A Distributionally Robust Formulation for Stochastic Quadratic Bi-level Programming

Pablo Adasme[1], Abdel Lisser[2] and Chen Wang[2]

[1] *Departamento de Ingenieria Electrica, Universidad de Santiago de Chile, Avenida Ecuador 3519, Santiago, Chile*

[2] *Laboratoire de Recherche en Informatique, Universite Paris-Sud XI, Batiment 650, 91405 Orsay Cedex, France*

Keywords: Distributionally Robust Optimization, Stochastic Programming, Binary Quadratic Bi-level Programming, Mixed Integer Programming.

Abstract: In this paper, we propose a distributionally robust model for a (0-1) stochastic quadratic bi-level programming problem. To this purpose, we first transform the stochastic bi-level problem into an equivalent deterministic formulation. Then, we use this formulation to derive a bi-level distributionally robust model (Liao, 2011). The latter is accomplished while taking into account the set of all possible distributions for the input random parameters. Finally, we transform both, the deterministic and the distributionally robust models into single level optimization problems (Audet et al., 1997). This allows comparing the optimal solutions of the proposed models. Our preliminary numerical results indicate that slight conservative solutions can be obtained when the number of binary variables in the upper level problem is larger than the number of variables in the follower.

1 INTRODUCTION

Bi-level programming (BP) is a hierarchical optimization framework. It consists in optimizing an objective function subject to a constrained set where another optimization problem is embedded. The first level optimization problem is referred to as the leader problem while the lower level, as the follower problem. Formally, a BP problem can be written as follows

$$\begin{aligned} \min_{\{x\in X, y\}} \quad & F(x,y) \\ \text{s.t.} \quad & G(x,y) \leq 0 \\ & \min_{\{y\}} f(x,y) \\ & \text{s.t.} \quad g(x,y) \leq 0 \end{aligned}$$

where $x \in R^{n_1}$, $y \in R^{n_2}$, $F : R^{n_1} \times R^{n_2} \rightarrow R$ and $f : R^{n_1} \times R^{n_2} \rightarrow R$ are the decision variables and the objective functions for the upper and lower level problems, respectively. The functions $G : R^{n_1} \times R^{n_2} \rightarrow R^{m_1}$ and $g : R^{n_1} \times R^{n_2} \rightarrow R^{m_2}$ denote upper and lower level constraints. The goal is to find an optimal point such that the leader and the follower minimizes their respective objective functions subject to their respective linking constraints (Audet et al., 1997). Applications of BP include transportation, network design, management and planning among others. For more application domains, see for instance (Floudas and Pardalos, 2001). It has been shown that bi-level problems are strongly NP-Hard, even for the simplest case where all the involved functions are affine (Audet et al., 1997).

As far as we know, robust optimization approaches have not yet been reported in the literature for bi-level programming. Some preliminary works concerning pure stochastic programming approaches can be found, for instance, in (Audestad et al., 2006; Özaltin et al., 2010; Carrion et al., 2009; Kalashnikov et al., 2010; Wynter, 2009). In (Carrion et al., 2009), an application for retailer futures market trading is considered whereas a natural gas cash-out problem is studied in (Kalashnikov et al., 2010).

Stochastic programming (SP) as well as robust optimization (RO) are well known optimization techniques to deal with mathematical problems involving uncertainty in the input parameters. In SP, it is usually assumed that the probability distributions are discrete and known or that they can be estimated (Shapiro et al., 2009). There are two well known scenario approaches in SP, the *recourse model* and the *probabilistic constrained approach*. See for instance (Schultz et al., 1996; Birge and Louveaux, 1997). Different from the SP approach, the RO framework assumes that the input random parameters lie within a convex uncertainty set and that the robust solutions must remain feasible for all possible realizations of the in-

Adasme P., Lisser A. and Wang C..
A Distributionally Robust Formulation for Stochastic Quadratic Bi-level Programming.
DOI: 10.5220/0004207100240031
In *Proceedings of the 2nd International Conference on Operations Research and Enterprise Systems* (ICORES-2013), pages 24-31
ISBN: 978-989-8565-40-2

put parameters. Thus, the optimization is performed over the worst case realization of the input parameters. In compensation, we obtain robust solutions which are protected from undesired fluctuations in the input parameters. In this case, the objective function provides more conservative solutions. We refer the reader to (Bertsimas and Sim, 2004) and (Bertsimas et al., 2010) for a more general understanding on RO.

In this paper, we propose a distributionally RO model for a (0-1) stochastic quadratic bi-level problem with expectation in the objective and probabilistic knapsack constraints in the leader. To this purpose, we first transform the stochastic problem into an equivalent deterministic problem (Gaivoronski et al., 2011). Subsequently, we apply a novel and simple distributionally robust approach proposed by (Liao, 2011) to derive a distributionally robust formulation for our stochastic bi-level problem. The latter allows optimizing the objective function over the set of all possible distributions in the input random parameters. Finally, we compute optimal solutions by transforming both problems, the deterministic as well as the distributionally models into single level optimization problems (Audet et al., 1997). Preliminary numerical comparisons are given. The paper is organized as follows. Section 2, presents the stochastic model under study and the equivalent deterministic formulation. In section 3, we derive the distributionally robust formulation. In section 4, we transform the deterministic and robust models into single level optimization problems. Then, in section 5, we provide preliminary numerical comparisons. Finally, section 6 concludes the paper.

2 PROBLEM FORMULATION

We consider the following (0-1) stochastic quadratic bi-level problem we denote hereby Q_0 as follows

$$\max_{\{x\}} \quad \mathbb{E}\left\{\sum_{i=1}^{n_1}\sum_{j=1}^{n_1} D_{i,j}(\xi)x_i x_j\right\} \tag{1}$$

$$\text{s.t.} \quad \mathbb{P}\left\{\sum_{j=1}^{n_1} a_j(\xi)x_j + \sum_{j=1}^{n_2} b_j(\xi)y_j \leq c(\xi)\right\} \geq (1-\alpha) \tag{2}$$

$$x_j \in \{0,1\}, \quad j = 1:n_1 \tag{3}$$

$$y \in \arg\max_{\{y\}}\{\sum_{j=1}^{n_2} d_j y_j\} \tag{4}$$

$$\text{s.t.} \quad \sum_{j=1}^{n_1} F_{i,j}x_j + \sum_{j=1}^{n_2} G_{i,j}y_j \leq h_i, \quad i = 1:m_2 \tag{5}$$

$$0 \leq y_j \leq 1, \quad j = 1:n_2 \tag{6}$$

where $x \in \{0,1\}^{n_1}$ and $y \in [0,1]^{n_2}$ are the leader and the follower decision variables respectively. In Q_0, (1)-(3) correspond to the leader problem while (4)-(6) represent the follower problem. The term $\mathbb{E}\{\cdot\}$ denotes mathematical expectation while $\mathbb{P}\{\cdot\}$ represents a probability imposed on the upper level knapsack constraint. This probability should be satisfied at least for $(1-\alpha)\%$ of the cases where $\alpha \in (0,0.5]$ represents the risk. The matrices D, F, G and vectors a, b, d, h, c are input nonnegative real matrices/vectors defined accordingly. We assume that the matrix $D = D(\xi)$, vectors $a = a(\xi), b = b(\xi)$ and $c = c(\xi)$ are random variables distributed according to a discrete probability distribution Ω. As such, one may suppose that $a_j(\xi)$, $b_j(\xi)$ and $c(\xi)$ are concentrated on a finite set of scenarios as $a_j(\xi) = \{a_j^1, .., a_j^K\}$, $b_j(\xi) = \{b_j^1, .., b_j^K\}$ and $c(\xi) = \{c^1, .., c^K\}$ respectively, with probability vector $q^T = (q_1, .., q_K)$ such that $\sum_{k=1}^K q_k = 1$ and $q_k \geq 0$. In (Gaivoronski et al., 2011), the authors propose a deterministic equivalent formulation for Q_0 by replacing the probabilistic constraint (2) with the following deterministic constraints

$$\sum_{j=1}^{n_1} a_j^k x_j + \sum_{j=1}^{n_2} b_j^k y_j \leq c^k + M_k z_k, \quad z_k \in \{0,1\} \forall k$$

$$\sum_{k=1}^{K} q_k z_k \leq \alpha \tag{7}$$

where M_k is defined for each $k = 1:K$ by $M_k = \sum_{j=1}^{n_1} a_j^k + \sum_{j=1}^{n_2} b_j^k - c^k$. The variable z_k for each k is a binary variable used to decide whether a particular constraint is discarded. This is handled by taking the risk α in constraint (7).

Analogously, the random variables $D_{i,j}(\xi)$ are discretely distributed, i.e. $D_{i,j}(\xi) = (D_{i,j}^1, \ldots, D_{i,j}^K), \forall i, j$ such that $\sum_{k=1}^K \rho_k = 1$ and $\rho_k \geq 0$ where ρ is the probability vector. Thus, the expectation in the objective function (1) can be written as

$$\max_{\{x\}} \quad \sum_{k=1}^{K} \rho_k \left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_1} D_{i,j}^k x_i x_j\right)$$

This yields the following deterministic equivalent problem we denote by Q_D as follows

$$\max_{\{x,z\}} \quad \sum_{k=1}^{K} \rho_k \left(\sum_{i=1}^{n_1}\sum_{j=1}^{n_1} D_{i,j}^k x_i x_j\right)$$

$$\text{s.t.} \quad \sum_{j=1}^{n_1} a_j^k x_j + \sum_{j=1}^{n_2} b_j^k y_j \leq c^k + M_k z_k, \quad \forall k$$

$$\sum_{k=1}^{K} q_k z_k \le \alpha$$
$$z_k \in \{0,1\} \forall k$$
$$x_j \in \{0,1\}, \quad j=1:n_1$$
$$y \in \arg\max_{\{y\}} \{\sum_{j=1}^{n_2} d_j y_j\}$$
$$\text{s.t.} \quad \sum_{j=1}^{n_1} F_{i,j} x_j + \sum_{j=1}^{n_2} G_{i,j} y_j \le h_i, \quad i=1:m_2$$
$$0 \le y_j \le 1, \quad j=1:n_2$$

This model is a deterministic equivalent formulation for Q_0 provided the assumption on the discrete probability space Ω holds.

3 THE DISTRIBUTIONALLY ROBUST FORMULATION

In this section, we derive a distributionally RO model for Q_D. For this, we assume that the probability distribution of the random vectors $\rho^T = (\rho_1,..,\rho_K)$ and $q^T = (q_1,..,q_K)$ are not known and that they can be estimated by some statistical mean from some available historical data. Thus, we consider the maximum likelihood estimator of the probability vectors ρ^T and q^T to be the observed frequency vectors.

3.0.1 The Distributionally Robust Model

In order to formulate a robust model for Q_D, we write its objective function as follows

$$\min_{\{x\}} \max_{\{\pi \in H_\beta\}} \sum_{k=1}^{K} \pi_k \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_1} -D_{i,j}^k x_i x_j \right) \tag{8}$$

and the left hand side of constraint (7) as the maximization problem

$$\max_{\{p \in H_\gamma\}} \sum_{k=1}^{K} p_k z_k \tag{9}$$

where the sets H_β and H_γ are defined respectively as

$$H_\beta = \left\{ \pi_k \ge 0, \forall k : \sum_{k=1}^{K} \pi_k = 1, \sum_{k=1}^{K} \frac{|\pi_k - \rho_k|}{\sqrt{\rho_k}} \le \beta \right\}$$

and

$$H_\gamma = \left\{ p_k \ge 0, \forall k : \sum_{k=1}^{K} p_k = 1, \sum_{k=1}^{K} \frac{|p_k - q_k|}{\sqrt{q_k}} \le \gamma \right\}$$

where $\beta, \gamma \in [0, \infty)$. Now, let $\delta_k = \pi_k - \rho_k$, then the inner max problem in (8) can be written as

$$\max_{\{\delta\}} \quad \sum_{k=1}^{K} (\delta_k + \rho_k) \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_1} -D_{i,j}^k x_i x_j \right)$$
$$\text{s.t.} \quad \sum_{k=1}^{K} \frac{|\delta_k|}{\sqrt{\rho_k}} \le \beta \tag{10}$$
$$\sum_{k=1}^{K} \delta_k = 0 \tag{11}$$
$$\delta_k \ge -\rho_k, \quad k=1:K \tag{12}$$

The associated dual problem is

$$\min_{\{w^1, \varphi^1, v^1\}} \sum_{k=1}^{K} \rho_k \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_1} -D_{i,j}^k x_i x_j \right) + \sum_{k=1}^{K} \rho_k w_k^1 + \beta \varphi^1$$
$$\text{s.t.} \quad \varphi^1 \ge \sqrt{\rho_k} \left(v^1 + w_k^1 - \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k x_i x_j \right), \forall k$$
$$\varphi^1 \ge -\sqrt{\rho_k} \left(v^1 + w_k^1 - \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k x_i x_j \right), \forall k$$
$$w_k^1 \ge 0, \quad \forall k$$

and φ^1, v^1, w^1 are Lagrangian multipliers for constraints (10)-(12), respectively. Similarly, we obtain a dual formulation for (9) as follows

$$\min_{\{w^2, \varphi^2, v^2\}} \quad \sum_{k=1}^{K} q_k z^k + \sum_{k=1}^{K} q_k w_k^2 + \gamma \varphi^2$$
$$\text{s.t.} \quad \varphi^2 \ge \sqrt{q_k} \left(v^2 + w_k^2 + z_k \right), \forall k$$
$$\varphi^2 \ge -\sqrt{q_k} \left(v^2 + w_k^2 + z_k \right), \forall k$$
$$w_k^2 \ge 0, \quad \forall k$$

where φ^2, v^2, w^2 are Lagrangian multipliers associated with its primal constraints. Now, replacing these dual problems in Q_D gives rise to the following distributionally robust formulation we denote by Q_D^R

$$\max_{\{w^1, \varphi^1, v^1, w^2, \varphi^2, v^2, x, z\}} \sum_{k=1}^{K} \rho_k \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k x_i x_j \right) - \sum_{k=1}^{K} \rho_k w_k^1 - \beta \varphi^1$$
$$\text{s.t.} \varphi^1 \ge \sqrt{\rho_k} \left(v^1 + w_k^1 - \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k x_i x_j \right), \forall k$$
$$\varphi^1 \ge -\sqrt{\rho_k} \left(v^1 + w_k^1 - \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k x_i x_j \right), \forall k$$
$$w_k^1 \ge 0, \quad \forall k \tag{13}$$

$$\sum_{j=1}^{n_1} a_j^k x_j + \sum_{j=1}^{n_2} b_j^k y_j \le c_k + M_k z_k, \quad k = 1:K$$
$$z_k \in \{0,1\} \quad k = 1:K$$
$$\sum_{k=1}^{K} q_k z_k + \sum_{k=1}^{K} q_k w_k^2 + \gamma\varphi^2 \le \alpha$$
$$\varphi^2 \ge \sqrt{q_k}(z_k + v^2 + w_k^2), \quad \forall k$$
$$\varphi^2 \ge -\sqrt{q_k}(z_k + v^2 + w_k^2), \quad \forall k$$
$$w_k^2 \ge 0, \quad \forall k \tag{14}$$
$$x_j \in \{0,1\}, \quad j = 1:n_1$$
$$y \in \arg\max_{\{y\}} \{\sum_{j=1}^{n_2} d_j y_j\}$$
$$\text{s.t.} \quad \sum_{j=1}^{n_1} F_{i,j} x_j + \sum_{j=1}^{n_2} G_{i,j} y_j \le h_i, \quad i = 1:m_2$$
$$0 \le y_j \le 1, \quad j = 1:n_2$$

In the next section we transform both models: Q_D and Q_D^R into single level optimization problems. More precisely, we obtain Mixed Integer Linear programming problems (MILP) (Audet et al., 1997).

4 EQUIVALENT MILP FORMULATIONS

Since the follower problem is the same for both Q_D and Q_D^R, we derive equivalent MILPs by replacing the follower problem with its primal, dual and complementarity slackness conditions. These conditions can be written as

$$\sum_{j=1}^{n_1} F_{i,j} x_j + \sum_{j=1}^{n_2} G_{i,j} y_j \le h_i, \quad i = 1:m_2 \tag{15}$$
$$0 \le y_j \le 1, \quad j = 1:n_2 \tag{16}$$

$$\sum_{i=1}^{m_2} \lambda_i G_{i,j} + \mu_j \ge d_j, \quad j = 1:n_2 \tag{17}$$
$$\lambda_i \ge 0, \quad i = 1:m_2 \tag{18}$$
$$\mu_j \ge 0, \quad j = 1:n_2 \tag{19}$$
$$\lambda_i \left(h_i - \sum_{j=1}^{n_1} F_{i,j} x_j - \sum_{j=1}^{n_2} G_{i,j} y_j \right) = 0, \quad i = 1:m_2 \tag{20}$$
$$\mu_j (1 - y_j) = 0, \quad j = 1:n_2 \tag{21}$$
$$\left(\sum_{i=1}^{m_2} \lambda_i G_{i,j} + \mu_j - d_j \right) y_j = 0, \quad j = 1:n_2 \tag{22}$$

where (15)-(16) and (17)-(19) are the primal and dual follower constraints, respectively. Note that constraints (20)-(22) are quadratic constraints. In (Audet et al., 1997), the authors propose a splitting scheme to linearize these complementarity constraints. The approach introduces binary variables as follows

$$h_i - \sum_{j=1}^{n_1} F_{i,j} x_j - \sum_{j=1}^{n_2} G_{i,j} y_j + \nu_i^1 L \le L, \quad i = 1:m_2 \tag{23}$$
$$\lambda_i \le \nu_i^1 L, \quad i = 1:m_2 \tag{24}$$
$$\nu_i^1 \in \{0,1\}, \quad i = 1:m_2 \tag{25}$$
$$1 - y_j + \nu_j^2 L \le L, \quad j = 1:n_2 \tag{26}$$
$$\mu_j \le \nu_j^2 L, \quad j = 1:n_2 \tag{27}$$
$$\nu_j^2 \in \{0,1\}, \quad j = 1:n_2 \tag{28}$$
$$\sum_{i=1}^{m_2} \lambda_i G_{i,j} + \mu_j - d_j + \nu_j^3 L \le L, \quad j = 1:n_2 \tag{29}$$
$$y_j \le \nu_j^3 L, \quad j = 1:n_2 \tag{30}$$
$$\nu_j^3 \in \{0,1\}, \quad j = 1:n_2 \tag{31}$$

where constraints (23)-(25), (26)-(28) and (29)-(31) replace the single constraints (20), (21) and (22), respectively. The parameter L is a large positive number.

Finally, let $\psi_{i,j} = x_i x_j$ be a linearization variable for each quadratic term in Q_D and Q_D^R (Fortet, 1960). Thus, a MILP formulation for Q_D can be written as

$$\max_{\{x,y,z,\psi,\lambda,\mu,\nu^1,\nu^2,\nu^3\}} \sum_{k=1}^{K} \rho_k \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k \psi_{i,j} \right)$$
$$\text{s.t.} \quad \sum_{j=1}^{n_1} a_j^k x_j + \sum_{j=1}^{n_2} b_j^k y_j \le c^k + M_k z_k, \quad \forall k$$
$$\sum_{k=1}^{K} q_k z_k \le \alpha$$
$$z_k \in \{0,1\} \forall k$$
$$\psi_{i,j} \le x_i, \quad i,j = 1:n_1 \tag{32}$$
$$\psi_{i,j} \le x_j, \quad i,j = 1:n_1 \tag{33}$$
$$\psi_{i,j} \ge x_j + x_i - 1, \quad i,j = 1:n_1 \tag{34}$$
$$\psi_{i,j} \in \{0,1\}, \quad i,j = 1:n_1 \tag{35}$$
$$x_j \in \{0,1\}, \quad j = 1:n_1$$
$$\sum_{j=1}^{n_1} F_{i,j} x_j + \sum_{j=1}^{n_2} G_{i,j} y_j \le h_i, \quad i = 1:m_2$$
$$0 \le y_j \le 1, \quad j = 1:n_2$$

$$\sum_{i=1}^{m_2} \lambda_i G_{i,j} + \mu_j \geq d_j, \quad j = 1:n_2$$
$$\lambda_i \geq 0, \quad i = 1:m_2$$
$$\mu_j \geq 0, \quad j = 1:n_2$$
$$h_i - \sum_{j=1}^{n_1} F_{i,j} x_j - \sum_{j=1}^{n_2} G_{i,j} y_j + \nu_i^1 L \leq L,$$
$$i = 1:m_2$$
$$\lambda_i \leq \nu_i^1 L, \quad i = 1:m_2$$
$$\nu_i^1 \in \{0,1\}, \quad i = 1:m_2$$
$$1 - y_j + \nu_j^2 L \leq L, \quad j = 1:n_2$$
$$\mu_j \leq \nu_j^2 L, \quad j = 1:n_2$$
$$\nu_j^2 \in \{0,1\}, \quad j = 1:n_2$$
$$\sum_{i=1}^{m_2} \lambda_i G_{i,j} + \mu_j - d_j + \nu_j^3 L \leq L,$$
$$j = 1:n_2$$
$$y_j \leq \nu_j^3 L, \quad j = 1:n_2$$
$$\nu_j^3 \in \{0,1\}, \quad j = 1:n_2$$

where constraints (32)-(35) are Fortet linearization constraints. We denote this model by MIP_D. Consequently, a MILP distributionally robust model for Q_D^R can be written as follows

$$\max_{\{\Upsilon\}} \quad \sum_{k=1}^{K} \rho_k \left(\sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k \psi_{i,j} \right)$$
$$- \sum_{k=1}^{K} \rho_k w_k^1 - \beta \varphi^1$$
$$\text{s.t.} \varphi^1 \geq \sqrt{\rho_k} \left(v^1 + w_k^1 - \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k \psi_{i,j} \right), \forall k$$
$$\varphi^1 \geq -\sqrt{\rho_k} \left(v^1 + w_k^1 - \sum_{i=1}^{n_1} \sum_{j=1}^{n_1} D_{i,j}^k \psi_{i,j} \right), \forall k$$
$$w_k^1 \geq 0, \quad \forall k$$

$$\sum_{j=1}^{n_1} a_j^k x_j + \sum_{j=1}^{n_2} b_j^k y_j \leq c_k + M_k z_k, \quad k = 1:K$$
$$\sum_{k=1}^{K} q_k z_k + \sum_{k=1}^{K} q_k w_k^2 + \gamma \varphi^2 \leq \alpha$$
$$z_k \in \{0,1\} \quad k = 1:K$$
$$\varphi^2 \geq \sqrt{q_k}(z_k + v^2 + w_k^2), \quad \forall k$$
$$\varphi^2 \geq -\sqrt{q_k}(z_k + v^2 + w_k^2), \quad \forall k$$
$$w_k^2 \geq 0, \quad \forall k$$
$$\psi_{i,j} \leq x_i, \quad i,j = 1:n_1$$
$$\psi_{i,j} \leq x_j, \quad i,j = 1:n_1$$
$$\psi_{i,j} \geq x_j + x_i - 1, \quad i,j = 1:n_1$$
$$\psi_{i,j} \in \{0,1\}, \quad i,j = 1:n_1$$
$$x_j \in \{0,1\}, \quad j = 1:n_1$$
$$\sum_{j=1}^{n_1} F_{i,j} x_j + \sum_{j=1}^{n_2} G_{i,j} y_j \leq h_i, \quad i = 1:m_2$$
$$0 \leq y_j \leq 1, \quad j = 1:n_2$$
$$\sum_{i=1}^{m_2} \lambda_i G_{i,j} + \mu_j \geq d_j, \quad j = 1:n_2$$
$$\lambda_i \geq 0, \quad i = 1:m_2$$
$$\mu_j \geq 0, \quad j = 1:n_2$$
$$h_i - \sum_{j=1}^{n_1} F_{i,j} x_j - \sum_{j=1}^{n_2} G_{i,j} y_j + \nu_i^1 L \leq L,$$
$$i = 1:m_2$$
$$\lambda_i \leq \nu_i^1 L, \quad i = 1:m_2$$
$$\nu_i^1 \in \{0,1\}, \quad i = 1:m_2$$
$$1 - y_j + \nu_j^2 L \leq L, \quad j = 1:n_2$$
$$\mu_j \leq \nu_j^2 L, \quad j = 1:n_2$$
$$\nu_j^2 \in \{0,1\}, \quad j = 1:n_2$$
$$\sum_{i=1}^{m_2} \lambda_i G_{i,j} + \mu_j - d_j + \nu_j^3 L \leq L,$$
$$j = 1:n_2$$
$$y_j \leq \nu_j^3 L, \quad j = 1:n_2$$
$$\nu_j^3 \in \{0,1\}, \quad j = 1:n_2$$

where $\Upsilon = \{w^1, \varphi^1, v^1, w^2, \varphi^2, v^2, x, y, z, \psi, \lambda, \mu, \nu^1, \nu^2, \nu^3\}$. We denote this model by MIP_D^R.

In the next section, we provide numerical comparisons between MIP_D and MIP_D^R. This allows measuring the conservatism level of MIP_D^R with respect to MIP_D. The conservatism level can be measured by the loss in optimality in exchange for a robust solution which is more protected against uncertainty (Bertsimas and Sim, 2004). This means, the less conservative the robust solutions are, the better the RO approach.

5 NUMERICAL RESULTS

In this section, we present preliminary numerical results. A Matlab program is developed using Cplex 12.3 for solving MIP_D and MIP_D^R. The numerical experiments have been carried out on a Pentium IV, 1.9 GHz with 2 GB of RAM under windows XP. The input data is generated as follows. The probability vectors ρ and q are uniformly distributed in $[0,1]$ such that the sums are equal to one. The parameter α is set to 0.1. Matrices F, G and vectors $a^k, b^k, \forall k$ are uniformly distributed in $[0,1]$. The symmetric matrices $D^k, \forall k$ and vector d are uniformly distributed in $[0,10]$. The scalars $c^k, \forall k$ and the vector h are generated respectively as

$$c^k = \frac{1}{2}\left(\sum_{j=1}^{n_1} a_j^k + \sum_{j=1}^{n_2} b_j^k\right), \quad \forall k$$

and

$$h_i = \frac{1}{2}\left(\sum_{j=1}^{n_1} F_{i,j} + \sum_{j=1}^{n_2} G_{i,j}\right), \quad \forall i = 1 : m_2$$

In table 1, columns 1-4 give the size of the instances. Columns 5-6 provide the average optimal solutions over 25 different sample instances. Finally, column 7 gives the average gaps we compute for each instance as $\frac{(\text{MIP}_D - \text{MIP}_D^R)}{\text{MIP}_D} \cdot 100\%$. These results are calculated for different values of β and γ. From table 1, we

Table 1: Average comparisons over 25 instances.

Instance size				Avg. Opt. Sol.		Avg. Gap$_R$
n_1	n_2	K	m_2	MIP$_D$	MIP$_D^R$	
$\beta = 50$ and $\gamma = 50$						
10	10	10	5	300.09	267.31	10.85 %
10	10	30	5	283.95	229.39	21.88 %
10	10	10	10	322.94	284.46	11.98 %
20	10	10	5	985.82	917.55	6.95 %
10	20	10	5	152.09	115.25	22.12 %
$\beta = 100$ and $\gamma = 50$						
10	10	10	5	313.29	258.47	17.74 %
10	10	30	5	272.49	212.07	22.05 %
10	10	10	10	320.94	290.30	9.29 %
20	10	10	5	990.99	931.64	5.93 %
10	20	10	5	138.99	100.97	27.53 %
$\beta = 50$ and $\gamma = 100$						
10	10	10	5	290.98	255.61	12.06 %
10	10	30	5	278.32	197.80	28.66 %
10	10	10	10	311.54	282.26	9.08 %
20	10	10	5	1013.41	958.89	5.23 %
10	20	10	5	169.78	89.12	47.33 %

mainly observe that the solutions tend to be more conservative when a) the number of scenarios K is larger than n_1, n_2 and m_2 and b) when the number of variables of the follower problem: n_2 is larger than n_1, K and m_2. On the opposite, we see slight conservative solutions when the number of binary variables: n_1 is larger than n_2, K and m_2. The variations of β and γ do not seem to affect these trends. However, they seem to affect the conservatism level in each case. For example, the average increases significantly up to 47.33% when $\beta < \gamma$ and n_2 is large. Same remarks when K is large.

In order to see how the parameters β and γ affect the conservatism levels, we solve one instance for each row in table 1 while varying only β and γ. These results are shown in tables, 2, 3, 4, 5 and 6, respectively. All columns in these tables provide the same information for each instance. In columns 1-2, we give the values of β and γ. Columns 3-4 give the optimal solutions for MIP$_D$ and MIP$_D^R$, respectively. Finally, in column 5, we give the gap we compute as $\frac{(\text{MIP}_D - \text{MIP}_D^R)}{\text{MIP}_D} \cdot 100\%$. In table 2, we observe that

Table 2: Instance # 1: $n_1 = n_2 = 10$, $m_2 = 5$, $K = 10$.

Robustness		Optimal Solutions		Gap$_R$
β	γ	MIP$_D$	MIP$_D^R$	
0	0	328.37	328.37	0 %
0	30		328.37	0 %
0	60		328.37	0 %
0	90		328.37	0 %
30	0		301.18	8.28 %
30	30		311.27	5.21 %
30	60		315.48	3.93 %
30	90		315.48	3.93 %
60	0		290.70	11.47 %
60	30		291.79	11.14 %
60	60		311.04	5.28 %
60	90		311.04	5.28 %
90	0		302.53	7.87 %
90	30		309.27	5.82 %
90	60		309.27	5.82 %
90	90		290.54	11.52 %

when $\beta = 0$, then augmenting the values of γ does not affect the optimal solutions. This is not the case when $\gamma = 0$ and $\beta > 0$. Next, when both $\beta > 0$ and $\gamma > 0$, the optimal solutions are affected. In particular, we observe that the parameter β affects more the optimal solutions than γ does. For example, when β goes from 30 to 60, we observe an increment of 5.93% while from 30 to 90, we observe an increment of 0.61%. This is not the case when γ increases. In this particular case, we observe a decrement of 1.28% in each case. The increase of γ seems to produce the opposite effect than incrementing β. For example, we notice that when $\beta = 30, 60, 90$ and γ goes from 0 to 30, 60 or 90, the gaps are decremented except in the worst case when both, $\beta = \gamma = 90$.

Similar observations are obtained for instances 3 and 5 in tables 4 and 6, respectively. Instances 2 and 4 in tables 3 and 5 respectively, provide additional information. Table 3 corresponds to the case where the number of scenarios K is larger compared to n_1, n_2 and m_2. In this case, increasing γ when $\beta = 0$ affects the optimal solutions. In particular, when $\beta = 0$ and γ goes from 60 to 90, we have a large increase of 31.04% in the conservatism level. This is repeated for

Table 3: Instance # 2: $n_1 = n_2 = 10$, $m_2 = 5$, $K = 30$.

Robustness		Optimal Solutions		Gap_R
β	γ	MIP_D	MIP_D^R	
0	0		181.14	0 %
0	30		181.03	0.06 %
0	60		179.85	0.71 %
0	90		123.63	31.75 %
30	0		178.82	1.28 %
30	30		177.12	2.22 %
30	60		177.12	2.22 %
30	90	181.14	123.67	31.73 %
60	0		176.63	2.49 %
60	30		176.63	2.49 %
60	60		175.07	3.35 %
60	90		123.08	32.05 %
90	0		174.60	3.61 %
90	30		173.15	4.41 %
90	60		173.15	4.41 %
90	90		121.96	32.67 %

Table 4: Instance # 3: $n_1 = n_2 = 10$, $m_2 = 10$, $K = 10$.

Robustness		Optimal Solutions		Gap_R
β	γ	MIP_D	MIP_D^R	
0	0		331.48	0 %
0	30		331.48	0 %
0	60		331.48	0 %
0	90		331.48	0 %
30	0		316.51	4.52 %
30	30		316.51	4.52 %
30	60		316.51	4.52 %
30	90	331.48	311.11	6.15 %
60	0		306.65	7.49 %
60	30		306.65	7.49 %
60	60		306.65	7.49 %
60	90		309.91	6.51 %
90	0		308.84	6.83 %
90	30		308.84	6.83 %
90	60		308.84	6.83 %
90	90		308.84	6.83 %

each value of $\beta = 0, 30, 60, 90$ when γ goes from 60 to 90. The worst gap occurs when $\beta = \gamma = 90$.

Finally, in table 5 we observe weak conservatism levels in all cases. In fact, they are lower than 10%. This instance corresponds to the case when the binary variables of the leader problem, i.e. n_1 are larger when compared to n_2, m_2 and K. Notice that when $\beta = 0$ and γ grows, then the optimal solutions are slightly affected.

6 CONCLUSIONS

In this paper, we proposed a distributionally robust model for a (0-1) stochastic quadratic bi-level pro-

Table 5: Instance # 4: $n_1 = 20, n_2 = 10$, $m_2 = 5$, $K = 10$.

Robustness		Optimal Solutions		Gap_R
β	γ	MIP_D	MIP_D^R	
0	0		982.24	0 %
0	30		965.06	1.75 %
0	60		973.95	0.84 %
0	90		982.24	0 %
30	0		923.13	6.02 %
30	30		934.96	4.81 %
30	60		940.78	4.22 %
30	90	982.24	940.78	4.22 %
60	0		940.38	4.26 %
60	30		943.63	3.93 %
60	60		931.84	5.13 %
60	90		902.04	8.16 %
90	0		936.32	4.67 %
90	30		926.40	5.68 %
90	60		929.28	5.39 %
90	90		895.58	8.82 %

Table 6: Instance # 5: $n_1 = 10, n_2 = 20$, $m_2 = 5$, $K = 10$.

Robustness		Optimal Solutions		Gap_R
β	γ	MIP_D	MIP_D^R	
0	0		257.00	0 %
0	30		257.00	0 %
0	60		257.00	0 %
0	90		257.00	0 %
30	0		241.17	6.16 %
30	30		241.17	6.16 %
30	60		241.17	6.16 %
30	90	257.00	241.17	6.16 %
60	0		230.29	10.39 %
60	30		230.29	10.39 %
60	60		230.29	10.39 %
60	90		230.29	10.39 %
90	0		223.45	13.06 %
90	30		223.45	13.06 %
90	60		223.45	13.06 %
90	90		223.45	13.06 %

gramming problem. To this end, we transformed the stochastic bi-level problem into an equivalent deterministic model. Afterward, we derived a bi-level distributionally robust model using the deterministic formulation. In particular, we applied a distributionally robust approach proposed in (Liao, 2011). This allows optimizing the problem when taking into account the set of all possible distributions of the input random parameters. Thus, we derived Mixed Integer Linear Programming formulations using Fortet linearization method (Fortet, 1960) and the approach proposed by (Audet et al., 1997). Finally, we compared the optimal solutions of this model to measure the conservatism level of the proposed robust model. Our preliminary numerical results show that slight conservative solutions are obtained for the case when the number of binary variables in the upper level problem is larger than the number of variables in the follower problem.

ACKNOWLEDGEMENTS

The first author is grateful for the financial support given by Conicyt Chilean government through the Insertion project number: 79100020.

REFERENCES

Audestad, J., Gaivoronski, A., and Werner, A. (2006). Extending the stochastic programming framework for the modeling of several decision makers: Pricing and competition in the telecommunication sector. *Annals of Operations Research*, 142:19–39.

Audet, C., Hansen, P., Jaumard, B., and Savard, G. (1997). Links between linear bilevel and mixed 01 programming problems. *Journal of Optimization Theory and Applications*, 93(2):273–300.

Bertsimas, D., Brown, D., and Caramanis, C. (2010). Theory and applications of robust optimization. *SIAM Review*, 53(3):464501.

Bertsimas, D. and Sim, M. (2004). The price of robustness. *Operations Research*, 52(1):35–53.

Birge, J. and Louveaux, F. (1997). *Introduction to stochastic programming*. Springer-Verlag, New York.

Carrion, M., Arroyo, J., and Conejo, A. (2009). A bilevel stochastic programming approach for retailer futures market trading. *IEEE Transactions on Power Systems*, 24(3):1446–1456.

Özaltin, O., Prokopyev, O., and Schaefer, A. (2010). The bilevel knapsack problem with stochastic right-hand sides. *Operations Research Letters*, 38(4):328–333.

Floudas, C. and Pardalos, P. (2001). *Encyclopedia of Optimization*. Kluwer Academic Publishers, Dordrecht. The Netherlands.

Fortet, R. (1960). Applications de l'algebre de boole en recherche operationelle. *Revue Francaise de Recherche Operationelle*, 4:17–26.

Gaivoronski, A., Lisser, A., and Lopez, R. (2011). Knapsack problem with probability constraints. *Journal of Global Optimization*, 49(3):397–413.

Kalashnikov, V., Perez-Valdes, G., Tomasgard, A., and Kalashnykova, N. (2010). Natural gas cash-out problem: Bilevel stochastic optimization approach. *European Journal of Operational Research*, 206(1):18–33.

Liao, S. (2011). Staffing a call center with uncertain non-stationary arrival rate and flexibility. *To appear in OR Spectrum*.

Schultz, R., Leen, S., and Vlerk, M. V. D. (1996). Two-stage stochastic integer programming: a survey. *Statistica Neerlandica*, 50:404–416.

Shapiro, A., Dentcheva, D., and Ruszczyński, A. (2009). *Lectures on Stochastic Programming: Modeling and Theory*. SIAM Philadelphia, Series on Optimization, 9 of MPS/SIAM, Philadelphia, 436 edition.

Wynter, L. (2009). *Encyclopedia of Optimization, chapter Stochastic Bilevel Programs*. Springer.