Robust Discrete Optimization and Downside Risk Measures

Dimitris Bertsimas *   Melvyn Sim †

Revised: January 2005

Abstract

We propose methods for robust discrete optimization in which the objective function has cost components that are subject to independent and bounded perturbations. Motivated by risk management practice, we approximate the problem of optimization of VaR, a widely used downside risk measure by introducing four approximating models that origininate in robust optimization. We show that all four models allow the flexibility of adjusting the level of conservativism such that the probability of the actual cost being less than a specified level, in the worst case distribution, is at least $1 - \alpha$. Under a robust model with ellipsoidal uncertainty set, we propose a Frank-Wolfe type algorithm that we show converges to a locally optimal solution, and in computational experiments is remarkably effective. We propose a robust model that is at most $1 + \varepsilon$ more conservative than the ellipsoidal uncertainty set model and show that we can reduce the robust model to solving a polynomial number of nominal problems. We generalize our earlier proposed framework and show that it results in decreased conservatism, while maintaining the complexity of the nominal problem. Among the proposed robust models, we show both theoretically and computationally that the robust model under ellipsoidal uncertainty set is the least conservative.

*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 has been proposed in the last decade as a tractable approach to optimization under uncertainty. By replacing probability distributions with uncertainty sets as the primitive information regarding uncertainty, researchers (Ben-Tal and Nemirovski [2, 3, 4] and El-Ghaoui et al. [11, 12]) have found efficient algorithms to solve certain classes of convex optimization problems under uncertainty. In designing such uncertainty sets two considerations are, in our opinion, important:

(a) Preserving the computational tractability both theoretically and most importantly practically of the nominal problem. From a theoretical perspective it is desirable that if the nominal problem is solvable in polynomial time, then the robust problem is also polynomially solvable.

(b) Being able to adjust the level of conservatism based on broad assumptions on the probability distributions. This is important, since from these guarantees we can select parameters that affect the tradeoff between robustness and optimality.

As an example of how these considerations are applied to design robust approaches, we refer to Bertsimas and Sim [9] who proposed a tractable approach for conic optimization that relates the robust solution with the desired level of conservatism in terms of feasibility guarantees.

Most proposals for robust discrete optimization do not address these issues, however. Kouvelis and Yu [20] 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. A related objective is the minimax-regret approach, which seeks to minimize the worst case loss in objective value that may occur. Again, under the minimax-regret notion of robustness, many of the polynomially solvable discrete optimization problems become NP-hard. Under the minimax-regret robustness approach, Averbakh [1] showed that polynomial solvability is preserved for a specific discrete optimization problem (optimization over a uniform matroid) when each cost coefficient can vary within an interval (interval representation of uncertainty); however, the approach does not seem to generalize to other discrete optimization problems. Furthermore, in all these approaches, it is unclear how one could articulate the level of conservatism from a probabilistic perspective. There have also been research efforts to apply stochastic programming methods to discrete optimization (see for example Schultz et. al. [23]), but the computational requirements are even more severe in this case.

There are few papers that addresses both criteria. Ishii et. al. [16] 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 of Ben-Tal and Nemirovski [2, 3, 4] and El-Ghaoui et al. [11, 12]. Bertsimas and Sim [7, 8] propose an approach for solving robust discrete optimization problems that has the flexibility of adjusting the level of conservatism 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.

Our primary goal in this paper is to understand (a) the tractability and (b) the degree of conservatism of alternative models for discrete optimization under uncertainty. Our secondary goal is to make a connection between robust optimization and optimization of quantile-based downside risk measures. In recent years the financial industry has extensively used quantile-based downside risk measures. Indeed one such measure, Value-at-Risk, or VaR, has been increasingly used as a risk management tool (see e.g. Jorion [17]) and has had significant impact in the application of risk management in practice. Rather than focusing on specific probability distributions (as classical stochastic optimization does) or uncertainty sets (as robust optimization does) we focus on a particular risk measure, VaR, as the primitive objective. More specifically our contributions in this paper include:

(a) We propose four models that originate in robust optimization as approximating problems for the optimization of VaR. Under rather general probabilistic assumptions, we provide probability bounds that allow the modeler to control the flexibility of adjusting the conservativism of the solution such that the probability of the cost being less than the specified level, in the worst case distribution, is at least $1 - \alpha$.

(b) For the first of the four approximating models that is based on an ellipsoidal uncertainty set, we propose a Frank-Wolfe type algorithm, which we prove converges to a locally optimal solution, and in computational experiments is remarkably effective. We show both theoretically and computationally that the robust model under ellipsoidal uncertainty is potentially the least conservative.

(c) For the second of the four approximating models we show that it is at most $1+\varepsilon$ more conservative than the ellipsoidal uncertainty set model and that we can reduce the robust model to solving a polynomial number of nominal problems, thus proving polynomial tractability.

(d) We contrast the conservatism of the earlier proposal of Bertsimas and Sim [7], which constitutes the third approximating model, with the robust model under an ellipsoidal uncertainty set and show that it can be $O(m^{1/4})$ more conservative, where m is the number of uncertain cost components in the objective.

(e) We also propose a fourth approximating model that generalizes the framework of Bertsimas and

Sim [7] and results in decreased conservatism by $O(m^{1/4})$, while maintaining the complexity of the nominal problem. However, we show that this model can still be $O(m^{1/4})$ more conservative than the robust model under an ellipsoidal uncertainty set.

Structure of the paper. In Section 2, we introduce the problem of optimization of quantile-based downside risk and outline the four approximating models that originate in robust optimization. In Section 3, we present algorithms for solving the approximating robust models. In Section 4, we compare theoretically the level of conservatism among the proposed robust models. In Section 5, we present some experimental findings comparing the robust models in solving the shortest path problem. The final section contains some concluding remarks.

2 Down-side Risk and Robust Discrete Optimization

A nominal discrete optimization problem is:

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

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 potentially uncertain and we want to minimize the objective that accounts for the variability of the cost components.

We define the α-quantile of a random variable $\mathcal{X}$ as

$$q_\alpha(\mathcal{X}) := \inf\{z | P(\mathcal{X} \geq z) \leq \alpha\}, \quad \alpha \in (0,1).$$

Our point of departure is the following theorem.

Theorem 1 (Levy and Kroll [22] and Levy [21]) *Let $\mathcal{X}$ and $\mathcal{Y}$ be random variables with continuous densities. Then $E[u(\mathcal{X})] \geq E[u(\mathcal{Y})]$ for all increasing utility functions $u(\cdot)$ if and only if $q_\alpha(\mathcal{X}) \leq q_\alpha(\mathcal{Y}), \quad \forall \alpha \in (0,1)$ and we have strict inequality for some α.*

Letting $\mathcal{X} = \tilde{\boldsymbol{c}}'\boldsymbol{x}$, and using the notation

$$q_\alpha(\boldsymbol{x}) = q_\alpha(\mathcal{X}) = \inf\{z | \Pr(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq z) \leq \alpha\} \tag{2}$$

it is thus natural to minimize $q_\alpha(\boldsymbol{x})$ for a specific α. Note that if $\tilde{c}$ is normally distributed, then $\mathcal{X}$ is normally distributed, and thus minimizing $q_\alpha(\boldsymbol{x})$ is equivalent in this case to minimize the variance of

$\mathcal{X}$. We thus propose to solve the following problem:

$$\begin{aligned} \text{minimize} \quad & q_\alpha(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in X. \end{aligned} \tag{3}$$

By definition, for any feasible solution $\boldsymbol{x}$ we have $\mathrm{P}(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq q_\alpha(\boldsymbol{x})) \leq \alpha$. Hence, the parameter α controls the level of conservatism of the solutions desired by the modeler such that the cost of at most $q_\alpha(\boldsymbol{x})$ is attainable with probability $1-\alpha$. In the absolute worst case scenario we set $\alpha = 0$, which can be over-conservative for most modelers.

There two main difficulties with model (3): (a) the distributions of the cost components are rarely known and (b) evaluating $q_\alpha(\boldsymbol{x})$ is not obvious. For this reason, we consider a fairly general cost uncertainty model.

Model of Cost Uncertainty C:

Each cost entry $\tilde{c}_j,\ j \in N$ is modeled as an independent bounded random variable with mean value c_j that takes values in $[c_j - \sigma_j, c_j + \sigma_j]$, $\sigma_j \geq 0$.

Remark : We note that Bertsimas and Sim [8] assume a more restrictive uncertainty model, in which each uncertain data entry is symmetrically distributed.

Without loss of generality, we assume that only the first $m \leq n$ cost components have positive deviations and that $\sigma_1 \geq \ldots \geq \sigma_m > 0$. Hence, the cost components, $\tilde{c}_j$, $j > m$ are deterministic, that is , $\sigma_j = 0$. For notational convenience, we define $M = \{1, \ldots, m\}$ to represent the set of indices with positive deviations and $\boldsymbol{e}$ denotes a vector with $e_j = 1$ for $j \in M$ and $e_j = 0$, otherwise. We also define $\sigma_{\min} = \sigma_m$ and $\sigma_{\max} = \sigma_1$.

As we do not know the exact distributions of the cost components, it would be impossible to determine the value $q_\alpha(\boldsymbol{x})$. In our proposed approach, we solve the following model,

$$\begin{aligned} \text{minimize} \quad & r_\alpha(\boldsymbol{x}) \\ \text{subject to} \quad & \boldsymbol{x} \in X, \end{aligned} \tag{4}$$

where under the model of cost uncertainty C, $r_\alpha(\boldsymbol{x})$ satisfies:

$$\mathrm{P}(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha(\boldsymbol{x})) \leq \alpha.$$

We propose four robust objective measures as follows:

Robust Objective Measures

(A)

$$r_\alpha^{(A)}(\boldsymbol{x}) = \boldsymbol{c}'\boldsymbol{x} + \Omega_\alpha \sqrt{\sum_{j \in M} \sigma_j^2 x_j^2}.$$

This robust objective measure corresponds to the ellipsoidal uncertainty sets proposed by Ben-Tal and Nemirovski [2, 3, 4] and El-Ghaoui et al. [11, 12].

(B)

$$r_\alpha^{(B)}(\boldsymbol{x}) = \boldsymbol{c'x} + g\left(\sum_{j\in M} \sigma_j^2 x_j^2\right),$$

where $g(w)$ is a piecewise linear concave function such that

$$g(w) \le \Omega_\alpha \sqrt{w} \le (1+\varepsilon) g(w), \qquad \forall w \in \left\{\sum_{j\in M} \sigma^2 x_j^2 \;:\; \boldsymbol{x} \in X\right\}$$

This objective measure is motivated from (A) from the perspective of computational tractability.

(C)

$$r_\alpha^{(C)}(\boldsymbol{x}) = \boldsymbol{c'x} + \max_{\substack{\boldsymbol{z}:\ \boldsymbol{0}\le\boldsymbol{z}\le\boldsymbol{e}\\ \boldsymbol{e'z}\le\Gamma_\alpha}} \sum_{j\in M} \sigma_j x_j z_j.$$

This objective measure is proposed in Bertsimas and Sim [8]. Bertsimas and Sim [8] show that with this robust objective measure, we can reduce the problem to solving at most $|M|+1$ nominal problems of different cost vectors. In other words, the robust counterpart is polynomially solvable if the nominal problem is polynomially solvable.

(D)

$$r_\alpha^{(D)}(\boldsymbol{x}) = \boldsymbol{c'x} + \max_{\substack{\boldsymbol{0}\le\boldsymbol{z}\le\boldsymbol{e}\\ \boldsymbol{e'z}\le\Omega_\alpha\sqrt{\boldsymbol{e'x}}}} \sum_{j\in M} \sigma_j x_j z_j.$$

This robust objective measure is similar to (C), except that the parameter Γ_α in (C) is now dependent on $\boldsymbol{e'x}$. We will show that this method is less conservative while retaining the computational complexity of the nominal problem.

We observe in feasible solutions with deterministic objective such that $x_j = 0$ for all $j \in M$, or $\boldsymbol{e'x} = 0$, all the robust objective measures take the value of $\boldsymbol{c'x}$. The next theorem establishes probability bounds that show that minimizing $r_\alpha(\boldsymbol{x})$ approximates minimizing $q_\alpha(\boldsymbol{x})$.

Theorem 2 *Let $\boldsymbol{x}$ be a feasible solution in $X \subseteq \{0,1\}^n$ such that the objective is uncertain, that is, $\boldsymbol{e'x} > 0$. Under the model of cost uncertainty C,*

(a)

$$\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \ge r_\alpha^{(A)}(\boldsymbol{x})\right) \le \exp\left(-\frac{7\Omega_\alpha^2}{16}\right)$$

(b)

$$\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \ge r_\alpha^{(B)}(\boldsymbol{x})\right) \le \exp\left(-\frac{7\Omega_\alpha^2}{16(1+\varepsilon)^2}\right)$$

(c)

$$\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(C)}(\boldsymbol{x})\right) \leq \exp\left(-\frac{7\Gamma_\alpha^2}{16m}\right)$$

(d)

$$\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(D)}(\boldsymbol{x})\right) \leq \exp\left(-\frac{7\Omega_\alpha^2}{16}\right)$$

Proof : Let $\tilde{s}_j = (\tilde{c}_j - c_j)/\sigma_j$ and $y_j = \sigma_j x_j$, $j \in M$. Observe that $\tilde{s}_j$ are independent random variables in $[-1, 1]$ and $E(\tilde{s}_j) = 0$.

(a) We have

$$\begin{aligned}
\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(A)}(\boldsymbol{x})\right) &= \mathrm{P}\left(\sum_{j\in M} \tilde{s}_j y_j \geq \Omega_\alpha \sqrt{\sum_{j\in M} y_j^2}\right) \\
&\leq \frac{\mathrm{E}\left(\exp\left(\theta \sum_{j\in M} \tilde{s}_j y_j\right)\right)}{\exp\left(\theta\Omega_\alpha\sqrt{\sum_{j\in M} y_j^2}\right)}, \qquad \text{[for some } \theta > 0] \qquad (5) \\
&= \frac{\prod_{j\in M} \mathrm{E}\left(\exp\left(\theta \tilde{s}_j y_j\right)\right)}{\exp\left(\theta\Omega_\alpha\sqrt{\sum_{j\in M} y_j^2}\right)}, \qquad (6)
\end{aligned}$$

where the inequality (5) follows from Markov inequality. We will now show that for any random variable, $\tilde{s} \in [-1, 1]$, $\mathrm{E}(\tilde{s}) = 0$ and $b > 0$,

$$\mathrm{E}\left(\exp(b\tilde{s})\right) \leq \exp\left(\frac{4b^2}{7}\right). \qquad (7)$$

Note that

$$\begin{aligned}
\mathrm{E}\left(\exp(b\tilde{s})\right) &= \mathrm{E}\left(1 + b\tilde{s} + \sum_{k=2} \frac{(b\tilde{s})^k}{k!}\right) \\
&\leq 1 + \sum_{k=2} \frac{b^k}{k!} \\
&= \exp(b) - b \\
&\leq \exp\left(\frac{4b^2}{7}\right),
\end{aligned}$$

where the last inequality is easily verified numerically. Therefore, continuing from inequality (6), we have

$$\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(A)}(\boldsymbol{x})\right) \leq \exp\left(\frac{4}{7}\theta^2 \sum_{j\in M} y_j^2 - \theta\Omega_\alpha\sqrt{\sum_{j\in M} y_j^2}\right).$$

Choosing $\theta\sqrt{\sum_{j\in M} y_j^2} = \frac{7}{8}\Omega_\alpha$, we derive the desired bound.

(b) We have

$$\begin{aligned}
\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(B)}(\boldsymbol{x})\right) &= \mathrm{P}\left(\sum_{j\in M} \tilde{s}_j y_j \geq g\left(\sum_{j\in M} y_j^2\right)\right) \\
&\leq \mathrm{P}\left(\sum_{j\in M} \tilde{s}_j y_j \geq \frac{\Omega_\alpha}{1+\varepsilon}\sqrt{\sum_{j\in M} y_j^2}\right) \\
&\leq \exp\left(-\frac{7\Omega_\alpha^2}{16(1+\varepsilon)^2}\right),
\end{aligned}$$

where the last inequality follows form part (a).

(c) From Proposition 2 of Bertsimas and Sim [8], we have

$$\mathrm{P}\left(\sum_{j\in M} \tilde{s}_j \sigma_j x_j \geq \max_{\substack{\boldsymbol{z}:\ \boldsymbol{0}\leq\boldsymbol{z}\leq\boldsymbol{e} \\ \boldsymbol{e}'\boldsymbol{z}\leq\Gamma_\alpha}} \sum_{j\in M} \sigma_j x_j z_j\right) \leq \mathrm{P}\left(\sum_{j\in M} \tilde{s}_j \gamma_j \geq \Gamma_\alpha\right),$$

where $0 \leq \gamma_j \leq 1$, $j \in M$. Therefore, we have

$$\begin{aligned}
\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(C)}(\boldsymbol{x})\right) &\leq \mathrm{P}\left(\sum_{j\in M} \tilde{s}_j \gamma_j \geq \Gamma\right) \\
&\leq \frac{\mathrm{E}\left(\exp\left(\theta \sum_{j\in M} \tilde{s}_j \gamma_j\right)\right)}{\exp(\theta\Gamma_\alpha)}, \qquad \text{[for some } \theta > 0\text{]} \\
&= \frac{\prod_{j\in M} \mathrm{E}\left(\exp(\theta \tilde{s}_j \gamma_j)\right)}{\exp(\theta\Gamma_\alpha)}, \\
&= \frac{\prod_{j\in M} \exp\left(\frac{4}{7}\theta^2\gamma_j^2\right)}{\exp(\theta\Gamma_\alpha)}, \\
&\leq \exp\left(\frac{4}{7}\theta^2|M| - \theta\Gamma_\alpha\right).
\end{aligned}$$

Choosing, $\theta = \frac{7\Gamma_\alpha}{8|M|}$, we have

$$\mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(C)}(\boldsymbol{x})\right) \leq \exp\left(-\frac{7\Gamma_\alpha^2}{16|M|}\right) = \exp\left(-\frac{7\Gamma_\alpha^2}{16m}\right).$$

(d) Let $M' = \{j \;:\; x_j = 1, j \in M\}$, $\Gamma = \Omega_\alpha \sqrt{\boldsymbol{e'x}}$ and observe that $|M'| = \boldsymbol{e'x}$ and

$$\begin{aligned}
& \mathrm{P}\left(\tilde{\boldsymbol{c}}'\boldsymbol{x} \geq r_\alpha^{(D)}(\boldsymbol{x})\right) \\
= \;& \mathrm{P}\left(\sum_{j \in M} \tilde{s}_j \sigma_j x_j \geq \max_{\substack{\boldsymbol{z}:\, \boldsymbol{0} \leq \boldsymbol{z} \leq \boldsymbol{e} \\ \boldsymbol{e}'\boldsymbol{z} \leq \Gamma}} \sum_{j \in M} \sigma_j x_j z_j\right) \\
= \;& \mathrm{P}\left(\sum_{j \in M'} \tilde{s}_j \sigma_j x_j \geq \max\left\{\sum_{j \in M'} \sigma_j x_j z_j \mid \boldsymbol{0} \leq \boldsymbol{z} \leq \boldsymbol{e}, \sum_{j \in M'} z_j \leq \Gamma\right\}\right) \\
\leq \;& \exp\left(-\frac{7\Gamma^2}{16|M'|}\right) \\
= \;& \exp\left(-\frac{7\Omega_\alpha^2}{16}\right),
\end{aligned}$$

where the second last inequality follows from part (c). ■

3 Algorithms for Solving Proposed Robust Models

Bertsimas and Sim [7] propose a tractable algorithm for solving the robust objective measure (C). In this section, we will focus on algorithms for solving the rest of the proposed robust models.

3.1 Solving Robust Model (A)

The robust objective measure (A) is nonlinear and hence there are difficulties in formulating it as a mixed-integer linear optimization problem that can be solved by off-the-shelf software.

Observe that if x_j is binary, then $x_j^2 = x_j$. Letting $d_j = \sigma_j^2$, we are thus naturally led to consider a more general problem

$$G^* = \min_{\boldsymbol{x} \in X} \boldsymbol{c'x} + f(\boldsymbol{d'x}) \tag{8}$$

with $f(\cdot)$ a concave function. In particular, with $f(w) = \Omega\sqrt{w}$, we have the robust model (A).

We first show that Problem (8) reduces to solving a finite number of nominal problems (1), with different costs vectors. Let $W = \{\boldsymbol{d'x} \mid \boldsymbol{x} \in \{0,1\}^n\}$ and $\eta(w)$ be a subgradient of the concave function $f(\cdot)$ evaluated at w, that is, $f(u) - f(w) \leq \eta(w)(u - w)\ \forall u \in R$. For the case of $f(w) = \Omega\sqrt{w}$, we have $f'(0) = \infty$ which corresponds to the set of feasible solutions, $\boldsymbol{x} \in X$ that yield deterministic objectives, that is, $\boldsymbol{e'x} = 0$. To avoid solving a problem with infinite cost components, we choose

$$\eta(w) = \begin{cases} f'(w) & \text{if } w \in W \backslash \{0\}, \\ \frac{f(d_{\min}) - f(0)}{d_{\min}} & \text{if } w = 0, \end{cases} \tag{9}$$

where $d_{\min} = \sigma_{\min}^2$.

Theorem 3 *Let*

$$Z(w) = \min_{\boldsymbol{x}\in X} (\boldsymbol{c} + \eta(w)\boldsymbol{d})'\boldsymbol{x} + f(w) - w\eta(w), \tag{10}$$

and $w^* = \arg\min_{w\in W} Z(w)$*. Then,* w^* *is an optimal solution to Problem (8) and* $G^* = Z(w^*)$*.*

Proof : We first show that $G^* \geq \min_{w\in W} Z(w)$. Let $\boldsymbol{x}^*$ be an optimal solution to Problem (8) and $w^* = \boldsymbol{d'x^*} \in W$. We have

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

Conversely, for any $w \in W$, let y_w be an optimal solution to Problem (10). We have

$$\begin{aligned} Z(w) &= (\boldsymbol{c} + \eta(w)\boldsymbol{d})'\boldsymbol{y_w} + f(w) - w\eta(w) \\ &= \boldsymbol{c'y_w} + f(\boldsymbol{d'y_w}) + \underbrace{\eta(w)(\boldsymbol{d'y_w} - w) - (f(\boldsymbol{d'y_w}) - f(w))}_{\geq 0} \\ &\geq \boldsymbol{c'y_w} + f(\boldsymbol{d'y_w}) \qquad (11) \\ &\geq \min_{\boldsymbol{x}\in X} \boldsymbol{c'x} + f(\boldsymbol{d'x}) = G^*, \end{aligned}$$

where inequality (11) for $w \in W\backslash\{0\}$ follows, since $\eta(w)$ is a subgradient. To see that inequality (11) 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 $\eta(0)(\boldsymbol{d'y_w} - 0) - (f(\boldsymbol{d'y_w}) - f(0)) \geq 0$. Therefore $G^* = \min_{w\in W} Z(w)$. ■

Note that when $d_j = \sigma^2$, then $W = \{0, \sigma^2, \ldots, m\sigma^2\}$, and thus $|W| = m + 1$, In this case, Problem (8) reduces to solving $m + 1$ nominal problems (10), that is, polynomial solvability is preserved.

Proposition 1 *If* $\sigma_j = \sigma$*, for* $j \in M$

$$\min_{\boldsymbol{x}\in X} r_\alpha^{(A)}(\boldsymbol{x}) = \min_{w=0,1,\ldots,m} Z(w),$$

where

$$Z(w) = \begin{cases} \min_{\boldsymbol{x}\in X} \quad \left(\boldsymbol{c} + \dfrac{\Omega_\alpha\sigma}{2\sqrt{w}}\boldsymbol{e}\right)'\boldsymbol{x} + \dfrac{\Omega_\alpha\sigma\sqrt{w}}{2} & \text{if } w = 1, \ldots, m, \\ \min_{\boldsymbol{x}\in X} \quad (\boldsymbol{c} + \Omega\sigma\boldsymbol{e})'\boldsymbol{x} & \text{if } w = 0. \end{cases} \tag{12}$$

Proof : With $f(w) = \Omega_\alpha \sigma \sqrt{w}$ and substituting $\eta(w) = f'(w) = \Omega_\alpha \sigma / (2\sqrt{w})$, $\forall\ w \in W \backslash \{0\}$ and $\eta(0) = (f(d_{\min}) - f(0))/d_{\min} = f(1) - f(0) = \Omega\sigma$ to Eq. (10) we obtain Eq. (12). ■

An immediate corollary of Theorem 3 is to consider a parametric approach as follows:

Corollary 1 *An optimal solution to Problem (8) coincides with one of the optimal solutions to the parametric problem:*

$$\begin{aligned} \text{minimize} \quad & (\boldsymbol{c} + \theta \boldsymbol{d})' \boldsymbol{x} \\ \text{subject to} \quad & \boldsymbol{x} \in X, \end{aligned} \tag{13}$$

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

This establishes a connection of Problem (8) with parametric discrete optimization (see Gusfield [15], Hassin and Tamir [18]). It turns out that if X is a matroid, the minimal set of optimal solutions to Problem (13) as θ varies is polynomial in size, see Eppstein [13] and Fern et. al. [14]. 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 $\binom{n}{2} + 1 = O(n^2)$ different orderings, the corresponding robust problem is also polynomially solvable.

For the case of shortest paths, Karp and Orlin [19] 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.

Unfortunately, in the general case, the size $|W|$ is potentially exponential with respect to n, hence, we will explore a method that only guarantees a local optimum.

3.1.1 A Frank-Wolfe Type Algorithm

Let $\boldsymbol{x}_\theta$ be an optimal solution of the parametric problem (13). We are interested in an approach that will yield a local optimal solution, $\boldsymbol{y}$ and a subset of parameters $\Theta \subseteq [\eta(\boldsymbol{e'd}), \eta(0)]$ such that

$$\boldsymbol{c'y} + f(\boldsymbol{d'y}) \leq \boldsymbol{c'x}_\theta + f(\boldsymbol{d'x}_\theta)$$

for all $\theta \in \Theta$. A natural method is to apply a Frank-Wolfe type algorithm, that is to successively linearize the function $f(\cdot)$.

Algorithm 1 *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 (8) and* Θ.

Algorithm:

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

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

3. *Output* $\boldsymbol{x_{k+1}}$ *and* $\Theta = [\min\{\theta, \eta(\boldsymbol{x_{k+1}})\}, \max\{\theta, \eta(\boldsymbol{x_{k+1}})\}]$

We next show that Algorithm 1 converges to a locally optimal solution.

Theorem 4 *Let* $\boldsymbol{x}$, $\boldsymbol{y}$ *and* $\boldsymbol{z}_\eta$ *be optimal solutions to the following problems:*

$$\boldsymbol{x} = \arg\min_{\boldsymbol{u}\in X}(\boldsymbol{c} + \theta\boldsymbol{d})'\boldsymbol{u}, \tag{14}$$

$$\boldsymbol{y} = \arg\min_{\boldsymbol{u}\in X}(\boldsymbol{c} + \eta(\boldsymbol{d'x})\boldsymbol{d})'\boldsymbol{u} \tag{15}$$

$$\boldsymbol{z}_\eta = \arg\min_{\boldsymbol{u}\in X}(\boldsymbol{c} + \eta\boldsymbol{d})'\boldsymbol{u}, \tag{16}$$

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

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

(b) *(Monotonicity) If* $\theta > \eta(\boldsymbol{d'x})$, *then* $\eta(\boldsymbol{d'x}) \geq \eta(\boldsymbol{d'y})$. *Likewise, if* $\theta < \eta(\boldsymbol{d'x})$, *then* $\eta(\boldsymbol{d'x}) \leq \eta(\boldsymbol{d'y})$. *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)*

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

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

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

and

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

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

Proof : (a) We have

$$\begin{aligned}
\boldsymbol{c'x} + f(\boldsymbol{d'x}) &= (\boldsymbol{c} + \eta(\boldsymbol{d'x})\boldsymbol{d})'\boldsymbol{x} - \eta(\boldsymbol{d'x})\boldsymbol{d'x} + f(\boldsymbol{d'x}) \\
&\geq \boldsymbol{c'y} + \eta(\boldsymbol{d'x})\boldsymbol{d'y} - \eta(\boldsymbol{d'x})\boldsymbol{d'x} + f(\boldsymbol{d'x}) \qquad \text{[follows from Eq. (15)]} \\
&= \boldsymbol{c'y} + f(\boldsymbol{d'y}) + \underbrace{\{\eta(\boldsymbol{d'x})(\boldsymbol{d'y} - \boldsymbol{d'x}) - (f(\boldsymbol{d'y}) - f(\boldsymbol{d'x})\}}_{\geq 0} \\
&\geq \boldsymbol{c'y} + f(\boldsymbol{d'y}),
\end{aligned}$$

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

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

$$\begin{aligned} \boldsymbol{c'y} + \eta(\boldsymbol{d'x})\boldsymbol{d'y} &\leq \boldsymbol{c'x} + \eta(\boldsymbol{d'x})\boldsymbol{d'x} && \text{[follows from Eq. (15)]} \\ -(\boldsymbol{c'y} + \theta\boldsymbol{d'y}) &\leq -(\boldsymbol{c'x} + \theta\boldsymbol{d'x}). && \text{[follows from Eq. (14)]} \end{aligned}$$

Adding the two inequalities we obtain

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

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

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

$$\begin{aligned} \boldsymbol{c'x} + \theta\boldsymbol{d'x} &\leq \boldsymbol{c'z}_\eta + \theta\boldsymbol{d'z}_\eta && \text{[follows from Eq. (14)]} \\ \boldsymbol{c'x} + \eta\boldsymbol{d'x} &\geq \boldsymbol{c'z}_\eta + \eta\boldsymbol{d'z}_\eta && \text{[follows from Eq. (16)]} \\ \boldsymbol{c'y} + \eta(\boldsymbol{d'x})\boldsymbol{d'y} &\leq \boldsymbol{c'z}_\eta + \eta(\boldsymbol{d'x})\boldsymbol{d'z}_\eta && \text{[follows from Eq. (15)]} \\ \boldsymbol{c'y} + \eta\boldsymbol{d'y} &\geq \boldsymbol{c'z}_\eta + \eta\boldsymbol{d'z}_\eta && \text{[follows from Eq. (16)]} \end{aligned}$$

From the first two inequalities we obtain

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

and from the last two we have

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

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

$$\begin{aligned} \boldsymbol{c'y} + f(\boldsymbol{d'y}) &= (\boldsymbol{c} + \eta(\boldsymbol{d'x})\boldsymbol{d})'\boldsymbol{y} - \eta(\boldsymbol{d'x})\boldsymbol{d'y} + f(\boldsymbol{d'y}) \\ &\leq (\boldsymbol{c} + \eta(\boldsymbol{d'x})\boldsymbol{d})'\boldsymbol{z}_\eta - \eta(\boldsymbol{d'x})\boldsymbol{d'y} + f(\boldsymbol{d'y}) \qquad \text{[follows from Eq. (15)]} \\ &= \boldsymbol{c'z}_\eta + f(\boldsymbol{d'z}_\eta) + \underbrace{\{f(\boldsymbol{d'y}) - f(\boldsymbol{d'z}_\eta) - \eta(\boldsymbol{d'x})(\boldsymbol{d'y} - \boldsymbol{d'z}_\eta)\}}_{=h(\boldsymbol{d'z}_\eta)} \\ &\leq \boldsymbol{c'z}_\eta + f(\boldsymbol{d'z}_\eta), \end{aligned} \qquad (17)$$

where inequality (17) follows from observing that the function $h(\alpha) = f(\boldsymbol{d'y}) - f(\alpha) - \eta(\boldsymbol{d'x})(\boldsymbol{d'y} - \alpha)$ is a convex function with $h(\boldsymbol{d'y}) = 0$ and $h(\boldsymbol{d'x}) \leq 0$. Since $\boldsymbol{d'z}_\eta$ is in the convex hull of $\boldsymbol{d'x}$ and $\boldsymbol{d'y}$, by convexity, $h(\boldsymbol{d'z}_\eta) \leq \mu h(\boldsymbol{d'y}) + (1-\mu)h(\boldsymbol{d'x}) \leq 0$, for some $\mu \in [0, 1]$. ■

Given a feasible solution, $\boldsymbol{x}$, Theorem 4(a) implies that we may improve the objective by solving a sequence of problems using Algorithm 1. Note that at each iteration, we are optimizing a linear function over X. Theorem 4(b) implies that the sequence of $\theta_k = \eta(\boldsymbol{d'x_k})$ is monotone and since it is bounded it converges. To avoid cycling at the local solutions, we set the termination condition to $\theta_k = \theta_{k+1}$. Since, X is finite, the monotone sequence θ_k will converge in a finite number of steps. Theorem 4(c) implies that at termination a locally optimal solution is found.

Suppose $\theta = \eta(\boldsymbol{e'd})$ and $\{\boldsymbol{x_1}, \ldots, \boldsymbol{x_k}\}$ be the sequence of solutions of Algorithm 1. From Theorem 4(b), we have

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

When Algorithm 1 terminates at the solution $\boldsymbol{x_k}$, then from Theorem 4(c),

$$\boldsymbol{c'x_k} + f(\boldsymbol{d'x_k}) \leq \boldsymbol{c'z}_\eta + f(\boldsymbol{d'z}_\eta), \tag{18}$$

where $\boldsymbol{z}_\eta$ is defined in Eq. (16) for all $\eta \in [\eta(\boldsymbol{e'd}), \eta(\boldsymbol{d'x_k})]$. Likewise, if we apply Algorithm 1 starting at $\bar{\theta} = \eta(0)$, and let $\{\boldsymbol{y_1}, \ldots, \boldsymbol{y_l}\}$ be the sequence of solutions of Algorithm 1, then we have

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

and

$$\boldsymbol{c'y_l} + f(\boldsymbol{d'y_l}) \leq \boldsymbol{c'z}_\eta + f(\boldsymbol{d'z}_\eta) \tag{19}$$

for all $\eta \in [\eta(\boldsymbol{d'y_l}), \eta(0)]$. If $\eta(\boldsymbol{d'x_k}) \geq \eta(\boldsymbol{d'y_l})$, we have $\eta(\boldsymbol{d'x_k}) \in [\eta(\boldsymbol{d'y_l}), \eta(0)]$ and $\eta(\boldsymbol{d'y_l}) \in [\eta(\boldsymbol{e'd}), \eta(\boldsymbol{d'x_k})]$. Hence, following from the inequalities (18) and (19), we conclude that

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

for all $\eta \in [\eta(\boldsymbol{e'd}), \eta(\boldsymbol{d'x_k})] \cup [\eta(\boldsymbol{d'y_l}), \eta(0)] = [\eta(\boldsymbol{e'd}), \eta(0)]$. Therefore, both $\boldsymbol{y_l}$ and $\boldsymbol{x_k}$ are globally optimal solutions. However, if $\eta(\boldsymbol{d'y_l}) > \eta(\boldsymbol{d'x_k})$, we are assured that the global optimal solution is $\boldsymbol{x_k}$, $\boldsymbol{y_l}$ or in $\{\boldsymbol{x} : \boldsymbol{x} = \arg\min_{\boldsymbol{u} \in X} (\boldsymbol{c} + \eta \boldsymbol{d})'\boldsymbol{u},\ \eta \in (\eta(\boldsymbol{d'x_k}), \eta(\boldsymbol{d'y_l}))\}$. Hence, we can systematically solve for better solutions by partitioning the interval $[\eta(\overline{w}), \eta(\underline{w})]$, with $\underline{w} = \boldsymbol{d'y_l}$, $\overline{w} = \boldsymbol{d'x_k}$ 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 1 in the intervals. Repeating this way until all intervals are covered, we can obtain the global solution. Unfortunately, the proposed method is not necessarily polynomial.

3.2 Solving Robust Model (B)

We can represent the function, $g(\cdot)$ in the robust objective measure (B) as follows:

$$g(w) = \min\{a_1 w + b_1, \ldots, a_k w + b_k\} \tag{20}$$

with k number linear pieces such that

$$g(w) \le \Omega_\alpha \sqrt{w} \le (1+\varepsilon) g(w) \qquad \forall w \in \{\boldsymbol{d}'\boldsymbol{x} \ : \ \boldsymbol{x} \in X\}. \tag{21}$$

Proposition 2

$$\min_{\boldsymbol{x}\in X} \boldsymbol{c}'\boldsymbol{x} + g(\boldsymbol{d}'\boldsymbol{x}) = \min\left\{ \min_{\boldsymbol{x}\in X} (\boldsymbol{c} + a_1\boldsymbol{d})'\boldsymbol{x} + b_1, \ldots, \min_{\boldsymbol{x}\in X} (\boldsymbol{c} + a_k\boldsymbol{d})'\boldsymbol{x} + b_k \right\}. \tag{22}$$

Proof : From Eq. (20) the proof is immediate. ■

Clearly, the computational complexity depends on the number of linear segments in the function, $g(\cdot)$, which is dependent on the parameter ϵ.

Theorem 5 *The inequality (21) is satisfied for the function, $g(w)$ with parameter,*

$$\left.\begin{aligned} a_i &= \frac{\Omega_\alpha\sqrt{y_i} - \Omega_\alpha\sqrt{y_{i-1}}}{y_i - y_{i-1}} \\ b_i &= \Omega_\alpha\sqrt{y_{i-1}} - y_{i-1}a_i \end{aligned}\right\} \quad \forall i = 1, \ldots, k,$$

where

$$\begin{aligned} y_0 &= 0 \\ y_i &= d_{\min}\zeta^{i-1} \qquad i = 1, \ldots, k, \end{aligned}$$

where $\zeta = \left((1+\varepsilon) + \sqrt{(1+\varepsilon)^2 - 1}\right)^4$ *and*

$$k = \left\lceil \frac{\ln\left(\frac{m d_{\max}}{d_{\min}}\right)}{\ln(\zeta)} \right\rceil + 1.$$

Proof : Observe that the parameters a_i and b_i describe respectively the gradient and vertical axis intercept of the line segment joining the points $(y_{i-1}, \Omega_\alpha\sqrt{y_{i-1}})$ and $(y_i, \Omega_\alpha\sqrt{y_i})$. Since $g(0) = 0$ (from $b_0 = 0$) and the set, $\{\boldsymbol{d}'\boldsymbol{x} \ : \ \boldsymbol{x} \in X\} \cap (0, d_{\min})$ is empty, the inequality (21) is satisfied for $w \in \{\boldsymbol{d}'\boldsymbol{x} \ : \ \boldsymbol{x} \in X\} \cap [0, d_{\min})$. We will hence focus on the domain for which $w \ge d_{min}$.

Observe that for all $\boldsymbol{x} \in X$, we have $y_k \ge m d_{\max} \ge \boldsymbol{d}'\boldsymbol{x}$. Hence, for every feasible solution $\boldsymbol{x}$ such that $w = \boldsymbol{d}'\boldsymbol{x} > 0$, we can find a line segment such that $w \in [y_{i-1}, y_i]$, $i \in \{2, \ldots, k\}$. Noting that

$y_i/y_{i-1} = \zeta$, for all $i \in \{2, \ldots, k\}$, we have

$$
\begin{aligned}
\frac{\Omega_\alpha \sqrt{w}}{g(w)} &= \frac{\Omega_\alpha \sqrt{w}}{a_i w + b_i} \\
&= \frac{\Omega_\alpha \sqrt{w}}{\left(\frac{\Omega_\alpha \sqrt{y_i} - \Omega_\alpha \sqrt{y_{i-1}}}{y_i - y_{i-1}}\right)(w - y_{i-1}) + \Omega_\alpha \sqrt{y_{i-1}}} \\
&= \frac{\sqrt{\frac{w}{y_{i-1}}}}{\left(\frac{\sqrt{\frac{y_i}{y_{i-1}}}-1}{\frac{y_i}{y_{i-1}}-1}\right)\left(\frac{w}{y_{i-1}} - 1\right) + 1} \\
&= \frac{1}{\left(\frac{1}{\sqrt{\zeta}-1}\right)\xi + \frac{1}{\xi}\left(1 - \left(\frac{1}{\sqrt{\zeta}-1}\right)\right)} && \text{[where } \xi = \sqrt{w/y_{i-1}}\text{]} \\
&\leq \frac{1}{2\sqrt{\left(\frac{1}{\sqrt{\zeta}-1}\right)\left(1 - \left(\frac{1}{\sqrt{\zeta}-1}\right)\right)}} && \text{[choosing optimal } \xi\text{]} \\
&= \frac{\zeta^{\frac{1}{2}} + 1}{2\zeta^{\frac{1}{4}}} \\
&= \frac{\left((1+\varepsilon) + \sqrt{(1+\varepsilon)^2 - 1}\right)^2 + 1}{2\left((1+\varepsilon) + \sqrt{(1+\varepsilon)^2 - 1}\right)} \\
&= 1 + \varepsilon.
\end{aligned}
$$

■

The practicality of the approach depends on the number of line segments, k, which is proportional to $1/\ln(\zeta) = O(\sqrt{1/\varepsilon})$ and hence, critically dependent on the tolerance level, ϵ. In Table 3.2, we show how the choice of ϵ affects the magnitude of $1/\ln(\zeta)$.

ε	$1/\ln(\zeta)$
0.1	0.564
0.01	1.769
0.001	5.591
0.0001	17.68

For a fixed approximation level, ε, the robust model of (B) can be solved in $O(\ln(md_{\max}) - \ln(d_{\min}))$ nominal problems. This is certainty an attractive consideration to solving the robust model (A) for which the complexity is unknown for many polynomially solvable discrete problems.

3.3 Solving Robust Model (D)

We consider solving a more general problem,

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c'x} + \max_{\substack{\boldsymbol{z}:\ \boldsymbol{0}\le\boldsymbol{z}\le\boldsymbol{e} \\ \boldsymbol{e'z}\le f(\boldsymbol{e'x})}} \sum_{j\in M} \sigma_j x_j \\ \text{subject to} \quad & \boldsymbol{x} \in X, \end{aligned}$$

where $f(w)$ being a concave function with subgradient, $\eta(w)$. The robust objective measure (D) corresponds to $f(w) = \Omega_\alpha\sqrt{w}$ and $\eta(w)$ to Eq. (9). This robust objective measure is similar to robust objective measure (C) except that the parameter Γ_α in (C) is now dependent on $\boldsymbol{e'x}$. We will show that this method retains the computational complexity of the nominal problem.

For notational convenience, we let $S_l = \{1, \ldots, l\}$ and define $\sigma_0 = 0$ and $S_0 = M$.

Theorem 6 *Problem (3.3) satisfies* $Z^* = \min_{(l,k):l,k\in M\cup\{0\}} Z_{lk}$, *where*

$$Z_{lk} = \min_{\boldsymbol{x}\in X} \boldsymbol{c'x} + \sum_{j\in S_l} (\sigma_j - \sigma_l)x_j + \eta(k)\sigma_l\boldsymbol{e'x} + \sigma_l(f(k) - k\eta(k)). \tag{23}$$

Proof : By strong duality of the inner maximization function with respect to $\boldsymbol{z}$, Problem (3.3) is equivalent to solving the following problem:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c'x} + \sum_{j\in M} p_j + f(\boldsymbol{e'x})\theta \\ \text{subject to} \quad & p_j \ge \sigma_j x_j - \theta \qquad \forall j \in M \\ & p_j \ge 0 \qquad \forall j \in M \\ & x \in X \\ & \theta \ge 0, \end{aligned} \tag{24}$$

We eliminate the variables p_j and express Problem (24) as follows:

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c'x} + \sum_{j\in M} \max\{\sigma_j x_j - \theta, 0\} + f(\boldsymbol{e'x})\theta \\ \text{subject to} \quad & x \in X \\ & \theta \ge 0. \end{aligned} \tag{25}$$

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

$$\max\{\sigma_j x_j - \theta, 0\} = \begin{cases} \sigma_j - \theta & \text{if } x_j = 1 \text{ and } \sigma_j \ge \theta \\ 0 & \text{if } x_j = 0 \text{ or } \sigma_j < \theta. \end{cases} \tag{26}$$

Recalling that $\sigma_1 \geq \ldots \geq \sigma_m$, by restricting θ into separate intervals, we obtain that

$$Z^* = \min_{\theta \geq 0} \min_{l=0,\ldots,m} Z_l(\theta)$$

where $Z_l(\theta)$, $l = 1, \ldots, m$, is defined for $\theta \in [\sigma_l, \sigma_{l+1}]$ is

$$Z_l(\theta) = \min_{\boldsymbol{x} \in X} \boldsymbol{c'x} + \sum_{j \in S_l} (\sigma_j - \theta) x_j + f(\boldsymbol{e'x})\theta \tag{27}$$

and for $\theta \in [\sigma_1, \infty)$:

$$Z_0(\theta) = \min_{\boldsymbol{x} \in X} \boldsymbol{c'x} + f(\boldsymbol{e'x})\theta. \tag{28}$$

For each interval, $\theta \in [\sigma_l, \sigma_{l+1}]$, the optimal solution of the function $Z_l(\theta)$ is achieved at either σ_l or σ_{l+1}. Hence, we can restrict θ to a finite set $\{\sigma_1, \ldots, \sigma_m, 0\}$ and establish that

$$Z^* = \min_{l \in M \cup \{0\}} \boldsymbol{c'x} + \sum_{j \in S_l} (\sigma_j - \sigma_l) x_j + f(\boldsymbol{e'x})\sigma_l. \tag{29}$$

Since $\boldsymbol{e'x} \in \{0, 1, \ldots, m\}$, we apply Theorem 3 to obtain the subproblem decomposition of (23). ■

Theorem 6 suggests that the robust problem remains polynomially solvable if the nominal problem is polynomially solvable, but at the expense of higher computational complexity. Analogously to Theorem 4, we provide a necessary condition for optimality the one can exploit in a local search algorithm.

Theorem 7 *An optimal solution $\boldsymbol{x}$ to Problem (3.3) is also an optimal solution to the following problem:*

$$\begin{aligned} \text{minimize} \quad & \boldsymbol{c'y} + \sum_{j \in S_{l^*}} (\sigma_j - \sigma_{l^*}) y_j + \eta(\boldsymbol{e'x})\sigma_{l^*} \boldsymbol{e'y} \\ \text{subject to} \quad & \boldsymbol{y} \in X, \end{aligned} \tag{30}$$

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

Proof : Suppose $\boldsymbol{x}$ is an optimal solution for Problem (3.3) but not for Problem (30). Let $\boldsymbol{y}$ be the optimal solution to Problem (30). Therefore,

$$\begin{aligned} & \boldsymbol{c'x} + \max_{\substack{\boldsymbol{z}:\ \boldsymbol{0} \leq \boldsymbol{z} \leq \boldsymbol{e} \\ \boldsymbol{e'z} \leq f(\boldsymbol{e'x})}} \left\{ \sum_{j \in M} \sigma_j x_j z_j \right\} \\ = \quad & \min_{l \in M \cup \{0\}} \boldsymbol{c'x} + \sum_{j \in S_l} (\sigma_j - \sigma_l) x_j + f(\boldsymbol{e'x})\sigma_l \\ = \quad & \boldsymbol{c'x} + \sum_{j \in S_{l^*}} (\sigma_j - \sigma_{l^*}) x_j + f(\boldsymbol{e'x})\sigma_{l^*} \\ = \quad & \boldsymbol{c'x} + \sum_{j \in S_{l^*}} (\sigma_j - \sigma_{l^*}) x_j + \eta(\boldsymbol{e'x})\sigma_{l^*}\boldsymbol{e'x} - \eta(\boldsymbol{e'x})\sigma_{l^*}\boldsymbol{e'x} + f(\boldsymbol{e'x})\sigma_{l^*} \end{aligned} \tag{31}$$

$$
\begin{aligned}
&> \boldsymbol{c'y} + \sum_{j \in S_{l^*}} (\sigma_j - \sigma_{l^*}) y_j + \eta(\boldsymbol{e'x})\sigma_{l^*}\boldsymbol{e'y} - \eta(\boldsymbol{e'x})\sigma_{l^*}\boldsymbol{e'x} + f(\boldsymbol{e'x})\sigma_{l^*} \\
&= \boldsymbol{c'y} + \sum_{j \in S_{l^*}} (\sigma_j - \sigma_{l^*}) y_j + f(\boldsymbol{e'y})\sigma_{l^*} + \underbrace{(\eta(\boldsymbol{e'x})(\boldsymbol{e'y} - \boldsymbol{e'x}) - (f(\boldsymbol{e'y}) - f(\boldsymbol{e'x})))\,\sigma_{l^*}}_{\geq 0} \\
&\geq \boldsymbol{c'y} + \sum_{j \in S_{l^*}} (\sigma_j - \sigma_{l^*}) y_j + f(\boldsymbol{e'y})\sigma_{l^*} \\
&\geq \min_{l \in M \cup \{0\}} \boldsymbol{c'y} + \sum_{j \in S_l} (\sigma_j - \sigma_l) y_j + f(\boldsymbol{e'y})\sigma_l \\
&= \boldsymbol{c'y} + \max_{\substack{\boldsymbol{z}:\ \boldsymbol{0} \leq \boldsymbol{z} \leq \boldsymbol{e} \\ \boldsymbol{e'z} \leq f(\boldsymbol{e'y})}} \left\{ \sum_{j \in M} \sigma_j y_j z_j \right\} \qquad (32)
\end{aligned}
$$

where the Eqs. (31) and (32) follows from Eq. (29). This contradicts that $\boldsymbol{x}$ is optimal. ■

4 On the Conservatism of Robust Solutions

In this section, we will provide some insights on the degree of conservatism among the robust approaches, (A) through (D). Since we do not know the actual cost distributions, it is impossible to compare the robust objective measures with the exact quantile $q_\alpha(\boldsymbol{x})$. Nevertheless, a reasonable approach is to compare the deviations of the robust objective measure from the mean,

$$
\kappa_\alpha(\boldsymbol{x}) \equiv r_\alpha(\boldsymbol{x}) - \boldsymbol{c'x}
$$

among different robust measures that yield the same probability bound α. We denote $\kappa^{(A)}(\boldsymbol{x})$ through $\kappa^{(D)}(\boldsymbol{x})$ as the deviation measure respectively for the robust models (A) through (D). For the worst case model, we define $\kappa^{(E)}(\boldsymbol{x}) = \boldsymbol{\sigma' x}$. From the fact that $\|\boldsymbol{a}\|_2 \leq \|\boldsymbol{a}\|_1 \leq \sqrt{m}\|\boldsymbol{a}\|_2$, for all $\boldsymbol{a} \in \Re^m$, we can easily establish the following tight inequalities:

$$
\frac{1}{\Omega_\alpha} \leq \frac{\kappa^{(E)}(\boldsymbol{x})}{\kappa_\alpha^{(A)}(\boldsymbol{x})} \leq \frac{\sqrt{m}}{\Omega_\alpha} \qquad (33)
$$

$$
\frac{1}{\Omega_\alpha} \leq \frac{\kappa^{(E)}(\boldsymbol{x})}{\kappa_\alpha^{(B)}(\boldsymbol{x})} \leq \frac{(1+\varepsilon)\sqrt{m}}{\Omega_\alpha} \qquad (34)
$$

$$
1 \leq \frac{\kappa^{(E)}(\boldsymbol{x})}{\kappa_\alpha^{(C)}(\boldsymbol{x})} \leq \frac{m}{\Gamma_\alpha} \qquad (35)
$$

$$
1 \leq \frac{\kappa^{(E)}(\boldsymbol{x})}{\kappa_\alpha^{(D)}(\boldsymbol{x})} \leq \frac{\sqrt{m}}{\Omega_\alpha} \qquad (36)
$$

for all $\kappa^{(E)}(\boldsymbol{x}) > 0$

Observe that in all the proposed robust measures, the probability bounds decreases exponentially with respect to Ω_α^2 (for models (A), (B) and (D)) and Γ_α^2/m (for model (C)). For a reasonable range of α from 20% to 0.01%, Ω_α and $\Gamma_\alpha/\sqrt{m}$ lies in the range $[1.9, 5.1]$. From the inequalities (33) through (36), we observe that all the proposed robust models have the potential of being far less conservative compared to the absolute worst case model. In fact, $\kappa^{(E)}(\boldsymbol{x})$ could be as large as $O(\sqrt{m})$ of $\kappa_\alpha^{(A)}(\boldsymbol{x})$, $\kappa_\alpha^{(B)}(\boldsymbol{x})$ (for say $\varepsilon = 0.01$), $\kappa_\alpha^{(C)}(\boldsymbol{x})$ or $\kappa_\alpha^{(D)}(\boldsymbol{x})$. The robust models of (C) and (D) are always less conservative than the worst case model, which is not necessarily true for the robust models of (A) and (B). In some situations, $\kappa_\alpha^{(A)}(\boldsymbol{x})$ and $\kappa_\alpha^{(B)}(\boldsymbol{x})$ can be as large as Ω_α of the worst case deviation, which is, however, practically a smaller deviation factor compared to the converse.

The measure with the lowest deviation, $\kappa_\alpha(\boldsymbol{x})$ for the same probability bound α is generally the least conservative. We benchmark the worse case deviation against $\kappa_\alpha^{(A)}(\boldsymbol{x})$ and $\kappa_\alpha^{(B)}(\boldsymbol{x})$ as follows:

$$1 \leq \frac{\kappa_\alpha^{(B)}(\boldsymbol{x})}{\kappa_\alpha^{(A)}(\boldsymbol{x})} \leq 1 + \varepsilon, \tag{37}$$

$$\min\left\{\frac{1}{\Omega_\alpha}, 1\right\} \leq \frac{\kappa_\alpha^{(C)}(\boldsymbol{x})}{\kappa_\alpha^{(A)}(\boldsymbol{x})} \leq \frac{1}{\Omega_\alpha}\sqrt{\lfloor \Omega_\alpha\sqrt{m}\rfloor + (\Omega_\alpha\sqrt{m} - \lfloor \Omega_\alpha\sqrt{m}\rfloor)^2}, \tag{38}$$

$$\min\left\{\frac{1}{\Omega_\alpha}, 1\right\} \leq \frac{\kappa_\alpha^{(C)}(\boldsymbol{x})}{\kappa_\alpha^{(A)}(\boldsymbol{x})} \leq \frac{1}{\Omega_\alpha}\sqrt{\lfloor \Omega_\alpha\sqrt{\boldsymbol{e}'\boldsymbol{x}}\rfloor + (\Omega_\alpha\sqrt{\boldsymbol{e}'\boldsymbol{x}} - \lfloor \Omega_\alpha\sqrt{\boldsymbol{e}'\boldsymbol{x}}\rfloor)^2}, \tag{39}$$

where the inequalities (38) and (39) follow from Proposition 5 of Bertsimas et al [6]. The inequality (38) is derived from substituting $\Gamma_\alpha = \Omega_\alpha\sqrt{m}$, which gives the desired quantile guarantee.

Therefore, for the same quantile guarantee, the value of $\kappa_\alpha^{(A)}(\boldsymbol{x})$ is at most Ω_α times of $\kappa_\alpha^{(C)}(\boldsymbol{x})$ and $\kappa_\alpha^{(D)}(\boldsymbol{x})$. However, the magnitude of $\kappa^{(C)}(\boldsymbol{x})$ (respectively, $\kappa^{(D)}(\boldsymbol{x})$) could be as large as $O(m^{1/4})$ (respectively, $O((\boldsymbol{e}'\boldsymbol{x})^{1/4})$) the magnitude of $\kappa_\alpha^{(A)}(\boldsymbol{x})$, suggesting that the robust method under the ellipsoidal uncertainty set has the potential of being least conservative. The robust method of (B) is an attractive approximation of (A). From the inequalities (37) through (39), we easily derive that

$$\frac{1}{1+\varepsilon}\min\left\{\frac{1}{\Omega_\alpha}, 1\right\} \leq \frac{\kappa_\alpha^{(C)}(\boldsymbol{x})}{\kappa_\alpha^{(B)}(\boldsymbol{x})} \leq \frac{1}{\Omega_\alpha}\sqrt{\lfloor \Omega_\alpha\sqrt{m}\rfloor + (\Omega_\alpha\sqrt{m} - \lfloor \Omega_\alpha\sqrt{m}\rfloor)^2} \tag{40}$$

$$\frac{1}{1+\varepsilon}\min\left\{\frac{1}{\Omega_\alpha}, 1\right\} \leq \frac{\kappa_\alpha^{(C)}(\boldsymbol{x})}{\kappa_\alpha^{(B)}(\boldsymbol{x})} \leq \frac{1}{\Omega_\alpha}\sqrt{\lfloor \Omega_\alpha\sqrt{\boldsymbol{e}'\boldsymbol{x}}\rfloor + (\Omega_\alpha\sqrt{\boldsymbol{e}'\boldsymbol{x}} - \lfloor \Omega_\alpha\sqrt{\boldsymbol{e}'\boldsymbol{x}}\rfloor)^2}. \tag{41}$$

Similar to the robust method of (A), the value of $\kappa_\alpha^{(B)}(\boldsymbol{x})$ is at most $\Omega_\alpha(1+\varepsilon)$ times of $\kappa_\alpha^{(C)}(\boldsymbol{x})$ and $\kappa_\alpha^{(D)}(\boldsymbol{x})$. However, the magnitude of $\kappa^{(C)}(\boldsymbol{x})$ (respectively, $\kappa^{(D)}(\boldsymbol{x})$) could be as large as $O(m^{1/4})$ (respectively, $O((\boldsymbol{e}'\boldsymbol{x})^{1/4})$) the magnitude of $\kappa_\alpha^{(B)}(\boldsymbol{x})$, which is far more conservative.

Finally, we compare the level of conservatism between the robust methods of (C) and (D). For the same probability bound, we have

$$1 \leq \frac{\kappa_\alpha^{(C)}(\boldsymbol{x})}{\kappa_\alpha^{(D)}(\boldsymbol{x})} \leq \frac{\min\{\boldsymbol{e}'\boldsymbol{x}, \Omega_\alpha\sqrt{m}\}}{\Omega_\alpha\sqrt{\boldsymbol{e}'\boldsymbol{x}}} \leq \frac{m^{\frac{1}{4}}}{\sqrt{\Omega_\alpha}},$$

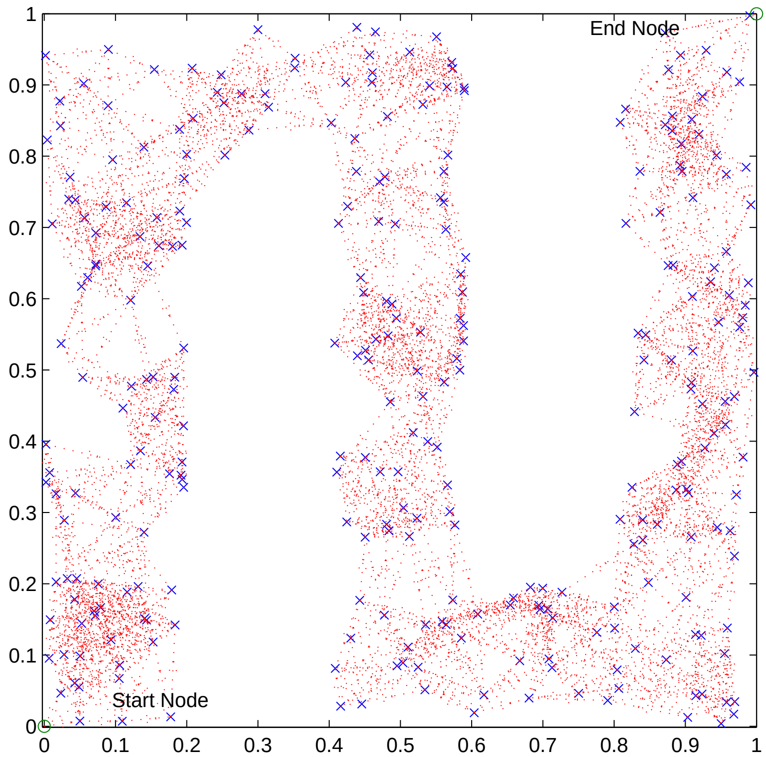

Figure 1: Randomly Generated Directed Graph.

for all $\kappa_\alpha^{(D)}(\boldsymbol{x}) > 0$. This indicates that the robust method of (D) can be far less conservative than (C).

5 Experimental Results

In this section, we solve the shortest path problem for which the arc costs are subject to uncertainty. The randomly generated directed graph (see Figure 1), $G = (V, E)$ has $|V| = 300$ nodes, and $|E| = 1475$ arcs. For each arc, $e \in E$, the average arc cost, c_e, $e \in E$ is randomly generated in $[0, 10]$, and the maximum arc cost deviation, σ_e is randomly generated in $[0, 8]$.

Using Dijkstra's algorithm [10], the shortest path problem can be solved in $O(|N|^2)$. We solve and compare the robust solutions using models (A) through (D). For solving robust model (A), we use Algorithm 1 to obtain locally optimal solutions. Starting with $\theta = \eta(\sum_{e\in E}\sigma_e^2)$ and $\theta = \eta(\sigma_{\min}^2)$, we obtain two locally optimal solutions, which surprisingly in all the experiments, were found in less than 5 iterations. Furthermore, in all the cases, the locally optimal solutions derived from both extreme values of θ are identical and hence, the locally optimal solutions are indeed globally optimal. For robust model (B), we set $\varepsilon = 0.1$ and it turns out that the optimal solutions are the same as the optimal solution of robust method (A). Therefore, in this experiment, the empirical results of robust model (A) apply to that of robust model (B).

Since $|E| = 1475$, in the robust model of (C), for $\alpha \leq 0.1$, the parameter Γ_α is at least 88. However, the average number of arcs in the robust shortest paths is generally less than 50. Hence, for all the range of α, the robust model of (C) yields the worst case solution for which each arc cost is at the maximum. Therefore, under the robust quantile framework, the robust model (C) is as conservative as the worst case model, which is uninteresting.

Therefore, in this experiment, we report and compare the findings for robust model (A), and (D). We also compare the robust solutions with the average case model (all costs are equal to their expected value) and with the worst case model (all costs are at their maximum value).

Using simulation, we estimate the distribution of the path cost perturbing costs of the arcs. We assume that every arc $e \in E$ has random cost component, $\tilde{c}_e$ that is independently perturbed, with probability ρ, from the lowest value $\underline{c}_e$ to $c_e + d_e$. The value $\underline{c}_e$ is chosen so that the mean arc cost remains at c_e. We let $\tilde{z}$ as the uncertain path cost, evaluated at the solution, $\boldsymbol{x}$, that is, $\tilde{z} = \tilde{\boldsymbol{c}}'\boldsymbol{x}$. From simulation, we can estimate the β quantile as follows:

$$\hat{q}_\beta = \tilde{z}_{(\lfloor T(1-\beta) \rfloor)},$$

where T is the number of simulations, and $\tilde{z}_{(i)}$ is the ith ordered statistic of $\tilde{z}$ such that $\tilde{z}_{(1)} \leq \tilde{z}_{(2)} \leq \ldots \leq \tilde{z}_{(T)}$. To avoid confusion, we use β to refer to the estimated β quantile of the path cost and α to refer to the quantile parameter corresponding to the inputs of the robust models.

Setting $\rho = 0.4$, we generate $1,000,000$ random scenarios and plot the distributions of the path cost for the robust solutions obtain using the robust models with parameters $\alpha = \{0.1, 0.01, 0.001, 0.0001\}$.

In Figures 2, and 3, we present the simulated cost distributions of the robust solutions generated by models (A) and (D).

In both robust models (A) and (D), we observe that as the parameter α of the robust models decreases, there is a general trend of decreasing the variance of the cost at the expense of increasing the expected cost. The main motivation of having robustness is to guarantee for most instances, a lower path cost. We illustrate this phenomenon in Figures 4 and 5 in which we plot the estimated β quantile of the path cost obtained using robust models (A) and (D), respectively. Within the plots, we draw vertical lines at $\beta = 0.001$ showing that the robust shortest paths can guarantee a lower cost for 99.9% of the instances compared to the solution resulting from minimizing the average path cost.

In Table 1, we compare the β quantile path cost among the solutions of robust model (A), robust (D) (both with parameters $\alpha = \beta$), average case model and the worst case model. We observe in the experiment that robust model (A) with parameter α generates a solution with the lowest α quantile value compared to the average, worst case and robust model (D). The performance of robust model (D) is less impressive compared to robust model (A). When β decreases to 0.001, robust model (D) yields the same solution as the worst case model. On the other hand, the robust model (A) is capable of generating solutions with better quantile values compared to the worst case model.

We observe a similar trend as we vary $\rho \in [0.1, 0.4]$. These experiments suggest that models (A) and (B) are potentially less conservative than models (C) and (D). This behavior is in agreement with our theoretical findings.

	$\beta = 0.1$	$\beta = 0.01$	$\beta = 0.001$	$\beta = 0.0001$
Robust Model (A)	261.1764	280.8574	295.2183	307.1875
Robust Model (D)	262.4384	293.044	303.5045	309.3071
Average Case Model	264.894	291.2963	309.4891	323.8137
Worst Case Model	281.4117	295.2095	303.5045	309.3071

Table 1: Comparison of the estimated β quantile for different robust solutions.

6 Conclusions

We show in this paper that the proposed robust models can change the distributions of the objective function to varying degrees of conservatism. In particular, the robust model under ellipsoidal uncertainty (robust model (A)) is potentially the least conservative both from a theoretical and possibly an empirical perspective. Based on this robust model, we believe the proposed Frank-Wolfe algorithm is useful in practice for iteratively improving the robustness of a solution.

References

[1] Averbakh, I. (2001): On the complexity of a class of combinatorial optimization problems with uncertainty, *Math. Prog. 90*, 263-272.

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

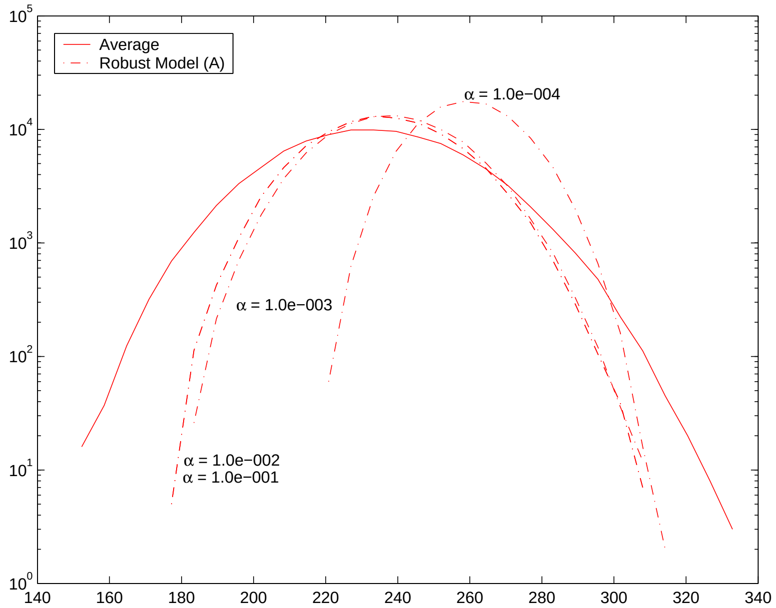

Figure 2: Simulated cost distributions for robust model (A).

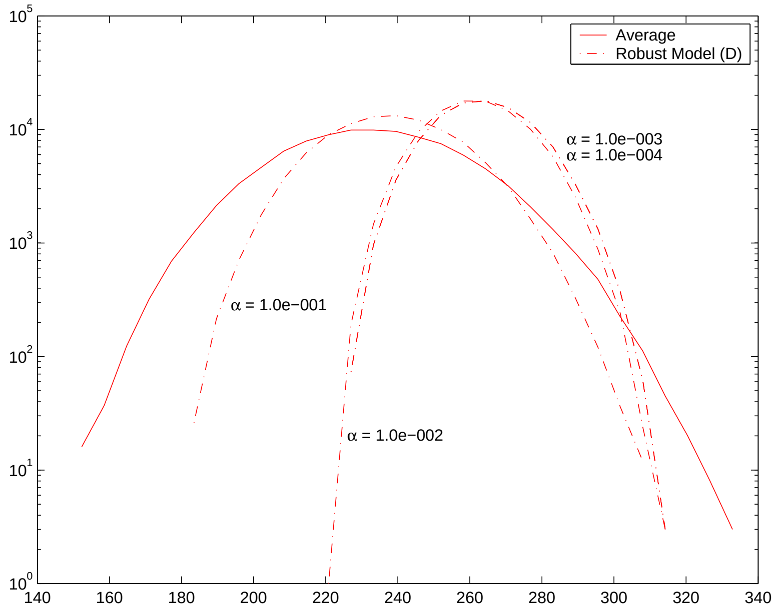

Figure 3: Simulated cost distributions for robust model (D).

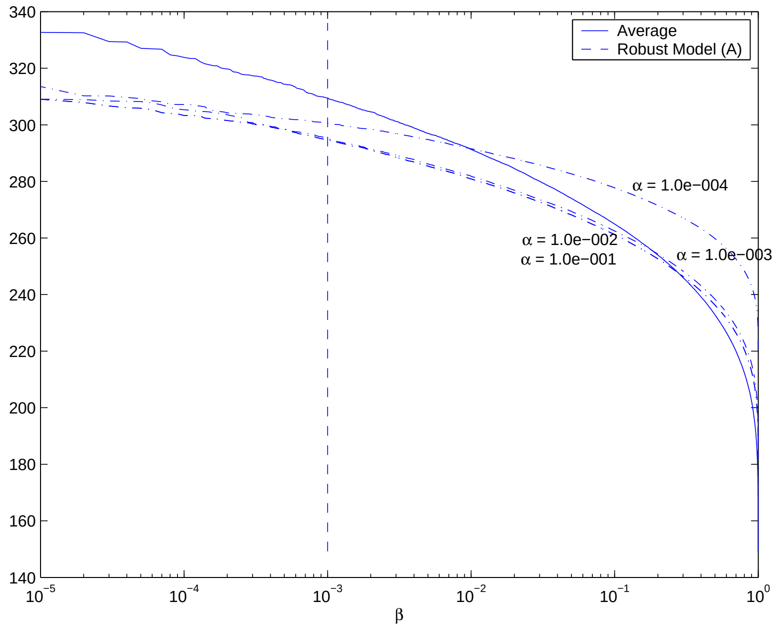

Figure 4: Plot of $\hat{q}_\beta$ of the path costs obtained using robust model (A).

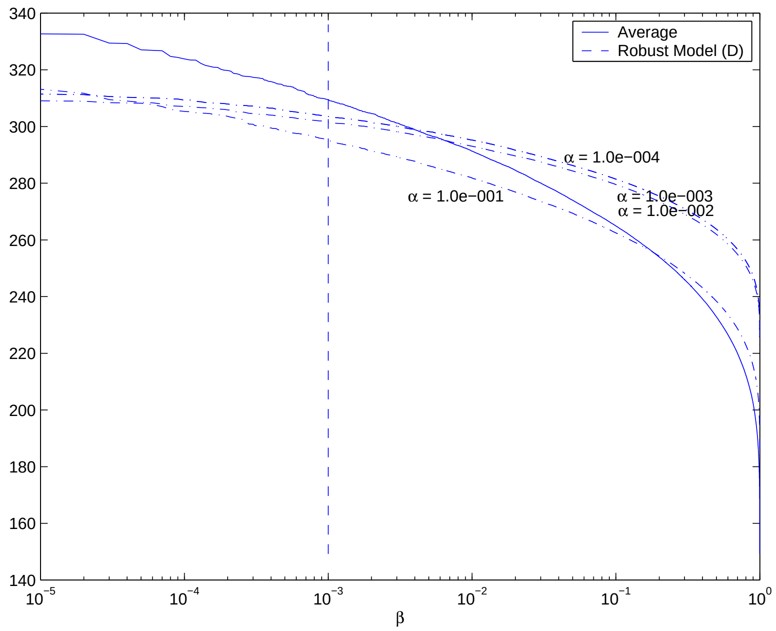

Figure 5: Plot of $\hat{q}_\beta$ of the path costs obtained using robust model (D).

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

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

[5] 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.

[6] Bertsimas, D., Pachamanova, D., Sim, M. (2003): Robust Linear Optimization under General Norms, *Oper. Res. Let.* 32, 510–516, 2004.

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

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

[9] Bertsimas, D., M. Sim (2004): Robust Conic Optimization, *Working Paper*, MIT.

[10] Dijkstra, E. W. (1959): A note on two problems in connection with graphs. *Numerische Mathematik*, 269-271.

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

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

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

[14] 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.

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

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

[17] Jorion, P. (1997): Value at Risk, McGraw-Hill, New York, NY.

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

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

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

[21] Levy, H. (1992): Stochastic Dominance and Expected Utility: Survey and Analysis, *Management Science* 38, 555-593.

[22] Levy, H. Kroll, Y. (1978): Ordering Uncertain Options with Borrowing and Lending, *The Journal of Finance*, 31.

[23] R. Schultz, L. Stougie and M. van der Vlerk (1998): Solving stochastic programs with complete integer recourse by enumeration: A framework using Grobner basis reductions, *Math. Prog.* 83, 1998, 229-252.