F$ with $S \subseteq E'$, it returns the answer that there is no feasible subset in E'. Since this test is done through the feasibility oracle, the time required to solve (5), which is denoted by $\tau(|E|)$, satisfies $\tau(|E|) \geq f(|E|)$.

The aim of this paper is to propose a fully polynomial time approximation scheme (FPAS) for the above minimum variance problem P under the above three assumptions. Especially, if F is the set of spanning trees in an undirected graph $G = (V,E)$ (such F clearly satisfies (A1)~(A3)), we shall show that there exists an $0(|E|\sqrt{|V|} \cdot \tau(|V|, |E|))$ algorithm for the minimum variance problem, where $\tau(|V|, |E|)$ is the time to solve the minimum cost spanning tree problem.

The techniques we use to develop an FPAS for P satisfying (A1)~(A3) are the parametric characterization for the quasiconcave program developed by Sniedovich [21, 22] and Katoh and Ibaraki [16], and the scaling technique which has been used to develop a fully polynomial time approximation scheme for the knapsack problems (see Lawler [18] for example), polynomial time algorithms for minimum cost circulation problems (see [8, 23] for example) and possibly others.

The parametric characterization of P states that an optimal solution of the parametric problem $P(\lambda)$ defined below provides an optimal solution of P, if an appropriate λ is chosen.

$$P(\lambda) : \underset{S \in F}{\text{minimize}} \sum_{j \in S} c_j^2 - \lambda \sum_{j \in S} c_j \quad , \tag{6}$$

where λ is a nonnegative parameter. Thus, solving P is reduced to finding a $\lambda = \lambda^*$ with which an optimal solution to $P(\lambda^*)$ is also optimal to P. Such characterization can be obtained by specializing the results obtained by Sniedovich [21, 22] and Katoh and Ibaraki [16] to our case. Similar characterization has also been reported (e.g., Kataoka [14], Ishii et al. [11], Ichimori et al. [10], and Katoh and Ibaraki [15] discuss some types of stochastic programs, Kawai and Katoh [17] discusses a type of markovian decision process and Dinkelbach [2] and Jagannathan [13] discuss the fractional program).

This characterization, however, does not tell how to find such λ^*. The straightforward approach for finding λ^* is to compute optimal solutions of $P(\lambda)$ over the entire range of λ. However, the number of such solutions is not polynomially bounded in most cases, e.g., see Carstensen [1]. One of exceptions is that F is the set of spanning trees in an undirected graph. For this case, based on the parametric characterization, a polynomial time algorithm is directly derived, which will be treated in Section **3**.

On the other hand, for example, if F is one of the sets of matchings in a bipartite graph, perfect matchings in an undirected graph or spanning trees in a directed graph, F satisfies (A1)~(A3) and the corresponding $P(\lambda)$ can be solved in polynomial time, but the number of optimal solutions of $P(\lambda)$ over the entire range of λ is not known to be polynomially bounded.

Therefore, in general, polynomial time algorithms for P seem to be difficult to develop, and we then focus on approximation schemes in this paper. A solution is said to be an *ϵ-approximate solution* if its *relative error* is bounded above by ϵ. An *approximate scheme* is an algorithm containing $\epsilon > 0$ as an parameter such that, for any given ϵ, it can provide an ϵ-approximate solution. If it runs in polynomial in both input size and $1/\epsilon$, the scheme is called a *fully polynomial time approximation* (FPAS) [6, 20].

An FPAS for P based on the parametric characterization is obtained by scaling the costs c_j. In other words, we use the costs $\lfloor c_j/_2t \rfloor$ instead of c_j for an appropriately chosen positive integer t and computes optimal solutions of $P(\lambda)$ over the entire range of λ, where $\lfloor a \rfloor$ denotes the largest integer not greater than a. Then it is shown that an ϵ-approximate solution is found among those obtained solutions as the one minimizing $var(S)$. We apply the Eisner and Severence method [4] to solve $P(\lambda)$ with scaled costs over the entire range of λ. With some modifications of their method, the required time is shown to be polynomial in $|E|$ and $1/\epsilon$ under assumptions (A1)~(A3).

An FPAS for the similar problems has been proposed by Katoh and Ibaraki [16]. Though their method is also based on the parametric characterization, it does not employ the scaling technique. In addition, [16] characterizes the class of problems for which their method becomes FPAS. However, our problem P does not belong to this class (especially the condition (A5) given in Section 5 of [16] does not hold for P).

The paper is organized as follows. Section 2 gives the relationship between P and $P(\lambda)$. Based on the relationship, Section 3 develops a polynomial time algorithm for the minimum variance spanning tree problem. Section 4 gives the properties necessary to develop an FPAS for P satisfying (A1)$\sim$(A3). Section 5 explains the outline of the FPAS. Section 6 describes the FPAS and analyzes the running time.

2. Relationship between P and $P(\lambda)$

Let S^* and $S(\lambda)$ be optimal to P and $P(\lambda)$ respectively. Katoh and Ibaraki [16] and Sniedovich [21, 22] considered the following problem Q:

$$Q : \underset{x \in X}{\text{minimize}}\, h(f_1(x), f_2(x)) \quad ,$$

where x denotes an n-dimensional decision vector and X denotes a feasible region. $f_i, i = 1,2$, are real-valued functions and $h(u_1, u_2)$ is quasiconcave over an appropriate region and differentiable in $u_i, i = 1,2$. They proved the following lemma.

Lemma 2.1 [16, 21, 22] Let x^* be optimal to Q and let $u^*_i = f_i(x^*), i = 1,2$. Define λ^* by

$$\lambda^* = \left\{ \frac{\partial h(u^*_1, u^*_2)}{\partial u_2} \right\} \Big/ \left\{ \frac{\partial h(u^*_1, u^*_2)}{\partial u_1} \right\} \quad . \tag{7}$$

Then any optimal solution to the following parametric problem $Q(\lambda)$ with $\lambda = \lambda^*$ is optimal to Q.

$$Q(\lambda) : \underset{x \in X}{\text{minimize}}\, f_1(x) + \lambda f_2(x) \quad . \qquad \square$$

The following lemma is obtained by specializing Lemma 2.1 to problem P.

Lemma 2.2 Let λ^* be defined by

$$\lambda^* = 2 \sum_{j \in S^*} c_j / p \quad . \tag{8}$$

Then $S(\lambda^*)$ is optimal to P.

Proof. First note that for any $S \in F$,

$$\begin{aligned} var(S) &= \frac{1}{|S|} \sum_{j \in S} (c_j - \frac{1}{|S|} \sum_{j \in S} c_j)^2 \\ &= \frac{1}{|S|} \sum_{j \in S} c_j^2 - \frac{1}{|S|^2} (\sum_{j \in S} c_j)^2 \\ &= \frac{1}{p} \sum_{j \in S} c_j^2 - \frac{1}{p^2} (\sum_{j \in S} c_j)^2 \quad \text{(by (4))} \quad . \end{aligned} \tag{9}$$

Associate $0-1$ characteristic vector $x(S) = (x_1(S), \ldots, x_{|E|}(S))$ with each $S \in F$ (i.e., $x_j(S) = 1$ if $j \in S$ and $x_j(S) = 0$ otherwise) and let X be the set of all such $x(S)$. Let

$$f_1(x(S)) = \sum_{j=1}^{|E|} c_j^2 x_j(S) \ , \quad f_2(x(S)) = \sum_{j=1}^{|E|} c_j x_j(S)$$

and

$$h(u_1,u_2) = \frac{1}{p}(u_1 - \frac{1}{p}(u_2)^2) \quad .$$

Then it is easy to see that for each S

$$var(S) = \frac{1}{p}[f_1(x(S)) - \frac{1}{p}\{f_2(x(S))\}^2] \quad .$$

Therefore P can be rewritten into

$$\underset{x(S)\in X}{\text{minimize}}\frac{1}{p}[f_1(x(S)) - \frac{1}{p}\{f_2(x(S))\}^2] \quad .$$

Since $h(u_1,u_2)$ is quasiconcave, it turns out that P is a special case of Q. As a result, by $\partial h(u_1,u_2)/\partial u_1 = 1/p$ and $\partial h(u_1,u_2)/\partial u_2 = -2u_2/p^2$, the lemma follows from Lemma 2.1. □

This lemma states that $P(\lambda)$ is not known unless P is solved. A straightforward to resolve this dilemma is to solve $P(\lambda)$ for all λ; the one with the minimum $var(S)$ is an optimal solution of P. This type of approach can sometimes provide polynomial time algorithms. One of such cases is that F is a set of spanning trees in an undirected graph, which will be treated in the next section. In general, however, the number of solutions generated over the entire range of λ is not polynomially bounded, and it is difficult to develop polynomial time algorithms by this approach. However, as will be seen in Sections 5 and 6, this approach is useful if we apply the scaling technique to costs c_j.

It is well known in the theory of parametric programming (see for example [1, 7, 9, 10]) that $z(\lambda)$ (the objective value of $P(\lambda)$) is a piecewise linear concave function as illustrated in Fig. 1, with a finite number of joint points $\lambda_{(1)},\lambda_{(2)},\ldots,\lambda_{(N)}$ with $0 < \lambda_{(1)} < \lambda_{(2)} < \ldots < \lambda_{(N)}$. Here N denotes the number of total joint points, and let $\lambda_{(0)} = 0$ and $\lambda_{(N+1)} = \infty$ by convention. In what follows, for two real numbers a,b with $a \le b, (a,b)$ and $[a,b]$ stand for the open interval $\{x|a < x < b\}$ and the closed interval $\{x|a \le x \le b\}$ respectively. The following two lemmas are also known in the parametric combinatorial programming.

Lemma 2.3 [9, 10] For any $\lambda' \in (\lambda_{(k-1)},\lambda_{(k)}), k = 1,\ldots,N+1$, $S(\lambda')$ is optimal to $P(\lambda)$ for all $\lambda \in [\lambda_{(k-1)},\lambda_{(k)}]$. □

Let for $k = 1,\ldots,N+1$

$$F^*_k \equiv \{S\in F|S \text{ is optimal to } P(\lambda) \text{ for all } \lambda \in [\lambda_{(k-1)},\lambda_{(k)}]\} \quad .$$

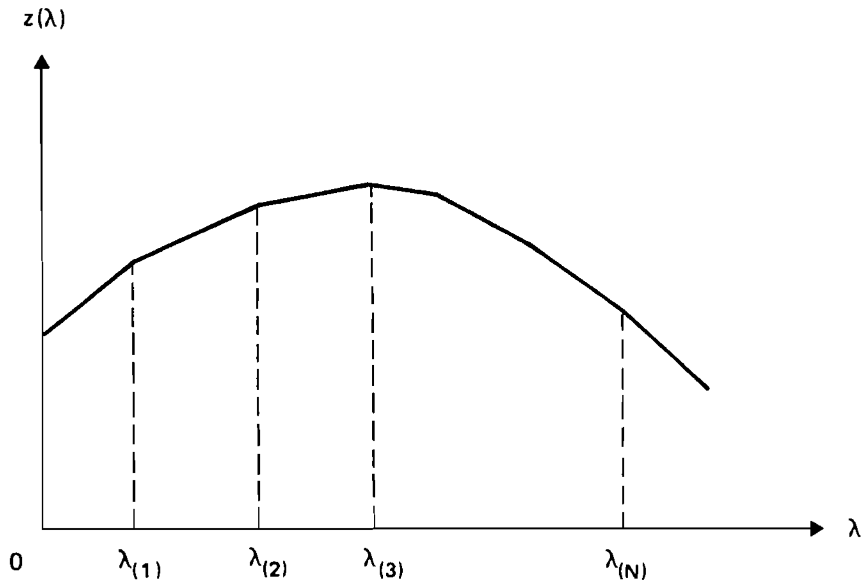

Figure 1. Illustration of $z(\lambda)$

Lemma 2.4 [9, 10] (i) For any two $S,S' \in F^*_k$ with $1 \le k \le N+1$,

$$\sum_{j \in S} c_j^2 = \sum_{j \in S'} c_j^2 \text{ and } \sum_{j \in S} c_j = \sum_{j \in S'} c_j$$

hold.

(ii) For any $S \in F^*_{k-1}$ and any $S' \in F^*_k$ with $2 \le k \le N+1$,

$$\sum_{j \in S} c_j < \sum_{j \in S'} c_j$$

holds. □

Lemmas 2.3 and 2.4 (i) imply that in order to determine $z(\lambda)$ for all $\lambda \ge 0$, it is sufficient to compute $S(\lambda')$ for an arbitrary $\lambda' \in (\lambda_{(k-1)},\lambda_{(k)})$ for each $k = 1,2,\ldots,N+1$. We shall use the notation S^*_k to stand for any $S \in F^*_k$.

Eisner and Severence [4] proposed an algorithm that determines $z(\lambda)$ for all $\lambda \ge 0$ and $S^*_k, k = 1,\ldots,N+1$ for a large class of combinatorial parametric problems including $P(\lambda)$ as a special case. They showed that the running time of their algorithm is proportional to (the number of joint points) $\times$ (the time required to solve $P(\lambda)$ for a given λ). Since $P(\lambda)$ for a given λ can be solved in $0(\tau(|E|))$time by assumption (A3), we have the following lemma.

Lemma 2.5 The Eisner and Severence method determines $z(\lambda)$ for all $\lambda \geq 0$ and computes S^*_k, $k = 1,\ldots,N+1$ in $0(N \cdot \tau(|E|))$ time. $\square$

3. A Minimum Variance Spanning Tree Problem

We shall concentrate on the case in which F is a set of all spanning trees in an undirected graph $G = (V,E)$, where V is a set of vertices and E is a set of edges. We call problem P with such F the ***minimum variance spanning tree problem.*** The following upper bound on N is known by Gusfield [9].

Lemma 3.1 [9]

$$N = 0(|E|\sqrt{|V|}) \quad . \qquad \square$$

Therefore, combining Lemmas 2.2, 2.5 and 3.1, we have the following theorem.

Theorem 3.1 Let $\tau(|E|, |V|)$ denote the time required to solve the minimum-cost spanning tree problem. Then the minimum variance spanning tree problem can be solved in $0(|E|\sqrt{V}\tau(|E|, |V|))$ time. $\square$

Since the best known algorithm for the minimum-cost spanning tree problem requires $0(|E|\beta(|E|, |V|))$ time, which was given by Fredman and Tarjan [5], where

$$\beta(|E|, |V|) = \min \{i | \log^{(i)}|V| \leq |E|/|V|\} \quad . \tag{10}$$

$\beta(|E|,|V|)$ is a very slowly growing function.

Corollary 3.1 The minimum variance spanning tree problem can be solved in $0(|E|^2\sqrt{|V|}\beta(|E|, |V|))$ time. $\square$

4. Basic Properties Necessary for Developing FPAS for P

We shall first give several results which are necessary to construct a fully polynomial time approximation scheme for P. Let

$$v_1 < v_2 < \cdots < v_m \tag{11}$$

be the sorted list of different values in $\{c_1, c_2, \ldots, c_{|E|}\}$. Then we have the following lemma.

Lemma 4.1

$$N \leq p \cdot (v_m - v_1) \tag{12}$$

Proof. By Lemma 2.4 (ii) and the integrality of c_j,

$$\sum_{j \in S^*_{k+1}} c_j - \sum_{j \in S^*_k} c_j \geq 1, \ \ k = 1, \ldots, N \tag{13}$$

holds. Since $p \cdot v_1 \leq \sum_{j \in S^*_1} c_j$ and $\sum_{j \in S^*_{N+1}} c_j \leq p \cdot v_m$ hold by (11), the lemma follows.

Lemma 4.2 An optimal solution of P can be obtained in $0(p \cdot (v_m - v_1) \cdot \tau(|E|))$ time.

Proof. Let λ^* defined in (8) belong to $[\lambda_{(k-1)}, \lambda_{(k)}]$ for some k with $1 \leq k \leq N+1$. Then S^*_k is optimal to $P(\lambda^*)$ by definition of S^*_k and is also optimal to P by Lemma 2.2. Therefore the lemma follows from Lemmas 2.5 and 4.1. □

Notice that $(v_m - v_1)$ is not polynomial in input size. However, this result is useful to obtain an FPAS for P as will be seen in the next section.

Lemma 4.3 [19] Problem BALANCE can be solved in $0(m \cdot f(|E|))$ time. □

Now we shall state the relationship between the objective values of BALANCE and P. Let S° be optimal to BALANCE.

Lemma 4.4 For any $S \in F$, we have

$$\frac{2(p-1)}{p^3} \cdot \{d(S)\}^2 \leq var(S) \leq \frac{p-1}{2p} \cdot \{d(S)\}^2 \ \ . \tag{14}$$

Proof. For the sake of simplicity, assume $S = \{1,2,\ldots,p\}$ with $c_1 \leq c_2 < \cdots \leq c_p$. Then $d(S) = c_p - c_1$ follows. It is easy to see that

$$var(S) = \frac{1}{p^2} \sum_{i=1}^{p-1} \sum_{j=i+1}^{p} (c_j - c_i)^2 \tag{15}$$

holds. By $|c_j - c_i| \leq c_p - c_1 (= d(S))$ with $1 \leq i, j \leq p$, the second inequality of (14) immediately follows. By the well known inequality $q\sum_{j=1}^{q} a_j^2 \geq (\sum_{j=1}^{q} a_j)^2$ for nonnegative numbers $a_1, a_2, \ldots, a_q$,

$$\frac{1}{p^2}\sum_{i=1}^{p-1}\sum_{j=i+1}^{p}(c_j - c_i)^2 \geq \frac{2}{p^3(p-1)}(\sum_{i=1}^{p-1}\sum_{j=i+1}^{p}|c_i - c_j|)^2 \tag{16}$$

holds. Since

$$\begin{aligned}\sum_{i=1}^{p-1}\sum_{j=i+1}^{p}|c_i - c_j| &= (p-1)(c_p - c_1) + (p-3)(c_{p-2} - c_2) + \cdots \\ &\geq (p-1)(c_p - c_1) \\ &= (p-1)d(S) \quad ,\end{aligned}$$

the first inequality of (14) follows from (15) and (16). □

Lemma 4.5

$$\frac{2(p-1)}{p^3}\cdot\{d(S^\circ)\}^2 \leq var(S^*) \leq \frac{p-1}{2p}\{d(S^\circ)\}^2 \tag{17}$$

holds.

Proof. Since $d(S^\circ) \leq d(S^*)$ holds by the optimality of S°, the first inequality of (17) follows from the first inequality of (14). Since $var(S^*) \leq var(S^\circ)$ holds by the optimality of S^*, the second inequality of (17) follows from the second inequality of (14).

□

5. The Outline of FPAS for P

First note that if $d(S^\circ) = 0$, it is obvious that $var(S^\circ) = 0$ and hence S° is optimal to P. By assumption (A3) and Lemma 4.3, S° can be found in polynomial time. As a result, P can be solved in polynomial time if $d(S^\circ) = 0$. Therefore assume $d(S^\circ) > 0$ in the following discussion.

An FPAS for P is constructed by applying the so-called "scaling technique". In other words, we replace the costs c_j for all j by

$$\tilde{c}_j = \lfloor \frac{c_j}{2^t} \rfloor \quad , \tag{18}$$

where t is determined by

$$t = \left\lfloor \log_2 \frac{2\epsilon \cdot d(S^\circ)}{p(\sqrt{p^2 + 4\epsilon} + p)} \right\rfloor \quad . \tag{19}$$

Let $\tilde{P}$ denote problem P with costs c_j replaced by $\tilde{c}_j$ for all j. The number t is chosen so that (i) an optimal solution of $\tilde{P}$ is an ϵ-approximate solution for P and (ii) the time required to solve $\tilde{P}$ is polynomial in $|E|$ and $1/\epsilon$. We first prove the first claim.

Lemma 5.1 Let $\tilde{S}^*$ be optimal to $\tilde{P}$. Then we have

$$\frac{var(\tilde{S}^*) - var(S^*)}{var(S^*)} \leq \epsilon \quad . \tag{20}$$

Proof. Let $\tilde{S}^\circ$ be optimal to BALANCE with the scaled costs $\tilde{c}_j$, and let $d(S,\tilde{c})$ denote the objective value of BALANCE with the scaled costs $\tilde{c}_j$. Define α_j by

$$c_j = 2^t \cdot \tilde{c}_j + \alpha_j \quad , \qquad j = 1,\ldots,|E| \tag{21}$$

where α_j satisfies $0 \leq \alpha_j < 2^t$. For the sake of simplicity, we use the notation $\sum_S \gamma_{ij}$ for a set of real numbers γ_{ij}, $i,j \in S$ to stand for

$$\sum_S \gamma_{ij} \equiv \sum_{\substack{i,j \in S \\ i<j}} \gamma_{ij} \quad . \tag{22}$$

Then by (15)

$$var(\tilde{S}^*) = \frac{1}{p^2}\sum_{\tilde{S}^*}(2^t \cdot \tilde{c}_i + \alpha_i - 2^t \cdot \tilde{c}_j - \alpha_j)^2$$

$$= \frac{2^{2t}}{p^2}\sum_{\tilde{S}^*}(\tilde{c}_i - \tilde{c}_j)^2 + \frac{2^{t+1}}{p^2}\sum_{\tilde{S}^*}(\alpha_i - \alpha_j)(\tilde{c}_i - \tilde{c}_j)$$

$$+ \frac{1}{p^2}\sum_{\tilde{S}^*}(\alpha_i - \alpha_j)^2 \quad .$$

By $|\alpha_i - \alpha_j| \le 2^t$, it follows that

$$var(\tilde{S}^*) \le \frac{2^{2t}}{p^2}\sum_{\tilde{S}^*}(\tilde{c}_i - \tilde{c}_j)^2 + \frac{2^{2t+1}}{p^2}\sum_{\tilde{S}^*}|\tilde{c}_i - \tilde{c}_j| + \frac{p-1}{2p} \cdot 2^{2t} \quad . \tag{23}$$

By the well-known inequality $(\sum_{j=1}^{q} a_j)^2 \le q\sum_{j=1}^{q} a_j^2$ for nonnegative numbers $a_1, a_2, \ldots, a_q$, we have

$$\sum_{\tilde{S}^*}|\tilde{c}_i - \tilde{c}_j| \le \sqrt{\frac{p(p-1)}{2} \cdot \sum_{\tilde{S}^*}(\tilde{c}_i - \tilde{c}_j)^2}$$

$$\le \sqrt{\frac{p(p-1)}{2} \cdot \sum_{\tilde{S}^\circ}(\tilde{c}_i - \tilde{c}_j)^2} \quad \text{(by the optimality of } \tilde{S}^*) \quad . \tag{24}$$

By $|\tilde{c}_i - \tilde{c}_j| \le d(\tilde{S}^\circ, \tilde{c})$ for $i,j \in \tilde{S}^\circ$, we have

$$\sum_{\tilde{S}^\circ}(\tilde{c}_i - \tilde{c}_j)^2 \le \frac{p(p-1)}{2} \cdot \{d(\tilde{S}^\circ, \tilde{c})\}^2 \quad . \tag{25}$$

Then we have

$$\sum_{\tilde{S}^*}|\tilde{c}_i - \tilde{c}_j| \le \frac{p(p-1)}{2} \cdot d(\tilde{S}^\circ, \tilde{c}) \qquad \text{(by (24) and (25))}$$

$$\le \frac{p(p-1)}{2} \cdot d(S^\circ, \tilde{c}) \qquad \text{(by the optimality of } \tilde{S}^\circ)$$

$$\le \frac{p(p-1)}{2 \cdot 2^t}(d(S^\circ) + 2^t) \quad \text{(by (21))}$$

$$= \frac{p(p-1)}{2^{t+1}} \cdot d(S^\circ) + \frac{p(p-1)}{2} \quad . \tag{26}$$

Next let us consider the term $\sum_{\tilde{S}^*}(\tilde{c}_i - \tilde{c}_j)^2$. We have

$$\sum_{\tilde{S}^*}(\tilde{c}_i - \tilde{c}_j)^2 \le \sum_{S^*}(\tilde{c}_i - \tilde{c}_j)^2 \quad \text{(by the optimality of } \tilde{S}^* \text{ to } \tilde{P})$$

$$= \sum_{S^*}\left(\frac{c_i-\alpha_i}{2^t} - \frac{c_j-\alpha_j}{2^t}\right)^2 \quad \text{(by (21))}$$

$$= \frac{1}{2^{2t}}\cdot\sum_{S^*}(c_i-c_j)^2 + \frac{2}{2^{2t}}\cdot\sum_{S^*}(c_i-c_j)(\alpha_j-\alpha_i)$$

$$+ \frac{1}{2^{2t}}\cdot\sum_{S^*}(\alpha_i-\alpha_j)^2$$

$$\leq \frac{p^2\cdot var(S^*)}{2^{2t}} + \frac{1}{2^{t-1}}\cdot\sum_{S^*}|c_i-c_j| + \frac{p(p-1)}{2} \quad . \tag{27}$$

(The last inequality is derived by $|\alpha_i-\alpha_j|\leq 2^t$.)

Again by using the inequality $(\sum_{j=1}^{q} a_j)^2 \leq q\sum_{j=1}^{q} a_j^2$ for nonnegative numbers $a_1,a_2,\ldots,a_q$, we have

$$\sum_{S^*}|c_i-c_j| \leq \sqrt{\frac{p(p-1)}{2}\cdot\sum_{S^*}(c_i-c_j)^2}$$

$$\leq \sqrt{\frac{p(p-1)}{2}\cdot\sum_{S^\circ}(c_i-c_j)^2} \quad \text{(by the optimality of } S^*)$$

$$\leq \frac{p(p-1)}{2}\cdot d(S^\circ) \quad (\text{by } |c_i-c_j|\leq d(S^\circ) \text{ for } i,j\in S^\circ) \quad . \tag{28}$$

By (23), (26), (27)and (28), it follows that

$$var(\tilde{S}^*) \leq var(S^*) + \frac{2^{t+1}(p-1)\cdot d(S^\circ)}{p} + \frac{2^{2t+1}(p-1)}{p} \quad . \tag{29}$$

By (17) and (29), it follows that

$$\frac{var(\tilde{S}^*)-var(S^*)}{var(S^*)} \leq \frac{p^2\{2^{t+1}(p-1)\cdot d(S^\circ)+2^{2t+1}(p-1)\}}{2(p-1)\{d(S^\circ)\}^2}$$

$$= \frac{p^2\{2^t\cdot d(S^\circ)+2^{2t}\}}{\{d(S^\circ)\}^2} \quad . \tag{30}$$

Since it holds by (19) that

$$2^t\cdot d(S^\circ) + 2^{2t} \leq \frac{2\epsilon\cdot d(S^\circ)^2}{p(\sqrt{p^2+4\epsilon}+p)} + \left\{\frac{2\epsilon\cdot d(S^\circ)}{p(\sqrt{p^2+4\epsilon}+p)}\right\}^2$$

$$= \frac{\epsilon}{p^2}\{d(S^\circ)\}^2 \quad , \tag{31}$$

(20) follows from (30). □

Now we shall show that we can solve $\tilde{P}$ is polynomial in $|E|$ and $1/\epsilon$. Let $\tilde{v}_1 < \tilde{v}_2 < \cdots < \tilde{v}_{\tilde{m}}$ be the sorted list of different values of $\tilde{c}_j, j = 1,\ldots,|E|$. By Lemma 4.2, if we apply the Eisner and Severence method to solve $\tilde{P}$ by solving $\tilde{P}(\lambda)$ over the entire range of λ, it requires

$$0(p \cdot (\tilde{v}_m - \tilde{v}_1) \cdot \tau(|E|)) \tag{32}$$

time. The term $\tilde{v}_m - \tilde{v}_1$ is estimated as follows.

$$\tilde{v}_m - \tilde{v}_1 \leq \frac{(v_m - v_1)}{2^t} + 1 \qquad \text{(by (15))}$$

$$\leq \frac{p(\sqrt{p^2+4\epsilon}+p)(v_m - v_1)}{\epsilon \cdot d(S^\circ)} + 1. \quad \left(\text{by } 2^{t+1} > \frac{2\epsilon \cdot d(S^\circ)}{p(\sqrt{p^2+4\epsilon}+p)} \text{ from (19)}\right)$$

However, $(v_m - v_1)$ is not in general bounded above by

$$d(S^\circ) \cdot g(|E|, 1/\epsilon) \tag{33}$$

for a certain function $g(|E|, 1/\epsilon)$ which is polynomial in $|E|$ and $1/\epsilon$. This implies that the direct application of the Eisner and Severence method to solve $\tilde{P}$ as in Lemma 4.2 does not lead to a fully polynomial time approximation scheme for P.

This difficulty is overcome as follows. We construct $\tilde{m}(\leq |E|)$ subproblems $\tilde{P}_1, \tilde{P}_2, \ldots, \tilde{P}_{\tilde{m}}$ of $\tilde{P}$ so that

(i) for some l with $1 \leq l \leq \tilde{m}$, an optimal solution of $\tilde{P}_l$ is optimal to $\tilde{P}$ and

(ii) Each $\tilde{P}_l, l = 1,\ldots,\tilde{m}$, can be solved in polynomial time in $|E|$ and $1/\epsilon$.

To define $\tilde{P}_l$, let, for any l,u satisfying $1 \leq l \leq u \leq \tilde{m}$,

$$E(l,u) = \{j \in E \mid \tilde{v}_l \leq \tilde{c}_j \leq \tilde{v}_u\} \quad . \tag{34}$$

For each l with $1 \leq l \leq \tilde{m}$, define u_l by

$$u_l \equiv \max\{u \mid l \leq u \leq \tilde{m}, \tilde{v}_u - \tilde{v}_l \leq \frac{p}{2} \cdot d(\tilde{S}^\circ, \tilde{c})\} \quad . \tag{35}$$

$\tilde{P}_l$ is defined as follows.

$$\tilde{P}_l : \text{minimize}\,\{var(S,\tilde{c}) | S \in F, S \subseteq E(l,u_l)\} \quad , \tag{36}$$

where $var(S,\tilde{c})$ is equal to $var(S)$ with c_j replaced by $\tilde{c}_j$ for all j. Note that if u_l does not exist for some l, $\tilde{P}_l$ is not defined.

Lemma 5.2 There exists l with $1 \leq l \leq \tilde{m}$ such that

$$\tilde{S}^* \subseteq E(l, u_l) \tag{37}$$

holds for any optimal solution $\tilde{S}^*$ to $\tilde{P}$.

Proof. Assume that the lemma does not hold. Let $\tilde{S}^*$ denote any optimal solution to $\tilde{P}$. Let

$$\tilde{c}^*{}_{\max} = \max_{j \in \tilde{S}^*} \tilde{c}_j \text{ and } \tilde{c}^*{}_{\min} = \min_{j \in \tilde{S}^*} \tilde{c}_j \quad .$$

Then

$$d(\tilde{S}^*, \tilde{c}) = \tilde{c}^*{}_{\max} - \tilde{c}^*{}_{\min} > \frac{p}{2} \cdot d(\tilde{S}^\circ, \tilde{c}) \tag{38}$$

is satisfied. By the first inequality of (14),

$$\frac{2(p-1)}{p^3} \cdot \{d(\tilde{S}^*, \tilde{c})\}^2 \leq var(\tilde{S}^*, \tilde{c}) \tag{39}$$

holds. Then it follows that

$$\begin{aligned} var(\tilde{S}^*, \tilde{c}) &\leq \frac{p-1}{2p} \{d(\tilde{S}^\circ, \tilde{c})\}^2 \text{ (by the second inequality of (17))} \\ &< \frac{p-1}{2p} \cdot \frac{4}{p^2} \{d(\tilde{S}^*, \tilde{c})\}^2 \text{ (by (38))} \\ &= \frac{2(p-1)}{p^3} \cdot \{d(\tilde{S}^*, \tilde{c})\}^2 \leq var(\tilde{S}^*, \tilde{c}) \text{ (by (39))} \quad . \end{aligned} \tag{40}$$

This is a contradiction. On the other hand, if $d(\tilde{S}^*, \tilde{c}) \leq \frac{p}{2} \cdot d(\tilde{S}^\circ, \tilde{c})$, let l satisfy $v_l = \tilde{c}^*{}_{\min}$. Then $\tilde{c}^*{}_{\max} \leq v_{u_l}$ clearly holds by definition of u_l. Therefore (37) follows. □

We shall show that each $\tilde{P}_l$ can be solved in polynomial time in $|E|$ and $1/\epsilon$. Consider the following parametric problem $\tilde{P}_l(\lambda)$ associated with each $\tilde{P}_l$.

$$\tilde{P}_l(\lambda) : \tilde{z}(\lambda) \equiv \text{minimize} \{ \sum_{j \in S} \tilde{c}_j^2 - \lambda \sum_{j \in S} \tilde{c}_j \mid S \in F, S \subseteq E(l, u_l) \} \quad . \tag{41}$$

By assumption (A3), for a given $\lambda, \tilde{P}_l(\lambda)$ can be solved in polynomial time. By applying Lemma 4.2 to $\tilde{P}_l$, it follows that $\tilde{P}_l$ can be solved in

$$O(p \cdot (\tilde{v}_{u_l} - \tilde{v}_l) \cdot \tau(|E|)) \tag{42}$$

time. By $p \leq |E|$, it is sufficient to show that $\tilde{v}_{u_l} - \tilde{v}_l$ is polynomial in $|E|$ and $1/\epsilon$. By (35), we have

$$\tilde{v}_{u_l} - \tilde{v}_l \leq \frac{p}{2} \cdot d(\tilde{S}^\circ, \tilde{c}) \quad . \tag{43}$$

By the optimality of $\tilde{S}^\circ$ to BALANCE with scaled costs $\tilde{c}_j$,

$$d(\tilde{S}^\circ, \tilde{c}) \leq d(S^\circ, \tilde{c}) \tag{44}$$

holds. Letting

$$c_{j_1} = \max\{c_j | j \in S^\circ\} \text{ and } c_{j_2} = \min\{c_j | j \in S^\circ\} \quad ,$$

we have

$$\begin{aligned} d(S^\circ, \tilde{c}) &= \tilde{c}_{j_1} - \tilde{c}_{j_2} \\ &= \left\lfloor \frac{c_{j_1}}{2^t} \right\rfloor - \left\lfloor \frac{c_{j_2}}{2^t} \right\rfloor \qquad \text{(by (18))} \\ &\leq \frac{1}{2^t}(c_{j_1} - c_{j_2}) + 1 \\ &= \frac{d(S^\circ)}{2^t} + 1 \quad . \end{aligned} \tag{45}$$

Note that

$$2^{t+1} > \frac{2\epsilon \cdot d(S^\circ)}{p(\sqrt{p^2 + 4\epsilon} + p)} \left[= \frac{2\epsilon \cdot d(S^\circ)}{p^2(\sqrt{1 + 4\epsilon/p^2} + 1)} \right] \tag{46}$$

holds by (19) and that for any $\delta > 0$,

$$\sqrt{1 + 4\delta} \leq \begin{cases} 6\sqrt{\delta} - 1 & \text{for } \delta \geq 1/4 \\ \sqrt{2} & \text{for } 0 < \delta \leq 1/4 \end{cases} \tag{47}$$

holds. Letting $\delta = \epsilon/p^2$, it follows from (46) and (47) that

$$2^t > \frac{\epsilon \cdot d(S^\circ)}{p^2(\sqrt{1 + 4\epsilon/p^2} + 1)} \geq \begin{cases} \dfrac{\sqrt{\epsilon} \cdot d(S^\circ)}{6p} & \text{for } \epsilon \geq 4p^2 \\ \dfrac{\epsilon \cdot d(S^\circ)}{p^2(\sqrt{2} + 1)} & \text{for } 0 < \epsilon \leq 4p^2 \quad . \end{cases} \tag{48}$$

Therefore by (45),

$$d(S^{\circ},\tilde{c}) \leq \begin{cases} \dfrac{6p}{\sqrt{\epsilon}} + 1 & \text{for } \epsilon \geq 4p^2 \\ \dfrac{p^2(\sqrt{2}+1)}{\epsilon} + 1 & \text{for } 0 < \epsilon \leq 4p^2 \end{cases} \tag{49}$$

follows. The following lemma is an immediate consequence of (42), (43), (44) and (49).

Lemma 5.3 Problem $\tilde{P}_l$ can be solved in $O((p^3/\sqrt{\epsilon} + p^2)\tau(|E|))$ time if $\epsilon \geq 4p^2$ and in $O((p^4/\epsilon + p^2)\tau(|E|))$ time otherwise. □

6. The Description of FPAS for P

We shall describe FPAS for P and then analyze its running time.

Procedure MVP

Input: The ground set E, a family of feasible subsets (which are implicitly given through an oracle explained in (A2)), a positive integer p of (4), integer costs $c_j, j \in E$, and a positive number ϵ.

Output: An ϵ-approximate solution for P.

Step 1. Solve problem BALANCE and let $d(S^\circ)$ be the optimal objective value. If $d(S^\circ) = 0$, output S° as an optimal solution of P and halt. Else compute t by (19) and $\tilde{c}_j$ for all j by (18).

Step 2. Compute $\tilde{v}_1, \tilde{v}_2, \ldots, \tilde{v}_{\tilde{m}}$ with $\tilde{v}_1 < \tilde{v}_2 < \cdots < \tilde{v}_{\tilde{m}}$ which are distinct numbers of $\tilde{c}_j, j \in E$. Solve problem BALANCE with scaled costs $\tilde{c}_j$ and let $\tilde{S}^\circ$ be its optimal solution.

Step 3. For each $l = 1, 2, \ldots, \tilde{m}$, do the following.

(i) Compute u_l by (35), if u_l exists, and the set $E(l, u_l)$ by (34). If u_l does not exist, return to the beginning of Step 3.

(ii) Solve $\tilde{P}_l(\lambda)$ of (41) for all $\lambda \in [2\tilde{v}_l, 2\tilde{v}_{u_l}]$ by applying the Eisner and Severence method.

(iii) Among solutions obtained in (ii), let $\tilde{S}^*_l$ be the one minimizing $var(S, \tilde{c})$ (i.e., $\tilde{S}^*_l$ is optimal to $\tilde{P}_l$).

Step 4. Find $\tilde{S}^*_{l^*}$ such that

$$\tilde{S}^*_{l^*} = \min \{ var(\tilde{S}^*_l, \tilde{c}) \,|\, 1 \le l \le \tilde{m},\ u_l \text{ exists} \} \quad .$$

Output $\tilde{S}^*_{l^*}$ as an ϵ- approximate solution to P. □

Theorem 6.1. Procedure MVP is an FPAS for the minimum variance problem P. Its running time is

$$\begin{array}{ll} 0(\log_2 \epsilon + \log_2(|v_1| + |v_m|) + |E|^2 + p^2|E|\tau(|E|) + p^4|E|/\sqrt{\epsilon} + p^2|E|\tau(|E|)/\sqrt{\epsilon}) & \text{if } \epsilon \ge 4p^2 \ , \\ 0(\log_2 \epsilon + \log_2(|v_1| + v_m|) + |E|^2 + p^2|E|\tau(|E|) + p^5|E|/\epsilon + p^3|E|\tau(|E|)/\epsilon) & \text{otherwise} \ . \end{array}$$

Proof. The correctness of MVP follows from the following four facts.

Fact 1. If $d(S^\circ) = 0, S^\circ$ is optimal to P.

Fact 2. An optimal solution to $\tilde{P}$ is an ϵ-approximate solution of P by Lemma 5.1.

Fact 3. An optimal solution to $\tilde{P}_l$ for some l with $1 \le l \le \tilde{m}$ is optimal to $\tilde{P}$ by Lemma 5.2.

Fact 4. An optimal solution to $\tilde{P}_l$ can be found as an optimal solution $\tilde{S}_l(\tilde{\lambda})$ to $\tilde{P}_l(\tilde{\lambda})$ for some $\tilde{\lambda}$. Since such $\tilde{\lambda}$ satisfies $\tilde{\lambda} = 2 \sum_{j \in \tilde{S}_l(\tilde{\lambda})} \tilde{c}_j / p$ by Lemma 2.2,

$$2\tilde{v}_l \le \tilde{\lambda} \le 2\tilde{v}_{u_l}$$

follows by $\tilde{v}_l \le \tilde{c}_j \le \tilde{v}_{u_l}$ for all $j \in E(l, u_l)$.

The running time is analyzed as follows. Step 1 is first analyzed. Solving BALANCE requires $O(m \cdot f(|E|))$ time by Lemma 4.3. By $m \le |E|$ and $f(|E|) \le \tau(|E|)$ by assumption (A3), $O(m \cdot f(|E|)) = O(|E| \cdot \tau(|E|))$ follows. The time required to compute t of (19) is estimated as follows. First note that

$$\sqrt{1+4\delta} \ge \begin{cases} 2\sqrt{\delta} - 1 & \text{for } \delta \ge 1/4 \\ 1 & \text{for } 0 < \delta \le 1/4 \end{cases} ,$$

and

$$2^t \le \frac{2\epsilon \cdot d(S^\circ)}{p(\sqrt{p^2+4\epsilon}+p)} \le \begin{cases} \dfrac{\sqrt{\epsilon} \cdot d(S^\circ)}{p} & \text{for } \epsilon \ge 4p^2 \\ \dfrac{\epsilon \cdot d(S^\circ)}{p^2} & \text{for } 0 < \epsilon \le 4p^2 \end{cases} . \tag{50}$$

Computing t is done by first setting $s = 1$ and increasing t by one every time

$$2^s < \frac{2\epsilon \cdot d(S^\circ)}{p(\sqrt{p^2+4\epsilon}+p)} \tag{51}$$

is satisfied. Let s^* be the first s for which (51) is not satisfied. Then $t = s^*$ holds by definition of t. This computation requires $O(t)$ times of comparisons of (51). By (19) and (50), we have

$$O(t) = O\left[\log_2 \frac{2\epsilon \cdot d(S^\circ)}{p(\sqrt{p^2+4\epsilon}+p)}\right]$$

$$= O(\log_2 \epsilon + \log_2 d(S^\circ)) \quad (\text{by } p \ge 2)$$

$$= O(\log_2 \epsilon + \log_2(|v_1| + |v_m|)) \quad (\text{by } d(S^\circ) \le |v_1| + |v_m|) \quad . \tag{52}$$

Computing $\tilde{c}_j$ for all $j \in E$ requires $O(|E|)$ time. Solving BALANCE with scaled costs in Step 2 requires $O(\tilde{m} \cdot f(|E|))(=O(|E| \cdot f(|E|)) = O(|E| \cdot \tau(|E|)))$ time.

Step 3 (i) requires $O(|E|)$ time for each l. By Lemma 5.3, Step 3 (ii) requires $O((p^3/\sqrt{\epsilon}+p^2) \cdot \tau(|E|))$ time if $\epsilon \geq 4p^2$ and $O((p^4/\epsilon+p^2) \cdot \tau(|E|))$ time otherwise. By the discussion prior to Lemma 4.2, the number of optimal solutions for $\tilde{P}_l(\lambda)$ for $\lambda \in [2\tilde{v}_l, 2\tilde{v}_{u_l}]$ is

$$O(p \cdot (\tilde{v}_{u_l} - \tilde{v}_l)) \quad .$$

By (43)~(49),

$$p \cdot (\tilde{v}_{u_l} - \tilde{v}_l) = \begin{cases} O(p^3/\sqrt{\epsilon}) & \text{if } \epsilon \geq 4p^2 \\ O(p^4/\epsilon) & \text{otherwise} \end{cases} .$$

Since the evaluation of $var(S,\tilde{c})$ for $S \in F$ requires $O(p)$ time, Step 3 (iii) requires

$$\begin{array}{ll} O(p^4/\sqrt{\epsilon}) & \text{if } 4\epsilon \geq 4p^2 \\ O(p^5/\epsilon) & \text{otherwise} \end{array} .$$

Since the loop of Step 3(i), (ii) and (iii) is repeated $O(\tilde{m})$ times and $\tilde{m} \leq |E|$ holds, the time required for Step 3 is

$$\begin{array}{ll} O(|E|^2 + p^2|E|\tau(|E|) + p^4|E|/\sqrt{\epsilon} + p^3|E|\tau(|E|)/\sqrt{\epsilon}) & \text{if } \epsilon \geq 4p^2 \\ O(|E|^2 + p^2|E|\tau(|E|) + p^5|E|/\epsilon + p^4|E|\tau(|E|)/\epsilon) & \text{otherwise} \end{array} .$$

Step 4 requires $O(\tilde{m}) = O(|E|)$ time. It follows from the above discussion that Procedure MVP requires

$$\begin{array}{ll} O(\log_2\epsilon + \log_2(|v_1|+|v_m|) + |E|^2 + p^2|E|\tau(|E|) + p^4|E|/\sqrt{\epsilon} + p^3|E|\tau(|E|)/\sqrt{\epsilon}) & \text{if } \epsilon \geq 4p^2 \\ O(\log_2\epsilon + \log_2(|v_1|+|v_m|) + |E|^2 + p^2|E|\tau(|E|) + p^5|E|/\epsilon + p^4|E|\tau(|E|)/\epsilon) & \text{otherwise} \end{array} .$$

This is clearly polynomial in input length and $1/\epsilon$. □

7. Conclusion

We first showed the relationship between the minimum variance problem P and the parametric problem $P(\lambda)$. Based on this relation, we showed that the minimum variance spanning tree problem can be solved in polynomial time. We mention here that the result can be directly generalized to the case where F is a set of bases in a matroid, assuming that (A2) is satisfied. In this case, (A1) follows from definition of a base in a matroid and (A3) follows from (A2) since the minimum-cost base problem can be solved by the greedy algorithm. Also notice that the number of joint points of the parametric minimum-cost base problem is $0(|E|\sqrt{p})$ (the proof is done in the same manner as in [9]).

Secondly, we developed a fully polynomial time approximation scheme for P satisfying assumptions (A1)~(A3). However the complexity issue for P has not been settled down. It is not known yet whether problem P is NP-complete or not under assumptions of (A1)~(A3). This is left for future research.

References

[1] P.J. Carstensen, The complexity of some problems in parametric linear and combinatorial programming, Ph.D. Thesis, Dept. of Mathematics, The University of Michigan, MI (1983).

[2] W. Dinkelbach, On some nonlinear fractional programming, *Management Science*, 13 (1967), 492-498.

[3] J. Edmonds and R. Karp, Theoretical improvements in the algorithmic efficiency for network flow problems, *J. ACM*, 19 (1972), 248-264.

[4] M.J. Eisner and D.G. Severence, Mathematical techniques for efficient record segmentation in large shared databases, *J. ACM*, 23 (1976), 619-635.

[5] M.L. Fredman and R.E. Tarjan, Fibonacci heaps and their uses in improved network optimization algorithms, *J. ACM*, 34 (1987), 596-615.

[6] M.R. Garey and D.S. Johnson, Computers and Intractability: A Guide to the Theory of NP-Completeness, Freeman, San Francisco, CA, 1979.

[7] T. Gal, Linear parametric programming - a brief survey, *Mathematical Programming Study*, 21 (1984), 43-68.

[8] Z. Galil and E. Tardos, An $0(n^2(m+n\log n)\log n)$ min-cost flow algorithm, Proceedings of 27th IEEE Symposium on Foundations of Computer Science, (1986), 1-9.

[9] D. Gusfield, Sensitivity analysis for combinatorial optimization, Memorandum No. UCB/ERL M 80/22, Electronics Research Laboratory, College of Engineering, University of California, Berkeley, 1980.

[10] D. Gusfield, Parametric combinatorial computing and a problem of program module distribution, *J. ACM*, 30 (1983), 551-563.

[11] T. Ichimori, S. Shiode, H. Ishii and T. Nishida, Minimum spanning tree with normal variables as weights, *J. Oper. Res. Soc. Japan*, 24 (1981), 61-65.

[12] H. Ishii, S. Shiode, T. Nishida and Y. Namasuya, Stochastic spanning tree problem, *Disc. Appl. Math.,* 3 (1981), 263-273.

[13] R. Jagannathan, On some properties of programming problems in parametric forms pertaining to fractional programming, *Management Science*, 12 (1966), 609-615.

[14] S. Kataoka, A stochastic programming model, *Econometrica*, 3 (1963), 181-196.

[15] N. Katoh and T. lbaraki, A polynomial time algorithm for a chance-constrained single machine scheduling problem, *Operations Research Letters*, 2 (1983), 62-65.

[16] N. Katoh and T. Ibaraki, A parametric characterization and an ϵ-approximation scheme for the minimization of a quasiconcave program, *Disc. Appl. Math.*, 17 (1987), 39-66.

[17] H. Kawai and N. Katoh, Variance constrained markov decision process, *J. Oper. Res. Soc. Japan*, 30 (1987), 88-100.

[18] E.L. Lawler, Fast approximation algorithms for knapsack problems, *Math. Oper. Res.*, 4 (1979), 339-356.

[19] S. Martello, W.R. Pulleyblank, P. Toth, D. de Werra, Balanced optimization problems,*Operations Research Letters,* 3 (1984), 275-278.

[20] C.H. Papadimitriou and K. Steiglitz, Combinatorial Optimization: Algorithms and Complexity, Prentice-Hall, Englewood Cliffs, New Jersey, 1982.

[21] M. Sniedovich, C-programming problems: a class of nonlinear optimization problems, *Disc. Appl. Math.*, 9 (1984), 301-305.

[22] M. Sniedovich, C-programming: an outline, *Operations Research Letters*, 4 (1985), 19-21.

[23] E. Tardos, A Strongly polynomial minimum cost circulation algorithm, *Combinatorica*, 5 (1985), 247-255.