New reformulations of distributionally robust shortest path problem

Jianqiang Cheng[a], Janny Leung[b], Abdel Lisser[a]

[a]*Laboratoire de Recherche en Informatique (LRI), Université Paris Sud - XI, Bât. 650, 91405 Orsay Cedex, France*
{cheng, lisser}@lri.fr
[b]*Department of Systems Engineering and Engineering Management, The Chinese University of Hong Kong, Shatin, New Territories, Hong Kong*
janny@se.cuhk.edu.hk

Abstract

This paper considers a stochastic version of the shortest path problem, namely the Distributionally Robust Stochastic Shortest Path Problem (DRSSPP) on directed graphs. In this model, each arc has a deterministic cost and a random delay. The mean vector and the second-moment matrix of the uncertain data are assumed to be known, but the exact information of the distribution is unknown. A penalty occurs when the given delay constraint is not satisfied. The objective is to minimize the sum of the path cost and the expected path delay penalty. As this problem is NP-hard, we propose new reformulations and approximations using a sequence of semidefinite programming problems which provide tight lower bounds. Finally, numerical tests are conducted to illustrate the tightness of the bounds and the value of the proposed distributionally robust approach.

Keywords: Stochastic programming, Shortest path, Distributionally robust optimization, Semidefinite programming.

1. Introduction

The Shortest Path (SP) problem is a well-known combinatorial optimization problem and has been extensively studied for the last decades [3, 8, 10]. The objective of SP is to find a path with minimum distance or cost between two specified vertices of a given graph. In the deterministic SP problem, all the parameters are assumed to be known. However, due to different kinds of real life uncertainties, it may be difficult to specify the parameters precisely. Assuming deterministic values for parameters could lead to infeasibilities when the prescribed deterministic solution is implemented. One way to address this issue is robust optimization where the constraints involving random parameters are satisfied for all realizations of the random events (see, e.g., Soyster [29], Ben-Tal and Nemirovski [5]). Moreover, the random parameters are defined within a given uncertainty set. For a comprehensive overview on robust optimization, we refer the reader to the book by Ben-Tal et al. [4], the survey by Gabrel et al. [12] and references herein.

The robust shortest path problem has been widely studied. For instance, Yu and Yang [33] studied the robust shortest path problem in a layered network under two robustness criteria; they proved that the problem is NP-complete and devised a pseudo-polynomial algorithm. Gabrel et al. [13] proposed an integer linear program formulation for the studied robust shortest path and analyzed the theoretical complexity of the resulting problems.

An alternative to robust optimization is to model the problem as a stochastic optimization problem. The stochastic shortest path problem (SSPP) has also been widely studied in the past decades [15, 18, 20, 22, 24]. Provan [25] and Polychronopoulos and Tsitsiklis [26] studied expected shortest paths in networks where information on arc cost values is accumulated as the graph is being traversed, while Nikolova [23] maximized the probability that the path length does not exceed a given threshold value. Nie and Wu [22] studied the problem of finding a priori shortest paths to guarantee a given likelihood of arriving on-time in a stochastic network and also provided a pseudo-polynomial approximation based on extreme-dominance.

In transportation management systems, stochastic optimization has been applied widely as well. Sen et al. [27] formulated a network flow multiobjective model where one objective function consists in minimizing the expected travel-time between given origin and destination nodes whereas the second objective function minimizes the variance of travel-time. Miller-Hooks and Mahmassani [17] addressed the problem of determining least expected time paths in stochastic, time-varying networks where the arc weights (arc travel times) are random variables with probability distribution functions that vary with time. Xing and Zhou [32] investigated a fundamental problem of finding the most reliable path under different spatial correlation assumptions, and a Lagrangian substitution approach is used to get a lower bound. Fu and Rilett [11] studied a dynamic and stochastic shortest path problem to come-up with the expected shortest path in a traffic network where the link travel times are modeled as a continuous-time stochastic process, and proposed a heuristic algorithm based on the k-shortest path algorithm. In a recent paper, Mokarami and Hashemi [19] considered both robust and stochastic versions of the constrained shortest path problem, where an uncertain transit time was associated to each arc in addition to the arc cost. Moreover, they presented tractable approaches for solving the corresponding robust and stochastic constrained shortest path problems.

Most formulations and solution algorithms that address the SSPP require the knowledge of the underlying probability distributions of the random data. When the probability distribution is not known in advance, distributionally robust optimization can be used to handle the uncertainty [14] where only a part of the uncertainty information is required, such as the first two moments and the uncertainty support [7, 9]. In addition, a wide range of disributionally robust optimization problems can be reformulated as SDP problems, and hence solved efficiently thanks to semidefinite programming (SDP) [9].

In this paper, we study the Distributionally Robust Stochastic Shortest Path Problem (DRSSPP) where only a part of the information on random data is assumed to be known. In this model, each arc has a deterministic cost and a random delay. Furthermore, we assume that only the first and the second moments of the delay are known. This problem has a simple recourse formulation, i.e., we deal with the delays of the path by introducing a penalty which is incurred when the delay constraint is not satisfied. The objective is to minimize the sum of the path cost and the expected path delay penalty. As the deterministic shortest path problem with delay is NP-hard [31], it follows that DRSSPP is also NP-hard by choosing all the arc variances equal to 0.

This paper is organized as follows. In Section 2, we give the mathematical formulation of DRSSPP. Two equivalent deterministic formulations are presented in Section 3. In Section 4, we present a copositive reformulation of DRSSPP when the support is nonnegative. In Section 5, two relaxed versions of DRSSPP are given to approximate the original problem. In Section 6, a numerical study is provided to evaluate the approximation and to illustrate the value of the proposed distributionally robust approach. The conclusions are given in the last section.

2. DRSSPP Formulation

Let $\mathcal{G} = (V, A)$ be a digraph with $n = |V|$ nodes and $m = |A|$ arcs. Each arc $a \in A$ has an associated cost $c(a) > 0$ as well as a random delay represented by the random variable $\tilde{\delta}(a)$. We assume w.l.o.g that $c_1, \dots, c_m$ denote the costs while $\tilde{\delta}_1, ..., \tilde{\delta}_m$ are the random delays. Let $c = \{c_1, ..., c_m\}$ and $\tilde{\delta} = \{\tilde{\delta}_1, ..., \tilde{\delta}_m\}$.

When the exact probability distribution of $\tilde{\delta}$ denoted by $\mathcal{F}$ is known, the *Stochastic Shortest Path Problem (SSPP)* consists in finding a directed path between two given vertices s and t such that the sum of the cost and the expected delay cost is minimal. The delay cost is based on a penalty per time unit $d > 0$ that has to be paid whenever the total delay exceeds a given threshold $D > 0$. In transportation applications, D may represent the preferred arrival time and d the unit cost of delay, so the last term of the objective represents the expected cost of delay.

Then, SSPP can be mathematically formulated as follows [6]:

$$\text{(SSPP)} \quad \min_{x \in \{0,1\}^m} \quad c^T x + d \cdot \mathbb{E}_{\mathcal{F}}[\tilde{\delta}^T x - D]^+ \tag{1a}$$

$$\text{s.t. } Mx = b \tag{1b}$$

where $[\cdot]^+ = max\{0, \cdot\}$, $\mathbb{E}[X]$ denotes the expectation of a random variable X, $M \in \mathbb{R}^{n \times m}$ is the *node-arc incidence matrix* and $b \in \mathbb{R}^n$, with all elements being 0 except the *s-th* and *t-th* elements, which are 1 and -1, respectively [1].

The objective function is composed of two terms, namely the total cost of the shortest path and the expectation cost related to the delay constraint. The second term can be interpreted as the expectation of individual penalization of excess delays of the arcs. This formulation is also known in stochastic programming as a simple recourse formulation.

2.1. Distributionally Robust Formulation

SSPP requires that the exact information of the distribution $\mathcal{F}$ is known. However, this is not often the case for many practical problems. Therefore, distributionally robust optimization can be used to handle the uncertainty. In this paper, we model the SSPP as distributionally robust SSPP as follows:

$$\text{(DRSSPP)} \quad \min_{x \in \{0,1\}^m} \quad c^T x + d \cdot \max_{\mathcal{F} \in \mathcal{D}} \mathbb{E}_{\mathcal{F}}[\tilde{\delta}^T x - D]^+ \tag{2a}$$

$$\text{s.t. } Mx = b \tag{2b}$$

where $\mathcal{D}$ is the collection of probability distributions of interest.

In the following, DRSSPP is considered under the following key assumption:

Assumption (A1): The distributional uncertainty set accounts for information about the support $\mathcal{S}$, mean μ, and an upper bound Σ on the covariance matrix of the random vector $\tilde{\delta}$

$$\mathcal{D}(\mathcal{S}, \mu, \Sigma) = \left\{ \mathcal{F} \in \mathcal{M} \,\middle|\, \begin{array}{l} \mathbb{P}(\tilde{\delta} \in \mathcal{S}) = 1 \\ \mathbb{E}_{\mathcal{F}}[\tilde{\delta}] = \mu \\ \mathbb{E}_{\mathcal{F}}[(\tilde{\delta} - \mu)(\tilde{\delta} - \mu)^T] \preceq \Sigma \end{array} \right\} .$$

where $\mathcal{M}$ is the set of all probability distributions on the measurable space $(\mathbb{R}^m, \mathcal{B})$, with $\mathcal{B}$ the Borel σ-algebra on $\mathbb{R}^m$.

3. Deterministic Formulations

In this section, we present two equivalent deterministic formulations of DRSSPP. The first formulation is a direct derivation similar in approach to previous work [9]. The second one is a new formulation with a smaller matrix constraint size than the first one, which is much more effective computationally, as shown in Section 6.

Delage and Ye [9] have previously studied the distributionally robust approach; they gave an equivalent deterministic formulation which we apply hereafter to DRSSPP.

Theorem 1. *Under assumption (A1), together with $\mathcal{S} = \mathbb{R}^m$, problem (2) is equivalent to the following deterministic problem.*

$$\text{(DRSSPP1):} \quad \min_{x\in\{0,1\}^m, t\in\mathbb{R}, \mathbf{q}\in\mathbb{R}^m, \mathbf{Q}\in\mathbb{R}^{m\times m}} \quad c^T x + d \cdot ((\Sigma + \mu\mu^T) \bullet \mathbf{Q} + \mu^T \mathbf{q} + t) \tag{3a}$$

$$\begin{bmatrix} t + D & \frac{(\mathbf{q}-x)^T}{2} \\ \frac{\mathbf{q}-x}{2} & \mathbf{Q} \end{bmatrix} \succeq 0 \tag{3b}$$

$$\begin{bmatrix} t & \frac{\mathbf{q}^T}{2} \\ \frac{\mathbf{q}}{2} & \mathbf{Q} \end{bmatrix} \succeq 0, \tag{3c}$$

$$Mx = b \tag{3d}$$

where $\bullet$ is the inner product defined by $A \bullet B = \sum_{i,j} A_{ij} B_{ij}$.

Proof. The main idea of the proof follows from [9], whereby the distributionally robust objective function is transformed into its equivalent deterministic reformulation. We note that $\max_{\mathcal{F}\in\mathcal{D}} \mathbb{E}_{\mathcal{F}}[\tilde{\delta}^T x - D]^+$ is equivalent to the following optimization problem:

$$\max_{\mathcal{F}\in\mathcal{M}} \quad \int_{\mathcal{S}} [\delta^T x - D]^+ d\mathcal{F}(\delta) \tag{4a}$$

$$\text{s.t.} \quad \int_{\mathcal{S}} d\mathcal{F}(\delta) = 1 \tag{4b}$$

$$\int_{\mathcal{S}} \delta d\mathcal{F}(\delta) = \mu \tag{4c}$$

$$\int_{\mathcal{S}} (\delta - \mu)(\delta - \mu)^T d\mathcal{F}(\delta) \preceq \Sigma \tag{4d}$$

Accordingly, the Lagrangian function of problem (4) is as follows:

$$
\begin{aligned}
Lag &= \int_{\mathcal{S}} [\delta^T x - D]^+ d\mathcal{F}(\delta) + r(1 - \int_{\mathcal{S}} d\mathcal{F}(\delta)) \\
&\quad -\mathbf{q}^T (\int_{\mathcal{S}} \delta d\mathcal{F}(\delta) - \mu) + \mathbf{Q} \bullet (\Sigma - \int_{\mathcal{S}} (\delta - \mu)(\delta - \mu)^T d\mathcal{F}(\delta))
\end{aligned}
$$

where $r \in \mathbb{R}, \mathbf{q} \in \mathbb{R}^m$ and $0 \preceq \mathbf{Q} \in \mathbb{R}^{m \times m}$ are the dual variables for constraints (4b), (4c) and (4d) respectively. Simplifying, we obtain

$$
\begin{aligned}
Lag &= r + \mu^T \mathbf{q} + \mathbf{Q} \bullet (\Sigma + \mu\mu^T) \\
&\quad + \int_{\mathcal{S}} \{[\delta^T x - D]^+ - r - \mathbf{q}^T \delta - \mathbf{Q} \bullet \delta\delta^T\} d\mathcal{F}(\delta)
\end{aligned}
$$

In order to maximize Lag with $\mathcal{F} \in \mathcal{M}$, we have

$$
\begin{aligned}
\max_{\mathcal{F} \in \mathcal{M}} Lag &= r + \mu^T \mathbf{q} + \mathbf{Q} \bullet (\Sigma + \mu\mu^T) + \min_{\mathbf{Q} \succeq 0, \mathbf{q} \in \mathbb{R}^m, r, t \in \mathbb{R}} t \\
\text{s.t.} &\quad t \geq [\delta^T x - D]^+ - r - \mathbf{q}^T \delta - \mathbf{Q} \bullet \delta\delta^T, \forall \delta \in \mathcal{S}
\end{aligned}
$$

Precisely, the optimal $\mathcal{F}$ is the dirac distribution with probability 1 on the point $\delta \in \mathcal{S}$ which maximizes $[\delta^T x - D]^+ - r - \mathbf{q}^T \delta - \mathbf{Q} \bullet \delta\delta^T$.

Then, the dual of problem (4) can be written as follows:

$$
\begin{aligned}
\min_{\mathbf{Q} \succeq 0, \mathbf{q} \in \mathbb{R}^m, r, t \in \mathbb{R}} \max_{\mathcal{F} \in \mathcal{M}} Lag = \min_{\mathbf{Q} \succeq 0, \mathbf{q} \in \mathbb{R}^m, r, t \in \mathbb{R}} &\quad r + \mu^T \mathbf{q} + \mathbf{Q} \bullet (\Sigma + \mu\mu^T) + t \\
\text{s.t.} &\quad t \geq [\delta^T x - D]^+ - r - \mathbf{q}^T \delta - \mathbf{Q} \bullet \delta\delta^T, \forall \delta \in \mathcal{S}
\end{aligned}
$$

Furthermore, it is equivalent to

$$
\begin{aligned}
\min_{\mathbf{Q} \succeq 0, \mathbf{q} \in \mathbb{R}^m, r \in \mathbb{R}, t \in \mathbb{R}} &\quad r + \mu^T \mathbf{q} + \mathbf{Q} \bullet (\Sigma + \mu\mu^T) + t \\
\text{s.t.} &\quad t \geq \delta^T x - D - r - \mathbf{q}^T \delta - \mathbf{Q} \bullet \delta\delta^T, \forall \delta \in \mathcal{S} \\
&\quad t \geq -r - \mathbf{q}^T \delta - \mathbf{Q} \bullet \delta\delta^T, \forall \delta \in \mathcal{S}
\end{aligned}
$$

By change of variables from $t + r$ to t, we have the following equivalent reformulation:

$$
\begin{aligned}
\min_{\mathbf{Q} \succeq 0, \mathbf{q} \in \mathbb{R}^m, t \in \mathbb{R}} &\quad t + \mu^T \mathbf{q} + \mathbf{Q} \bullet (\Sigma + \mu\mu^T) \\
\text{s.t.} &\quad (1; \delta)^T \begin{bmatrix} t + D & \frac{(\mathbf{q} - x)^T}{2} \\ \frac{\mathbf{q} - x}{2} & \mathbf{Q} \end{bmatrix} (1; \delta) \geq 0 \\
&\quad (1; \delta)^T \begin{bmatrix} t & \frac{\mathbf{q}^T}{2} \\ \frac{\mathbf{q}}{2} & \mathbf{Q} \end{bmatrix} (1; \delta) \geq 0
\end{aligned}
$$

As $\mathcal{S} = \mathbb{R}^m$, thus the dual is equivalent to

$$\begin{aligned} \min_{\mathbf{Q}\succeq 0, \mathbf{q}\in\mathbb{R}^m, t\in\mathbb{R}} \quad & t + \mu^T \mathbf{q} + \mathbf{Q} \bullet (\Sigma + \mu\mu^T) \\ \text{s.t.} \quad & \begin{bmatrix} t + D & \frac{(\mathbf{q}-x)^T}{2} \\ \frac{\mathbf{q}-x}{2} & \mathbf{Q} \end{bmatrix} \succeq 0 \\ & \begin{bmatrix} t & \frac{\mathbf{q}^T}{2} \\ \frac{\mathbf{q}}{2} & \mathbf{Q} \end{bmatrix} \succeq 0 \end{aligned}$$

It is easy to show that the conditions on the support, first and second moments ensure that the dirac distribution lies in the relative interior set of problem (4). Furthermore, based on the results of Proposition 3.4 in Shapiro [28], the strong duality holds.

□

Apart from the binary constraint, problem (3) is an SDP problem which is theoretically solvable in polynomial time. However, in practice, solving the SDP problem is very time-consuming. In order to overcome this drawback, we propose a new efficient formulation without imposing any additional assumption.

Theorem 2. *Under assumption (A1) and $\mathcal{S} = \mathbb{R}^m$, problem (2) is equivalent to*

$$\textit{(DRSSPP2):} \quad \min_{x\in\{0,1\}^m, p_0, q_0, t\in\mathbb{R}} \quad c^T x + d \cdot ((x^T \Sigma x + (\mu^T x)^2) \cdot p_0 + \mu^T x \cdot q_0 + t) \tag{5a}$$

$$\begin{bmatrix} t + D & \frac{q_0 - 1}{2} \\ \frac{q_0-1}{2} & p_0 \end{bmatrix} \succeq 0 \tag{5b}$$

$$\begin{bmatrix} t & \frac{q_0}{2} \\ \frac{q_0}{2} & p_0 \end{bmatrix} \succeq 0, \tag{5c}$$

$$Mx = b \tag{5d}$$

Remark 1. In problem (5), the dimension of the linear matrix inequality is 2×2, compared to $(m + 1) \times (m + 1)$ for the linear matrix inequality of problem (3).

Proof. The proof consists of two parts: first, we establish the primal-dual relationship between problem (2) and problem (5). Second, we show that the strong duality holds.

Part 1. We introduce a new random variable $\tilde{\delta}_0 = \tilde{\delta}^T x$. The mean and the upper bound of the variance of $\tilde{\delta}_0$ are $\mu_0 = \mu^T x$ and $\sigma_0 = x^T \Sigma x$ respectively. We denote the support of $\tilde{\delta}_0$ by $\mathcal{S}_0 \subseteq \mathbb{R}$. Accordingly, the stochastic part of problem (2), i.e, $\max_{\mathcal{F}\in\mathcal{D}} \mathbb{E}_{\mathcal{F}}[\tilde{\delta}^T x - D]^+$, can be described as the following moment problem:

$$
\begin{aligned}
(\mathrm{P}): \quad \max_{\mathcal{F}\in\mathcal{M}} \quad & \int_{\mathcal{S}_0} [\delta_0 - D]^+ d\mathcal{F}(\delta_0) && \text{(6a)}\\
s.t. \quad & \int_{\mathcal{S}_0} d\mathcal{F}(\delta_0) = 1 && \text{(6b)}\\
& \int_{\mathcal{S}_0} \delta_0 d\mathcal{F}(\delta_0) = \mu_0 && \text{(6c)}\\
& \int_{\mathcal{S}_0} (\delta_0)^2 d\mathcal{F}(\delta_0) - \mu_0^2 \le \sigma_0 && \text{(6d)}
\end{aligned}
$$

Further, the Lagrangian function of problem (6) can be written as follows:

$$
\begin{aligned}
Lag \quad = \quad & \int_{\mathcal{S}_0} [\delta_0 - D]^+ d\mathcal{F}(\delta_0) + \dot{r}(1 - \int_{\mathcal{S}_0} d\mathcal{F}(\delta_0)) \\
& + q_0(\mu_0 - \int_{\mathcal{S}_0} \delta_0 d\mathcal{F}(\delta_0)) + p_0 \cdot (\sigma_0 + \mu_0^2 - \int_{\mathcal{S}_0} \delta_0^2 d\mathcal{F}(\delta_0))
\end{aligned}
$$

where $r, q_0, p_0 \in \mathbb{R}$ with $p_0 \geq 0$ are the dual variables for constraints (6b), (6c) and (6d) respectively. After simplifying, we have

$$
\begin{aligned}
Lag \quad = \quad & r + q_0\mu_0 + p_0(\sigma_0 + \mu_0^2) \\
& + \int_{\mathcal{S}_0} \{[\delta_0 - D]^+ - r - q_0\delta_0 - p_0\delta_0^2\} d\mathcal{F}(\delta_0)
\end{aligned}
$$

To maximize Lag over $\mathcal{F} \in \mathcal{M}$, we have

$$
\begin{aligned}
\max_{\mathcal{F}\in\mathcal{M}} Lag \quad = \quad & r + q_0\mu_0 + p_0(\sigma_0 + \mu_0^2) + \min_{p_0\geq 0, q_0, t, r\in\mathbb{R}} t \\
& \text{s.t.} \quad t \geq [\delta_0 - D]^+ - r - q_0\delta_0 - p_0\delta_0^2, \forall\, \delta_0 \in \mathcal{S}_0.
\end{aligned}
$$

Precisely, the optimal $\mathcal{F}$ is the dirac distribution with probability 1 on the point $\delta_0 \in \mathcal{S}_0$ which maximizes $[\delta_0 - D]^+ - r - q_0\delta_0 - p_0\delta_0^2$.

Then, the Lagrangian dual of problem (P) is as follows:

$$
\begin{aligned}
\min_{p_0\geq 0, q_0, t, r\in\mathbb{R}} \max_{\mathcal{F}\in\mathcal{M}} Lag \quad = \quad \min \quad & r + q_0\mu_0 + p_0(\sigma_0 + \mu_0^2) + t \\
\text{s.t.} \quad & t \geq [\delta_0 - D]^+ - r - q_0\delta_0 - p_0\delta_0^2, \forall\, \delta_0 \in \mathcal{S}_0
\end{aligned}
$$

Furthermore, it is equivalent to

$$
\begin{aligned}
\min_{p_0 \geq 0, p_0, q_0, t \in \mathbb{R}} \quad & q_0\mu_0 + p_0(\sigma_0 + \mu_0^2) + t \\
\text{s.t.} \quad & t \geq \delta_0 - D - q_0\delta_0 - p_0\delta_0^2, \forall\, \delta_0 \in \mathcal{S}_0 \\
& t \geq -q_0\delta_0 - p_0\delta_0^2, \forall\, \delta_0 \in \mathcal{S}_0
\end{aligned}
$$

which can be re-written as

$$
\begin{aligned}
\min_{p_0 \geq 0, p_0, q_0, t \in \mathbb{R}} \quad & q_0\mu_0 + p_0(\sigma_0 + \mu_0^2) + t \\
\text{s.t.} \quad & (1; \delta_0)^T \begin{bmatrix} t + D & \frac{q_0 - 1}{2} \\ \frac{q_0 - 1}{2} & p_0 \end{bmatrix} (1; \delta_0) \geq 0 \\
& (1; \delta_0)^T \begin{bmatrix} t & \frac{q_0{}^T}{2} \\ \frac{q_0}{2} & p_0 \end{bmatrix} (1; \delta_0) \geq 0
\end{aligned}
$$

As the support of $\tilde{\delta}$ is $\mathbb{R}^m$ and $x \neq 0$, the support of $\tilde{\delta}_0 = \tilde{\delta}^T x$ is $\mathbb{R}$, i.e., $\mathcal{S}_0 = \mathbb{R}$. Thus, the dual is equivalent to

$$
\begin{aligned}
\min_{p_0 \geq 0, p_0, q_0, t \in \mathbb{R}} \quad & q_0\mu_0 + p_0(\sigma_0 + \mu_0^2) + t \\
\text{s.t.} \quad & \begin{bmatrix} t + D & \frac{q_0 - 1}{2} \\ \frac{q_0 - 1}{2} & p_0 \end{bmatrix} \geq 0, \\
& \begin{bmatrix} t & \frac{q_0}{2} \\ \frac{q_0}{2} & p_0 \end{bmatrix} \geq 0
\end{aligned}
$$

Part 2. Analogous to the proof of Theorem 1, the strong duality holds due to the results of Shapiro [28]. Therefore, combining with the other terms and constraints of problem (2), we show that problem (2) is equivalent to problem (5). □

3.1. Simplified DRSSPP Reformulation

In the case where the mean μ and the covariance matrix Σ are positively proportional to the cost of the arc and to $\mu\mu^T$ respectively, DRSSPP can be significantly simplified. There assumptions are realistic for some real world problems, e.g., in transportation networks the traffic delay is proportional to the length of the roads.

Theorem 3. *If there exists $K \geq 0$ such that $\Sigma = K\mu\mu^T$. Then problem (2) is equivalent to*

$$\min_{x\in\{0,1\}^m, t,p,q\in\mathbb{R}} \quad c^T x + d \cdot ((K+1)\cdot p + q + t) \tag{7a}$$

$$\begin{bmatrix} t + D & \frac{q-\mu^T x}{2} \\ \frac{q-\mu^T x}{2} & p \end{bmatrix} \succeq 0 \tag{7b}$$

$$\begin{bmatrix} t & \frac{q}{2} \\ \frac{q}{2} & p \end{bmatrix} \succeq 0 \tag{7c}$$

$$Mx = b \tag{7d}$$

Proof. We prove that problem (7) is equivalent to problem (5). Let Opt_1 and Opt_2 to be the optimal objective values of problems (5) and (7) respectively. First we prove that $Opt_1 \geq Opt_2$. Suppose that (x^*, p_0^*, q_0^*, t^*) is an optimal solution of problem (5). As (x^*, p_0^*, q_0^*, t^*) is a feasible solution of problem (5), then we have

$$\begin{bmatrix} t^* + D & \frac{q_0^*-1}{2} \\ \frac{q_0^*-1}{2} & p_0^* \end{bmatrix} \succeq 0, \quad \begin{bmatrix} t & \frac{q_0^*}{2} \\ \frac{q_0^*}{2} & p_0^* \end{bmatrix} \succeq 0$$

Let $p_1 = p_0^*(\mu^T x^*)^2$ and $q_1 = q_0^* \mu^T x^*$. When $\mu^T x^* \neq 0$, the above conditions can be transformed as follows:

$$\begin{bmatrix} t^* + D & \frac{q_1-\mu^T x^*}{2\mu^T x^*} \\ \frac{q_1-\mu^T x^*}{2\mu^T x^*} & \frac{p_1}{(\mu^T x^*)^2} \end{bmatrix} \succeq 0 \quad \begin{bmatrix} t & \frac{q_1}{2\mu^T x^*} \\ \frac{q_1}{2\mu^T x^*} & \frac{p_1}{(\mu^T x^*)^2} \end{bmatrix} \succeq 0$$

which is equivalent to

$$\begin{bmatrix} t^* + D & \frac{q_1-\mu^T x^*}{2} \\ \frac{q_1-\mu^T x^*}{2} & p_1 \end{bmatrix} \succeq 0 \quad \begin{bmatrix} t^* & \frac{q_1}{2} \\ \frac{q_1}{2} & p_1 \end{bmatrix} \succeq 0$$

Therefore (x^*, p_1, q_1, t^*) is a feasible solution of problem (7). When $\mu^T x^* = 0$, $p_1 = 0, q_1 = 0$, so $(x^*, p_1, q_1, t^*) = (x^*, 0, 0, t^*)$ is also a feasible solution. Furthermore, its objective value is the same as Opt_1. Therefore, we have $Opt_1 \geq Opt_2$.

Similarly, we can prove that $Opt_1 \leq Opt_2$. By Theorem 2, problem (5) is equivalent to problem (2). Hence the conclusion follows. □

Remark 2. In problem (7), we can observe that the dimension of its linear matrix inequality is 2×2, but also its objective function is linear.

4. Special case : nonnegative support

In the above sections, we presented two equivalent deterministic formulations of DRSSPP where the support of $\tilde{\delta}$ is the whole space, namely $\mathbb{R}^m$. In this section, we consider the case where the support is nonnegative, i.e., $\mathcal{S} = \mathbb{R}^m_+$. Before presenting the deterministic reformulation, we introduce the following lemma:

Lemma 1. *Given* $\mathbf{Q} \in \mathbb{R}^{m\times m}, \mathbf{q} \in \mathbb{R}^m, t \in \mathbb{R}$. *If matrix* $\mathbf{Q}$ *is positive semidefinite, then* $P_1 := \begin{bmatrix} t & \mathbf{q}^T \\ \mathbf{q} & \mathbf{Q} \end{bmatrix} \in \mathcal{COP}^{m+1}$ *is equivalent to*

$$\exists \mathbf{p} \in \mathbb{R}^m_+, P_2 := \begin{bmatrix} t & (\mathbf{q}-\mathbf{p})^T \\ \mathbf{q}-\mathbf{p} & \mathbf{Q} \end{bmatrix} \succeq 0,$$

where $\mathcal{COP}^m$ *is the cone of copositive matrices:*

$$\mathcal{COP}^m = \{M \in \mathcal{S}^m : x^T M x \geq 0 \, \textit{for all } x \in \mathbb{R}^m_+\}.$$

Proof. First, suppose that there exists a $\mathbf{p} \in \mathbb{R}^m_+$ such that $P_2 \succeq 0$. Then for any $(\xi_0; \xi) \in \mathbb{R}^{m+1}_+$, we have

$$(\xi_0;\xi)^T P_2 (\xi_0;\xi) = t\xi_0^2 + 2\xi_0(\mathbf{q}-\mathbf{p})^T\xi + \xi^T\mathbf{Q}\xi \geq 0.$$

Furthermore, $t\xi_0^2 + 2\xi_0\mathbf{q}^T\xi + \xi^T\mathbf{Q}\xi \geq 2\xi_0\mathbf{p}^T\xi \geq 0$, where the latter inequality holds because of the nonnegativity of $\mathbf{p}$ and ξ. Thus, we conclude that $P_1 \in \mathcal{COP}^{m+1}$. Conversely, when $P_1 \in \mathcal{COP}^{m+1}$, then we consider

$$f^* := \min_{\xi \geq 0} t + 2\mathbf{q}^T\xi + \xi^T\mathbf{Q}\xi \geq 0, \tag{11}$$

whose dual problem is $\max_{\lambda\geq 0}\inf_{\xi} \mathcal{L}(\lambda,\xi) := t + 2\mathbf{q}^T\xi + \xi^T\mathbf{Q}\xi - \lambda^T\xi$. Further because $\mathbf{Q}$ is positive semidefinite and $\xi = 1$ (all elements of ξ are equal to 1) is a Slater point for the primal problem, then the strong duality holds and thus $\max_{\lambda\geq 0}\inf_{\xi} \mathcal{L}(\lambda,\xi) \geq 0$, which is equivalent to

$$\exists \lambda \geq 0, \inf_{\xi} \mathcal{L}(\lambda,\xi) = \inf_{\xi} t + 2\mathbf{q}^T\xi + \xi^T\mathbf{Q}\xi - \lambda^T\xi \geq 0.$$

Therefore, $\begin{bmatrix} t & (\mathbf{q}-\lambda/2)^T \\ \mathbf{q}-\lambda/2 & \mathbf{Q} \end{bmatrix} \succeq 0$. Furthermore, we conclude that $P_2 \succeq 0$ by setting $\mathbf{p} = \frac{\lambda}{2} \geq 0$. Thus, the lemma holds. □

Remark 3. It is well known that testing whether a given matrix $\begin{bmatrix} t & q^T \\ q & \mathbf{Q} \end{bmatrix}$ is copositive is co-NP-complete [21]. However, according to Lemma 1, it becomes polynomial time solvable when $\mathbf{Q}$ is positive semidefinite.

Theorem 4. *Under assumption (A1), together with $\mathcal{S} = \mathbb{R}_+^m$, problem (2) is equivalent to the following deterministic problem*

$$\min_{\mathbf{Q}\geq 0, p, q\in\mathbb{R}^m, t\in\mathbb{R}, x\in\{0,1\}^m} \quad c^T x + d\cdot((\Sigma+\mu\mu^T)\bullet\mathbf{Q}+\mu^T\mathbf{q}+t) \tag{12a}$$

$$\begin{bmatrix} t+D & \frac{(\mathbf{q}-x-\mathbf{p})^T}{2} \\ \frac{\mathbf{q}-x-\mathbf{p}}{2} & \mathbf{Q} \end{bmatrix} \geq 0, \tag{12b}$$

$$\begin{bmatrix} t & \frac{\mathbf{q}-\lambda^T}{2} \\ \frac{\mathbf{q}-\lambda}{2} & \mathbf{Q} \end{bmatrix} \geq 0, \quad \lambda\in\mathbb{R}_+^m \tag{12c}$$

$$Mx = b \tag{12d}$$

Proof. The proof consists of two parts. First, we derive the deterministic formulation by applying the results of Lemma 1 in [9]. The second part of the proof relies on the results of Lemma 1. First, according to the results of Lemma 1 in [9], problem (2) is equivalent to the following deterministic problem

$$\min \quad c^T x + d\cdot((\Sigma+\mu\mu^T)\bullet\mathbf{Q}+\mu^T\mathbf{q}+t) \tag{13a}$$

$$\min_{\tilde{\delta}\geq 0}\ t+\mathbf{q}^T\tilde{\delta}+\tilde{\delta}^T\mathbf{Q}\tilde{\delta}\geq 0 \tag{13b}$$

$$\min_{\tilde{\delta}\geq 0}\ t+\mathbf{q}^T\tilde{\delta}+\tilde{\delta}^T\mathbf{Q}\tilde{\delta}-\tilde{\delta}^T x+D\geq 0 \tag{13c}$$

$$Mx = b \tag{13d}$$

$$\mathbf{Q}\geq 0, q\in\mathbb{R}^m, t\in\mathbb{R}, x\in\{0,1\}^m \tag{13e}$$

It is straightforward to show that constraint (13b) has the same formulation as constraint (11) in Lemma 1. Thus by applying the same technique as in the proof of Lemma 1, constraint (13b) is equivalent to the following constraint:

$$\begin{bmatrix} t & \frac{\mathbf{q}-\lambda^T}{2} \\ \frac{\mathbf{q}-\lambda}{2} & \mathbf{Q} \end{bmatrix} \geq 0, \quad \mathbf{p}, \lambda\in\mathbb{R}_+^m.$$

By applying the same proof for constraint (13c), the conclusion follows. □

As mentioned in Section 3, SDP problems are time-consuming. In parallel to the case where the support is the whole space, i.e., $\mathcal{S} = \mathbb{R}^m$, we present another deterministic formulation for problem (2) when $\mathcal{S} = \mathbb{R}_+^m$.

Theorem 5. *Under assumption (A1) and $\mathcal{S} = \mathbb{R}^m_+$, problem (2) is equivalent to*

$$\min_{x\in\{0,1\}^m, p_0,\ q_0\ t\in\mathbb{R}} \quad c^T x + d \cdot ((x^T \Sigma x + (\mu^T x)^2) \cdot p_0 + \mu^T x \cdot q_0 + t) \tag{14a}$$

$$\begin{bmatrix} t + D & \frac{q_0 - 1 - \lambda_0}{2} \\ \frac{q_0 - 1 - \lambda_0}{2} & p_0 \end{bmatrix} \succeq 0,\ \lambda_0, \lambda \in \mathbb{R}_+ \tag{14b}$$

$$\begin{bmatrix} t & \frac{q_0 - \lambda}{2} \\ \frac{q_0 - \lambda}{2} & p_0 \end{bmatrix} \succeq 0, \tag{14c}$$

$$Mx = b \tag{14d}$$

Proof. First, we introduce a new random variable $\tilde{\delta}_0 = \tilde{\delta}^T x$. Accordingly, the mean and the upper bound of the variance of $\tilde{\delta}_0$ are $\mu_0 = \mu^T x$ and $\sigma_0 = x^T \Sigma x$ respectively. As $x \neq 0$ and $\tilde{\delta} \geq 0$, thus the support of $\tilde{\delta}_0$ is $\mathbb{R}_+$. It is easy to show that the distributional uncertainty set of $\tilde{\delta}_0$ is as follows

$$\mathcal{D}(\mathbb{R}_+, \mu_0, \sigma_0) = \left\{ \mathcal{F} \in \mathcal{M} \,\middle|\, \begin{array}{l} \mathbb{P}(\tilde{\delta}_0 \in \mathbb{R}_+) = 1 \\ \mathbb{E}_{\mathcal{F}}[\tilde{\delta}_0] = \mu_0 \\ \mathbb{E}_{\mathcal{F}}[(\tilde{\delta}_0 - \mu_0)(\tilde{\delta}_0 - \mu_0)^T] \preceq \sigma_0 \end{array} \right\}.$$

where $\mathcal{M}$ is the set of all probability distributions on the measurable space $(\mathbb{R}, \mathcal{B})$, with $\mathcal{B}$ is the Borel σ-algebra on $\mathbb{R}$. Then, with the results of Theorem 4, the conclusion follows. □

5. Relaxed Approximation

As DRSSPP is NP-hard, special interest is given to its relaxations. We present the relaxed approximations of the two deterministic formulations of problem (2) when the support of $\tilde{\delta}$ is the whole space, i.e., $\mathbb{R}^m$. For the first deterministic formulation DRSSPP1, there is a natural linear relaxation on binary variables x which is as follows:

$$\text{(DRSSPP1-SDP):} \min_{x\geq 0, t\in\mathbb{R}, \mathbf{Q}\in\mathbb{R}^{m\times m}, \mathbf{q}\in\mathbb{R}^m} \quad c^T x + d \cdot ((\Sigma + \mu\mu^T) \bullet \mathbf{Q} + \mu^T \mathbf{q} + t) \tag{15a}$$

$$\begin{bmatrix} t + D & \frac{(\mathbf{q} - x)^T}{2} \\ \frac{\mathbf{q} - x}{2} & \mathbf{Q} \end{bmatrix} \succeq 0 \tag{15b}$$

$$\begin{bmatrix} t & \frac{\mathbf{q}^T}{2} \\ \frac{\mathbf{q}}{2} & \mathbf{Q} \end{bmatrix} \succeq 0 \tag{15c}$$

$$Mx = b \tag{15d}$$

In this case, problem (15) is an SDP problem. It is easy to check that the optimal objective value of DRSSPP1-SDP is a lower bound of DRSSPP.

If we take variables p_0, q_0 and t as parameters, the second deterministic formulation of DRSSPP becomes a quadratic problem with binary constraints. Thus, we solve the problem by applying SDP relaxation methods. By introducing redundant constraints, we get the following SDP approximation of DRSSPP2 as follows:

$$\text{(DRSSPP2-SDP):} \min_{t,p_0,q_0\in\mathbb{R},x,\mathbf{X}} \quad c^T x + d \cdot (p_0(\Sigma + \mu\mu^T) \bullet \mathbf{X} + \mu^T x \cdot q_0 + t) \tag{16a}$$

$$\begin{bmatrix} t + D & \frac{q_0-1}{2} \\ \frac{q_0-1}{2} & p_0 \end{bmatrix} \succeq 0 \tag{16b}$$

$$\begin{bmatrix} t & \frac{q_0}{2} \\ \frac{q_0}{2} & p_0 \end{bmatrix} \succeq 0 \tag{16c}$$

$$M_i x = b_i, \;\; i = 1, \dots, n \tag{16d}$$

$$M_i^T \mathbf{X} M_i = b_i^2, \quad X_{ii} = x_i, i = 1, \dots, n \tag{16e}$$

$$\begin{bmatrix} 1 & x^T \\ x & \mathbf{X} \end{bmatrix} \succeq 0 \tag{16f}$$

where M_i is the i-th row vector of the matrix M. As the binary quadratic terms are replaced by an SDP relaxation, then the optimal objective value of DRSSPP2-SDP is a lower bound of DRSSPP as well.

When the variables p_0, q_0 and t are fixed in DRSSPP2-SDP, we obtain an SDP problem which can be solved in polynomial time. When x and $\mathbf{X}$ are fixed, DRSSPP2-SDP gives rise to another SDP problem. Thus, we can apply the alternating direction method which provides in this case a conservative approximation of DRSSPP2-SDP.

5.1. Alternating Direction Method

Let $p_0 = \bar{p}_0$, $q_0 = \bar{q}_0$ and $t = \bar{t}$ such that constraints (16b) and (16c) are feasible, then DRSSPP2-SDP can be written as

$$(P(\bar{p}_0, \bar{q}_0, \bar{t})) : \min_{x,\mathbf{X}} \quad c^T x + d \cdot (\bar{p}_0(\Sigma + \mu\mu^T) \bullet \mathbf{X} + \mu^T x \cdot \bar{q}_0 + \bar{t}) \tag{17a}$$

$$M_i x = b_i, \;\; i = 1, \dots, n \tag{17b}$$

$$M_i^T \mathbf{X} M_i = b_i^2, \quad X_{ii} = x_i, i = 1, \dots, n \tag{17c}$$

$$\begin{bmatrix} 1 & x^T \\ x & \mathbf{X} \end{bmatrix} \succeq 0 \tag{17d}$$

If we consider $x = \bar{x}$ and $\mathbf{X} = \bar{X}$ such that constraints (16d), (16e) and (16f) are feasible, then the second SDP problem formulation of DRSSPP2-SDP is

$$(P(\bar{x}, \bar{X})) : \min_{t, p_0, q_0 \in \mathbb{R}} \quad c^T \bar{x} + d \cdot (p_0(\Sigma + \mu\mu^T) \bullet \bar{X} + \mu^T \bar{x} \cdot q_0 + t) \quad (18a)$$

$$\begin{bmatrix} t + D & \frac{\mathbf{q}_0 - 1}{2} \\ \frac{\mathbf{q}_0 - 1}{2} & p_0 \end{bmatrix} \succeq 0 \quad (18b)$$

$$\begin{bmatrix} t & \frac{q_0}{2} \\ \frac{q_0}{2} & p_0 \end{bmatrix} \succeq 0 \quad (18c)$$

Hence an approximate solution of DRSSPP2-SDP can be found by the following Alternating Direction Procedure.

Algorithm 1: *Alternating Direction Method*

- **Step 0.** Let $\epsilon \geq 0$ be a given numerical precision parameter and choose initial parameters $p_0 = p^0$, $q_0 = q^0$ and $t = t^0$ such that constraints (16b) and (16c) are satisfied. Set the iteration counter $k = 0$ and $f^0 = -\inf$.
- **Step 1.** Solve the subproblem $P(p^k, q^k, t^k)$ and let (x^{k+1}, X^{k+1}) be the obtained optimal solutions while the optimal objective value is denoted by f^{k+1} .
- **Step 2.** If $f^k - f^{k+1} \leq \epsilon$, return $(x^{k+1}, X^{k+1}, p^k, q^k, t^k, f^{k+1})$ and stop.
- **Step 3.** Solve the subproblem $P(x^{k+1}, X^{k+1})$ to obtain an optimal solution $(p^{k+1}, q^{k+1}, t^{k+1})$.
- **Step 4.** Set $k := k + 1$ and go to **Step 1**.

Theorem 6. *If the problem DRSSPP2-SDP is bounded and has a feasible solution for the initial values of p_0,q_0 and t, then the sequence of the objective values $\{f^k\}$ generated by Algorithm 1 is nonincreasing. Moreover, the sequence $\{f^k\}$ converges to a finite limit and f^k is an upper bound of DRSSPP2-SDP.*

Proof. We first show that the sequence of values f^k produced by Algorithm 1 is nonincreasing. For any step k, the optimal solution (p^k, q^k, t^k) given by Step 3 is a feasible solution of problem $P(x^{k+1}, X^{k+1})$. Thus the optimal objective value of problem $P(x^{k+1}, X^{k+1})$ is less than or equal to f^{k+1}. Moreover, the optimal solution (x^{k+1}, X^{k+1}) given by Step 1 is a feasible solution of problem $P(p^{k+1}, q^{k+1}, t^{k+1})$. Then f^{k+2} is less than or equal to the optimal objective value of problem $P(x^{k+1}, X^{k+1})$. Above all, we have $f^{k+2} \leq f^{k+1}$. Thus the sequence of $\{f^t\}$ is nonincreasing. Furthermore, since the solution sequence (x^k, X^k) is bounded and the objective function of DRSSPP2-SDP is continuous, the monotonicity of the objective value sequence implies that $\{f^k\}$ has a finite limit. $\square$

Remark 4. The aforementioned approximation approaches in this section can be easily applied to the two deterministic formulations of problem (2) when the support of $\tilde{\delta}$ is nonnegative, i.e., $\mathcal{S} = \mathbb{R}^m_+$.

6. Numerical study

The objective of this section is twofold: firstly, we demonstrate numerically the performances of our methods by comparing two lower bounds for the DRSSPP. Secondly, we show how the distributionally robust approach can protect against distribution ambiguity when the probability distribution on the random variables is not known.

All the considered models are solved by Sedumi 1.3 [30] and CPLEX [16] with their default parameters on an Intel Core 2 Duo @ 2.26 GHz with 4.0 GB RAM.

6.1. Numerical results for DRSSPP Bounds

We focus on the two relaxed approximations aforementioned in Section 5. We consider three directed graphs for our numerical tests with $(|V|, |A|)$ equal to $(21, 39)$, $(30, 68)$ and $(40, 112)$ respectively. The input data for the models are randomly generated as follows. The cost c is uniformly generated from $[0, 10]$. The mean μ is uniformly generated from the interval $[5, 10]$ and the covariance matrix Σ is generated by the MATLAB function "gallery('randcorr',n)*2". The penalty d is set to 5 and D is set to the mean of the delay of the shortest path. We set the initial parameters for Algorithm 1 as follows: $\epsilon = 0.1$, $p_0 = \frac{1}{4D}$, $q_0 = 0$ and $t = 0$.

For the sake of simplicity, the problems DRSSPP1-SDP and DRSSPP2-SDP are called hereafter original and modified approximations respectively. In order to compare the quality of our two relaxations, we use the branch-and-bound method [6] to come up with the integer optimal solutions. The bound used in the branch-and-bound method corresponds to the original SDP relaxations. We denote the optimal values of the two SDP relaxations and the optimal value obtained with the branch-and-bound method by V^{SDP1}, V^{SDP2} and V^{OPT}, respectively.

Numerical results are given in Table 1, where column one gives the name of the instances and columns two and three present the size of the instances. Columns four and five show the optimal value of the branch-and-bound method and the corresponding CPU time respectively. Columns six to eight report the optimal value of the original approximation, the corresponding CPU time and the gap with the optimal value of the branch-and-bound method respectively. The last three columns report the optimal value of the modified approximation, the corresponding CPU time and the gap with the optimal value of the branch-and-bound method respectively. The gap is defined by Gap = $\frac{V^{OPT}-V^{SDP}}{V^{OPT}}100\%$.

DATA			B& B		Original			Modified		
Name	n	m	V^{OPT}	CPU (s)	V^{SDP1}	CPU (s)	Gap(%)	V^{SDP2}	CPU (s)	Gap(%)
Inst1	21	39	38.17	37.1	36.15	6.3	5.29	38.17	3.0	0.00
Inst2	21	39	32.03	36.7	31.40	4.7	1.97	32.03	1.7	0.00
Inst3	21	39	38.38	35.7	36.06	4.7	6.04	38.28	2.7	0.26
Inst4	21	39	35.37	36.9	33.67	4.8	4.81	35.37	2.2	0.00
Inst5	21	39	31.24	31.6	30.73	4.6	1.63	31.24	3.3	0.00
Inst6	30	68	137.63	1055.2	137.63	198.6	0.00	137.63	5.0	0.00
Inst7	30	68	133.06	1159.7	133.06	238.8	0.00	133.06	2.8	0.00
Inst8	30	68	140.81	1004.4	140.81	187.5	0.00	140.81	2.9	0.00
Inst9	30	68	132.25	1148.3	132.25	167.0	0.00	132.25	2.9	0.00
Inst10	30	68	131.79	1066.9	131.79	167.8	0.00	131.79	2.9	0.00
Inst11	40	112	167.33	41383.4	166.65	4230.9	0.41	167.26	23.0	0.04
Inst12	40	112	170.73	42696.1	168.18	3821.8	1.49	169.70	25.3	0.60
Inst13	40	112	170.75	42284.9	170.75	4620.6	0.00	170.75	7.6	0.00
Inst14	40	112	170.41	42326.7	170.41	4617.5	0.00	170.41	7.7	0.00
Inst15	40	112	170.50	42631.2	167.84	4025.8	1.56	169.27	33.1	0.72

Table 1: DRSSPP Computational results

Table 1 shows that the modified approximation outperforms significantly the original one in terms of the quality of the bounds and the CPU time. The latter is at most 34 seconds for the $(40, 112)$ instances for the modified approximation while the minimum CPU time for the original one is more than 3800 seconds for the largest instances. This performance shows the efficiency of the modified approach for solving large size instances.

We also consider a set of large size instances which are obtained by modifying graphs taken from the OR-library ([2]). As the size of these graphs is very large for the capability of SDP solvers, we consider only subgraphs of appropriate size for our experiments. The input data for the models are randomly generated as follows. The cost c is the same as in the original graph. The remaining input data are generated similarly to the previous randomly generated graphs, i.e, the mean μ is uniformly generated from $[5, 10]$. The parameters of Algorithm 1 are initialized to the values stated previously. The branch-and-bound method is highly demanding in terms of computing time when the size of the graphs is very large. Therefore, we compare only our two SDP relaxations for the large size instances.

Numerical results are given in Table 2 where column one gives the name of the instances and columns two and three present the size of the instances. Columns four and five show the optimal value of the original approximation and the corresponding CPU time. The last three columns report the optimal value of the modified approximation, the corresponding CPU time and the gap with the optimal value of the original approximation respectively. The gap is defined by Gap = $\frac{V^{SDP2} - V^{SDP1}}{V^{SDP1}} 100\%$.

DATA			Original		Modified		
Name	n	m	V^{SDP1}	CPU (s)	V^{SDP2}	CPU (s)	Gap(%)
Inst1	30	177	5.90	15510	5.90	7.7	0.00
Inst2	45	190	10.96	22496	10.96	10.3	0.00
Inst3	65	199	194.16	25118	194.16	14.9	0.00
Inst4	65	206	214.23	29487	214.23	15.8	0.00
Inst5	100	223	770.98	49513	770.98	47.5	0.00
Inst6	100	481	–	–	144.43	127.3	–
Inst7	100	753	–	–	17.09	365.0	–
Inst8	100	999	–	–	10.85	1319.4	

Table 2: DRSSPP Computational results; "–" indicates that no solution was found because of lack of memory

Table 2 shows that our modified approximation outperforms significantly the original one in terms of the size of solved instances as well as the CPU time.

6.2. Numerical results for the Distributionally Robust Method

In this section, we compare the solution of our proposed distributionally robust approach with the solution of a stochastic programming approach. We recall the

Stochastic Shortest Path Problem (SSPP) as follows:

$$\text{(SSPP)} \quad \min_{x\in\{0,1\}^m} \quad c^T x + d \cdot \mathbb{E}_{\mathcal{F}}[\tilde{\delta}^T x - D]^+ \tag{19a}$$

$$\text{s.t. } Mx = b \tag{19b}$$

where $\mathcal{F}$ has a known mean and covariance matrix structure. We consider either a normal distribution or a log-normal distribution.

For SSPP, we propose to apply the Sample Average Approximation (SAA) method to solve it. For more details on the SAA method, we refer the reader to [31]. Accordingly, the SAA problem is

$$OPT_N = \min_{x\in\{0,1\}^m} \quad c^T x + d \cdot \frac{\sum_{k=1}^{N} [\delta^{k^T} x - D]^+}{N} \tag{20a}$$

$$\text{s.t. } Mx = b \tag{20b}$$

which is equivalent to the following mixed integer linear programming problem

$$OPT_N = \min_{x\in\{0,1\}^m} \quad c^T x + d \cdot \frac{\sum_{k=1}^{N} s_k}{N} \tag{21a}$$

$$\text{s.t. } s_k \geq \delta^{k^T} x - D,\ s_k \geq 0,\ k = 1, \ldots, N \tag{21b}$$

$$Mx = b \tag{21c}$$

where the scenarios $\delta^1, \ldots, \delta^N$ are independent and sampled from the distribution $\mathcal{F}$. In our numerical tests, we set the number of scenarios N to 1000 for the SAA method.

To compare the robustness of the proposed solutions, we generated a random set of 1000 instances based on the graph of size $(n, m) = (21, 39)$. The cost c is the same as the cost vector in section 6.1. For the sake of simplicity, $\tilde{\delta}_i, i = 1, \ldots, m$, are assumed to be independent with known mean μ_i and variance σ_i^2. The mean μ_i is uniformly drawn on the interval $[0, 10]$, and the variance σ_i^2 is drawn on the interval $[0, 4]$. The penalty parameter d is set to 100 while we set the threshold D to the mean of the delay of the shortest path. In Table 3, the average performance of stochastic programming, where the probability distribution is assumed to be normal or log-normal, and distributionally robust optimization is compared over the 1000 test instances. We assume that the mean and the covariance matrix of the stochastic programming problems are the same as the robust ones.

Table 3 also reports between parentheses the additional cost in percentage of the distributionally robust and stochastic approaches with respect to the stochastic optimal costs for the current distribution, which is either normal or log-normal. When the distribution used and the current distribution are the same, the percentage for the stochastic optimal costs is zero.

Notice that the normal distribution is only defined based on the mean and the covariance information. In the case of the log-normal distribution, we set $\tilde{\delta}_i = 0.9\mu_i + \zeta_i$ with $\zeta_i \sim \ln \mathcal{N}\left(\ln\left(\frac{\bar{\mu}_i^2}{\sqrt{\bar{\mu}_i^2+\sigma_i^2}}\right), \ln\left(1 + \frac{\sigma_i^2}{\bar{\mu}_i^2}\right)\right)$, where $\bar{\mu}_i = 0.1\mu_i$. The obtained distribution satisfies the mean and covariance information with heavier tail in the direction of large values.

		Stochastic solutions		Robust solutions
		Normal	Log-Normal	Robust
Current distribution	Normal dist.	29.32(0%)	35.02(19%)	32.80(12%)
	Log-normal dist.	43.17(11%)	38.73(0%)	40.70(5%)

Table 3: Comparison between stochastic and distributionally robust solutions.

We observe that the distributionally robust approach is a conservative approximation of the stochastic optimization problem. The distributionally robust costs are higher than the stochastic ones for the normal and log-normal distributions by 5% and 12%, respectively. However, the stochastic program does not protect against distribution changes. For instance, if the current distribution is log-normal and we assume that it is normal, the solution obtained through the normal distribution costs 11% more than the real optimal solution while the one obtained through the robust approach costs only 5% more. To sum up, the robust approach is a good candidate solution when the exact information about the distribution is unknown.

7. Conclusions

In this paper, we consider a distributionally robust shortest path problem on directed graphs. Two equivalent deterministic formulations are given: one is based on the existing results and the other one is proposed for the first time in this paper. As DRSSPP is NP-hard, we approximate it by two relaxations through two deterministic formulations. The first relaxation is an SDP problem and can be solved in polynomial time while the second one is solved by applying the alternating direction method. Moreover, we also present two deterministic reformulations of DRSSPP when the support is non-negative. For these cases, we show that testing the copositivity of a matrix under certain conditions can be performed in polynomial time whereas the problem for general matrices is NP-hard. Our numerical experiments indicate that our proposed relaxed approximation outperforms the standard formulation. Finally, an extensive set of experiments were conducted to illustrate the value of the distributionlly robust approach. In addition, our approach can be applied to a large number of other stochastic optimization problems with binary variables, e.g., the stochastic knapsack problem with simple recourse.

Acknowledgements

This research benefited from the support of the FMJH Program Gaspard Monge in optimization and operation research, and from the support to this program from EDF. PGMO/IROE grant No. 2012-042H.

References

[1] Ahuja, R. K., T. L. Magnanti and J. B. Orlin, "Network Flows: Theory, Algorithms and Applications", Prentice Hall, New Jersey, 1993.

[2] Beasley, J., *Or-library*, 2010. `http://people.brunel.ac.uk/~mastjjb/jeb/info.html`.

[3] Bellman, R. E., *On a routing problem*, Quarterly of Applied Mathematic 16 (1958), 87–90.

[4] Ben-Tal, A., L. El Ghaoui and A. Nemirovski, "Robust optimization", Princeton University Press, 2009.

[5] Ben-Tal, A., and A. Nemirovski, *Robust solutions of linear programming problems contaminated with uncertain data*, Mathematical Programming, 88(3) (2000), 411–424.

[6] Cheng, J., S. Kosuch and A. Lisser, *Stochastic Shortest Path Problem with Uncertain Delays*, ICORES (2012), 256–264.

[7] Cheng, J., E. Delage and A. Lisser, *Distributionally robust stochastic knapsack problem authors*, SIAM Journal on Optimization 24(3) (2014), 1485–1506.

[8] Dijkstra, E. W., *A note on two problems in connection with graphs*, Numerische Mathematik 1 (1959), 269–271.

[9] Delage, E. and Y. Ye, *Distributionally Robust Optimization Under Moment Uncertainty with Application to Data-Driven Problems*, Operations Research 58(3) (2010), 595–612.

[10] Ford, L. R. and D. R. Fulkerson, "Flows in Networks", Princeton University Press, Princeton, 1962.

[11] Fu, L. and L.R. Rilett, *Expected shortest paths in dynamic and stochastic transportation networks*, Transportation Research Part B 32 (1998), 499–516.

[12] Gabrel, V., C. Murat and A. Thiele, *Recent Advances in Robust Optimization and Robustness: An Overview*, working paper, 2013.

[13] Gabrel, V., C. Murat and L. Wu, *New models for the robust shortest path problem: complexity, resolution and generalization*, Annals of Operations Research 207(1) (2013), 97–120.

[14] Goh, J. and M. Sim, *Distributionally robust optimization and its tractable approximations*, Operations Research 58(4) (2010), 902–917.

[15] Hutsona, K. R. and D. R. Shierb, *Extended dominance and a stochastic shortest path problem*, Computers and Operations Research 36 (2009), 584–596.

[16] IBM ILOG, *CPLEX High-performance mathematical programming engine* `http://www.ibm.com/software/integration/optimization/cplex/`.

[17] Miller-Hooks, E.D. and H.S. Mahmassani, *Least expected time paths in stochastic time-varying transportation networks*, Transportation Science 34 (2002), 198–215.

[18] Mirchandani, P. B. and H. Soroush, *Optimal paths in probabilistic networks: a case with temporary preferences*, Computers and Operations Research 12 (1985), 365–81.

[19] Mokarami, S. and S. M. Hashemi, *Constrained shortest path with uncertain transit times*, Journal of Global Optimization 63 (2015), 149–163.

[20] Murthy, I. and S. Sarkar, *A relaxation-based pruning technique for a class of stochastic shortest path problems*, Transportation Science 30 (1996), 220–236.

[21] Murty, K.G. and S.N. Kabadi, *Some NP-complete problems in quadratic and nonlinear programming*, Mathematical Programming 39(2) (1987), 117–129.

[22] Nie, Y. and X. Wu, *Shortest path problem considering on-time arrival probability*, Transportation Research Part B 43 (2009), 597–613.

[23] Nikolova, E., J. Kelner, M. Brand and M. Mitzenmacher, *Stochastic shortest paths via quasi-convex maximization*, Proceedings of European Symposium of Algorithms (2006), 552–563.

[24] Ohtsubo, Y., *Minimization risk models in stochastic shortest path problems*, Mathematical Methods of Operations Research 57 (2003), 79–88.

[25] Provan, J.S., *A polynomial-time algorithm to find shortest paths with recourse*, Networks 41 (2003), 115–125.

[26] Polychronopoulos, G.H. and J.N. Tsitsiklis, *Stochastic shortest path problems with recourse*, Networks 27 (1996), 133–143.

[27] Sen, S., R. Pillai, S. Joshi and A. Rathi, *A mean-variance model for route guidance in advanced traveler information systems*, Transportation Science 35 (2001), 37–49.

[28] Shapiro, A., ”On duality theory of conic linear problems,” Semi-Infinite Programming: Recent Advances, M.A. Goberna and M.A. López (Editors), Kluwer Academic Publishers, 2001, pp. 135–165.

[29] Soyster, A.L., *Convex programming with set-inclusive constraints and applications to inexact linear programming*, Operations Research, 21(5) (1973), 1154–1157.

[30] Sturm, J. F., *Using SeDuMi 1.02, a Matlab toolbox for optimization over symmetric cones*, Optimization Methods and Software 11-12 (1999), 625–653.

[31] Verweij, B., S. Ahmed, A. J. Kleywegt, G. Nemhauser and A. Shapiro, *The sample average approximation method applied to stochastic routing problems: A computational study*, Computational Optimization and Applications 24(2-3) (2003), 289–333.

[32] Xing, T. and X. Zhou, *Finding the most reliable path with and without link travel time correlation: A lagrangian substitution based approach*, Transportation Research Part B: Methodological 45 (2011), 1660–1679.

[33] Yu, G. and J. Yang, *On the Robust Shortest Path Problem*, Computers & Operations Research 25(6) (1998), 457–468.