Robust Discrete Optimization Under Ellipsoidal Uncertainty Sets

Dimitris Bertsimas *　　　Melvyn Sim †

March, 2004

Abstract

We address the complexity and practically efficient methods for robust discrete optimization under ellipsoidal uncertainty sets. Specifically, we show that the robust counterpart of a discrete optimization problem with correlated objective function data is NP-hard even though the nominal problem is polynomially solvable. For uncorrelated and identically distributed data, however, we show that the robust problem retains the complexity of the nominal problem. For uncorrelated, but not identically distributed data we propose an approximation method that solves the robust problem within arbitrary accuracy. We also propose a Frank-Wolfe type algorithm for this case, which we prove converges to a locally optimal solution, and in computational experiments is remarkably effective. Finally, we propose a generalization of the robust discrete optimization framework we proposed earlier that (a) allows the key parameter that controls the tradeoff between robustness and optimality to depend on the solution and (b) results in increased flexibility and decreased conservatism, while maintaining the complexity of the nominal problem.

*Boeing Professor of Operations Research, Sloan School of Management and Operations Research Center, Massachusetts Institute of Technology, E53-363, Cambridge, MA 02139, dbertsim@mit.edu. The research of the author was partially supported by the Singapore-MIT alliance.

†NUS Business School, National University of Singapore, dscsimm@nus.edu.sg.

1 Introduction

Robust optimization as a method to address uncertainty in optimization problems has been in the center of a lot of research activity. Ben-Tal and Nemirovski [1, 2, 3] and El-Ghaoui et al. [7, 8] propose efficient algorithms to solve certain classes of convex optimization problems under data uncertainty that is described by ellipsoidal sets.

Kouvelis and Yu [15] propose a framework for robust discrete optimization, which seeks to find a solution that minimizes the worst case performance under a set of scenarios for the data. Unfortunately, under their approach, the robust counterpart of a polynomially solvable discrete optimization problem can be NP-hard. Bertsimas and Sim [5, 6] propose an approach in solving robust discrete optimization problems that has the flexibility of adjusting the level of conservativeness of the solution while preserving the computational complexity of the nominal problem. This is attractive as it shows that adding robustness does not come at the price of a change in computational complexity. Ishii et. al. [12] consider solving a stochastic minimum spanning tree problem with costs that are independently and normally distributed leading to a similar framework as robust optimization with an ellipsoidal uncertainty set. However, to the best of our knowledge, there has not been any work or complexity results on robust discrete optimization under ellipsoidal uncertainty sets.

It is thus natural to ask whether adding robustness in the cost function of a given discrete optimization problem under an ellipsoidal uncertainty set leads to a change in computational complexity and whether we can develop practically efficient methods to solve robust discrete optimization problems under ellipsoidal uncertainty sets.

Our objective in this paper is to address these questions. Specifically our contributions include:

(a) Under a general ellipsoidal uncertainty set that models correlated data, we show that the robust counterpart can be NP-hard even though the nominal problem is polynomially solvable in contrast with the uncertainty sets proposed in Bertsimas and Sim [5, 6].

(b) Under an ellipsoidal uncertainty set with uncorrelated data, we show that the robust problem can be reduced to solving a collection of nominal problems with different linear objectives. If the distributions are identical, we show that we only require to solve $r+1$ nominal problems, where r is the number of uncertain cost components, that is in this case the computational complexity is preserved. Under uncorrelated data, we propose an approximation method that solves the robust problem within an additive ϵ. The complexity of the method is $O((nd_{\max})^{1/4}\epsilon^{-1/2})$, where $d_{\max}$ is the largest number in the data describing the ellipsoidal set. We also propose a Frank-Wolfe type algorithm for this case, which we prove converges to a locally optimal solution, and in computational experiments is remarkably effective. We also link the robust problem with uncorrelated data to classical problems in parametric discrete optimization.

(c) We propose a generalization of the robust discrete optimization framework in Bertsimas and Sim [5] that allows the key parameter that controls the tradeoff between robustness and optimality to depend on the solution. This generalization results in increased flexibility and decreased conservatism, while maintaining the complexity of the nominal problem.

Structure of the paper. In Section 2, we formulate robust discrete optimization problems under ellipsoidal uncertainty sets (correlated data) and show that the problem is NP-hard even for nominal problems that are polynomially solvable. In Section 3, we present structural results and establish that the robust problem under ball uncertainty (uncorrelated and identically distributed data) has the same complexity as the nominal problem. In Sections 4 and 5, we propose approximation methods for the robust problem under ellipsoidal uncertainty sets with uncorrelated but not identically distributed data. In Section 6, we present the generalization of the robust discrete optimization framework in Bertsimas and Sim [5]. In Section 7, we present some experimental findings relating to the computation speed and the quality of robust solutions. The final section contains some concluding remarks.

2 Formulation of Robust Discrete Optimization Problems

A nominal discrete optimization problem is:

$$\begin{array}{rl} \text{minimize} & \boldsymbol{c}'\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in X, \end{array} \tag{1}$$

with $X \subseteq \{0,1\}^n$. We are interested in problems where each entry $\tilde{c}_j,\ j \in N = \{1,2,\ldots,n\}$ is uncertain and described by an uncertainty set C. Under the robust optimization paradigm, we solve

$$\begin{array}{rl} \text{minimize} & \displaystyle\max_{\tilde{\boldsymbol{c}} \in C} \tilde{\boldsymbol{c}}'\boldsymbol{x} \\ \text{subject to} & \boldsymbol{x} \in X. \end{array} \tag{2}$$

Writing $\tilde{\boldsymbol{c}} = \boldsymbol{c} + \tilde{\boldsymbol{s}}$, where $\boldsymbol{c}$ is the nominal value and the deviation $\tilde{\boldsymbol{s}}$ is restricted to the set $D = C - \boldsymbol{c}$, Problem (2) becomes:

$$\begin{array}{rl} \text{minimize} & \boldsymbol{c}'\boldsymbol{x} + \xi(\boldsymbol{x}) \\ \text{subject to} & \boldsymbol{x} \in X, \end{array} \tag{3}$$

where $\xi(\boldsymbol{x}) = \max_{\tilde{\boldsymbol{s}} \in D} \tilde{\boldsymbol{s}}'\boldsymbol{x}$. Special cases of Formulation (3) include:

(a) $D = \{\boldsymbol{s} : \tilde{s_j} \in [0, d_j]\}$, leading to $\xi(\boldsymbol{x}) = \boldsymbol{d}'\boldsymbol{x}$.

(b) $D = \{\boldsymbol{s} : \|\boldsymbol{\Sigma}^{-1/2}\boldsymbol{s}\|_2 \leq \Omega\}$ that models ellipsoidal uncertainty sets proposed by Ben-Tal and Nemirovski [1, 2, 3] and El-Ghaoui et al. [7, 8]. It easily follows that $\xi(\boldsymbol{x}) = \Omega\sqrt{\boldsymbol{x}'\boldsymbol{\Sigma}\boldsymbol{x}}$, where $\boldsymbol{\Sigma}$ is the covariance matrix of the random cost coefficients. For the special case that $\boldsymbol{\Sigma} = \mathrm{diag}(d_1, \ldots, d_n)$, i.e., the random cost coefficients are uncorrelated, we obtain that $\xi(\boldsymbol{x}) = \Omega\sqrt{\sum_{j \in N} d_j x_j^2} = \Omega\sqrt{\boldsymbol{d}'\boldsymbol{x}}$.

(c) $D = \{\boldsymbol{s} : 0 \leq s_j \leq d_j \ \forall j \in J, \sum_{k \in N} \frac{s_k}{d_k} \leq \Gamma\}$ proposed in Bertsimas and Sim [6]. It follows that in this case $\xi(\boldsymbol{x}) = \max_{\{S : |S| = \Gamma, S \subseteq J\}} \sum_{j \in S} d_j x_j$, where J is the set of random cost components. Bertsimas and Sim [6] show that Problem (3) reduces to solving at most $|J| + 1$ nominal problems for different cost vectors. In other words, the robust counterpart is polynomially solvable if the nominal problem is polynomially solvable.

Under models (a) and (c), robustness preserves the computational complexity of the nominal problem. Our objective in this paper is to investigate the price (in increased complexity) of robustness under ellipsoidal uncertainty sets (model (b)) and propose effective algorithmic methods to tackle models (b), (c).

Our first result is unfortunately negative. Under ellipsoidal uncertainty sets with general covariance matrices, the price of robustness is an increase in computational complexity. The robust counterpart may become NP-hard even though the nominal problem is polynomially solvable.

Theorem 1 *The robust problem (3) with $\xi(\boldsymbol{x}) = \Omega\sqrt{\boldsymbol{x}'\Sigma\boldsymbol{x}}$ (Model (b)) is NP-hard, for the following classes of polynomially solvable nominal problems: shortest path, minimum cost assignment, resource scheduling, minimum spanning tree.*

Proof : Kouvelis and Yu [15] prove that the problem

$$\begin{array}{rl} \text{minimize} & \max\{\boldsymbol{c}_1'\boldsymbol{x}, \boldsymbol{c}_2'\boldsymbol{x}\} \\ \text{subject to} & \boldsymbol{x} \in X, \end{array} \tag{4}$$

is NP-hard for the polynomially solvable problems mentioned in the statement of the theorem. We show a simple transformation of Problem (4) to Problem (3) with $\xi(\boldsymbol{x}) = \Omega\sqrt{\boldsymbol{x}'\Sigma\boldsymbol{x}}$ as follows:

$$\begin{aligned} \max\{\boldsymbol{c}_1'\boldsymbol{x}, \boldsymbol{c}_2'\boldsymbol{x}\} &= \max\left\{\frac{\boldsymbol{c}_1'\boldsymbol{x} + \boldsymbol{c}_2'\boldsymbol{x}}{2} + \frac{\boldsymbol{c}_1'\boldsymbol{x} - \boldsymbol{c}_2'\boldsymbol{x}}{2}, \frac{\boldsymbol{c}_1'\boldsymbol{x} + \boldsymbol{c}_2'\boldsymbol{x}}{2} - \frac{\boldsymbol{c}_1'\boldsymbol{x} - \boldsymbol{c}_2'\boldsymbol{x}}{2}\right\} \\ &= \frac{\boldsymbol{c}_1'\boldsymbol{x} + \boldsymbol{c}_2'\boldsymbol{x}}{2} + \max\left\{\frac{\boldsymbol{c}_1'\boldsymbol{x} - \boldsymbol{c}_2'\boldsymbol{x}}{2}, -\frac{\boldsymbol{c}_1'\boldsymbol{x} - \boldsymbol{c}_2'\boldsymbol{x}}{2}\right\} \end{aligned}$$

$$
\begin{aligned}
&= \frac{\boldsymbol{c_1'x}+\boldsymbol{c_2'x}}{2}+\left|\frac{\boldsymbol{c_1'x}-\boldsymbol{c_2'x}}{2}\right| \\
&= \frac{\boldsymbol{c_1'x}+\boldsymbol{c_2'x}}{2}+\frac{1}{2}\sqrt{\boldsymbol{x}'(\boldsymbol{c_1}-\boldsymbol{c_2})(\boldsymbol{c_1}-\boldsymbol{c_2})'\boldsymbol{x}}.
\end{aligned}
$$

The NP-hard Problem (4) is transformed to Problem (3) with $\xi(\boldsymbol{x}) = \Omega\sqrt{\boldsymbol{x}'\Sigma\boldsymbol{x}}$, $\boldsymbol{c} = (\boldsymbol{c_1}+\boldsymbol{c_2})/2$, $\Omega = 1/2$ and $\Sigma = (\boldsymbol{c_1}-\boldsymbol{c_2})(\boldsymbol{c_1}-\boldsymbol{c_2})'$. Thus, Problem (3) with $\xi(\boldsymbol{x}) = \Omega\sqrt{\boldsymbol{x}'\Sigma\boldsymbol{x}}$ is NP-hard. ■

We next would like to propose methods for model (b) with $\boldsymbol{\Sigma} = \mathrm{diag}(d_1,\ldots,d_n)$. We are thus naturally led to consider the problem

$$
G^* = \min_{\boldsymbol{x}\in X} \boldsymbol{c}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x}) \tag{5}
$$

with $f(\cdot)$ a concave function. In particular, $f(x) = \Omega\sqrt{x}$ models ellipsoidal uncertainty sets with uncorrelated random cost coefficients (model (b)).

3 Structural Results

We first show that Problem (5) reduces to solving a number of nominal problems (1). Let $W = \{\boldsymbol{d}'\boldsymbol{x} \mid \boldsymbol{x} \in \{0,1\}^n\}$ and $\eta(w)$ be a subgradient of the concave function $f(\cdot)$ evaluated at w, that is, $f(u) - f(w) \le \eta(w)(u-w)\ \forall u \in R$. If $f(w)$ is a differentiable function and $f'(0)=\infty$, we choose

$$
\eta(w) = \begin{cases} f'(w) & \text{if } w \in W\backslash\{0\}, \\ \frac{f(d_{\min})-f(0)}{d_{\min}} & \text{if } w = 0, \end{cases}
$$

where $d_{\min} = \min_{\{j:\ d_j>0\}} d_j$.

Theorem 2 *Let*

$$
Z(w) = \min_{\boldsymbol{x}\in X} (\boldsymbol{c}+\eta(w)\boldsymbol{d})'\boldsymbol{x} + f(w) - w\eta(w), \tag{6}
$$

and $w^* = \arg\min_{w\in W} Z(w)$. *Then,* w^* *is an optimal solution to Problem (5) and* $G^* = Z(w^*)$.

Proof : We first show that $G^* \ge \min_{w\in W} Z(w)$. Let $\boldsymbol{x}^*$ be an optimal solution to Problem (5) and $w^* = \boldsymbol{d}'\boldsymbol{x}^* \in W$. We have

$$
\begin{aligned}
G^* &= \boldsymbol{c}'\boldsymbol{x}^* + f(\boldsymbol{d}'\boldsymbol{x}^*) = \boldsymbol{c}'\boldsymbol{x}^* + f(w^*) = (\boldsymbol{c}+\eta(w^*)\boldsymbol{d})'\boldsymbol{x}^* + f(w^*) - w^*\eta(w^*) \\
&\ge \min_{\boldsymbol{x}\in X} (\boldsymbol{c}+\eta(w^*)\boldsymbol{d})'\boldsymbol{x} + f(w^*) - w^*\eta(w^*) = Z(w^*) \ge \min_{w\in W} Z(w).
\end{aligned}
$$

Conversely, for any $w \in W$, let y_w be an optimal solution to Problem (6). We have

$$\begin{aligned} Z(w) &= (\boldsymbol{c} + \eta(w)\boldsymbol{d})'\boldsymbol{y_w} + f(w) - w\eta(w) \\ &= \boldsymbol{c}'\boldsymbol{y_w} + f(\boldsymbol{d}'\boldsymbol{y_w}) + \eta(w)(\boldsymbol{d}'\boldsymbol{y_w} - w) - (f(\boldsymbol{d}'\boldsymbol{y_w}) - f(w)) \\ &\geq \boldsymbol{c}'\boldsymbol{y_w} + f(\boldsymbol{d}'\boldsymbol{y_w}) \qquad (7) \\ &\geq \min_{\boldsymbol{x} \in X} \boldsymbol{c}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x}) = G^*, \end{aligned}$$

where inequality (7) for $w \in W\backslash\{0\}$ follows, since $\eta(w)$ is a subgradient. To see that inequality (7) follows for $w = 0$ we argue as follows. Since $f(v)$ is concave and $v \geq d_{\min}$ $\forall v \in W\backslash\{0\}$, we have

$$f(d_{\min}) \geq \frac{v - d_{\min}}{v} f(0) + \frac{d_{\min}}{v} f(v), \quad \forall v \in W\backslash\{0\}.$$

Rearranging, we have

$$\frac{f(v) - f(0)}{v} \leq \frac{f(d_{\min}) - f(0)}{d_{\min}} = \eta(0) \quad \forall v \in W\backslash\{0\},$$

leading to $\eta(0)(\boldsymbol{d}'\boldsymbol{y_w} - 0) - (f(\boldsymbol{d}'\boldsymbol{y_w}) - f(0)) \geq 0$. Therefore $G^* = \min_{w \in W} Z(w)$. ■

Note that when $d_j = \sigma^2$, then $W = \{0, \sigma^2, \ldots, n\sigma^2\}$, and thus $|W| = n + 1$, In this case, Problem (5) reduces to solving $n + 1$ nominal problems (6), i.e., polynomial solvability is preserved. Specifically, for the case of an ellipsoidal uncertainty set $\boldsymbol{\Sigma} = \sigma^2 \boldsymbol{I}$, leading to $\xi(\boldsymbol{x}) = \Omega\sqrt{\sum_j \sigma^2 x_j^2} = \Omega\sigma\sqrt{\boldsymbol{e}'\boldsymbol{x}}$, we derive explicitly the subproblems involved.

Proposition 1 *Under an ellipsoidal uncertainty set with* $\xi(\boldsymbol{x}) = \Omega\sigma\sqrt{\boldsymbol{e}'\boldsymbol{x}}$,

$$G^* = \min_{w=0,1,\ldots,n} Z(w),$$

where

$$Z(w) = \begin{cases} \min_{\boldsymbol{x} \in X} \quad \left(\boldsymbol{c} + \dfrac{\Omega\sigma}{2\sqrt{w}}\boldsymbol{e}\right)' \boldsymbol{x} + \dfrac{\Omega\sigma\sqrt{w}}{2} & \text{if } w = 1, \ldots, n, \\ \min_{\boldsymbol{x} \in X} \quad (\boldsymbol{c} + \Omega\sigma\boldsymbol{e})'\boldsymbol{x} & \text{if } w = 0. \end{cases} \qquad (8)$$

Proof : With $\xi(\boldsymbol{x}) = \Omega\sigma\sqrt{\boldsymbol{e}'\boldsymbol{x}}$, we have $f(w) = \Omega\sigma\sqrt{w}$. Substituting $\eta(w) = f'(w) = \Omega\sigma/(2\sqrt{w})$, $\forall\, w \in W\backslash\{0\}$ and $\eta(0) = (f(d_{\min}) - f(0))/d_{\min} = f(1) - f(0) = \Omega\sigma$ to Eq. (6) we obtain Eq. (8). ■

An immediate corollary of Theorem 2 is to consider a parametric approach as follows:

Corollary 1 *An optimal solution to Problem (5) coincides with one of the optimal solutions to the parametric problem:*

$$\begin{aligned} \text{minimize} \quad & (\boldsymbol{c} + \theta\boldsymbol{d})'\boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{x} \in X, \end{aligned} \qquad (9)$$

for $\theta \in [\eta(\boldsymbol{e}'\boldsymbol{d}), \eta(0)]$.

This establishes a connection of Problem (5) with parametric discrete optimization (see Gusfield [11], Hassin and Tamir [13]). It turns out that if X is a matroid, the minimal set of optimal solutions to Problem (9) as θ varies is polynomial in size, see Eppstein [9] and Fern et. al. [10]. For optimization over a matroid, the optimal solution depends on the ordering of the cost components. Since, as θ varies, it is easy to see that there are at most $\binom{n}{2}+1$ different orderings, the corresponding robust problem is also polynomially solvable.

For the case of shortest paths, Karp and Orlin [14] provide a polynomial time algorithm using the parametric approach when all d_j's are equal. In contrast, the polynomial reduction in Proposition 1 applies to all discrete optimization problems.

More generally, $|W| \leq d_{\max} n$ with $d_{\max} = \max_j d_j$. If $d_{\max} \leq n^\alpha$ for some fixed α, then Problem (5) reduces to solving $n^\alpha(n+1)$ nominal problems (6). However, when $d_{\max}$ is exponential in n, an approach that enumerates all elements of W does not preserve polynomiality. For this reason, as well as deriving more practical algorithms even in the case that $|W|$ is polynomial in n we develop in the next section new algorithms.

4 Approximation via Piecewise Linear Functions

In this section, we develop a method for solving Problem (5) that is based on approximating the function $f(\cdot)$ with a piecewise linear concave function. We first show that if $f(\cdot)$ is a piecewise linear concave function with a polynomial number of segments, we can also reduce Problem (5) to solving a polynomial number of subproblems.

Proposition 2 *If $f(w), w \in [0, \boldsymbol{e}'\boldsymbol{d}]$ is a continuous piecewise linear concave function of k segments, then*

$$\min_{\boldsymbol{x} \in X} (\boldsymbol{c} + \eta_j \boldsymbol{d})' \boldsymbol{x}, \tag{10}$$

where η_j is the gradient of the jth linear piece of the function $f(\cdot)$.

Proof : The proof follows directly from Theorem 2 and the observations that if $f(w), w \in [0, \boldsymbol{e}'\boldsymbol{d}]$ is a continuous piecewise linear concave function of k linear pieces, the set of subgradients of each of the linear pieces constitutes the minimal set of subgradients for the function f. ■

We next show that approximating the function $f(\cdot)$ with a piecewise linear concave function leads to an approximate solution to Problem (5).

Theorem 3 *For $W = [\underline{w}, \overline{w}]$ such that $\boldsymbol{d}'\boldsymbol{x} \in W \ \forall \boldsymbol{x} \in X$, let $g(w)$, $w \in W$ be a piecewise linear concave function approximating the function $f(w)$ such that $-\epsilon_1 \leq f(w) - g(w) \leq \epsilon_2$ with $\epsilon_1, \epsilon_2 \geq 0$. Let $\boldsymbol{x}_{\mathrm{H}}$ be an optimal solution of the problem:*

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c}'\boldsymbol{x} + g(\boldsymbol{d}'\boldsymbol{x}) \\ \textit{subject to} & \boldsymbol{x} \in X \end{array} \tag{11}$$

and let $G_{\mathrm{H}} = \boldsymbol{c}'\boldsymbol{x}_{\mathbf{H}} + f(\boldsymbol{d}'\boldsymbol{x}_{\mathbf{H}})$. Then,

$$G^* \leq G_{\mathrm{H}} \leq G^* + \epsilon_1 + \epsilon_2.$$

Proof : We have that

$$\begin{aligned} G^* &= \min_{\boldsymbol{x} \in X} \{\boldsymbol{c}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x})\} \\ &\leq G_{\mathrm{H}} = \boldsymbol{c}'\boldsymbol{x}_{\mathrm{H}} + f(\boldsymbol{d}'\boldsymbol{x}_{\mathrm{H}}) \\ &\leq \boldsymbol{c}'\boldsymbol{x}_{\mathrm{H}} + g(\boldsymbol{d}'\boldsymbol{x}_{\mathrm{H}}) + \epsilon_2 \qquad (12) \\ &= \min_{\boldsymbol{x} \in X} \{\boldsymbol{c}'\boldsymbol{x} + g(\boldsymbol{d}'\boldsymbol{x})\} + \epsilon_2 \\ &\leq \min_{\boldsymbol{x} \in X} \{\boldsymbol{c}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x})\} + \epsilon_1 + \epsilon_2 \qquad (13) \\ &= G^* + \epsilon_1 + \epsilon_2, \end{aligned}$$

where inequalities (12) and (13) follow from $-\epsilon_1 \leq f(w) - g(w) \leq \epsilon_2$. ■

We next apply the approximation idea to the case of ellipsoidal uncertainty sets. Specifically, we approximate the function $f(w) = \Omega\sqrt{w}$ in the domain $[\underline{w}, \overline{w}]$ with a piecewise linear concave function $g(w)$ satisfying $0 \leq f(w) - g(w) \leq \epsilon$ using the least number of linear pieces.

Proposition 3 *For $\epsilon > 0$, w_0 given, let $\phi = \epsilon/\Omega$ and for $i = 1, \ldots, k$ let*

$$w_i = \phi^2 \left\{ 2\left(i + \sqrt{\frac{\sqrt{w_0}}{2\phi} + \frac{1}{4}} \right)^2 - \frac{1}{2} \right\}^2. \tag{14}$$

Let $g(w)$ be a piecewise linear concave function on the domain $w \in [w_0, w_k]$, with breakpoints $(w, g(w)) \in \{(w_0, \Omega\sqrt{w_0}), \ldots, (w_k, \Omega\sqrt{w_k})\}$. Then, for all $w \in [w_0, w_k]$

$$0 \leq \Omega\sqrt{w} - g(w) \leq \epsilon.$$

Proof : Since at the breakpoints w_i, $g(w_i) = \Omega\sqrt{w_i}$, $g(w)$ is a concave function with $g(w) \le \Omega\sqrt{w}$, $\forall w \in [w_0, w_k]$. For $w \in [w_{i-1}, w_i]$, we have

$$\begin{aligned}
\Omega\sqrt{w} - g(w) &= \Omega\sqrt{w} - \left\{\Omega\sqrt{w_{i-1}} + \frac{\Omega\sqrt{w_i} - \Omega\sqrt{w_{i-1}}}{w_i - w_{i-1}}(w - w_{i-1})\right\} \\
&= \Omega\left\{\sqrt{w} - \sqrt{w_{i-1}} - \frac{w - w_{i-1}}{\sqrt{w_i} + \sqrt{w_{i-1}}}\right\}.
\end{aligned}$$

The maximum value of $\Omega\sqrt{w} - g(w)$ is attained at $\sqrt{w^*} = (\sqrt{w_i} + \sqrt{w_{i-1}})/2$. Therefore,

$$\begin{aligned}
\Omega\sqrt{w} - g(w) &\le \Omega\left\{\sqrt{w^*} - \sqrt{w_{i-1}} - \frac{w^* - w_{i-1}}{\sqrt{w_i} + \sqrt{w_{i-1}}}\right\} \\
&= \Omega\left\{\frac{\sqrt{w_i} - \sqrt{w_{i-1}}}{2} - \frac{\left(\frac{\sqrt{w_i}+\sqrt{w_{i-1}}}{2}\right)^2 - w_{i-1}}{\sqrt{w_i} + \sqrt{w_{i-1}}}\right\} \\
&= \Omega\left\{\frac{\sqrt{w_i} - \sqrt{w_{i-1}}}{2} - \frac{\left(\frac{\sqrt{w_i}+3\sqrt{w_{i-1}}}{2}\right)\left(\frac{\sqrt{w_i}-\sqrt{w_{i-1}}}{2}\right)}{\sqrt{w_i} + \sqrt{w_{i-1}}}\right\} \\
&= \frac{\Omega(\sqrt{w_i} - \sqrt{w_{i-1}})^2}{4(\sqrt{w_i} + \sqrt{w_{i-1}})} \\
&= \Omega\phi = \epsilon, \qquad (15)
\end{aligned}$$

where Eq. (15) follows by substituting Eq. (14). Since

$$\max_{w \in [w_{i-1}, w_i]} \{\Omega\sqrt{w} - g(w)\} = \epsilon,$$

the proposition follows. ■

Propositions 2, 3 and Theorem 3 lead to Algorithm 1.

Algorithm 1 *Approximation by piecewise linear concave functions.*

Input: $\boldsymbol{c}, \boldsymbol{d}, \underline{w}, \overline{w}, \Omega, \epsilon$, $f(x) = \Omega\sqrt{x}$ *and a routine that optimizes a linear function over the set* $X \subseteq \{0,1\}^n$.

Output: *A solution* $\boldsymbol{x}_{\mathrm{H}} \in X$ *for which* $G^* \le \boldsymbol{c}'\boldsymbol{x}_{\mathrm{H}} + f(\boldsymbol{d}'\boldsymbol{x}_{\mathrm{H}}) \le G^* + \epsilon$, *where* $G^* = \min_{\boldsymbol{x} \in X} \boldsymbol{c}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x})$.

Algorithm:

1. *(Initialization) Let* $\phi = \epsilon/\Omega$; *Let* $w_0 = \underline{w}$; *Let*

$$k = \left\lceil \sqrt{\frac{\Omega\sqrt{\overline{w}}}{2\epsilon} + \frac{1}{4}} - \sqrt{\frac{\Omega\sqrt{\underline{w}}}{2\epsilon} + \frac{1}{4}} \right\rceil = O\left(\sqrt{\frac{\Omega}{\epsilon}}(nd_{\max})^{\frac{1}{4}}\right)$$

where $d_{\max} = \max_j d_j$ *and for* $i = 1, \dots, k$ *let*

$$w_i = \phi^2 \left\{ 2\left(i + \sqrt{\frac{\sqrt{w}}{2\phi} + \frac{1}{4}} \right)^2 - \frac{1}{2} \right\}^2 .$$

2. *For* $i = 1, \dots, k$ *solve the problem*

$$Z_i = \min_{\boldsymbol{x} \in X} \left(\boldsymbol{c} + \frac{\Omega}{\sqrt{w_i} + \sqrt{w_{i-1}}} \boldsymbol{d} \right)' \boldsymbol{x}. \tag{16}$$

Let $\boldsymbol{x_i}$ *be an optimal solution to Problem (16).*

3. *Output* $G^*_{\mathrm{H}} = Z_{i^*} = \min_{i=1,\dots,k} Z_i$ *and* $\boldsymbol{x}_{\mathrm{H}} = \boldsymbol{x}_{i^*}$.

Theorem 4 *Algorithm 1 finds a solution* $\boldsymbol{x}_{\mathrm{H}} \in X$ *for which* $G^* \le \boldsymbol{c'}\boldsymbol{x}_{\mathrm{H}} + f(\boldsymbol{d'}\boldsymbol{x}_{\mathrm{H}}) \le G^* + \epsilon$.

Proof : Using Proposition 3 we find a piecewise linear concave function $g(w)$ that approximates within a given tolerance $\epsilon > 0$ the function $\Omega\sqrt{w}$. From Proposition 2 and since the gradient of the ith segment of the function $g(w)$ for $w \in [w_{i-1}, w_i]$ is

$$\eta_i = \Omega \frac{\sqrt{w_i} - \sqrt{w_{i-1}}}{w_i - w_{i-1}} = \frac{\Omega}{\sqrt{w_i} + \sqrt{w_{i-1}}}.$$

we solve the Problems for $i = 1, \dots, k$

$$Z_i = \min_{\boldsymbol{x} \in X} \left(\boldsymbol{c} + \frac{\Omega}{\sqrt{w_i} + \sqrt{w_{i-1}}} \boldsymbol{d} \right)' \boldsymbol{x}$$

Taking $G^*_{\mathrm{H}} = \min_i Z_i$ and using Theorem 3 it follows that Algorithm 1 produces a solution within ϵ. ■

Although the number of subproblems solved in Algorithm 1 is not polynomial with respect to the bit size of the input data, the computation involved is reasonable from a practical point of view. For example, in Table 1 we report the number of subproblems we need to solve for $\Omega = 4$, as a function of ϵ and $\boldsymbol{d'e} = \sum_{j=1}^{n} d_j$.

5 A Frank-Wolfe Type Algorithm

A natural method to solve Problem (5) is to apply a Frank-Wolfe type algorithm, that is to successively linearize the function $f(\cdot)$.

Algorithm 2 *The Frank-Wolfe type algorithm.*

Input: $\boldsymbol{c}, \boldsymbol{d}, \Omega$, $\theta \in [\eta(\boldsymbol{d'e}), \eta(0)]$, $f(w)$, $\eta(w)$ *and a routine that optimizes a linear function over the set* $X \subseteq \{0,1\}^n$.

Output: *A locally optimal solution to Problem (5).*

Algorithm:

ϵ	$\boldsymbol{d'e}$	k
0.01	10	25
0.01	100	45
0.01	1000	80
0.01	10000	121
0.001	10	80
0.001	100	141
0.001	1000	251
0.001	10000	447

Table 1: Number of subproblems, k as a function of the desired precision ϵ, size of the problem $\boldsymbol{d'e}$ and $\Omega = 4$.

1. *(Initialization)* $k = 0; \boldsymbol{x_0} := \arg\min_{\boldsymbol{y}\in X} (\boldsymbol{c} + \theta\boldsymbol{d})'\boldsymbol{y}$

2. *Until* $\boldsymbol{d'x_{k+1}} = \boldsymbol{d'x_k}$, $\boldsymbol{x_{k+1}} := \arg\min_{\boldsymbol{y}\in X} (\boldsymbol{c} + \eta(\boldsymbol{d'x_k})\boldsymbol{d})'\boldsymbol{y}$.

3. *Output* $\boldsymbol{x_{k+1}}$.

We next show that Algorithm 2 converges to a locally optimal solution.

Theorem 5 *Let $\boldsymbol{x}$, $\boldsymbol{y}$ and $\boldsymbol{z}_\eta$ be optimal solutions to the following problems:*

$$\begin{aligned} \boldsymbol{x} &= \arg\min_{\boldsymbol{u}\in X} (\boldsymbol{c} + \theta\boldsymbol{d})'\boldsymbol{u}, & (17) \\ \boldsymbol{y} &= \arg\min_{\boldsymbol{u}\in X} (\boldsymbol{c} + \eta(\boldsymbol{d'x})\boldsymbol{d})'\boldsymbol{u} & (18) \\ \boldsymbol{z}_\eta &= \arg\min_{\boldsymbol{u}\in X} (\boldsymbol{c} + \eta\boldsymbol{d})'\boldsymbol{u}, & (19) \end{aligned}$$

for some η strictly between θ and $\eta(\boldsymbol{d'x})$.

(a) *(Improvement)* $\boldsymbol{c'y} + f(\boldsymbol{d'y}) \le \boldsymbol{c'x} + f(\boldsymbol{d'x})$.

(b) *(Monotonicity) If $\theta > \eta(\boldsymbol{d'x})$, then $\eta(\boldsymbol{d'x}) \ge \eta(\boldsymbol{d'y})$. Likewise, if $\theta < \eta(\boldsymbol{d'x})$, then $\eta(\boldsymbol{d'x}) \le \eta(\boldsymbol{d'y})$. Hence, the sequence $\theta_k = \eta(\boldsymbol{d'x_k})$ for which $\boldsymbol{x_k} = \arg\min_{x\in X} (\boldsymbol{c}+\eta(\boldsymbol{d'x_{k-1}})\boldsymbol{d})'\boldsymbol{x}$ is either non-decreasing or non-increasing.*

(c) *(Local optimality)*

$$\boldsymbol{c'y} + f(\boldsymbol{d'y}) \le \boldsymbol{c'z}_\eta + f(\boldsymbol{d'z}_\eta),$$

for all η strictly between θ and $\eta(\boldsymbol{d}'\boldsymbol{x})$. Moreover, if $\boldsymbol{d}'\boldsymbol{y} = \boldsymbol{d}'\boldsymbol{x}$, then the solution $\boldsymbol{y}$ is locally optimal, that is

$$\boldsymbol{y} = \arg\min_{\boldsymbol{u}\in X} (\boldsymbol{c} + \eta(\boldsymbol{d}'\boldsymbol{y})\boldsymbol{d})'\boldsymbol{u}$$

and

$$\boldsymbol{c}'\boldsymbol{y} + f(\boldsymbol{d}'\boldsymbol{y}) \le \boldsymbol{c}'\boldsymbol{z}_\eta + f(\boldsymbol{d}'\boldsymbol{z}_\eta),$$

for all η between θ and $\eta(\boldsymbol{d}'\boldsymbol{y})$.

Proof : (a) We have

$$\begin{aligned} \boldsymbol{c}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x}) &= (\boldsymbol{c} + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d})'\boldsymbol{x} - \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x}) \\ &\ge \boldsymbol{c}'\boldsymbol{y} + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{y} - \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{x} + f(\boldsymbol{d}'\boldsymbol{x}) \\ &= \boldsymbol{c}'\boldsymbol{y} + f(\boldsymbol{d}'\boldsymbol{y}) + \{\eta(\boldsymbol{d}'\boldsymbol{x})(\boldsymbol{d}'\boldsymbol{y} - \boldsymbol{d}'\boldsymbol{x}) - (f(\boldsymbol{d}'\boldsymbol{y}) - f(\boldsymbol{d}'\boldsymbol{x})\} \\ &\ge \boldsymbol{c}'\boldsymbol{y} + f(\boldsymbol{d}'\boldsymbol{y}), \end{aligned}$$

since $\eta(\cdot)$ is a subgradient.

(b) From the optimality of $\boldsymbol{x}$ and $\boldsymbol{y}$, we have

$$\begin{aligned} \boldsymbol{c}'\boldsymbol{y} + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{y} &\le \boldsymbol{c}'\boldsymbol{x} + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{x} \\ -(\boldsymbol{c}'\boldsymbol{y} + \theta\boldsymbol{d}'\boldsymbol{y}) &\le -(\boldsymbol{c}'\boldsymbol{x} + \theta\boldsymbol{d}'\boldsymbol{x}). \end{aligned}$$

Adding the two inequalities we obtain

$$(\boldsymbol{d}'\boldsymbol{x} - \boldsymbol{d}'\boldsymbol{y})(\eta(\boldsymbol{d}'\boldsymbol{x}) - \theta) \ge 0.$$

Therefore, if $\eta(\boldsymbol{d}'\boldsymbol{x}) > \theta$ then $\boldsymbol{d}'\boldsymbol{y} \le \boldsymbol{d}'\boldsymbol{x}$ and since $f(w)$ is a concave function, i.e., $\eta(w)$ is non-increasing, $\eta(\boldsymbol{d}'\boldsymbol{y}) \ge \eta(\boldsymbol{d}'\boldsymbol{x})$. Likewise, if $\eta(\boldsymbol{d}'\boldsymbol{x}) < \theta$ then $\eta(\boldsymbol{d}'\boldsymbol{y}) \le \eta(\boldsymbol{d}'\boldsymbol{x})$. Hence, the sequence $\theta_k = \eta(\boldsymbol{d}'\boldsymbol{x_k})$ is monotone.

(c) We first show that $\boldsymbol{d}'\boldsymbol{z}_\eta$ is in the convex hull of $\boldsymbol{d}'\boldsymbol{x}$ and $\boldsymbol{d}'\boldsymbol{y}$. From the optimality of $\boldsymbol{x}$, $\boldsymbol{y}$, and $\boldsymbol{z}_\eta$ we obtain

$$\begin{aligned} \boldsymbol{c}'\boldsymbol{x} + \theta\boldsymbol{d}'\boldsymbol{x} &\le \boldsymbol{c}'\boldsymbol{z}_\eta + \theta\boldsymbol{d}'\boldsymbol{z}_\eta \\ \boldsymbol{c}'\boldsymbol{x} + \eta\boldsymbol{d}'\boldsymbol{x} &\ge \boldsymbol{c}'\boldsymbol{z}_\eta + \eta\boldsymbol{d}'\boldsymbol{z}_\eta \\ \boldsymbol{c}'\boldsymbol{y} + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{y} &\le \boldsymbol{c}'\boldsymbol{z}_\eta + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{z}_\eta \\ \boldsymbol{c}'\boldsymbol{y} + \eta\boldsymbol{d}'\boldsymbol{y} &\ge \boldsymbol{c}'\boldsymbol{z}_\eta + \eta\boldsymbol{d}'\boldsymbol{z}_\eta \end{aligned}$$

From the first two inequalities we obtain

$$(\boldsymbol{d}'\boldsymbol{z}_\eta - \boldsymbol{d}'\boldsymbol{x})(\theta - \eta) \geq 0,$$

and from the last two we have

$$(\boldsymbol{d}'\boldsymbol{z}_\eta - \boldsymbol{d}'\boldsymbol{y})(\eta(\boldsymbol{d}'\boldsymbol{x}) - \eta) \geq 0.$$

As η is between θ and $\eta(\boldsymbol{d}'\boldsymbol{x})$, then if $\theta < \eta < \eta(\boldsymbol{d}'\boldsymbol{x})$, we conclude since $\eta(\cdot)$ is non-increasing that $\boldsymbol{d}'\boldsymbol{y} \leq \boldsymbol{d}'\boldsymbol{z}_\eta \leq \boldsymbol{d}'\boldsymbol{x}$. Likewise, if $\eta(\boldsymbol{d}'\boldsymbol{x}) < \eta < \theta$, we have $\boldsymbol{d}'\boldsymbol{x} \leq \boldsymbol{d}'\boldsymbol{z}_\eta \leq \boldsymbol{d}'\boldsymbol{y}$., i.e., $\boldsymbol{d}'\boldsymbol{z}_\eta$ is in the convex hull of $\boldsymbol{d}'\boldsymbol{x}$ and $\boldsymbol{d}'\boldsymbol{y}$. Next, we have

$$\begin{aligned}
\boldsymbol{c}'\boldsymbol{y} + f(\boldsymbol{d}'\boldsymbol{y}) &= (\boldsymbol{c} + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d})'\boldsymbol{y} - \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{y} + f(\boldsymbol{d}'\boldsymbol{y}) \\
&\leq (\boldsymbol{c} + \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d})'\boldsymbol{z}_\eta - \eta(\boldsymbol{d}'\boldsymbol{x})\boldsymbol{d}'\boldsymbol{y} + f(\boldsymbol{d}'\boldsymbol{y}) \\
&= \boldsymbol{c}'\boldsymbol{z}_\eta + f(\boldsymbol{d}'\boldsymbol{z}_\eta) + \{f(\boldsymbol{d}'\boldsymbol{y}) - f(\boldsymbol{d}'\boldsymbol{z}_\eta) - \eta(\boldsymbol{d}'\boldsymbol{x})(\boldsymbol{d}'\boldsymbol{y} - \boldsymbol{d}'\boldsymbol{z}_\eta)\} \\
&= \boldsymbol{c}'\boldsymbol{z}_\eta + f(\boldsymbol{d}'\boldsymbol{z}_\eta) + h(\boldsymbol{d}'\boldsymbol{z}_\eta) \\
&\leq \boldsymbol{c}'\boldsymbol{z}_\eta + f(\boldsymbol{d}'\boldsymbol{z}_\eta), \qquad (20)
\end{aligned}$$

where inequality (20) follows from observing that the function $h(\alpha) = f(\boldsymbol{d}'\boldsymbol{y}) - f(\alpha) - \eta(\boldsymbol{d}'\boldsymbol{x})(\boldsymbol{d}'\boldsymbol{y} - \alpha)$ is a convex function with $h(\boldsymbol{d}'\boldsymbol{y}) = 0$ and $h(\boldsymbol{d}'\boldsymbol{x}) \leq 0$. Since $\boldsymbol{d}'\boldsymbol{z}_\eta$ is in the convex hull of $\boldsymbol{d}'\boldsymbol{x}$ and $\boldsymbol{d}'\boldsymbol{y}$, by convexity, $h(\boldsymbol{d}'\boldsymbol{z}_\eta) \leq \mu h(\boldsymbol{d}'\boldsymbol{y}) + (1-\mu)h(\boldsymbol{d}'\boldsymbol{x}) \leq 0$, for some $\mu \in [0, 1]$. ■

Given a feasible solution, $\boldsymbol{x}$, Theorem 5(a) implies that we may improve the objective by solving a sequence of problems using Algorithm 2. Note that at each iteration, we are optimizing a linear function over X. Theorem 5(b) implies that the sequence of $\theta_k = \eta(\boldsymbol{d}'\boldsymbol{x_k})$ is monotone and since it is bounded it converges. Since X is finite, then the algorithm converges in a finite number of steps. Theorem 5(c) implies that at termination (recall that the termination condition is $\boldsymbol{d}'\boldsymbol{y} = \boldsymbol{d}'\boldsymbol{x}$) Algorithm 2 finds a locally optimal solution.

Suppose $\theta = \eta(\boldsymbol{e}'\boldsymbol{d})$ and $\{\boldsymbol{x_1}, \ldots, \boldsymbol{x_k}\}$ be the sequence of solutions of Algorithm 2. From Theorem 5(b), we have

$$\theta = \eta(\boldsymbol{e}'\boldsymbol{d}) \leq \theta_1 = \eta(\boldsymbol{d}'\boldsymbol{x_1}) \leq \ldots \leq \theta_k = \eta(\boldsymbol{d}'\boldsymbol{x_k}).$$

When Algorithm 2 terminates at the solution $\boldsymbol{x_k}$, then from Theorem 5(c),

$$\boldsymbol{c}'\boldsymbol{x_k} + f(\boldsymbol{d}'\boldsymbol{x_k}) \leq \boldsymbol{c}'\boldsymbol{z}_\eta + f(\boldsymbol{d}'\boldsymbol{z}_\eta), \qquad (21)$$

where $\boldsymbol{z}_\eta$ is defined in Eq. (19) for all $\eta \in [\eta(\boldsymbol{e}'\boldsymbol{d}), \eta(\boldsymbol{d}'\boldsymbol{x_k})]$. Likewise, if we apply Algorithm 2 starting at $\bar{\theta} = \eta(0)$, and let $\{\boldsymbol{y_1}, \ldots, \boldsymbol{y_l}\}$ be the sequence of solutions of Algorithm 2, then we have

$$\bar{\theta} = \eta(0) \geq \bar{\theta}_1 = \eta(\boldsymbol{d}'\boldsymbol{y_1}) \geq \ldots \geq \bar{\theta}_l = \eta(\boldsymbol{d}'\boldsymbol{y_l}),$$

and

$$\boldsymbol{c}'\boldsymbol{y_l} + f(\boldsymbol{d}'\boldsymbol{y_l}) \leq \boldsymbol{c}'\boldsymbol{z}_\eta + f(\boldsymbol{d}'\boldsymbol{z}_\eta) \tag{22}$$

for all $\eta \in [\eta(\boldsymbol{d}'\boldsymbol{y_l}), \eta(0)]$. If $\eta(\boldsymbol{d}'\boldsymbol{x_k}) \geq \eta(\boldsymbol{d}'\boldsymbol{y_l})$, we have $\eta(\boldsymbol{d}'\boldsymbol{x_k}) \in [\eta(\boldsymbol{d}'\boldsymbol{y_l}), \eta(0)]$ and $\eta(\boldsymbol{d}'\boldsymbol{y_l}) \in [\eta(\boldsymbol{e}'\boldsymbol{d}), \eta(\boldsymbol{d}'\boldsymbol{x_k})]$. Hence, following from the inequalities (21) and (22), we conclude that

$$\boldsymbol{c}'\boldsymbol{y_l} + f(\boldsymbol{d}'\boldsymbol{y_l}) = \boldsymbol{c}'\boldsymbol{x_k} + f(\boldsymbol{d}'\boldsymbol{x_{k\eta}}) \leq \boldsymbol{c}'\boldsymbol{z}_\eta + f(\boldsymbol{d}'\boldsymbol{z}_\eta)$$

for all $\eta \in [\eta(\boldsymbol{e}'\boldsymbol{d}), \eta(\boldsymbol{d}'\boldsymbol{x_k})] \cup [\eta(\boldsymbol{d}'\boldsymbol{y_l}), \eta(0)] = [\eta(\boldsymbol{e}'\boldsymbol{d}), \eta(0)]$. Therefore, both $\boldsymbol{y_l}$ and $\boldsymbol{x_k}$ are globally optimal solutions. However, if $\eta(\boldsymbol{d}'\boldsymbol{y_l}) > \eta(\boldsymbol{d}'\boldsymbol{x_k})$, we are assured that the global optimal solution is $\boldsymbol{x_k}$, $\boldsymbol{y_l}$ or in $\{\boldsymbol{x} : \ \boldsymbol{x} = \arg\min_{\boldsymbol{u}\in X}(\boldsymbol{c} + \eta\boldsymbol{d})'\boldsymbol{u},\ \eta \in (\eta(\boldsymbol{d}'\boldsymbol{x_k}), \eta(\boldsymbol{d}'\boldsymbol{y_l}))\}$. We next determine an error bound between the optimal objective and the objective of the best local solution, which is either $\boldsymbol{x_k}$ or $\boldsymbol{y_l}$.

Theorem 6 **(a)** *Let* $W = [\underline{w}, \overline{w}]$, $\eta(\underline{w}) > \eta(\overline{w})$, $X' = X \cap \{\boldsymbol{x} : \boldsymbol{d}'\boldsymbol{x} \in W\}$, *and*

$$\boldsymbol{x_1} = \arg\min_{\boldsymbol{y}\in X'}(\boldsymbol{c} + \eta(\underline{w})\boldsymbol{d})'\boldsymbol{y}, \tag{23}$$

$$\boldsymbol{x_2} = \arg\min_{\boldsymbol{y}\in X'}(\boldsymbol{c} + \eta(\overline{w})\boldsymbol{d})'\boldsymbol{y}. \tag{24}$$

Then

$$G' \leq \min\{\boldsymbol{c}'\boldsymbol{x_1} + f(\boldsymbol{d}'\boldsymbol{x_1}), \boldsymbol{c}'\boldsymbol{x_2} + f(\boldsymbol{d}'\boldsymbol{x_2})\} \leq G' + \varepsilon,$$

where

$$G' = \min_{\boldsymbol{y}\in X'} \boldsymbol{c}'\boldsymbol{y} + f(\boldsymbol{d}'\boldsymbol{y}), \tag{25}$$

$$\varepsilon = \eta(\underline{w})(w^* - \underline{w}) + f(\underline{w}) - f(w^*),$$

and

$$w^* = \frac{f(\overline{w}) - f(\underline{w}) + \eta(\underline{w})\underline{w} - \eta(\overline{w})\overline{w}}{\eta(\underline{w}) - \eta(\overline{w})}.$$

(b) *Suppose the feasible solutions* $\boldsymbol{x_1}$ *and* $\boldsymbol{x_2}$ *satisfy*

$$\boldsymbol{x_1} = \arg\min_{\boldsymbol{y}\in X}(\boldsymbol{c} + \eta(\boldsymbol{d}'\boldsymbol{x_1})\boldsymbol{d})'\boldsymbol{y}, \tag{26}$$

$$\boldsymbol{x_2} = \arg\min_{\boldsymbol{y}\in X}(\boldsymbol{c} + \eta(\boldsymbol{d}'\boldsymbol{x_2})\boldsymbol{d})'\boldsymbol{y}, \tag{27}$$

such that $\eta(\underline{w}) > \eta(\overline{w})$, *with* $\underline{w} = \boldsymbol{d}'\boldsymbol{x_1}$, $\overline{w} = \boldsymbol{d}'\boldsymbol{x_2}$ *and there exists an optimal solution* $\boldsymbol{x}^* = \arg\min_{y\in X}(\boldsymbol{c} + \eta\boldsymbol{d})'\boldsymbol{y}$ *for some* $\eta \in (\eta(\overline{w}), \eta(\underline{w}))$, *then*

$$G^* \leq \min\{\boldsymbol{c}'\boldsymbol{x_1} + f(\boldsymbol{d}'\boldsymbol{x_1}), \boldsymbol{c}'\boldsymbol{x_2} + f(\boldsymbol{d}'\boldsymbol{x_2})\} \leq G^* + \varepsilon, \tag{28}$$

where $G^* = \boldsymbol{c}'\boldsymbol{x}^* + f(\boldsymbol{d}'\boldsymbol{x}^*)$.

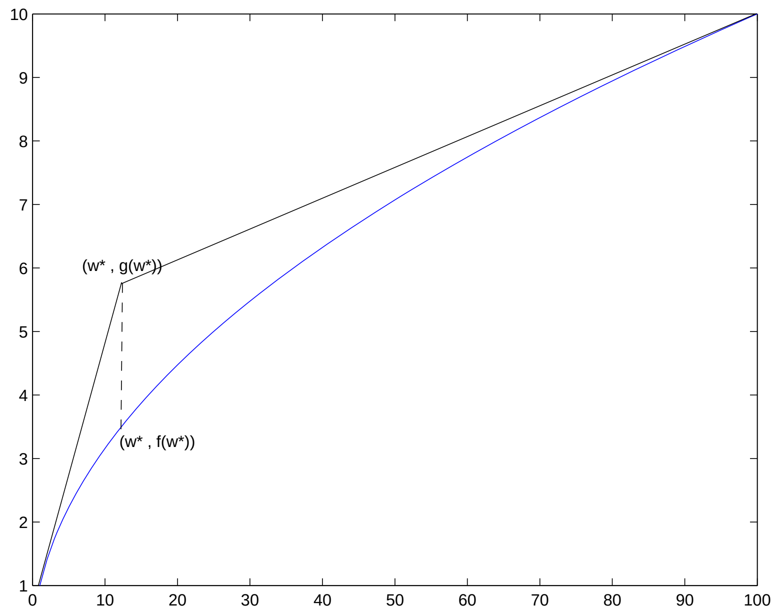

Figure 1: Illustration of the maximum gap between the functions $f(w)$ and $g(w)$.

Proof : (a) Let $g(w)$, $w \in W$ be a piecewise concave function comprising of two line segments through $(\underline{w}, f(\underline{w}))$, $(\overline{w}, f(\overline{w}))$ with respective subgradients $\eta(\underline{w})$ and $\eta(\overline{w})$. Clearly $f(w) \leq g(w)$ for $w \in W$, and hence, we have $-\varepsilon \leq f(w) - g(w) \leq 0$, where $\varepsilon = \max_{w \in W}(g(w) - f(w)) = g(w^*) - f(w^*)$, noting that the maximum difference occurs at the intersection of the line segments (see Figure 1). Therefore,

$$g(w^*) = \eta(\underline{w})(w^* - \underline{w}) + f(\underline{w}) = \eta(\overline{w})(w^* - \overline{w}) + f(\overline{w}).$$

Solving for w^*, we have

$$w^* = \frac{f(\overline{w}) - f(\underline{w}) + \eta(\underline{w})\underline{w} - \eta(\overline{w})\overline{w}}{\eta(\underline{w}) - \eta(\overline{w})}.$$

Applying Proposition 2 with X' instead of X and $k = 2$, we obtain

$$\min_{y \in X'} \boldsymbol{c'y} + g(\boldsymbol{d'y}) = \min \left\{ \boldsymbol{c'x_1} + g(\boldsymbol{d'x_1}), \boldsymbol{c'x_2} + g(\boldsymbol{d'x_2}) \right\}.$$

Finally, from Theorem 3, we have

$$G'^* \leq \min \left\{ \boldsymbol{c'x_1} + f(\boldsymbol{d'x_1}), \boldsymbol{c'x_2} + f(\boldsymbol{d'x_2}) \right\} \leq G'^* + \varepsilon.$$

(b) Under the stated conditions, observe that the optimal solutions of the problems (26) and (27) are respectively the same as the optimal solutions of the problems (23) and (24). Let $\eta \in (\eta(\boldsymbol{d'x_2}), \eta(\boldsymbol{d'x_1}))$

such that $\boldsymbol{x}^* = \arg\min_{y\in X}(\boldsymbol{c}+\eta\boldsymbol{d})'\boldsymbol{y}$. We establish that

$$\begin{aligned}
\boldsymbol{c}'\boldsymbol{x}^* + \eta\boldsymbol{d}'\boldsymbol{x}^* &\le \boldsymbol{c}'\boldsymbol{x_1} + \eta\boldsymbol{d}'\boldsymbol{x_1}\\
\boldsymbol{c}'\boldsymbol{x}^* + \eta(\boldsymbol{d}'\boldsymbol{x_1})\boldsymbol{d}'\boldsymbol{x}^* &\ge \boldsymbol{c}'\boldsymbol{x_1} + \eta(\boldsymbol{d}'\boldsymbol{x_1})\boldsymbol{d}'\boldsymbol{x_1}\\
\boldsymbol{c}'\boldsymbol{x}^* + \eta\boldsymbol{d}'\boldsymbol{x}^* &\le \boldsymbol{c}'\boldsymbol{x_2} + \eta\boldsymbol{d}'\boldsymbol{x_2}\\
\boldsymbol{c}'\boldsymbol{x}^* + \eta(\boldsymbol{d}'\boldsymbol{x_2})\boldsymbol{d}'\boldsymbol{x}^* &\ge \boldsymbol{c}'\boldsymbol{x_2} + \eta(\boldsymbol{d}'\boldsymbol{x_2})\boldsymbol{d}'\boldsymbol{x_2}.
\end{aligned}$$

Since $\eta(\boldsymbol{d}'\boldsymbol{x_2}) < \eta < \eta(\boldsymbol{d}'\boldsymbol{x_1})$, it follows that $\boldsymbol{d}'\boldsymbol{x}^* \in [\boldsymbol{d}'\boldsymbol{x_1}, \boldsymbol{d}'\boldsymbol{x_2}]$ and hence, $G^* = G'$ and the bounds of (25) follows from part (a). ∎

Remark: If $\eta(\boldsymbol{d}'\boldsymbol{y_l}) > \eta(\boldsymbol{d}'\boldsymbol{x_k})$, Theorem 6(b) provides a guarantee on the quality of the best solution of the two locally optimal solution $\boldsymbol{x_k}$ and $\boldsymbol{y_l}$ relative to the global optimum. Moreover, we can improve the error bound by partitioning the interval $[\eta(\overline{w}), \eta(\underline{w})]$, with $\underline{w} = \boldsymbol{d}'\boldsymbol{y_l}$, $\overline{w} = \boldsymbol{d}'\boldsymbol{x_k}$ into two subintervals, $[\eta(\overline{w}), (\eta(\overline{w}) + \eta(\underline{w}))/2]$ and $[(\eta(\overline{w}) + \eta(\underline{w}))/2, \eta(\underline{w})]$ and applying Algorithm 2 in the intervals. Using Theorem 6(a), we can obtain improved bounds. Continuing this way, we can find the globally optimal solution.

6 Generalized Bertsimas and Sim Robust Formulation

Bertsimas and Sim [6] propose the following model for robust discrete optimization:

$$\begin{aligned}
Z^* &= \min_{\boldsymbol{x}\in X} \boldsymbol{c}'\boldsymbol{x} + \max_{\{S\cup\{t\}|\ S\subseteq J, |S|=\lfloor\Gamma\rfloor, t\in J\setminus S\}} \left\{ \sum_{j\in S} d_j x_j + (\Gamma - \lfloor\Gamma\rfloor) d_t x_t \right\}\\
&= \min_{\boldsymbol{x}\in X} \boldsymbol{c}'\boldsymbol{x} + \max_{\{\boldsymbol{z}:\boldsymbol{e}'\boldsymbol{z}\le\Gamma, \boldsymbol{0}\le\boldsymbol{z}\le\boldsymbol{e}\}} \left\{ \sum_{j\in J} d_j x_j z_j \right\}
\end{aligned} \tag{29}$$

They show:

Theorem 7 *(Bertsimas and Sim [6]) Let $\boldsymbol{x}^*$ be an optimal solution of Problem (29). If each $\tilde{c}_j$ is a random variable, independently and symmetrically distributed in $[c_j - d_j, c_j + d_j]$, then*

$$\Pr\left(\sum_j \tilde{c}_j x_j^* > Z^*\right) \le B(r,\Gamma) = \frac{1}{2^r}\left\{(1-\mu)\sum_{l=\lfloor\nu\rfloor}^{n}\binom{r}{l} + \mu\sum_{l=\lfloor\nu\rfloor+1}^{r}\binom{r}{l}\right\}, \tag{30}$$

where $\nu = \frac{\Gamma_i + r}{2}$ and $\mu = \nu - \lfloor\nu\rfloor$. Moreover, for $\Gamma = \Lambda\sqrt{r}$,

$$\lim_{r\to\infty} B(r,\Gamma) = 1 - \Phi(\Lambda), \tag{31}$$

where $\Phi(\Lambda)$ is the cumulative distribution function of a standard normal.

Intuitively, if we select $\Gamma = \Lambda\sqrt{r}$, then the probability that the robust solution exceeds Z^* is approximately $1 - \Phi(\Lambda)$. Since in this case feasible solutions are restricted to binary values, we can achieve a less conservative solution by replacing r by $\sum_{j\in J} x_j^* = \boldsymbol{e'_J x}$, i.e., the parameter Γ in the robust problem (29) depends on $\boldsymbol{e'_J x}$. We write $\Gamma = f(\boldsymbol{e'_J x})$, where $f(\cdot)$ is a concave function. Thus, we propose to solve the following problem:

$$Z^* = \min_{\boldsymbol{x}\in X} \boldsymbol{c'x} + \max_{\{\boldsymbol{z}:\boldsymbol{e'z}\le f(\boldsymbol{e'_J x}),\boldsymbol{0}\le\boldsymbol{z}\le\boldsymbol{e}\}} \left\{ \sum_{j\in J} d_j x_j z_j \right\}. \tag{32}$$

Without loss of generality, we assume that $d_1 \ge d_2 \ge \ldots \ge d_r$. We define $d_{r+1} = 0$ and let $S_l = \{1,\ldots,l\}$. For notational convenience, we also define $d_0 = 0$ and $S_0 = J$.

Theorem 8 *Let $\eta(w)$ be a subgradient of the concave function $f(\cdot)$ evaluated at w. Problem (32) satisfies $Z^* = \min\limits_{(l,k):l,k\in J\cup\{0\}} Z_{lk}$, where*

$$Z_{lk} = \min_{\boldsymbol{x}\in X} \boldsymbol{c'x} + \sum_{j\in S_l} (d_j - d_l)x_j + \eta(k) d_l \boldsymbol{e'_J x} + d_l(f(k) - k\eta(k)). \tag{33}$$

Proof : By strong duality of the inner maximization function with respect to $\boldsymbol{z}$, Problem (32) is equivalent to solving the following problem:

$$\begin{array}{lll} \text{minimize} & \boldsymbol{c'x} + \sum\limits_{j\in J} p_j + f(\boldsymbol{e'_J x})\theta & \\ \text{subject to} & p_j \ge d_j x_j - \theta & \forall j \in J \\ & p_j \ge 0 & \forall j \in J \\ & x \in X & \\ & \theta \ge 0, & \end{array} \tag{34}$$

We eliminate the variables p_j and express Problem (34) as follows:

$$\begin{array}{ll} \text{minimize} & \boldsymbol{c'x} + \sum\limits_{j\in J} \max\{d_j x_j - \theta, 0\} + f(\boldsymbol{e'_J x})\theta \\ \text{subject to} & x \in X \\ & \theta \ge 0. \end{array} \tag{35}$$

Since $\boldsymbol{x} \in \{0,1\}^n$, we observe that

$$\max\{d_j x_j - \theta, 0\} = \begin{cases} d_j - \theta & \text{if } x_j = 1 \text{ and } d_j \ge \theta \\ 0 & \text{if } x_j = 0 \text{ or } d_j < \theta. \end{cases} \tag{36}$$

By restricting the interval of θ can vary we obtain that $Z^* = \min_{\theta\geq 0}\min_{l=0,\ldots,r} Z_l(\theta)$ where $Z_l(\theta)$, $l = 1,\ldots,r$, is defined for $\theta \in [d_l, d_{l+1}]$ is

$$Z_l(\theta) = \min_{\boldsymbol{x}\in X} \boldsymbol{c'x} + \sum_{j\in S_l}(d_j - \theta)x_j + f(\boldsymbol{e'_J x})\theta \tag{37}$$

and for $\theta \in [d_1, \infty)$:

$$Z_0(\theta) = \min_{\boldsymbol{x}\in X} \boldsymbol{c'x} + f(\boldsymbol{e'_J x})\theta. \tag{38}$$

Since each function $Z_l(\theta)$ is optimized over the interval $[d_l, d_{l+1}]$, the optimal solution is realized in either d_l or d_{l+1}. Hence, we can restrict θ from the set $\{d_1,\ldots,d_r,0\}$ and establish that

$$Z^* = \min_{l\in J\cup\{0\}} \boldsymbol{c'x} + \sum_{j\in S_l}(d_j - d_l)x_j + f(\boldsymbol{e'_J x})d_l. \tag{39}$$

Since $\boldsymbol{e'_J x} \in \{0,1,\ldots,r\}$, we apply Theorem 2 to obtain the subproblem decomposition of (33). ■

Theorem 8 suggests that the robust problem remains polynomially solvable if the nominal problem is polynomially solvable, but at the expense of higher computational complexity. We next explore faster algorithms that are only guarantee local optimality. In this spirit and analogously to Theorem 5, we provide a necessary condition for optimality, which can be exploited in a local search algorithm.

Theorem 9 *An optimal solution, $\boldsymbol{x}$ to the Problem (32) is also an optimal solution to the following problem:*

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c'y} + \sum_{j\in S_{l^*}}(d_j - d_{l^*})y_j + \eta(\boldsymbol{e'_J x})d_{l^*}\boldsymbol{e'_J y} \\ \text{subject to} \quad & \boldsymbol{y} \in X, \end{aligned} \tag{40}$$

where $l^* = \arg\min_{l\in J\cup\{0\}} \sum_{j\in S_l}(d_j - d_l)x_j + f(\boldsymbol{e'_J x})d_l$.

Proof : Suppose $\boldsymbol{x}$ is an optimal solution for Problem (32) but not for Problem (40). Let $\boldsymbol{y}$ be the optimal solution to Problem (40). Therefore,

$$\begin{aligned} & \boldsymbol{c'x} + \max_{\{\boldsymbol{z}:\boldsymbol{e'z}\leq f(\boldsymbol{e'_J x}),\boldsymbol{0}\leq\boldsymbol{z}\leq\boldsymbol{e}\}}\left\{\sum_{j\in J} d_j x_j z_j\right\} \\ = \; & \min_{l\in J\cup\{0\}} \boldsymbol{c'x} + \sum_{j\in S_l}(d_j - d_l)x_j + f(\boldsymbol{e'_J x})d_l \\ = \; & \boldsymbol{c'x} + \sum_{j\in S_{l^*}}(d_j - d_{l^*})x_j + f(\boldsymbol{e'_J x})d_{l^*} \\ = \; & \boldsymbol{c'x} + \sum_{j\in S_{l^*}}(d_j - d_{l^*})x_j + \eta(\boldsymbol{e'_J x})d_{l^*}\boldsymbol{e'_J x} - \eta(\boldsymbol{e'_J x})d_{l^*}\boldsymbol{e'_J x} + f(\boldsymbol{e'_J x})d_{l^*} \end{aligned} \tag{41}$$

$$\begin{aligned}
&> \boldsymbol{c'y} + \sum_{j\in S_{l^*}} (d_j - d_{l^*})y_j + \eta(\boldsymbol{e'_J x}) d_{l^*}\boldsymbol{e'_J y} - \eta(\boldsymbol{e'_J x}) d_{l^*}\boldsymbol{e'_J x} + f(\boldsymbol{e'_J x}) d_{l^*} \\
&= \boldsymbol{c'y} + \sum_{j\in S_{l^*}} (d_j - d_{l^*})y_j + f(\boldsymbol{e'_J y}) d_{l^*} + \left(\eta(\boldsymbol{e'_J x})(\boldsymbol{e'_J y} - \boldsymbol{e'_J x}) - (f(\boldsymbol{e'_J y}) - f(\boldsymbol{e'_J x}))\right) d_{l^*} \\
&\geq \boldsymbol{c'y} + \sum_{j\in S_{l^*}} (d_j - d_{l^*})y_j + f(\boldsymbol{e'_J y}) d_{l^*} \\
&\geq \min_{l\in J\cup\{0\}} \boldsymbol{c'y} + \sum_{j\in S_l} (d_j - d_l)y_j + f(\boldsymbol{e'_J y}) d_l \\
&= \boldsymbol{c'y} + \max_{\{\boldsymbol{z}:\boldsymbol{e'z}\leq f(\boldsymbol{e'_J y}), \boldsymbol{0}\leq \boldsymbol{z}\leq \boldsymbol{e}\}} \left\{ \sum_{j\in J} d_j y_j z_j \right\} \qquad (42)
\end{aligned}$$

where the Eqs. (41) and (42) follows from Eq. (39). This contradicts that $\boldsymbol{x}$ is optimal. ■

7 Experimental Results

In this section, we provide experimental evidence on the effectiveness of Algorithm 2. We apply Algorithm 2 as follows. We start with two initial solutions $\boldsymbol{x_1}$ and $\boldsymbol{x_2}$. Starting with $\boldsymbol{x_1}$ ($\boldsymbol{x_2}$) Algorithm 2 finds a locally optimal solution $\boldsymbol{y_1}$ ($\boldsymbol{y_2}$). If $\boldsymbol{y_1} = \boldsymbol{y_2}$, by Theorem 5, the optimum solution is found. Otherwise, we report the optimality gap ε derived from Theorem 6. If we want to find the optimal solution, we partition into smaller search regions (see Remark after Theorem 6) using Theorem 4 and repeatedly apply Algorithm 2 until all regions are covered.

We apply the proposed approach to the binary knapsack and the uniform matroid problems.

7.1 The Robust Knapsack Problem

The binary knapsack problem is:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i\in N} \tilde{c}_i x_i \\
\text{subject to} \quad & \sum_{i\in N} w_i x_i \leq b \\
& \boldsymbol{x} \in \{0,1\}^n.
\end{aligned}$$

We assume that the costs $\tilde{c}_i$ are random variables that are independently distributed with mean c_i and variance $d_i = \sigma_i^2$. Under the ellipsoidal uncertainty set, the robust model is:

$$\begin{aligned}
\text{maximize} \quad & \sum_{i\in N} c_i x_i + \Omega\sqrt{\boldsymbol{d'x}} \\
\text{subject to} \quad & \sum_{i\in N} w_i x_i \leq b \\
& \boldsymbol{x} \in \{0,1\}^n.
\end{aligned}$$

The instance of the robust knapsack problem is generated randomly with $|N| = 200$ and capacity limit, b equals $20,000$. The nominal weight w_i is randomly chosen from the set $\{100, \ldots, 1500\}$, the cost c_i is randomly chosen from the set $\{10,000, \ldots, 15,000\}$, and the standard deviation σ_j is dependent on c_j such that $\sigma_j = \delta_j c_j$, where δ_j is uniformly distributed in $[0, 1]$. We vary the parameter Ω from 1 to 5 and report in Table 2 the best attainable objective, Z_H, the number of instance of nominal problem solved, as well as the optimality gap ε.

Ω	Z_H	Iterations	ε	ε/Z_H
1.0	1965421.36	4	0	0
2.0	2054638.82	6	0	0
2.5	2097656.46	6	0	0
3.0	2140207.75	6	3.1145	1.45523×10^{-6}
3.05	2144317.00	5	0	0
3.5	2182235.78	5	0	0
4.0	2224365.19	6	3.4046	1.53059×10^{-6}
4.5	2266054.21	7	0	0
5.0	2307475.12	8	0	0

Table 2: Performance of Algorithm 2 on the robust knapsack problem.

It is surprising that in all of the instances, we can obtain the optimal solution of the robust problem using a small number of iterations. Even for the cases, $\Omega = 3, 4$, where the Algorithm 2 terminates with more than one local minimum solutions, the resulting optimality gap is very small, which is usually acceptable in practical settings.

7.2 The Robust Minimum Cost over a Uniform Matroid

We consider the problem of minimizing the total cost of selecting k items out of a set of n items that can be expressed as the following integer programming problem:

$$\begin{aligned} \text{minimize} \quad & \sum_{i \in N} \tilde{c}_i x_i \\ \text{subject to} \quad & \sum_{i \in N} x_i = k \\ & \boldsymbol{x} \in \{0,1\}^n. \end{aligned} \tag{43}$$

In this problem, the cost components are subjected to uncertainty. If the model is deterministic, we can easily solve the problem in $O(n \log n)$ by sorting the costs in ascending order and choosing the first k items. In the robust framework under the ellipsoidal uncertainty set, we solve the following problem:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c}'\boldsymbol{x} + \Omega\sqrt{\boldsymbol{d}'\boldsymbol{x}} \\ \text{subject to} \quad & \sum_{i \in N} x_i = k \\ & \boldsymbol{x} \in \{0,1\}^n. \end{aligned} \tag{44}$$

Since the underlying set is a matroid, it is well known that Problem (44) can be solved in strongly polynomial time using parametric programming. Instead, we apply Algorithm 2 and observe the number of iterations needed before converging to a local minimum solution. Setting $|k| = |N|/2$, c_j and $\sigma_j = \sqrt{d_j}$ being uniformly distributed in $[5000, 20000]$ and $[500, 5000]$ respectively, we study the convergence properties as we vary $|N|$ from 200 to 20,000 and Ω from 1 to 3. For a given $|N|$ and Ω, we generate $\boldsymbol{c}$ and $\boldsymbol{d}$ randomly and solve 100 instances of the problem. Aggregating the results from solving the 100 instances, we report in Table 3 the average number of iterations before finding a local solution, the maximum relative optimality gap, ε/Z_H and the percentage of the local minimum solutions that are global, i.e., $\varepsilon = 0$.

The overall performance of Algorithm 2 is surprisingly good. It also suggests scalability, as the number of iterations is marginally affected by an increase in $|N|$. In fact, in most of the problems tested, we obtain the optimal solution by solving less than 10 iterations of the nominal problem. Even in cases when local solutions are found, the corresponding optimality gap is negligible. In summary, Algorithm 2 seems practically promising.

8 Conclusions

A message of the present paper is that the complexity of robust discrete optimization is affected by the choice of the uncertainty set. For ellipsoidal uncertainty sets, we have shown an increase in complexity for the robust counterpart of a discrete optimization problem for general covariance matrices $\boldsymbol{\Sigma}$ (correlated data), a preservation of complexity when $\boldsymbol{\Sigma} = \sigma \boldsymbol{I}$ (uncorrelated and identically distributed data), while we have left open the complexity when the matrix $\boldsymbol{\Sigma}$ is diagonal (uncorrelated data). In the latter case, we proposed two algorithms that in computational experiments have excellent empirical performance.

Ω	$\lvert N\rvert$	Ave. Iter.	$\max(\varepsilon/Z_H)$	Opt. Sol. %
1	200	5.73	7.89×10^{-7}	98%
1	500	5.91	3.71×10^{-8}	99%
1	1000	6.18	5.80×10^{-9}	99%
1	2000	6.43	0	100%
1	5000	6.72	0	100%
1	10000	6.92	0	100%
1	20000	6.98	0	100%
2	200	6.24	0	100%
2	500	6.50	0	100%
2	1000	6.80	0	100%
2	2000	6.95	0	100%
2	5000	6.98	0	100%
2	10000	7.01	0	100%
2	20000	7.02	0	100%
3	200	6.55	1.62×10^{-6}	94%
3	500	6.85	7.95×10^{-8}	97%
3	1000	6.92	0	100%
3	2000	7.01	1.08×10^{-9}	99%
3	5000	7.06	5.13×10^{-10}	98%
3	10000	7.07	0	100%
3	20000	7.07	0	100%

Table 3: Performance of Algorithm 2 on the robust minimum cost problem over a uniform matroid.

References

[1] Ben-Tal, A., A. Nemirovski (1998): Robust convex optimization, *Math. Oper. Res.* 23, 769-805.

[2] Ben-Tal, A., A. Nemirovski (1999): Robust solutions to uncertain programs, *Oper. Res. Lett.* 25, 1-13.

[3] Ben-Tal, A., A. Nemirovski (2000): Robust solutions of linear programming problems contaminated with uncertain data, *Math. Prog.* 88, 411-424.

[4] Ben-Tal, A., L. El-Ghaoui, A. Nemirovski (2000): Robust semidefinite programming, in Saigal, R., Vandenberghe, L., Wolkowicz, H., eds., *Semidefinite programming and applications*, Kluwer Academic Publishers, Norwell, Ma.

[5] Bertsimas, D., M. Sim (2003): Robust Discrete Optimization and Network Flows, *Math. Prog.* 98, 49-71.

[6] Bertsimas, D., M. Sim (2004): The Price of Robustness, *Oper. Res.* 52, 1, 35–53.

[7] El-Ghaoui, H. Lebret (1997): Robust solutions to least-square problems to uncertain data matrices, *SIAM J. Matrix Anal. Appl.* 18, 1035-1064.

[8] El-Ghaoui, L., Oustry, F. and H. Lebret (1998): Robust solutions to uncertain semidefinite programs, *SIAM J. Optim.* 9, 33-52.

[9] Eppstein, D. (1995): Geometric lower bounds for parametric matroid optimization, *Proc. 27th ACM Symposium on the Theory of Computing*, 662-671.

[10] Fern andez-Baca, D., G. Slutzki, and D. Eppstein. (1996): Using sparsification for parametric minimum spanning tree problems, *Proc. 5th Scand. Workshop in Algorithm Theory. Springer-Verlag, Lecture Notes in Computer Science*, 1097-1100.

[11] Gusfield, D. (1980): Sensitivity Analysis for combinatorial optimization, *Technical Report* UCB/ERL M80/22, University of Califonia, Berkeley.

[12] Ishii, H., S. Shiode, T. Nishida, Y. Namasuya. (1981) : Stochastic Spanning Tree Problem, *Discrete Appl. Math.* 3, 263-271.

[13] Hassin, R., A. Tamir (1989) Maximizing classes of two-parameter objectives over matriods, *Math. Oper. Res.* 14, 362-375.

[14] Karp, R. M., J. B. Orlin (1981): Parametric shortest path algorithms with an applications to cyclic stuffing, *Discrete Appl. Math.* 3, 37-45.

[15] Kouvelis, P., G. Yu (1997): Robust discrete optimization and its applications, *Kluwer Academic Publishers*, Norwell, Ma.