OPERATIONS RESEARCH
Vol. 61, No. 2, March–April 2013, pp. 386–397
ISSN 0030-364X (print) | ISSN 1526-5463 (online)

http://dx.doi.org/10.1287/opre.1120.1093

A General Framework for Designing Approximation Schemes for Combinatorial Optimization Problems with Many Objectives Combined into One

Shashi Mittal
Amazon.com, Seattle, Washington 98109, mshashi@alum.mit.edu

Andreas S. Schulz
Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139, schulz@mit.edu

In this paper, we present a general framework for designing approximation schemes for combinatorial optimization problems in which the objective function is a combination of more than one function. Examples of such problems include those in which the objective function is a product or ratio of two linear functions, parallel machine scheduling problems with the makespan objective, robust versions of weighted multiobjective optimization problems, and assortment optimization problems with logit choice models. The main idea behind our approximation schemes is the construction of an approximate Pareto-optimal frontier of the functions that constitute the given objective. Using this idea, we give the first fully polynomial-time approximation schemes for the max-min resource allocation problem with a fixed number of agents, combinatorial optimization problems in which the objective function is the sum of a fixed number of ratios of linear functions, or the product of a fixed number of linear functions, and assortment optimization problems with logit choice model.

Subject classifications: analysis of algorithms; combinatorial optimization; networks/graphs; production/scheduling.
Area of review: Optimization.
History: Received July 2011; revision received March 2012; accepted June 2012. Published online in *Articles in Advance* February 5, 2013.

1. Introduction

In many combinatorial optimization problems, the objective function is a combination of more than one function, or it can be interpreted in this way. For example, consider the problem of finding a spanning tree in a graph $G = (V, E)$ with edge weights $c_1\colon E \to \mathbb{Z}_+$ and $c_2\colon E \to \mathbb{Z}_+$, where c_1 may correspond to the failure probability of the edges, and c_2 to the cost of the edges. The goal is to find a spanning tree T of the graph for which $c_1(T) \cdot c_2(T)$ is minimized (Kuno 1999, Kern and Woeginger 2007). In this problem, the objective function is a combination of two linear objective functions combined together using the product function.

Another example of a problem whose objective function subsumes more than one function is the max-min resource allocation problem (Asadpour and Saberi 2007). Here, there are several resources that have to be distributed among agents. The utility of each agent is the sum of the utilities of the resources assigned to the agent. The objective is to maximize the utility of the agent with the lowest utility. In this problem, one can look at the utility of each agent as a separate objective function. Thus, the objective function of the problem is a combination of the objective functions of the individual agents using the minimum function.

This paper presents a unified approach for solving combinatorial optimization problems in which the objective function is a combination of more than one (but a fixed number) of objective functions. Usually, these problems turn out to be NP-hard. We show that under very general conditions, we can obtain a fully polynomial-time approximation scheme (FPTAS) for such problems. An FPTAS for a minimization (respectively, maximization) problem is a family of algorithms parametrized by ϵ such that for all $\epsilon > 0$, there is an algorithm A_ϵ that for any instance of the problem returns a solution whose objective function value is no more then $(1+\epsilon)$ (respectively, no less than $(1-\epsilon)$) times that of the optimal solution, and whose running time is polynomial in the input size of the problem as well as in $1/\epsilon$. On the other hand, if the running time of A_ϵ for each $\epsilon > 0$ is polynomial in the input size of the problem but not necessarily in $1/\epsilon$, then the family of algorithms A_ϵ is called a polynomial-time approximation scheme (PTAS). For an NP-hard problem, the existence of an FPTAS is typically the best possible approximation result one can obtain. Our technique turns out be surprisingly versatile: it can be applied to a variety of scheduling problems (e.g., unrelated parallel machine scheduling and vector scheduling), combinatorial optimization problems with nonlinear objective functions such as a product or a ratio of two linear functions, robust versions of weighted multiobjective optimization problems, and assortment optimization problems with logit choice models. We first give examples of some of the problems for which we can get FPTASes using our framework.

1.1. Examples of Problems

1.1.1. Combinatorial Optimization Problems with a Rational Objective. Consider a combinatorial optimization problem in which the objective function is a ratio of discrete linear functions:

$$\text{minimize} \quad g(x) = \frac{f_1(x)}{f_2(x)} = \frac{a_0 + a_1x_1 + \cdots + a_dx_d}{b_0 + b_1x_1 + \cdots + b_dx_d}, \qquad (1)$$
$$\text{s.t.} \quad x \in X \subseteq \{0,1\}^d.$$

We assume that $f_1(x) > 0$, $f_2(x) > 0$ for all $x \in X$. In this case, there are two linear objective functions that have been combined by using the quotient function. A well-known example is the computation of a minimum mean cost circulation in graphs. It is a well-known result that any polynomial-time algorithm for the corresponding linear objective problem can be used to obtain a polynomial-time algorithm for the same problem with a rational objective function (see, e.g., Megiddo 1979 and Radzik 1992). Extensions of this result have been given by Hashizume et al. (1987), Billionnet (2002), and Correa et al. (2010) to obtain approximation algorithms for the case where the corresponding linear objective problem is NP-hard. The main idea behind all these approaches is to convert the problem with a rational objective function to a parametric combinatorial optimization problem with a linear objective function, and then perform a binary search on the parameter to get an approximate solution for the problem. The main drawback of parametric methods is that they do not generalize in a straightforward way to the case where we have a sum of ratios of linear functions.

In §7, we give a fairly general sufficient condition for the existence of an FPTAS for this problem. It can be used to obtain an FPTAS, for example, for the knapsack problem with a rational objective function. In contrast to the methods described, our algorithm uses a nonparametric approach to find an approximate solution. One advantage of our technique is that it easily generalizes to more general rational functions as well, for example the sum of a fixed number of ratios of linear functions. Such a form often arises in assortment optimization in the context of retail management. In §7.1, we show how to obtain FPTASes using our framework for assortment optimization problems under two different choice models.

1.1.2. Resource Allocation and Scheduling Problems. The best-known approximation algorithm for the general max-min resource allocation problem has an approximation ratio of $\Omega(1/(\sqrt{m}\log^3 m))$, where m is the number of agents (Asadpour and Saberi 2007). In §4.1, we obtain the first FPTAS for this problem when the number of agents is fixed.

Scheduling problems can be thought of as an inverse of the resource allocation problem, where we want to assign jobs to machines and attempt to minimize the load on individual machines. Corresponding to the max-min resource allocation problem, we have the problem of scheduling jobs on unrelated parallel machines to minimize the makespan (i.e., the time at which the last job finishes its execution). When the number of machines m is fixed, this problem is referred to as the $Rm||C_{\max}$ problem (Graham et al. 1979). Another objective function that has been considered in the literature is the one in which the total load on different machines are combined together using an l_p norm (Azar et al. 2004). In §4.1, we give FPTASes for both of these scheduling problems. It should be noted that approximation schemes for the $Rm||C_{\max}$ problem already exist in the literature (e.g., Horowitz and Sahni 1976 and Lenstra et al. 1990).

A generalization of the $Rm||C_{\max}$ problem is the vector scheduling problem. In this problem, a job uses multiple resources on each machine, and the objective is to assign the jobs to the machines so as to minimize the maximum load over all the resources of the machines. A practical situation where such a problem arises is query optimization in parallel database systems (Garofalakis and Ioannidis 1996). In this case, a job is a database query, which uses multiple resources on a computer—for example, multiple cores of a processor, memory, hard disk, etc. A query can be assigned to any one of the multiple servers in the database system. Because the overall performance of a system is governed by the resource with the maximum load, the objective is to minimize the maximum load over all resources. Chekuri and Khanna (2004) give a PTAS for the problem when the number of resources on each machine is fixed. Moreover, they only consider the case where each job has the same requirement for a particular resource on all the machines. In §4.2, we show that when both the number of machines and resources are fixed, we can get an FPTAS for the problem, even when each job can use different amounts of a resource on different machines.

1.1.3. Combinatorial Optimization Problems with a Product Objective. For the product versions of the minimum spanning tree problem and the shortest s-t path problem, Kern and Woeginger (2007) and Goyal et al. (2011) give an FPTAS. Both these methods are based on linear programming techniques, and do not generalize to the case where we have more than two functions in the product. Moreover, their techniques do not extend to the case where the corresponding problem with a linear objective function is NP-hard (§6).

In §5.3 of this paper, we give FPTASes for the product version of the s-t path problem and the spanning tree problem using our framework. One advantage of our method is that it easily generalizes to the case where the objective function is a product of a fixed number of linear functions. It can also be used to design approximation schemes for the product version of certain NP-hard problems, such as the knapsack problem.

1.1.4. Robust Weighted Multiobjective Optimization Problems. Consider once again the spanning tree problem with two cost functions c_1 and c_2 on the edges, as given in the introduction. Another way to combine the two costs

is to find a spanning tree T that minimizes the weighted sum $w_1c_1(T) + w_2c_2(T)$ for some positive weights w_1 and w_2. However, in many cases it is not clear a priori which weights should be used to combine the two objectives. An alternative is to allow the weights $w = (w_1, w_2)$ to take values in a set W, and find a spanning tree that minimizes the cost of the weighted objective for the worst-case scenario weight in the set $W \subseteq \mathbb{R}^2_+$. More generally, we consider the following robust version of a weighted multiobjective optimization problem:

$$\text{minimize } g(x) = \max_{w \in W} w^T f(x), \quad x \in X \subseteq \{0,1\}^d. \tag{2}$$

Here, $f(x) = (f_1(x), \ldots, f_m(x))$ is a vector of m function values and $W \subseteq \mathbb{R}^m$ is a compact convex weight set. For the spanning tree problem and the shortest path problem, this robust version is NP-hard even for the case of two objectives (§6).

The robust version of weighted multiobjective optimization problems has been studied by Hu and Mehrotra (2010) for the case when each f_i is a continuous function. For discrete optimization problems, this formulation is a generalization of the robust discrete optimization model with a fixed number of scenarios (see, e.g., Kouvelis and Yu 1997) and with interval data (see, e.g., Kasperski and Zielinski 2007). This problem is NP-hard but admits an FPTAS for the robust version of many problems when the number of scenarios is fixed (Aissi et al. 2007). In §6, we generalize this result and show that we can get FPTASes for the robust version of weighted multiobjective optimization problems when the number of objectives is fixed, for the case of the spanning tree problem, the shortest path problem, and the knapsack problem.

1.2. Related Work

There are two well-known general methods for obtaining approximation schemes for combinatorial optimization problems. In the first method, the input parameters of the problem are rounded and then dynamic programming is applied on the modified instance to find an approximate solution for the original problem. This idea has been extensively used to find approximation schemes for a number of machine scheduling problems (e.g., Sahni 1976, Horowitz and Sahni 1976, Hochbaum and Shmoys 1987, Lenstra et al. 1990). The other method is shrinking the state space of the dynamic programs that solve the problem in pseudo-polynomial time. This idea was first used by Ibarra and Kim (1975) to obtain an approximation scheme for the knapsack problem. Woeginger (2000) gives a very general framework where such dynamic programs can be converted to an FPTAS, and using this framework he derives FPTASes for several scheduling problems. Another example is the work of Halman et al. (2009), who adopt the same methodology to get FPTASes for inventory management problems.

Approximate Pareto-optimal frontiers have previously been used in the design of approximation algorithms for concrete combinatorial optimization problems. For example, Koltun and Papadimitriou (2007) use this approach to find approximate answers to database queries, in which the objective is to retrieve an object that is simultaneously good with respect to more than one criterion. Another example is the work of Ackermann et al. (2007), who use a similar approach to design approximation schemes for the shortest path and the spanning tree problem, where the objective function is a combination of more than one linear objective function that are combined together using a well-behaved monotone function. However, the idea of using approximate Pareto-optimal frontiers to design approximation schemes turns out to be a much more general approach, and in this paper we present a unified approach for solving a diverse category of combinatorial optimization problems using this technique.

Several methods exist for designing approximations schemes for combinatorial optimization problems with multiple objectives. Safer and Orlin (1995a) give necessary and sufficient conditions for the existence of fully polynomial-time approximation schemes in multicriteria combinatorial optimization and then use it for designing fast approximation schemes for multicriteria flow, knapsack, and scheduling problems (Safer and Orlin 1995b, Safer et al. 2004). Papadimitriou and Yannakakis (2000) propose an efficient procedure to construct an approximate Pareto-optimal frontier for discrete multiobjective optimization problems, and we use their procedure in designing approximation schemes for the previously mentioned problems, as explained in the next section.

1.3. Overview of Our Framework

We present a general framework that we use to design FPTASes for the problems given in §1.1. The main idea behind this framework is to identify the individual functions that constitute the given objective function and then compute an approximate Pareto-optimal frontier corresponding to these individual functions. It is possible to get an approximate Pareto-optimal frontier for many combinatorial optimization problems under the general condition that the corresponding "exact" problem is solvable in pseudo-polynomial time. The exact problem, for example, for the spanning tree problem is, given a vector of nonnegative integer edge weights c and a nonnegative integer K, does there exist a spanning tree T such that $c(T) = K$? For many combinatorial optimization problems (e.g., the spanning tree problem, the shortest path problem, and the knapsack problem) the exact problem can indeed be solved in pseudo-polynomial time. For the resource allocation and scheduling problems, the exact problem is a variant of the partition problem, and we show that it is also solvable in pseudo-polynomial time.

Our framework works in the following three stages:

1. Show that the optimal solution of the problem lies on the *Pareto-optimal front* of the corresponding multiobjective optimization problem.

2. Show that there is at least one solution in the *approximate Pareto-optimal front* that is an approximate solution of the given optimization problem.

3. Construct an approximate Pareto-optimal frontier. (To do so, in many cases it is sufficient to be able to solve the *exact problem* corresponding to the original single-objective optimization problem in polynomial time.)

In §2, we formally introduce the concept of Pareto-optimal frontier and approximate Pareto-optimal frontier and also revisit previous results on constructing an approximate Pareto-optimal frontier for multiobjective optimization problems, which we use in the later sections for designing our approximation schemes. We then give our general framework for designing FPTASes for combinatorial optimization problems in which several objective functions are combined into one and state the conditions needed for the framework to work in the main theorem of this paper in §3. Subsequently, we derive FPTASes for the problems mentioned in §1.1 as corollaries to the main theorem.

2. Preliminaries on Multiobjective Optimization

An instance π of a multiobjective optimization problem Π is given by a set of m functions $f_1, \ldots, f_m$. Each $f_i\colon X \to \mathbb{R}_+$ is defined over the same set of feasible solutions, X. Let $|\pi|$ denote the binary-encoding size of the instance π. Assume that each f_i takes values in the range $[2^{-p(|\pi|)}, 2^{p(|\pi|)}]$ for some polynomial p. We first define the Pareto-optimal frontier for multiobjective minimization problems.

Definition 2.1. Let π be an instance of a multiobjective minimization problem. A *Pareto-optimal frontier*, denoted by $P(\pi)$, is a set of solutions $x \in X$, such that there is no $x' \in X$ such that $f_i(x') \leqslant f_i(x)$ for all i, with strict inequality for at least one i.

In other words, $P(\pi)$ consists of all undominated solutions. In many cases, it may not be tractable to compute $P(\pi)$ (e.g., determining whether a point belongs to the Pareto-optimal frontier for the two-objective shortest path problem is NP-hard (Papadimitriou and Yannakakis 2000)), or the number of undominated solutions can be exponential in $|\pi|$ (e.g., for the two-objective shortest path problem). One way of getting around this problem is to look at an approximate Pareto-optimal frontier, which is defined next.

Definition 2.2. Let π be an instance of a multiobjective minimization problem. For $\epsilon > 0$, an *ϵ-approximate Pareto-optimal frontier*, denoted by $P_\epsilon(\pi)$, is a set of solutions, such that for all $x \in X$, there is $x' \in P_\epsilon(\pi)$ such that $f_i(x') \leqslant (1+\epsilon) f_i(x)$, for all i.

In the rest of the paper, whenever we refer to an (approximate) Pareto-optimal frontier, we mutually refer to both its set of solutions and their vectors of objective function values.

Even though $P(\pi)$ may contain exponentially many points, it is possible to construct an approximate Pareto-optimal frontier $P_\epsilon(\pi)$ for all $\epsilon > 0$ that has a polynomial number of points. The following theorem gives one possible way to construct such an approximate Pareto-optimal frontier in polynomial time.

Theorem 2.3 (Papadimitriou and Yannakakis 2000). *Let m be fixed and $\epsilon > 0$. One can determine a $P_\epsilon(\pi)$ in time polynomial in $|\pi|$ and $1/\epsilon$ if the following "gap problem" can be solved in polynomial time*: *Given an m-vector of values $(v_1, \ldots, v_m)$, either*

(i) *return a solution $x \in X$ such that $f_i(x) \leqslant v_i$ for all $i = 1, \ldots, m$; or*

(ii) *assert that there is no $x \in X$ such that $f_i(x) \leqslant (1-\epsilon') v_i$ for all $i = 1, \ldots, m$, where $(1-\epsilon')(1+\epsilon)^{1/2} = 1$.*

We sketch the proof because our approximation schemes are based on it.

Proof. Suppose we can solve the gap problem in polynomial time. An approximate Pareto-optimal frontier can then be constructed as follows. Consider the box in $\mathbb{R}^m$ of possible function values given by $\{(v_1, \ldots, v_m)\colon 2^{-p(|\pi|)} \leqslant v_i \leqslant 2^{p(|\pi|)}$ for all $i\}$. We divide this box into smaller boxes, such that in each dimension, the ratio of successive divisions is equal to $1+\epsilon''$, where $\epsilon'' = (1+\epsilon)^{1/2} - 1$. For each corner point of all such smaller boxes, we call the gap problem. Among all solutions returned by solving the gap problems, we keep only those solutions that are not Pareto dominated by any other solution. This is the required $P_\epsilon(\pi)$. Because there are $O((p(|\pi|)/\epsilon)^m)$ many smaller boxes, this can be done in polynomial time. □

Thus, it suffices to solve the gap problem to compute an approximate Pareto-optimal frontier. We give a procedure here for solving the gap problem with respect to minimization problems, but it can be extended to maximization problems as well (see §7). This procedure is due to Papadimitriou and Yannakakis (2000).

We restrict our attention to the case when $X \subseteq \{0,1\}^d$, because many combinatorial optimization problems can be framed as 0/1-integer programming problems. Further, we consider linear objective functions; that is, $f_i(x) = \sum_{j=1}^{d} a_{ij} x_j$, and each a_{ij} is a nonnegative integer. Suppose we want to solve the gap problem for the m-vector $(v_1, \ldots, v_m)$. Let $r = \lceil d/\epsilon' \rceil$, where ϵ' satisfies $(1-\epsilon') \cdot (1+\epsilon)^{1/2} = 1$. We first define a "truncated" objective function. For all $j = 1, \ldots, d$, if for some i, $a_{ij} > v_i$, we set $x_j = 0$, and drop the variable x_j from each of the objective functions. Let V be the index set of the remaining variables. Thus, the coefficients in each objective function are now less than or equal to v_i. Next, we define a new objective function $f_i'(x) = \sum_{j \in V} a_{ij}' x_j$, where $a_{ij}' = \lceil a_{ij} r / v_i \rceil$. In the new objective function, the maximum value of a coefficient is now r. For $x \in X$, by Lemma A.1 (see appendix) the following two statements hold:

- If $f_i'(x) \leqslant r$, then $f_i(x) \leqslant v_i$.
- If $f_i(x) \leqslant v_i(1-\epsilon')$, then $f_i'(x) \leqslant r$.

Therefore, to solve the gap problem, it suffices to find an $x \in X$ such that $f_i'(x) \leqslant r$, for $i = 1, \ldots, m$, or assert that no such

x exists. Because all the coefficients of $f_i'(x)$ are nonnegative integers, there are $r+1$ possible values that the function $f_i(x)$ can take so that the inequality is satisfied. Hence there are $(r+1)^m$ ways overall in which all inequalities $f_i'(x) \leqslant r$ can be simultaneously satisfied. Suppose we want to find if there is an $x \in X$ such that $f_i'(x) = b_i$ for $i = 1, \ldots, m$. This is equivalent to finding an x such that $\sum_{i=1}^m M^{i-1} f_i'(x) = \sum_{i=1}^m M^{i-1} b_i$, where $M = dr+1$ is a number greater than the maximum value that $f_i'(x)$ can take.

Given an instance π of a multiobjective linear optimization problem over a discrete set $X \subseteq \{0,1\}^d$, the exact version of the problem is the following: Given a nonnegative integer C and a vector $(c_1, \ldots, c_d) \in \mathbb{Z}_+^d$, does there exist a solution $x \in X$ such that $\sum_{j=1}^d c_j x_j = C$? The following theorem establishes the connection between solving the exact problem and the construction of an approximate Pareto-optimal frontier.

THEOREM 2.4 (PAPADIMITRIOU AND YANNAKAKIS 2000). *Suppose we can solve the exact version of the problem in pseudo-polynomial time, then there is an algorithm for computing an approximate Pareto-optimal frontier $P_\epsilon(\pi)$ in time polynomial in the input size of the problem as well as $1/\epsilon$.*

PROOF. The gap problem can be solved by making at most $(r+1)^m$ calls to the pseudo-polynomial time algorithm for the exact problem, and the input to each call has numerical values of order $O((d^2/\epsilon)^{m+1})$. Therefore, all calls to the algorithm take polynomial time, hence the gap problem can be solved in polynomial time. The result now follows from Theorem 2.3. □

3. The General Formulation of the FPTAS

In this section, we present a general formulation of the FPTAS based on the ideas given in §2. We then show how this general framework can be adapted to obtain FPTASes for the problems given in §1.1.

Let $f_1, \ldots, f_m$, for m fixed, be functions that satisfy the conditions given in the beginning of §2. Let $h\colon \mathbb{R}_+^m \to \mathbb{R}_+$ be any function that satisfies the following two conditions:

1. $h(y) \leqslant h(y')$ for all $y, y' \in \mathbb{R}_+^m$ such that $y_i \leqslant y_i'$ for all $i = 1, \ldots, m$; and
2. $h(\lambda y) \leqslant \lambda^c h(y)$ for all $y \in \mathbb{R}_+^m$ and $\lambda > 1$, for some fixed $c > 0$.

In particular, h includes all the l_p norms (with $c = 1$), and the product of a fixed number (say, k) of linear functions (with $c = k$). We denote by $f(x)$ the vector $(f_1(x), \ldots, f_m(x))$.

Consider the following general optimization problem:

$$\text{minimize } g(x) = h(f(x)), \quad x \in X. \tag{3}$$

We show that if we can solve the corresponding exact problem (with a singe linear objective function) in pseudo-polynomial time, then there is an FPTAS to solve this general optimization problem as well. Also, even though we state all our results for minimization problems, there is a straightforward extension of the method to maximization problems as well.

LEMMA 3.1. *There is at least one optimal solution x^* to* (3) *such that $x^* \in P(\pi)$.*

PROOF. Let $\hat{x}$ be an optimal solution of (3). Suppose $\hat{x} \notin P(\pi)$. Then there exists $x^* \in P(\pi)$ such that $f_i(x^*) \leqslant f_i(\hat{x})$ for $i = 1, \ldots, m$. By Property 1 of h, $h(f(x^*)) \leqslant h(f(\hat{x}))$. Thus x^* minimizes the function g and is in $P(\pi)$. □

LEMMA 3.2. *Let $\hat{\epsilon} = (1+\epsilon)^{1/c} - 1$. Let $\hat{x}$ be a solution in $P_{\hat{\epsilon}}(\pi)$ that minimizes $g(x)$ over all the points $x \in P_{\hat{\epsilon}}(\pi)$. Then $\hat{x}$ is a $(1+\epsilon)$-approximate solution of* (3)*; that is, $g(\hat{x})$ is at most $(1+\epsilon)$ times the value of an optimal solution to* (3)*.*

PROOF. Let x^* be an optimal solution of (3) that is in $P(\pi)$. By the definition of an ϵ-approximate Pareto-optimal frontier, there exists $x' \in P_{\hat{\epsilon}}(\pi)$ such that $f_i(x') \leqslant (1+\hat{\epsilon}) f_i(x^*)$, for all $i = 1, \ldots, m$. Therefore,

$$\begin{aligned} g(x') &= h(f_1(x'), \ldots, f_m(x')) \\ &\leqslant h((1+\hat{\epsilon}) f_1(x^*), \ldots, (1+\hat{\epsilon}) f_m(x^*)) \\ &\leqslant (1+\hat{\epsilon})^c h(f_1(x^*), \ldots, f_m(x^*)) = (1+\epsilon) g(x^*), \end{aligned}$$

where the first inequality follows from Property 1 and the second inequality follows from Property 2 of h. Because $\hat{x}$ is a minimizer of $g(x)$ over all the solutions in $P_{\hat{\epsilon}}(\pi)$, $g(\hat{x}) \leqslant g(x') \leqslant (1+\epsilon) g(x^*)$. □

From these two lemmata and Theorem 2.4, we get the main theorem of this paper regarding the existence of an FPTAS for solving (3).

THEOREM 3.3. *Suppose the exact problem corresponding to the functions given in* (3) *can be solved in pseudo-polynomial time. Then there is an FPTAS for solving the general optimization problem* (3) *when m is fixed.*

The FPTAS can now be summarized as follows.

1. Construct an approximate Pareto-optimal frontier $P_{\hat{\epsilon}}(\pi)$ of the m objective functions that constitute the given objective function, where $\hat{\epsilon} = (1+\epsilon)^{1/c} - 1$.
2. Among all the solutions in $P_{\hat{\epsilon}}(\pi)$, return the one with the minimum function value.

Finally, we establish the running time of the above algorithm, when $P_{\hat{\epsilon}}(\pi)$ is computed using the procedure given in §2.

LEMMA 3.4. *The running time of the algorithm is polynomial in $|\pi|$ and $1/\epsilon$.*

PROOF. There are $O((p(|\pi|)/\epsilon)^m)$ corner points for which we need to solve the gap problem. Solving each gap problem requires calling the algorithm for solving the exact problem $O(r^m)$ times, which is $O((d/\epsilon)^m)$. The magnitude of the largest number input to the algorithm for the exact problem is $O((d^2/\epsilon)^{m+1})$. Hence the running time of the algorithm is $O(((p(|\pi|)d)/\epsilon^2) \cdot PP((d^2/\epsilon)^{m+1}, m, d))$, where $PP(M, m, d)$ is the running time of the pseudo-polynomial time algorithm for the exact problem with maximum magnitude of an input number equal to M. □

It should be noted that the algorithm is very generic and not necessarily the most efficient. For many special cases, there are more efficient ways to construct an approximate Pareto-optimal frontier, either by solving the gap problem more efficiently, or by making fewer queries to the gap problem solver. For example, Diakonikolas and Yannakakis (2009) and Vassilvitskii and Yannakakis (2005) give more efficient algorithms for computing approximate Pareto-optimal curves for multiobjective optimization problems when the number of objectives is two or three. Similarly, Warburton (1987) presents an algorithm for computing an approximate Pareto-optimal frontier for the multiobjective shortest path problem whose running time is better than that of the algorithm presented in §2. Therefore, even though the running times of the approximation schemes we present in this paper may appear to be prohibitively expensive, in many cases it is possible to engineer the algorithm to obtain approximation schemes with improved running times.

4. FPTAS for Scheduling and Resource Allocation Problems

Using the framework presented in §3, we give FPTASes for the max-min resource allocation problem, the $Rm||C_{\max}$ problem, and the vector scheduling problem.

4.1. The $Rm||C_{\max}$ Problem and the Max-Min Resource Allocation Problem

Recall the $Rm||C_{\max}$ scheduling problem defined in the introduction. There are m machines and n jobs, and the processing time of job k on machine i is p_{ik}. The objective is to schedule the jobs to minimize the makespan. The max-min resource allocation problem is similar to this scheduling problem, except that the objective here is to maximize the minimum completion time over all the machines. Observe that this corresponds to h being the l_∞-norm with $c = 1$ in the formulation given by (3).

We first give an integer programming formulation for the two problems. Let x_{ik} be the variable that is one if job k is assigned to machine i, zero otherwise. The m objective functions in this case (corresponding to each agent/machine) are given by $f_i(x) = \sum_{k=1}^{n} p_{ik}x_{ik}$, and the set X is given by

$$\sum_{i=1}^{m} x_{ik} = 1 \quad \text{for } k = 1, \ldots, n, \tag{4a}$$

$$x_{ik} \in \{0, 1\} \quad \text{for } i = 1, \ldots, m,\ k = 1, \ldots, n. \tag{4b}$$

The exact version for both problems is the following: Given an integer C and nonnegative integer coefficients c_{ik}, does there exist a 0/1-vector x such that $\sum_{i=1}^{m}\sum_{k=1}^{n} c_{ik}x_{ik} = C$, subject to the constraints (4a) and (4b)? The following lemma establishes that the exact problem can be solved in pseudo-polynomial time.

Lemma 4.1. *The exact problem for the max-min resource allocation problem and the $Rm||C_{\max}$ problem can be solved in pseudo-polynomial time.*

Proof. The exact problem can be viewed as a reachability problem in a directed graph. The graph is an $(n+1)$-partite directed graph; let us denote the partitions of this digraph by $V_0, \ldots, V_n$. The partition V_0 has only one node, labeled as $v_{0,0}$ (the source node), all other partitions have $C+1$ nodes. The nodes in V_k for $1 \leqslant k \leqslant n$ are labeled as $v_{k,0}, \ldots, v_{k,C}$. The arcs in the digraph are from nodes in V_k to nodes in V_{k+1} only, for $0 \leqslant k \leqslant n-1$. For all $c \in \{c_{1,k+1}, \ldots, c_{m,k+1}\}$, there is an arc from $v_{k,j}$ to $v_{k+1,j+c}$, if $j + c \leqslant C$. Then there is a solution to the exact version if and only if there is a directed path from the source node $v_{0,0}$ to the node $v_{n,C}$. Finding such a path can be accomplished by doing a depth-first search from the node $v_{0,0}$. The corresponding solution for the exact problem (if it exists) can be obtained using the path found by the depth-first search algorithm. □

Therefore, we obtain FPTASes for both the $Rm||C_{\max}$ problem as well as the max-min resource allocation problem. For the max-min resource allocation problem with a fixed number of agents, we give the first FPTAS, though approximation schemes for the $R||C_{\max}$ problem already exist in the literature (e.g., Horowitz and Sahni 1976 and Lenstra et al. 1990). Further, Theorem 3.3 implies that we get an FPTAS even when the objectives for different agents/machines are combined together using a function that satisfies Properties 1 and 2 given in §3. We therefore have the following corollary to Theorem 3.3.

Corollary 4.2. *There is an FPTAS for the max-min resource allocation problem with a fixed number of agents. Further, we also get an FPTAS for the max-min resource allocation problem with a fixed number of agents and the unrelated parallel machine problem when the objectives for different agents/machines are combined together using a function that satisfies Properties* 1 *and* 2 *given in* §3.

4.2. The Vector Scheduling Problem

The vector scheduling problem is a generalization of the $Rm||C_{\max}$ problem. In this problem, each job requires d resources for execution on each machine. Job k consumes an amount p_{ik}^j of a resource j on machine i. Suppose J_i is the set of jobs assigned to machine i. Thus the total usage of resource j on machine i is $\sum_{k\in J_i} p_{ik}^j$. The objective is to minimize over all the machines i and all the resources j, the value $\sum_{k\in J_i} p_{ik}^j$. We assume that both d and m are fixed.

Similar to the $Rm||C_{\max}$ problem, let x_{ik} be a variable that is one if job k is assigned to machine i, zero otherwise. In this case, we have a total of md functions and $f_{ij}(x) = \sum_{k=1}^{n} p_{ik}^j x_{ik}$, for $i = 1, \ldots, m$ and $j = 1, \ldots, d$. The md objective functions are combined together using the l_∞ norm in this problem. The underlying set of constraints is the same as given by (4a)–(4b). Therefore, the exact algorithm for the $Rm||C_{\max}$ problem works for the vector scheduling problem as well, and because we have a fixed number of objective functions, we get an FPTAS for the vector scheduling problem as well. Hence we have the following corollary to Theorem 3.3.

COROLLARY 4.3. *There is an FPTAS for the vector scheduling problem when the number of machines as well as the number of resources are fixed, even for the case when each job can use a different amount of a particular resource on different machines.*

5. FPTAS for Minimizing the Product of Two Linear Objective Functions

In this section, we give a general framework for designing FPTASes for problems in which the objective is to minimize the product of two linear cost functions. We then apply this technique to some product combinatorial optimization problems on graphs and then extend it to the case where the objective function is a product of a fixed number of linear functions.

5.1. Formulation of the FPTAS

Consider the following optimization problem:

$$\text{minimize } g(x) = f_1(x) \cdot f_2(x), \quad x \in X, \tag{5}$$

where $f_i\colon X \to \mathbb{Z}_+$ are linear functions and $X \subseteq \{0,1\}^d$. In our general formulation given by (3), the corresponding function h for this case is $h(y_1, y_2) = y_1 \cdot y_2$, and so $c = 2$. Thus, if we can construct an approximate Pareto-optimal frontier for $f_1(x)$ and $f_2(x)$ in polynomial time, we will be able to design an FPTAS for the product optimization problem. Therefore, we get the following corollary to Theorem 3.3.

COROLLARY 5.1. *There is an FPTAS for the problem given by* (5) *if the following exact problem can be solved in pseudo-polynomial time*: *Given* $(c_1, \ldots, c_d) \in \mathbb{Z}_+^d$ *and* $K \in \mathbb{Z}_+$, *does there exist* $x \in X$ *such that* $\sum_{i=1}^d c_i x_i = K$?

5.2. FPTASes for Some Problems with the Product Objective Function

Using Corollary 5.1, we now construct FPTASes for several combinatorial optimization problems involving the product of two objective functions.

1. *Spanning tree problem.* In this case, the exact problem is the following: given a graph $G = (V, E)$ with cost function $c\colon E \to \mathbb{Z}_+$ and a positive integer K, does there exist a spanning tree $T \subseteq E$ whose cost is equal to exactly K? Barahona and Pulleyblank (1987) give an $O((n^3 + p^2)p^2 \log p)$ algorithm for solving the exact problem, where n is the number of vertices in the graph and $p = n \cdot \max_e (c(e))$. Thus we have an FPTAS for the spanning tree problem with the product of two cost functions as the objective.

2. *Shortest s-t path problem.* The exact problem in this case is the following: given a graph $G = (V, E)$, vertices $s, t \in V$, a distance function $d\colon E \to \mathbb{Z}_+$, and an integer K, is there an s-t path with length equal to exactly K? Note that for the shortest path problem, the exact problem is strongly NP-complete, because it includes the Hamiltonian path problem as a special case. To circumvent this issue, we relax the original problem to that of finding a walk (in which a vertex can be visited more than once) between the vertices s and t that minimizes the product objective. The optimal solution of the relaxed problem will have the same objective function value as that of an optimal solution of the original problem, because any s-t walk can be truncated to get an s-t path. Therefore, it suffices to get an approximate solution for the relaxed problem.

The corresponding exact s-t walk problem is the following: Does there exist an s-t walk in the graph whose length is equal to exactly a given number $K \in \mathbb{Z}_+$? Because we are dealing with nonnegative weights, this problem can be solved in $O(mnK)$ time by dynamic programming, where n is the number of vertices and m is the number of edges in the graph. If the solution given by the algorithm is a walk instead of a path, we remove the cycles from the walk to get a path. Hence we obtain an FPTAS for the shortest s-t path problem with the product of two distance functions as the objective.

3. *Knapsack problem.* The exact problem for the knapsack problem is the following: given a set I of items with profit $p\colon I \to \mathbb{Z}_+$, size $s\colon I \to \mathbb{Z}_+$ and a given capacity C, does there exist a subset of I satisfying the capacity constraint and having total profit exactly equal to a given integer K? Again, this exact problem can be solved in $O(nK)$ time by dynamic programming, where n is the number of objects. Therefore we get an FPTAS for the product version of the knapsack problem.

4. *Minimum cost flow problem.* The problem we have is the following: given a directed graph $G = (V, A)$, vertices $s, t \in V$, an amount of flow $d \in \mathbb{Z}_+$ to send from s to t, capacities $u\colon A \to \mathbb{Z}_+$, and two cost functions $c_1, c_2\colon A \to \mathbb{Z}_+$, find a feasible s-t flow x of total amount d such that $c_1(x) \cdot c_2(x)$ is minimized. The minimum cost flow problem is different from the other three problems, because it suffices to find fractional solutions instead of integer solutions. In this case, the gap problem as stated in Theorem 2.3 can be solved directly using linear programming. Therefore we obtain an FPTAS for the minimum cost flow problem with the product objective function as well.

Note that in this case, the approximate solution that we obtain may not necessarily be integral. This is because when we solve the gap problem, we introduce constraints of the form $f_i(x) \leqslant (1 - \epsilon')v_i$ corresponding to each of the two objectives, in addition to the flow conservation and capacity constraints. This means that the constraint set may not be totally unimodular, and hence the solution obtained can possibly be nonintegral.

One advantage of our method is that it can be used to get an approximation scheme for the product version of an optimization problem even if the original problem is NP-hard, for example in the case of the knapsack problem, whereas previously existing methods cannot handle this case.

5.3. Products of More Than Two Linear Objective Functions

Another advantage of our technique over existing methods for designing FPTASes for product optimization problems (Kern and Woeginger 2007, Goyal et al. 2011) is that it can be easily extended to the case where the objective function is a product of more than two linear functions, as long as the total number of functions involved in the product is a constant. Consider the following generalization of the problem given by (5):

$$\text{minimize } g(x) = f_1(x) \cdot f_2(x) \cdot \cdots \cdot f_m(x), \quad x \in X, \tag{6}$$

where $f_i\colon X \to \mathbb{Z}_+$ are linear functions, for $i = 1, \ldots, m$, $X \subseteq \{0,1\}^d$ and m is a fixed number. This again fits into our framework given by (3), with $c = m$. Thus our technique yields an FPTAS for the problem given by (6) as well. We have therefore established the following corollary to Theorem 3.3.

Corollary 5.2. *There is an FPTAS for the problem given by* (6) *if m is fixed and if the following exact problem can be solved in pseudo-polynomial time*: *Given $(c_1, \ldots, c_d) \in \mathbb{Z}_+^d$ and $K \in \mathbb{Z}_+$, does there exist $x \in X$ such that $\sum_{i=1}^d c_i x_i = K$?*

6. FPTASes for Robust Weighted Multiobjective Optimization Problems

Consider the following robust version of a weighted multiobjective optimization problem given by Hu and Mehrotra (2010):

$$\text{minimize } g(x) = \max_{w \in W} w^T f(x), \quad x \in X \subseteq \{0,1\}^d. \tag{7}$$

Here, $f(x) = (f_1(x), \ldots, f_m(x)) \in \mathbb{R}_+^m$ is a vector of m function values, $W \subseteq W_f$, where $W_f = \{w \in \mathbb{R}_+^m\colon w_1 + \cdots + w_m = 1\}$ (i.e., the weights are nonnegative and they sum up to one) and W is a compact convex set. We assume that we can optimize a linear function over the set W in polynomial time; this ensures that the function $g(x)$ can be computed efficiently. Examples of some of the forms that the weight set W can take are as follows:

1. Simplex weight set: $W = W_f = \{w \in \mathbb{R}_+^m\colon w_1 + \cdots + w_m = 1\}$.

2. Ellipsoidal weight set: $W = \{w \in W_f\colon (w - \hat{w})^T S^{-1}(w - \hat{w}) \leqslant \gamma^2\}$, where $\hat{w}$, $\gamma > 0$, and S is a $m \times m$ positive semi-definite matrix.

3. Box weight set: $W = \{w \in W_f\colon \|w\|_\infty \leqslant k\}$, where $k > 0$.

In particular, the model with the simplex weight set can be considered to be a generalization of the robust optimization model with a fixed number of scenarios. The robust optimization problem with a fixed number of scenarios has the following form:

$$\text{minimize } h(x) = \max_{c \in \{c_1, \ldots, c_m\}} c^T x, \quad x \in X \subseteq \{0,1\}^d. \tag{8}$$

The connection between the problems given by (7) and (8) is established in the following lemma.

Lemma 6.1. *The problem given by* (8) *is equivalent to the problem given by* (7) *when $f_i(x) = c_i^T x$ for $i = 1, \ldots, m$ and the weight set is the simplex weight set W_f.*

Proof. For a given solution $x \in X$, its objective function value $h(x)$ in the formulation (8) is given by

$$\begin{aligned} h(x) &= \max\{c^T x\colon c \in \{c_1, \ldots, c_m\}\} \\ &= \max\{c^T x\colon c \in \text{conv}(\{c_1, \ldots, c_m\})\} \\ &= \max\{w_1 c_1^T x + \cdots + w_m c_m^T x\colon w \in W_f\} \\ &= \max\{w^T f(x)\colon w \in W_f\} \\ &= g(x), \end{aligned}$$

where $g(x)$ is the objective function value in the formulation given by (7), $\text{conv}(\{c_1, \ldots, c_m\})$ denotes the convex hull of the m points $c_1, \ldots, c_m$, and $f_i(x) = c_i^T x$ for $i = 1, \ldots, m$. This establishes the equivalence between the optimization problem given by (7) with the simplex weight set and the optimization problem given by (8). □

Using this observation, we establish the NP-hardness of the optimization problem given by (7).

Lemma 6.2. *The optimization problem given by* (7) *is NP-hard for the shortest path problem and the spanning tree problem.*

Proof. The following two-scenario robust optimization problem is known to be NP-hard for the shortest path problem and the spanning tree problem (Kouvelis and Yu 1997):

$$\text{minimize } h(x) = \max_{c \in \{c_1, c_2\}} c^T x, \quad x \in X \subseteq \{0,1\}^d. \tag{9}$$

Problem (9) is equivalent to the form given by (7) with $f_1(x) = c_1^T x$, $f_2(x) = c_2^T x$ and $W = \{(w_1, w_2) \in \mathbb{R}_+^2\colon w_1 + w_2 = 1\}$. Therefore the optimization problem given by (7) is also NP-hard in general. □

The function h that combines the functions $f_1(x), \ldots, f_m(x)$ in this case is

$$h(y_1, \ldots, y_m) = \max_{w \in W}(w_1 y_1 + \cdots + w_m y_m).$$

It can be easily verified that the function h satisfies Properties 1 and 2 given in §3, with $c = 1$. Therefore we get the following corollary to Theorem 3.3, which establishes the existence of FPTASes for the robust version of the shortest path problem, the spanning tree problem, and the knapsack problem.

Corollary 6.3. *There is an FPTAS for the problem given by* (7) *when m is fixed if the following exact problem can be solved in pseudo-polynomial time*: *Given $(c_1, \ldots, c_d) \in \mathbb{Z}_+^d$ and $K \in \mathbb{Z}_+$, does there exist $x \in X$ such that $\sum_{i=1}^d c_i x_i = K$?*

7. FPTASes for Problems with Rational Objective Functions

In this section, we consider combinatorial optimization problems involving a ratio of two linear objectives as given in the introduction:

$$\begin{aligned} \text{minimize } & g(x) = \frac{f_1(x)}{f_2(x)} = \frac{a_0 + a_1x_1 + \cdots + a_dx_d}{b_0 + b_1x_1 + \cdots + b_dx_d} \\ \text{s.t. } & x \in X \subseteq \{0,1\}^d. \end{aligned} \tag{10}$$

We assume that $f_1(x) > 0$, $f_2(x) > 0$ for all $x \in X$. The situation here is different from the problems we have considered previously, because in this case, in some sense, we are attempting to minimize f_1, while simultaneously maximizing f_2. Therefore we cannot use Theorem 3.3 directly for obtaining an FPTAS. We need to modify the definition of the Pareto-optimal frontier and the approximate Pareto-optimal frontier for this problem and restate the gap theorem for the modified definition. We first give the appropriate definition of the Pareto-optimal and the approximate Pareto-optimal frontier for this problem.

Definition 7.1. Consider the problem given by (10). For this problem, the *Pareto-optimal frontier* $P(\pi)$ is the set of all points x for which there is no x' such that $f_1(x') \leqslant f_1(x)$ and $f_2(x') \geqslant f_2(x)$ with strict inequality for at least one of them.

Definition 7.2. For the problem given by (10), for $\epsilon > 0$, an *approximate Pareto-optimal frontier* $P_\epsilon(\pi)$ is a set of solutions such that for all $x \in X$, there is $x' \in P_\epsilon(\pi)$ such that $f_1(x') \leqslant (1+\epsilon)f_1(x)$ and $f_2(x') \geqslant f_2(x)/(1+\epsilon)$.

We now state the modified gap theorem for this problem, which is a straightforward extension of the gap theorem of Papadimitriou and Yannakakis (2000). The proof of this theorem is very similar to the one for Theorem 2.3, so we omit it.

Theorem 7.3 (Modified Gap Theorem). *Let $\epsilon > 0$. One can determine a $P_\epsilon(\pi)$ in time polynomial in $|\pi|$ and $1/\epsilon$ if the following "gap problem" can be solved in polynomial time*: *Given a vector of values (v_1, v_2), either*

(i) *return a solution $x \in X$ such that $f_1(x) \leqslant v_1$ and $f_2(x) \geqslant v_2$, or*

(ii) *assert that there is no $x \in X$ such that $f_1(x) \leqslant (1-\epsilon_1')v_1$ and $f_2(x) \geqslant (1+\epsilon_2')v_2$, where $(1-\epsilon_1')(1+\epsilon)^{1/2} = 1$ and $(1+\epsilon_2') = (1+\epsilon)^{1/2}$.*

It is easy to see that Lemma 3.1 holds in this case, with the modified definition of the Pareto-optimal frontier. The analog of Lemma 3.2 is given next.

Lemma 7.4. *Let $P_\epsilon(\pi)$ denote the approximate Pareto-optimal frontier of the functions f_1 and f_2 corresponding to minimizing f_1 and maximizing f_2. Let $\hat{x}$ be the solution in $P_\epsilon(\pi)$ that minimizes $g(x)$ over all points $x \in P_\epsilon(\pi)$. Then $\hat{x}$ is a $(1+\epsilon)^2$-approximate solution for* (10).

Proof. Let x^* be an optimal solution of (10) that is in $P(\pi)$. By the definition of an ϵ-approximate Pareto-optimal frontier, there exists $x' \in P_\epsilon(\pi)$ such that $f_1(x') \leqslant (1+\epsilon)f_1(x^*)$ and $f_2(x') \geqslant f_2(x^*)/(1+\epsilon)$. Therefore,

$$g(x') \leqslant \frac{(1+\epsilon)f_1(x^*)}{(1+\epsilon)^{-1}f_2(x^*)} = (1+\epsilon)^2 g(x^*).$$

Because $\hat{x}$ is a minimizer of $g(x)$ over all the solutions in $P_\epsilon(\pi)$, $g(\hat{x}) \leqslant g(x') \leqslant (1+\epsilon)^2 g(x^*)$. □

The following theorem is an analog of Theorem 3.3 for this case.

Theorem 7.5. *There is an FPTAS for the problem given by* (10) *if the following exact problem can be solved in pseudo-polynomial time*: *Given $(c_1, \ldots, c_d) \in \mathbb{Z}_+^d$ and $K \in \mathbb{Z}_+$, does there exist $x \in X$ such that $\sum_{i=1}^d c_i x_i = K$?*

We give a proof of this theorem here, as it involves both maximization and minimization of the underlying objective functions.

Proof. From Theorem 7.3, it suffices to give a polynomial-time algorithm to solve the gap problem. Suppose we want to solve the gap problem for the two-vector (v_1, v_2). Let $r_1 = \lceil d/\epsilon_1' \rceil$. We first define a "truncated" objective function. For all $j = 1, \ldots, d$, if for some j, $a_j > v_1$, we set $x_j = 0$, and drop the variable x_j from each of the objective functions. Let V be the index set of the remaining variables. Thus, the remaining coefficients in f_1 are now less than or equal to v_1. Next, we define a new objective function $f_1'(x) = \sum_{j \in V} a_j' x_j$, where $a_j' = \lceil a_j r_1 / v_1 \rceil$. In the new objective function, the maximum value of a coefficient is now r_1. For $x \in X$, the following two statements hold by Lemma A.1:

- If $f_1'(x) \leqslant r_1$, then $f_1(x) \leqslant v_1$.
- If $f_1(x) \leqslant v_1(1-\epsilon_1')$, then $f_1'(x) \leqslant r_1$.

For f_2, we do the following. Let $r_2 = \lfloor d/\epsilon_2' \rfloor$. Let $f_2'(x) = \sum_{j \in V} b_j' x_j$, where $b_j' = \min(r_2, \lfloor b_j r_2 / v_2 \rfloor)$. So in f_2', all the coefficients are no more than r_2. The following two statements hold by Lemma A.2:

- If $f_2'(x) \geqslant r_2$ then $f_2(x) \geqslant v_2$.
- If $f_2(x) \geqslant (1+\epsilon_2')v_2$ then $f_2'(x) \geqslant r_2$.

Therefore, to solve the gap problem, it suffices to find an $x \in X$ such that $f_1'(x) \leqslant r_1$ and $f_2'(x) \geqslant r_2$, or assert that no such x exists. There are $r_1 + 1$ possible values that the function $f_1'(x)$ can take to satisfy the inequality $f_1'(x) \leqslant r_1$, and there are at most $r_2 d$ possible values that the function $f_2'(x)$ can take to satisfy the inequality $f_2'(x) \geqslant r_2$. Hence there are $O(r_1 r_2 d)$ ways overall in which both the inequalities can be simultaneously satisfied. Suppose we want to find if there is an $x \in X$ such that $f_i'(x) = b_i$ for $i = 1, 2$. This is equivalent to finding an x such that $f_1'(x) + M f_2'(x) = b_1 + M b_2$, where $M = d \cdot \max(r_1, r_2) + 1$ is a number greater than the maximum value that $f_i'(x)$ can take, for $i = 1, 2$. Hence, if we have a pseudo-polynomial time algorithm for solving the exact problem, we can solve the gap problem in polynomial time. □

This theorem implies than we can use our framework to get an FPTAS, for example, for the knapsack problem with a rational objective. In fact, it is not hard to see that the method can be extended to functions g having the form f_1f_2/f_3f_4, or $f_1/f_2 + f_3/f_4$ as well. As long as the number of functions f_i is fixed, we will get an FPTAS for the problem using our framework.

7.1. FPTAS for Assortment Optimization Problems

The problem of minimizing a sum-of-ratios form often arises in assortment optimization in the context of retail management. In this section, we obtain FPTASes for two models of the assortment optimization problem: the mixture of logits choice model and the nested logit choice model. For an introduction to the logit choice models, see the survey paper by Kök et al. (2009).

In the assortment optimization problem with the mixture of logits choice model, we have a set of n products indexed by $N = \{1, \ldots, n\}$ and m customer classes indexed by $C = \{1, \ldots, m\}$. The demand of a customer in a customer class $i \in C$ is modeled using a multinomial logit choice model with parameters $(v_{i0}, v_{i1}, \ldots, v_{in}) \in \mathbb{R}_+^{n+1}$. Here, v_{i0} denotes the preference of a customer in class i for purchasing no item, and v_{ij} is the preference of the customer to purchase product $j \in N$. If an assortment $S \subseteq N$ is offered to a customer in the class $i \in C$, the probability that the customer purchases a product $j \in N$ is given by

$$p_{ij}(S) = \begin{cases} \dfrac{v_{ij}}{v_{i0} + \sum_{k\in S} v_{ik}} & j \in S, \\ 0 & \text{otherwise.} \end{cases}$$

The profit corresponding to the purchase of an item j is w_j. Therefore the total profit from customer class $i \in C$ when an assortment $S \subseteq N$ is offered to the customers in class i is given by

$$f_i(S) = \sum_{j\in S} p_{ij}(S)w_j = \frac{\sum_{j\in S} w_j v_{ij}}{v_{i0} + \sum_{j\in S} v_{ij}}.$$

Let λ_i denote the fraction of the customers in class i, where $\sum_{i\in C} \lambda_i = 1$. The optimization problem is to find an assortment S that maximizes the objective function

$$g(S) = \sum_{i\in C} \lambda_i f_i(S).$$

Thus, in this case the objective function is a sum of m ratios. This problem is NP-hard even when there are only $m = 2$ customer classes but admits a PTAS if m is fixed (Rusmevichientong et al. 2010). Using our framework, we can get an FPTAS for the case when m is fixed as follows. Let x_j be the variable that is one if product $j \in N$ is offered in an assortment, zero otherwise. The objective function is $g(x) = \sum_{i=1}^m f_{i1}(x)/f_{i2}(x)$, where

$$f_{i1}(x) = \lambda_i \sum_{j=1}^n w_j v_{ij} x_j,$$

$$f_{i2}(x) = v_{i0} + \sum_{j=1}^n v_{ij} x_j.$$

There are no constraints in this problem. The exact problem in this case is, given a vector $c \in \mathbb{Z}_+^n$ and a nonnegative integer C, does there exist $x \in \{0, 1\}^n$ such that $\sum_{j=1}^n c_j x_j = C$? This is the subset-sum problem that can be solved in pseudo-polynomial time by dynamic programming. Hence we get an FPTAS for the assortment optimization problem with the mixture of logits choice model. Indeed, we have the following corollary to Theorem 2.3.

Corollary 7.6. *The assortment optimization problem with the mixture of logits choice model admits an FPTAS when the number of customer classes is fixed.*

In the mixture of logits choice model, the likelihood of choosing between two alternative products is independent of the assortment offered to the customer. This may not necessarily be true in practice. An alternative model that takes care of this anomaly is the nested logit choice model. In this model, there are G partitions of the product set N given by $H_1, \ldots, H_G$, where G is fixed. Assuming that there is only one class of customers, the probability that a customer purchases a product $j \in N$ when offered an assortment $S \subseteq N$ is given by

$$p_j(S) = \begin{cases} \dfrac{v_j}{(\sum_{l\in H_g\cap S} v_l)^{\alpha_g}} \cdot \dfrac{1}{1 + \sum_{k=1}^G (\sum_{l\in H_k\cap S} v_l)^{1-\alpha_k}} \\ \qquad \text{if } j \in H_g \cap S \text{ for some } g, \\ 0 \quad \text{otherwise.} \end{cases}$$

Here, $0 \leqslant \alpha_g \leqslant 1$ for all $g = 1, \ldots, G$ and $v_l \in \mathbb{Z}_{\geqslant 0}$ for all $l \in N$. In this model, the likelihood of choosing between two products is independent of the assortment offered if they are in the same partition, but depends on the assortment if they are in different partitions. The probability of not purchasing any product is $p_0(S) = 1/(1 + \sum_{k=1}^G (\sum_{l\in H_k\cap S} v_l)^{1-\alpha_k})$.

In the capacitated version of this problem, we also have a constraint $\sum_{i\in S} c_i \leqslant K$, where $c_i \in \mathbb{Z}_{\geqslant 0}$ corresponds to the capacity taken up by the product i and $K \in \mathbb{Z}_+$ corresponds to the total capacity available. The objective is to find an assortment S that maximizes $\sum_{j\in S} p_j(S)w_j$ subject to the capacity constraint. Rusmevichientong et al. (2009) show that this problem is NP-hard but admits a PTAS when G is fixed. They also prove that to get an approximate solution of this problem, it suffices to find an approximate solution of

the following sum-of-ratios problem:

$$\begin{aligned}\text{maximize}\quad & g(S_1,\ldots,S_G)=\sum_{i=1}^{G}\frac{\sum_{l\in S_i}u_l}{(\sum_{l\in S_i}v_l)^{\alpha_i}}\\ \text{s.t.}\quad & \sum_{i=1}^{G}\sum_{l\in S_i}c_l\leqslant K,\\ & S_i\subseteq H_i,\quad \text{for all } i=1,\ldots,G.\end{aligned}$$

Here, $u_l\in\mathbb{Z}_{\geqslant 0}$ for all $l\in N$. Let x_l be the indicator variable that is one if an item $l\in N$ is selected, zero otherwise. The objective function is $g(x)=\sum_{i=1}^{G}f_{i1}(x)/(f_{i2}(x))^{\alpha_i}$, where

$$f_{i1}(x)=\sum_{l\in H_i}u_lx_l,$$

$$f_{i2}(x)=\sum_{l\in H_i}v_lx_l.$$

Moreover, if $S_i=\varnothing$, then we count zero for the term $f_{i1}(x)/(f_{i2}(x))^{\alpha_i}$ in the objective function. Note that the denominator in each of the ratios in this problem is nonlinear. However, because each exponent α_i is upper bounded by one, we can still use our framework to get an FPTAS for this problem. First, we choose some k of the G sets $S_1,\ldots,S_G$ to be nonempty and the rest of the sets to be empty. Because there are G groups, we will need to do this 2^G times to cover all the possible cases. This does not affect the polynomial running time of our algorithm as G is fixed. Once we choose the k sets, say $S_{i_1},\ldots,S_{i_k}$ to be nonempty, we construct the Pareto-optimal frontier corresponding to maximizing the k linear functions $f_{i_11},\ldots,f_{i_k1}$ and minimizing the k linear functions $f_{i_12},\ldots,f_{i_k2}$. To ensure that each of the k sets $S_{i_1},\ldots,S_{i_k}$ is nonempty, we set the lower bound for the numerator function corresponding to these groups to be one when solving the gap problem (see the proof of Theorem 2.3). The underlying set of constraints is given by

$$\sum_{l\in N}c_lx_l\leqslant K,$$

$$x_l\in\{0,1\},\quad l\in N.$$

This is the knapsack constraint, and the corresponding exact problem can be solved in pseudo-polynomial time by dynamic programming. Hence we get an FPTAS for the assortment optimization problem in the nested logit choice model with capacity constraints. We therefore have the following corollary to Theorem 7.5.

Corollary 7.7. *The capacitated assortment optimization problem with nested logit choice model admits an FPTAS when the number of partitions G of the set of products N is fixed.*

8. Conclusion

The main contribution of this paper is a novel framework for designing approximation schemes for combinatorial optimization problems in which several functions are combined into one objective, or can be viewed as such. Using this framework, we design FPTASes for problems arising in scheduling and resource allocation, combinatorial optimization problems with a rational or a product objective function, robust weighted multiobjective optimization problems, and assortment optimization problems with logit choice model. Given the versatility of our technique, we believe that it will be applicable in many other situations as well.

Appendix

Lemma A.1. *Suppose* $f(x)=\sum_{j=1}^{d}a_jx_j$, $0\leqslant a_j\leqslant v$, $x_j\in\{0,1\}$ *and* $r=\lceil d/\epsilon\rceil$. *Let* $f'(x)=\sum_{j=1}^{d}a'_jx_j$, *where* $a'_j=\lceil a_jr/v\rceil$. *Then,*

1. *if* $f'(x)\leqslant r$, *then* $f(x)\leqslant v$;
2. *if* $f(x)\leqslant v(1-\epsilon)$, *then* $f'(x)\leqslant r$.

Proof. 1. Given $f'(x)\leqslant r$,

$$f(x)=\sum_{j=1}^{d}a_jx_j=\frac{v}{r}\sum_{j=1}^{d}\frac{a_jr}{v}x_j\leqslant\frac{v}{r}\sum_{j=1}^{d}\left\lceil\frac{a_jr}{v}\right\rceil x_j=\frac{v}{r}f'(x)\leqslant v.$$

2. Because $f(x)\leqslant v(1-\epsilon)$,

$$\sum_{j=1}^{d}\frac{a_jr}{v}x_j\leqslant r(1-\epsilon).$$

Rounding up each of the d numbers on the left-hand side, we get

$$\begin{aligned}&\sum_{j=1}^{d}\left\lceil\frac{a_jr}{v}\right\rceil x_j\leqslant r(1-\epsilon)+d\\ &\Rightarrow\quad f'(x)\leqslant r-\left\lceil\frac{d}{\epsilon}\right\rceil\epsilon+d\\ &\qquad\qquad\leqslant r.\quad\square\end{aligned}$$

Lemma A.2. *Suppose* $f(x)=\sum_{j=1}^{d}b_jx_j$, $0\leqslant b_j\leqslant v$, $x_j\in\{0,1\}$ *and* $r=\lceil d/\epsilon\rceil$. *Let* $f'(x)=\sum b'_jx_j$, *where* $b'_j=\min(r,\lfloor b_jr/v\rfloor)$. *Then,*

1. *if* $f'(x)\geqslant r$, *then* $f(x)\geqslant v$;
2. *if* $f(x)\geqslant(1+\epsilon)v$, *then* $f'(x)\geqslant r$.

Proof. 1. Given $f'(x)\geqslant r$,

$$\begin{aligned}f(x)=\sum_{j=1}^{d}b_jx_j=\frac{v}{r}\sum_{j=1}^{d}\frac{b_jr}{v}x_j&\geqslant\frac{v}{r}\sum_{j=1}^{r}\left\lfloor\frac{b_jr}{v}\right\rfloor x_j\\ &\geqslant\frac{v}{r}\sum_{j=1}^{r}b'_jx_j=\frac{v}{r}f'(x)\geqslant v.\end{aligned}$$

2. Let V be the index set of all the variables x_j such that $x_j=1$. Suppose $j\in V$ and $b'_j=r$. Then clearly $f'(x)\geqslant r$. Now assume that for all $j\in V$, $b'_j=\lfloor b_jr/v\rfloor$. Then,

$$\sum_{j\in V}\frac{b_jr}{v}x_j\geqslant(1+\epsilon)r.$$

Rounding down each of the numbers on the left-hand side and together with the assumption that $b'_j = \lfloor b_j r/v \rfloor$, we get

$$\sum_{j\in V}\left\lfloor \frac{b_j r}{v}\right\rfloor x_j \geqslant (1+\epsilon)r - d$$

$$\Rightarrow \quad f'(x) \geqslant r + \epsilon\left\lceil \frac{d}{\epsilon}\right\rceil - d \geqslant r. \quad \square$$

Acknowledgments

The authors thank Vineet Goyal and Ola Svensson for interesting discussions, and Claudio Telha for his feedback on a preliminary version of this paper. The authors also thank the associate editor and two anonymous referees for their constructive comments on the paper.

References

Ackermann H, Newman A, Röglin H, Vöcking B (2007) Decision-making based on approximate and smoothed Pareto curves. *Theoret. Comput. Sci.* 378(3):253–270.

Aissi H, Bazgan C, Vanderpooten D (2007) Approximation of min-max and min-max regret versions of some combinatorial optimization problems. *Eur. J. Oper. Res.* 179(2):281–290.

Asadpour A, Saberi A (2007) An approximation algorithm for max-min fair allocation of indivisible goods. Johnson DS, Feige U, eds. *Proc. 39th Annual ACM Sympos. Theory Comput.* (ACM, New York), 114–121.

Azar Y, Epstein L, Richter Y, Woeginger GJ (2004) All-norm approximation algorithms. *J. Algorithms* 52(1):120–133.

Barahona F, Pulleyblank W (1987) Exact arborescences, matchings and cycles. *Discrete Appl. Math.* 16(2):91–99.

Billionnet A (2002) Approximation algorithms for fractional knapsack problems. *Oper. Res. Lett.* 30(5):336–342.

Chekuri C, Khanna S (2004) On multidimensional packing problems. *SIAM J. Comput.* 33(4):837–851.

Correa JR, Fernandes CG, Wakabayashi Y (2010) Approximating a class of combinatorial problems with rational objective function. *Math. Programming B* 124(1–2):255–269.

Diakonikolas I, Yannakakis M (2009) Small approximate Pareto sets for biobjective shortest paths and other problems. *SIAM J. Comput.* 39(4):1340–1371.

Garofalakis MN, Ioannidis YE (1996) Multi-dimensional resource scheduling for parallel queries. Jagadish HV, Mumick IS, eds. *Proc. ACM SIGMOD Internat. Conf. Management of Data* (ACM, New York), 365–376.

Goyal V, Genc-Kaya L, Ravi R (2011) An FPTAS for minimizing the product of two non-negative linear cost functions. *Math. Programming* 126(2):401–405.

Graham RL, Lawler EL, Lenstra JK, Rinnooy Kan AHG (1979) Optimization and approximation in determinstic sequencing and scheduling: A survey. *Ann. Discrete Math.* 5:287–326.

Halman N, Klabjan D, Mostagir M, Orlin JB, Simchi-Levi D (2009) A fully polynomial-time approximation scheme for single-item stochastic inventory control with discrete demand. *Math. Oper. Res.* 34(3):674–685.

Hashizume S, Fukushima M, Katoh N, Ibaraki T (1987) Approximation algorithms for combinatorial fractional programming problems. *Math. Programming* 37(3):255–267.

Hochbaum DS, Shmoys DB (1987) Using dual approximation algorithms for scheduling problems: Theoretical and practical results. *J. Assoc. Comput. Machinery* 34(1):144–162.

Horowitz E, Sahni S (1976) Exact and approximate algorithms for scheduling nonidentical processors. *J. Assoc. Comput. Machinery* 23(2):317–327.

Hu J, Mehrotra S (2010) Robust and stochastically weighted multi-objective optimization models and reformulations. Technical report, Department of Industrial Engineering and Management Sciences, Northwestern University, Evanston, IL.

Ibarra OH, Kim CE (1975) Fast approximation algorithms for the knapsack and sum of subset problems. *J. Assoc. Comput. Machinery* 22(4):463–468.

Kasperski A, Zielinski P (2007) On the existence of an FPTAS for minmax regret combinatorial optimization problems with interval data. *Oper. Res. Lett.* 35(4):525–532.

Kern W, Woeginger GJ (2007) Quadratic programming and combinatorial minimum weight product problems. *Math. Programming* 110(3):641–649.

Kök AG, Fisher ML, Vaidyanathan R (2009) Assortment planning: Review of literature and industry practice. Agrawal N, Smith SA, eds. *Retail Supply Chain Management*. International Series in Operations Research and Management Science, Vol. 122 (Springer, New York), 99–153.

Koltun V, Papadimitriou CH (2007) Approximately dominating representatives. *Theoret. Comput. Sci.* 371(3):148–154.

Kouvelis P, Yu G (1997) *Robust Discrete Optimization and Its Applications* (Kluwer Academic Publishers, Boston).

Kuno T (1999) Polynomial algorithms for a class of minimum rank-two cost path problems. *J. Global Optim.* 15(4):405–417.

Lenstra JK, Shmoys DB, Tardos É (1990) Approximation algorithms for scheduling unrelated parallel machines. *Math. Programming* 46:259–271.

Megiddo N (1979) Combinatorial optimization with rational objective functions. *Math. Oper. Res.* 4(4):414–424.

Papadimitriou CH, Yannakakis M (2000) On the approximability of trade-offs and optimal access of Web sources. Blum A, ed. *Proc. 41st Annual IEEE Sympos. Foundations Comput. Sci.* (IEEE, Piscataway, NJ), 86–92.

Radzik T (1992) Newton's method for fractional combinatorial optimization. *Proc. 33rd Annual IEEE Sympos. Foundations Comput. Sci.* (IEEE, Piscataway, NJ), 659–669.

Rusmevichientong P, Shen Z-JM, Shmoys DB (2009) A PTAS for capacitated sum-of-ratios optimization. *Oper. Res. Lett.* 37(4):230–238.

Rusmevichientong P, Shmoys DB, Topaloglu H (2010) Assortment optimization with mixtures of logits. Technical report, School of Operations Research and Information Engineering, Cornell University, Ithaca, NY.

Safer HM, Orlin JB (1995a) Fast approximation schemes for multi-criteria combinatorial optimization. Technical report, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

Safer HM, Orlin JB (1995b) Fast approximation schemes for multi-criteria flow, knapsack, and scheduling problems. Technical report, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

Safer HM, Orlin JB, Dror M (2004) Fully polynomial approximation in multi-criteria combinatorial optimization. Technical report, Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA.

Sahni S (1976) Algorithms for scheduling independent tasks. *J. Assoc. Comput. Machinery* 23(1):116–127.

Vassilvitskii S, Yannakakis M (2005) Efficiently computing succinct trade-off curves. *Theoret. Comput. Sci.* 348(2–3):334–356.

Warburton A (1987) Approximation of Pareto optima in multiple-objective, shortest-path problems. *Oper. Res.* 35(1):70–79.

Woeginger GJ (2000) When does a dynamic programming formulation guarantee the existence of a fully polynomial time approximation scheme (FPTAS)? *INFORMS J. Comput.* 12(1):57–74.

Shashi Mittal works as an operations research scientist in the Global Transportation Strategy group at Amazon.com. His research interests are in combinatorial optimization, service operations, transportation, and business analytics.

Andreas S. Schulz is the Patrick J. McGovern Professor of Management and Professor of Mathematics of Operations Research in the Sloan School of Management at the Massachusetts Institute of Technology. His current research interests include algorithmic game theory, approximation algorithms, combinatorial optimization, computational complexity, integer programming, partial order theory, polyhedral combinatorics, and scheduling theory.