Dual approach for two-stage robust nonlinear optimization

Frans J.C.T. de Ruiter
Operations Research and Logistics Group, Wageningen University and Research, Wageningen, The Netherlands

Jianzhe Zhen*
Risk Analytics and Optimization Chair, École Polytechnique Fédérale de Lausanne, Switzerland

Dick den Hertog
Amsterdam Business School, University of Amsterdam, Amsterdam, The Netherlands

Adjustable robust minimization problems in which the adjustable variables appear in a convex way are difficult to solve. For example, if we substitute linear decision rules for the adjustable variables, then the model becomes convex in the uncertain parameters, whereas for computational tractability we need concavity in the uncertain parameters. In this paper we reformulate the original adjustable robust nonlinear problem with a polyhedral uncertainty set into an equivalent adjustable robust linear problem, for which all existing approaches for adjustable robust linear problems can be used. The reformulation is obtained by first dualizing over the adjustable variables and then over the uncertain parameters. The polyhedral structure of the uncertainty set then appears in the linear constraints of the dualized problem, and the nonlinear functions of the adjustable variables in the original problem appear in the uncertainty set of the dualized problem. We show how to recover the linear decision rule to the original primal problem. This paper also describes how to effectively obtain lower bounds (for minimization problems) on the optimal objective value by linking the realizations in the uncertainty set of the dualized problem to realizations in the original uncertainty set.

Key words: Adjustable robust optimization, nonlinear inequalities, duality, linear decision rules.

---

1. Introduction

1.1. Problem formulation

We consider the following general two-stage robust nonlinear minimization problem:

$$\inf_{x\in\mathcal{X}}\sup_{\zeta\in\mathcal{U}}\inf_{y}\left\{f_0(x)+g_0(y)\;\middle|\;\zeta^\top F_{i\cdot}(x)+f_i(x)+g_i(y)\leq 0,\ i=1,\ldots,m_1,\ A(x)\zeta+By=b(x)\right\}. \tag{1}$$

Here $\mathcal{X}\subseteq\mathbb{R}^{n_x}$, the functions $f_i:\mathbb{R}^{n_x}\to\mathbb{R}$, $g_i:\mathbb{R}^{n_y}\to\mathbb{R}$ are convex for all $i=0,\ldots,m_1$, $F_{i\cdot}(x)=(F_{i1}(x),\ldots,F_{in_\zeta}(x))$ and $F_{ij}:\mathbb{R}^{n_x}\to\mathbb{R}$ are real valued functions for all $i=1,\ldots,m_1$, and $j=$

* Corresponding author: trevorzhen@gmail.com

$1,\dots,n_\zeta$. The matrices $A(x)\in\mathbb{R}^{m_2\times n_\zeta}$ and the vector $b(x)\in\mathbb{R}^{m_2}$ depend on $x\in\mathbb{R}^{n_x}$ in an affine way:

$$A(x)=A^0+\sum_{l=1}^{n_x}A^l x_l,\qquad b(x)=b^0+\sum_{l=1}^{n_x}b^l x_l, \tag{2}$$

with $A^l\in\mathbb{R}^{m_2\times n_\zeta}$ and $b^l\in\mathbb{R}^{m_2}$ for all $l=0,\dots,n_\zeta$. Note that Problem (1) has *fixed recourse* because the functions g_i, $i=0,...,m_1$, and the matrix B do not depend on ζ. Therefore, there are no direct interaction terms between ζ and y, such as products $\zeta^\top y$. Throughout this paper we focus on nonempty polyhedral uncertainty sets:

$$\mathcal{U}=\{\zeta\geq 0: D\zeta=d\}, \tag{3}$$

where $D\in\mathbb{R}^{p\times n_\zeta}$ and $d\in\mathbb{R}^p$.

1.2. Literature review

Problem (1) is generally intractable even if all the objective and constraint functions are linear. Adjustable robust optimization techniques in the literature such as nonlinear decision rules, Benders decomposition, column-and-constraint generation method (Zeng and Zhao, 2013), copositive approach (Hanasusanto and Kuhn, 2018; Xu and Burer, 2018), and Fourier-Motzkin elimination (Zhen et al., 2017), are developed for linear adjustable problems and are not applicable for (1). Furthermore, even if we impose linear decision rules

$$y(\zeta)=y_0+\sum_{j=1}^{n_\zeta}y_j\zeta_j, \tag{4}$$

where $y_0,\dots,y_{n_\zeta}\in\mathbb{R}^{n_y}$, to the wait-and-see decision variables, the resulting conservative approximation of (1):

$$\inf_{\substack{x\in\mathcal{X}\\ y_0,y_j}}\ \sup_{\zeta_0\in\mathcal{U}}\left\{f_0(x)+g_0(y(\zeta))\ \middle|\ \forall\zeta\in\mathcal{U}: \begin{array}{l}\zeta^\top F_{i\cdot}(x)+f_i(x)+g_i(y(\zeta))\leq 0,\ i=1,\dots,m_1\\ A(x)\zeta+By(\zeta)=b(x)\end{array}\right\} \tag{5}$$

is still difficult to solve. This difficulty is due to the fact that the objective and constraint functions contain terms $g_i(y_0+\sum_{j=1}^{n_\zeta}y_j\zeta_j)$, $i=0,\dots,m_1$, which are convex in the uncertain parameters if g_i is nonlinear and convex. The inner maximization problem in (5) tries to maximize a convex function over a polyhedron, which is in general NP-hard.

There are only a few papers on adjustable robust nonlinear optimization known to the authors. Pınar and Tütüncü (2005) study a two-period adjustable robust portfolio problem to identify robust arbitrage opportunities when the uncertainty is ellipsoidal. They derive optimal decision rules from exploiting the explicit structure of their formulation, but it is unclear how this can be generalized to problems with more constraints, other uncertainty sets or other model formulations.

Takeda et al. (2008) consider an adjustable robust nonlinear model with a polyhedral uncertainty set, similar to the models considered in this paper. They solve a sampled model, while enumerating all vertices of the polytope uncertainty set. This quickly becomes unviable for even medium sized problems as the number of extreme points of the uncertainty set is exponential in the dimension of the uncertain parameter. Boni and Ben-Tal (2008) consider adjustable robust optimization models with conic quadratic constraints with ellipsoidal uncertainty sets. They approximate the model with linear decision rules and finally end up with a semidefinite optimization model.

Our paper significantly extends the approach of Bertsimas and de Ruiter (2016) where only linear problems are considered. Note that in the linear case, the original adjustable robust optimization models could already be solved with techniques as Fourier-Motzkin elimination, linear and nonlinear decision rules, Benders decomposition, and the column-and-constraint generation method of Zeng and Zhao (2013). This is not the case (at least, not directly) for the nonlinear problems where the original formulation cannot be solved with these techniques. However, in this paper we show that the dual of the nonlinear problem is linear in the adjustable variables. For this dual problem, the above mentioned well-known adjustable linear robust optimization techniques can be used.

1.3. Contributions

This paper uses the consecutive dualization scheme in Bertsimas and de Ruiter (2016) for linear problems, and extends it to two-stage robust *nonlinear* problems that have a polyhedral uncertainty set. The two major contributions of this paper are:

1. We apply a new relaxation technique to establish a close relation between linear decision rules for the original nonlinear problem and its equivalent dual (linear) reformulation.

2. Since linear decision rules are in general conservative, we need to provide lower bounds on the optimal objective value. We show how binding scenarios from the original uncertainty set can be obtained from binding scenarios in the dual formulation. This new technique also considerably improves the lower bounds proposed in Bertsimas and de Ruiter (2016) for the linear case.

We show that we can use our method to efficiently solve a distribution problem on a network with commitments, to solve the same problem but then without commitments, and to find the equilibrium of a system with piecewise-linear springs.

1.4. Paper organization and notation

The rest of this paper is organized as follows. In §2 we present our framework and derive our dualized formulation and linear decision rule model. We recover the linear decision rule for the original primal problem in §3. In §4 we explain how we obtain lower bounds on the optimal objective value to assess the quality of our solutions. Our numerical examples are presented in respectively §3 and §4 of EC.

The function g^* is the convex conjugate of the function $g:\mathbb{R}^{n_\nu}\to\mathbb{R}$ and is defined by:

$$g^*(z)=\sup_{\nu\in\mathrm{dom}(g)}\left\{\nu^\top z-g(\nu)\right\},$$

where $\mathrm{dom}(g)$ is the domain of the function g. The perspective $h:\mathbb{R}^{n_\nu}\times\mathbb{R}_+\to\mathbb{R}$ of a real-valued convex function $f:\mathbb{R}^{n_\nu}\to\mathbb{R}$ is defined for all $\nu\in\mathbb{R}^{n_\nu}$ and $t\in\mathbb{R}_+$ as $h(\nu,t)=tf(\nu/t)$ if $t>0$, and $h(\nu,0)=\liminf_{(\nu',t')\to(\nu,0)} t'f(\nu'/t')$ (Rockafellar, 1970, p.67). For ease of exposition, we use $tf(\nu/t)$ to denote the perspective function $h(\nu,t)$ in the rest of this paper.

2. The dual formulation

We first use the consecutive dualization approach of Bertsimas and de Ruiter (2016) to derive an equivalent linear reformulation of (1). Linear decision rules are then applied to the linear reformulation to obtain a conservative approximation. This constitutes a convex program that can be efficiently solved using off-the-shelf solvers. To this end, we first assume that (1) has a *relatively complete recourse.*

Assumption 1 **(Relatively complete recourse)**. *For all $x\in\mathcal{X}$ and all $\zeta\in\mathcal{U}$ there exists a $y\in\mathbb{R}^{n_y}$, such that*

$$\begin{cases}\zeta^\top F_{i\cdot}(x)+f_i(x)+g_i(y)\le 0 & i=1,\dots,m_1\\ A(x)\zeta+By=b(x)\end{cases}$$

and for all $i=1,\dots,m_1$ for which g_i is nonlinear we have $\zeta^\top F_{i\cdot}(x)+f_i(x)+g_i(y)<0$. □

This assumption implies that each here-and-now decision is strictly feasible. This assumption can in practice be satisfied by ensuring that $\mathcal{X}$ contains only feasible decisions. In the following theorem, we introduce a two-stage robust linear reformulation of (1).

Theorem 1 **(Dual formulation)**. *Let $\mathcal{U}$ be a polyhedral set as in* (3) *and assume that Assumption 1 holds. The here-and-now decision x is feasible for* (1) *if and only if x is feasible for the following dualized model:*

$$\begin{aligned}\inf_{x\in\mathcal{X}}\sup_{(u,v,w,z)\in\mathcal{V}}\inf_{\lambda}\Bigg\{&\sum_{i=0}^{m_1}v_if_i(x)+d^\top\lambda-w^\top b(x)-\sum_{i=0}^{m_1}z_i\;\Bigg|\\ &\sum_{k=1}^{p}D_{kj}\lambda_k\ge w^\top A_{\cdot j}(x)+\sum_{i=1}^{m_1}v_iF_{ij}(x),\quad j=1,\dots,n_\zeta\Bigg\},\end{aligned}\tag{6}$$

where $u=(u_0,\dots,u_{m_1})\in\mathbb{R}^{(m_1+1)n_y}$, $u_i\in\mathbb{R}^{n_y}$ for $i=0,...,m_1$, and

$$\mathcal{V}=\left\{(u,v,w,z):\ v\ge 0,\ v_0=1,\ v_i\,(g_i)^*\left(\tfrac{u_i}{v_i}\right)\le z_i,\ i=0,...,m_1,\ \sum_{i=0}^{m_1}u_i=-B^\top w\right\}.$$

Moreover, the infimum of (1) *coincides with that of* (6).

Proof. See §1 of EC. □

Note that the linear structure of the uncertainty set appears in the constraints of the dual formulation (6) and the convex structure of the adjustable variables is in the new uncertainty set $\mathcal{V}$. When $F_{i\cdot}(x), f_i(x)$ and $g_i(y)$ are affine functions, Theorem 1 coincides with the result in Theorem 1 of Bertsimas and de Ruiter (2016).

The obtained two-stage robust linear reformulation (6) can be conservatively approximated via linear decision rules. We impose the following linear decision rules to the wait-and-see variable λ,

$$\lambda(u,v,w,z) = \sum_{i=0}^{m_1} \Psi_i^\top u_i + \sum_{i=0}^{m_1} t_i v_i + \Phi^\top w + \sum_{i=0}^{m_1} \eta_i z_i,$$

where $\Psi_i \in \mathbb{R}^{n_y \times p}$, $t_i, \eta_i \in \mathbb{R}^p$, for all $i = 0, ..., m_1$, and $\Phi \in \mathbb{R}^{m_2 \times p}$. The resulting conservative approximation of (6) constitutes a robust optimization problem:

$$\begin{aligned}
&\inf_{\substack{x\in\mathcal{X},t_i\\ \Psi_i,\Phi,\eta_i}} \sup_{(u,v,w,z)\in\mathcal{V}} \; v^\top F_0(x) + d^\top \lambda(u,v,w,z) - w^\top b(x) - \sum_{i=0}^{m_1} z_i \\
&\text{s.t. } \forall (u,v,w,z)\in\mathcal{V}: D_{\cdot j}^\top \lambda(u,v,w,z) \geq w^\top A_{\cdot j}(x) + \sum_{i=1}^{m_1} v_i F_{ij}(x) \qquad j = 1,\ldots,n_\zeta,
\end{aligned} \tag{7}$$

where $F_0(x) = (f_0(x),\ldots,f_{m_1}(x))^\top \in \mathbb{R}^{m_1+1}$, $F_{\cdot j}(x) \in \mathbb{R}^{m_1}$ and $D_{\cdot j} \in \mathbb{R}^p$ are the j-th column vectors of $F(x)$ and D, respectively. It follows from Theorem 1 that (7) constitutes a conservative approximation of (1). Since the uncertain parameters appear linearly in (7), and $\mathcal{V}$ is convex, one can use standard robust optimization techniques to obtain the following tractable reformulation:

$$\begin{aligned}
\inf_{\substack{x\in\mathcal{X}\\ y_0,y_j,\Psi_i\\ \gamma\geq 0,t,\eta,\Phi}} \quad & f_0(x) + \gamma_{00} g_0\left(\frac{y_0 + \Psi_0 d}{\gamma_{00}}\right) + d^\top t_0 \\
\text{s.t.} \quad & \gamma_{0j} g_0\left(\frac{y_j - \Psi_0 D_{\cdot j}}{\gamma_{0j}}\right) \leq D_{\cdot j}^\top t_0 && j = 1,\ldots,n_\zeta \\
& f_i(x) + \gamma_{i0} g_i\left(\frac{y_0 + \Psi_i d}{\gamma_{i0}}\right) + d^\top t_i \leq 0 && i = 1,\ldots,m_1 \\
& F_{ij}(x) + \gamma_{ij} g_i\left(\frac{y_j - \Psi_i D_{\cdot j}}{\gamma_{ij}}\right) \leq D_{\cdot j}^\top t_i && i = 1,\ldots,m_1,\; j = 1,\ldots,n_\zeta \\
& \gamma_{i0} + d^\top \eta_i = 1 && i = 0,\ldots,m_1 \\
& \gamma_{ij} = D_{\cdot j}^\top \eta_i && i = 0,\ldots,m_1,\; j = 1,\ldots,n_\zeta \\
& By_0 + \Phi d = b(x) \\
& A_{\cdot j}(x) + By_j = \Phi D_{\cdot j} && j = 1,\ldots,n_\zeta,
\end{aligned} \tag{8}$$

where $y_j \in \mathbb{R}^{n_y}$ for all $j = 0, ..., n_\zeta$, $\Psi_i \in \mathbb{R}^{n_y \times p}$, $t_i, \eta_i \in \mathbb{R}^p$, for all $i = 0, ..., m_1$, $\gamma \in \mathbb{R}^{(m_1+1)\times(n_\zeta+1)}$ and $\Phi \in \mathbb{R}^{m_2 \times p}$.

Due to the introduction of the additional optimization variables, i.e., $y_j \in \mathbb{R}^{n_y}$ for all $j = 1, ..., n_\zeta$, $\Psi_i \in \mathbb{R}^{n_y \times p}$, $t_i, \eta_i \in \mathbb{R}^p$, for all $i = 0, ..., m_1$, $\gamma \in \mathbb{R}^{(m_1+1)\times(n_\zeta+1)}$ and $\Phi \in \mathbb{R}^{m_2 \times p}$, there are significantly more optimization variables in (8) than in the original problem (1). The original problem has only $(n_x + n_y)$ variables, but is intractable due to its nature. The tractability of (8) relies on the functions f_0, g_0, f_i, F_{ij} and g_i for all $i = 1, \dots, m_1$, and $j = 1, \dots, n_\zeta$. For example, if all these functions are conic quadratic functions, then (8) simply constitutes a conic quadratic program. More generally, the perspective function of a conically representable function can be represented in the same cone (Roos et al., 2018, Theorem 8). Therefore, the perspective functions do not lift model (8) to a higher complexity class if the original functions admit a conic representation.

Finally, we remark that if the uncertainty set $\mathcal{U}$ in (1) is non-polyhedral, one can outer approximate $\mathcal{U}$ by a polyhedral set before applying the approach developed in this section. For instance, if $\mathcal{U}$ is a conic representable set, one can use the developed scheme in Ben-Tal and Nemirovski (2001) to outer approximate $\mathcal{U}$ efficiently via a bounded polyhedron.

3. Recover a primal linear decision rule

We show that we can obtain a feasible linear decision rule for the primal model (1) from the linear decision rule model used to solve the dual model (8). For the proof, we need a novel perspective relaxation, which can be seen as an extended version of Jensen's inequality.

LEMMA 1 **(Perspective relaxation)**. *If the function $f : \mathbb{R}^{n_x} \to \mathbb{R}$ is convex, then for any $x_1, \dots, x_N \in \mathbb{R}^{n_x}$, $\alpha \in \mathbb{R}^N_+$ and $\gamma \in \mathbb{R}^N_+$ such that $\sum_{i=1}^N \alpha_i \gamma_i = 1$, we have:*

$$f\left(\sum_{i=1}^N \alpha_i x_i\right) \le \sum_{i=1}^N \alpha_i \gamma_i f\left(\frac{x_i}{\gamma_i}\right). \tag{9}$$

Proof. For any $x_1, \dots, x_N \in \mathbb{R}^{n_x}$ and $\alpha \in \mathbb{R}^N_+$, let $\gamma \in \mathbb{R}^N_+$ satisfy $\sum_{i=1}^N \alpha_i \gamma_i = 1$. Then, we have:

$$f\left(\sum_{i=1}^N \alpha_i x_i\right) = f\left(\sum_{i=1}^N \frac{\alpha_i \gamma_i}{\gamma_i} x_i\right) \le \sum_{i=1}^N \alpha_i \gamma_i f\left(\frac{x_i}{\gamma_i}\right),$$

where the inequality follows from Jensen's inequality, which applies because f is convex and $\sum_{i=1}^N \alpha_i \gamma_i = 1$, where $\alpha_i \gamma_i \in [0, 1]$, for all $i = 1, ..., N$. □

We now show that a feasible primal linear decision rule is directly obtained from variables that constitute a solution to (8).

THEOREM 2 **(Primal linear decision rule)**. *If x, y_j, $j = 0, \dots, n_\zeta$ are feasible for (8), then x, $y(\zeta) = y_0 + \sum_{j=1}^{n_\zeta} y_j \zeta_j$ is feasible for the primal linear decision rule model (1) and its objective value is at most the objective value of (8).*

Proof. Suppose x, y_j, $j=0,\ldots,n_\zeta$ are feasible for (8). We first show that x, together with linear decision rule $y(\zeta)=y_0+\sum_{j=1}^{n_\zeta} y_j\zeta_j$, results in an objective value that is at most as high as the solution for (8). Let $\zeta\in\mathcal{U}$, $\Psi_0\in\mathbb{R}^{n_y\times p}$ and let $\gamma_{0j}\geq 0$, $j=0,\ldots,n_\zeta$ such that $\gamma_{00}+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_j=1$. Then we have

$$\begin{aligned}
f_0(x)+g_0\left(y(\zeta)\right) &= f_0(x)+g_0\left(y_0+\sum_{j=1}^{n_\zeta} y_j\zeta_j\right)\\
&= f_0(x)+g_0\left(y_0+\Psi_0 d+\sum_{j=1}^{n_\zeta}(y_j-\Psi_0 D_{\cdot j})\zeta_j\right)\\
&\leq f_0(x)+\gamma_{00}g_0\left(\frac{y_0+\Psi_0 d}{\gamma_{00}}\right)+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_{0j}g_0\left(\frac{y_j-\Psi_0 D_{\cdot j}}{\gamma_{0j}}\right).
\end{aligned}$$

For the second equality we used the fact that for any $\Psi_0\in\mathbb{R}^{n_y\times p}$ we have $\Psi_0 d-\sum_{j=1}^{n_\zeta}\Psi_0 D_{\cdot j}\zeta_j=0$ since for any $\zeta\in\mathcal{U}$ we have $D\zeta=d$. The last inequality follows from Lemma 1. Using this relation we can further derive that

$$\begin{aligned}
&\sup_{\zeta\in\mathcal{U}}\left\{f_0(x)+g_0\left(y(\zeta)\right)\right\}\\
&\leq \sup_{\zeta\in\mathcal{U}}\inf_{\gamma_0\geq 0}\left\{f_0(x)+\gamma_{00}g_0\left(\frac{y_0+\Psi_0 d}{\gamma_{00}}\right)+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_j g_0\left(\frac{y_j-\Psi_0 D_{\cdot j}}{\gamma_{0j}}\right)\,\middle|\,\gamma_{00}+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_j=1\right\}\\
&\leq \inf_{\gamma_0\geq 0}\sup_{\zeta\in\mathcal{U}}\left\{f_0(x)+\gamma_{00}g_0\left(\frac{y_0+\Psi_0 d}{\gamma_{00}}\right)+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_j g_0\left(\frac{y_j-\Psi_0 D_{\cdot j}}{\gamma_{0j}}\right)\,\middle|\,\gamma_{00}+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_j=1\right\},
\end{aligned}$$

where in the second inequality we used weak duality. The obtained minimax problem is still intractable. However, it can be conservatively approximated by the following robust optimization problem:

$$\begin{aligned}
&\inf_{\gamma_0\geq 0}\sup_{\zeta_0\in\mathcal{U}}\left\{f_0(x)+\gamma_{00}g_0\left(\frac{y_0+\Psi_0 d}{\gamma_{00}}\right)+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_{0j}g_0\left(\frac{y_j-\Psi_0 D_{\cdot j}}{\gamma_{0j}}\right)\,\middle|\,\forall\zeta\in\mathcal{U}:\gamma_{00}+\sum_{j=1}^{n_\zeta}\gamma_{0j}\zeta_j=1\right\}\\
&=\inf_{\gamma_0\geq 0,t_0,\eta_0}\left\{f_0(x)+\gamma_{00}g_0\left(\frac{y_0+\Psi_0 d}{\gamma_{00}}\right)+d^\top t_0\,\middle|\,\gamma_{0j}g_0\left(\frac{y_j-\Psi_0 D_{\cdot j}}{\gamma_{0j}}\right)\leq D_{\cdot j}^\top t_0,\ j=1,\ldots,n_\zeta\right.\\
&\qquad\left.\gamma_{00}+\sum_{k=1}^{p}\eta_{0k}d_k=1,\ \gamma_{0j}=D_{\cdot j}^\top\eta_0,\ j=1,\ldots,n_\zeta\right\}.
\end{aligned}$$

The resulting objective function and constraint are contained in (8). One can apply the same approximation steps to show feasibility of the the constraints. That is, analogous it can be derived that for $i=1,\ldots,m_1$, the i-th constraint

$$\zeta^\top F_{i\cdot}(x)+f_i(x)+g_i(y(\zeta))\leq 0$$

is satisfied because the following set of constraints is satisfied in (8) for some $\Psi_i \in \mathbb{R}^{n_y \times p}$ and $\eta_i \in \mathbb{R}^{n_\zeta}$:

$$\begin{aligned}
& f_i(x) + \gamma_{i0} g_i\left(\frac{y_0 + \Psi_i d}{\gamma_{i0}}\right) + d^\top t_i \leq 0, \\
& F_{ij}(x) + \gamma_{ij} g_i\left(\frac{y_j - \Psi_i D_{\cdot j}}{\gamma_{ij}}\right) \leq D_{\cdot j}^\top t_i && j = 1, \ldots, n_\zeta, \\
& \gamma_{i0} + d^\top \eta_i = 1, \\
& \gamma_{ij} = D_{\cdot j}^\top \eta_i && j = 1, \ldots, n_\zeta.
\end{aligned}$$

Finally, using standard techniques in robust optimization, without perspective relaxation, one can show that $A(x)\zeta + By(\zeta) = b(x)$ is satisfied whenever there exists $\Phi \in \mathbb{R}^{m_2 \times p}$ that satisfies the remaining constraints of (8):

$$\begin{aligned}
& By_0 + \Phi d = b(x), \\
& A_{\cdot j}(x) + By_j = \Phi D_{\cdot j} && j = 1, \ldots, n_\zeta.
\end{aligned}$$

$\square$

To relate the objective value of the primal linear decision rule model (1) to the dual linear decision rule model (8) we use several conservative approximations in the proof of Theorem 2. Hence, the true objective value of the primal linear decision rule model could be lower than the value obtained of (8); see Remark 1 in §3 of EC for a numerical demonstration.

4. Lower bounds on the optimal value

A model with a finite sample of scenarios can provide a lower bound on the optimal value of (1). The sampled version of the dualized model is:

$$\begin{aligned}
\inf_{\substack{\tau, x \in \mathcal{X} \\ \lambda^1, \ldots, \lambda^S \geq 0}} \quad & \tau \\
\text{s.t.} \quad & f_0(x) + \sum_{i=1}^{m_1} v_i^s f_i(x) + d^\top \lambda^s - (w^s)^\top b(x) - \sum_{i=0}^{m_1} z_i^s \leq \tau \qquad \forall s = 1, \ldots, S \\
& \sum_{s=1}^{p} D_{kj} \lambda_k^s \geq (w^s)^\top A_{\cdot j}(x) + \sum_{i=1}^{m_1} (v_i^s) F_{i,l}(x) \qquad \forall j = 1, \ldots, n_\zeta, \ s = 1, \ldots, S,
\end{aligned} \tag{10}$$

where $\{(u^1, w^1, v^1, z^1), \ldots, (u^S, w^S, v^S, z^S)\}$ is a finite subset $\mathcal{V}$ with a single optimization variable λ^s for each scenario $s = 1, \ldots, S$. Note that this is a standard convex optimization model, but only guarantees feasibility of the here-and-now decisions for a small set of scenarios. The question is of course how to choose scenarios to get strong lower bounds. One way to obtain an effective finite set is described by Hadjiyiannis et al. (2011).

If we have a set of scenarios $\{(u^1,w^1,v^1,z^1),\dots,(u^S,w^S,v^S,z^S)\}$ for the sampled version of the dualized model, we can link and recover primal scenarios $\{\zeta^1,\dots,\zeta^S\}$ to obtain stronger lower bounds. To establish the link, we first dualize over $\lambda_1,\dots,\lambda_K$ in (10), which yields

$$\inf_{x\in\mathcal{X}} \sup_{\zeta\in\mathcal{U}} \sup_{1\le s\le S} \; f_0(x)+\sum_{i=1}^{m_1} v_i^s\left(\zeta^\top F_{i\cdot}(x)+f_i(x)\right)+\left(A(x)\zeta-b(x)\right)^\top w^s-\sum_{i=0}^{m_1} z_i^s. \tag{11}$$

For a fixed x we can now obtain primal scenarios ζ^s for each s as the maximizers of model (11):

$$\zeta^s\in\arg\max_{\zeta\in\mathcal{U}}\left\{\sum_{i=1}^{m_1} v_i^s\left(\zeta^\top F_{i\cdot}(x)\right)+(w^s)^\top\left(A(x)\zeta-b(x)\right)\right\}. \tag{12}$$

The resulting set of scenarios $\{\zeta^1,\dots,\zeta^s\}$ can then be used in a sampled model of (1).

A special case arises for *right-hand-side* uncertainty, where primal scenarios obtained by (12) provide stronger bounds than the dual scenarios. We say that there is only right-hand-side uncertainty if there is no direct interaction between the here-and-now decisions x and ζ. The more formal definition is given below.

Definition 1 (Right-hand-side uncertainty). Model (1) has right-hand-side uncertainty if there exist $\bar F_{i\cdot}\in\mathbb{R}^{n_\zeta}$ and $\bar A\in\mathbb{R}^{m_2\times n_\zeta}$ such that $A(x)=\bar A$ and $F_{i\cdot}(x)=\bar F_{i\cdot}$ for all $x\in\mathcal{X}$, $i=1,\dots,m_1$.

Using this definition, we can now formally prove that primal scenarios obtained from dual scenarios yield stronger lower bounds for right-hand-side uncertainty.

Theorem 3 **(Primal-dual scenarios)**. *Let $\{(u^1,w^1,v^1,z^1),\dots,(u^S,w^S,v^S,z^S)\}$ be a finite set of dual scenarios and $\{\zeta^1,\dots,\zeta^S\}$ be a set of primal scenarios obtained from* (12)*. If there is only right-hand-side uncertainty in model* (1)*, then the objective value of*

$$\begin{aligned}
\inf_{\substack{\tau,x\in\mathcal{X}\\ y^1,\dots,y^S}} \quad & \tau \\
\text{s.t.} \quad & f_0(x)+g_0(y^s)\le\tau \qquad \forall s=1,\dots,S \\
& (\zeta^s)^\top\bar F_{i\cdot}+f_i(x)+g_i(y^s)\le 0 \qquad \forall i=1,\dots,m_1,\; s=1,\dots,S \\
& \bar A\zeta^s+By^s=b(x) \qquad \forall s=1,\dots,S
\end{aligned} \tag{13}$$

is at least as high as the objective value of (10)*.*

Proof. By duality for linear programming, (10) is equivalent to (11). The latter formulation can be written as

$$\inf_{x\in\mathcal{X}}\sup_{s\in\{1,\dots,S\}}\left\{f_0(x)+\sum_{i=1}^{m_1} v_i^s\left((\zeta^s)^\top\bar F_{i\cdot}+f_i(x)\right)+(w^s)^\top\left(\bar A\zeta^s-b(x)\right)-\sum_{i=0}^{m_1} z_i^s\right\}, \tag{14}$$

where ζ^s are the primal scenarios obtained by (12). Since (u^s,w^s,v^s,z^s) are in $\mathcal{V}$ for all $s=1,\dots,S$, the value of (14) must be smaller than or equal to

$$\inf_{x\in\mathcal{X}}\sup_{s}\sup_{(u^s,w^s,v^s,z^s)\in\mathcal{V}}\left\{f_0(x)+\sum_{i=1}^{m_1} v_i^s\left((\zeta^s)^\top\bar F_{i\cdot}+f_i(x)\right)+(w^s)^\top\left(\bar A\zeta^s-b(x)\right)-\sum_{i=0}^{m_1} z_i^s\right\},$$

since we maximize over (u^s, w^s, v^s, z^s) in the full $\mathcal{V}$, instead of a subset. The value of this optimization problem is, by dualizing over (u^s, w^s, v^s, z^s), equivalent to (13). Hence, the optimal objective value is at least as high as the optimal objective value of (10). $\square$

The intuition behind the strength of the primal scenarios for right-hand-side uncertainty can be found in the fact that primal scenarios have no direct interaction with here-and-now decisions. That is, for right-hand-side uncertainty only, there are no terms in which both x and ζ appear. The dual model always includes the interaction terms with here-and-now decisions via the terms $\sum_{i=1}^{m_1} v_i^s f_i(x)$ and $(w^s)^\top b(x)$, even with right-hand-side uncertainty in the primal sampled model. Therefore, dual scenarios could be strong for some here-and-now decision x, but very weak for other here-and-now decisions. In that case, the feasible region of the dual sampled model is larger and therefore results in a lower objective value and thus a weaker lower bound.

For linear adjustable robust optimization models, Theorem 3 can also significantly improve lower bounds. In §5 of EC, we evaluate the performance of the lower bounding scheme proposed in this subsection using the same numerical experiment considered in Bertsimas and de Ruiter (2016). Original optimality gaps reported for the larger instances were more than halved when primal scenarios were obtained using (12). For the largest instance the linked primal scenarios reduced the gap from 10.7% to 5.2%. We do note that the numerical examples all satisfied the assumption of right-hand-side uncertainty. If there is no right-hand-side uncertainty, then a dual sampled model can yield tighter lower bounds than its primal sampled counterpart. A very small example showing this is given in §2 of EC. Therefore, the assumption of right-hand-side uncertainty assumption is crucial in Theorem 3.

Acknowledgments

The research of the first author is partially supported by the Netherlands Organisation for Scientific Research (NWO) Talent Grant 406-14-067. The second author is partially supported by the NWO Grant 613.001.208.

References

A. Ben-Tal and A. Nemirovski. On polyhedral approximations of the second-order cone. *Mathematics of Operations Research*, 26(2):193–205, 2001.

D. Bertsimas and F. de Ruiter. Duality in two-stage adaptive linear optimization: faster computation and stronger bounds. *INFORMS Journal on Computing*, 28(3):500–511, 2016.

O. Boni and A. Ben-Tal. Adjustable robust counterpart of conic quadratic problems. *Mathematical Methods of Operations Research*, 68(2):211–233, 2008.

M. Hadjiyiannis, P. Goulart, and D. Kuhn. A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. In *Proceedings IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, pages 7386–7391, 2011.

G. Hanasusanto and D. Kuhn. Conic programming reformulations of two-stage distributionally robust linear programs over wasserstein balls. *Operations Research*, 66(3):849–869, 2018.

M. Pınar and R. Tütüncü. Robust profit opportunities in risky financial portfolios. *Operations Research Letters*, 33(4):331–340, 2005.

R. Rockafellar. *Convex Analysis.* Princeton University Press, 1970.

E. Roos, D. den Hertog, A. Ben-Tal, F.J.C.T. de Ruiter, and J. Zhen. Approximation of hard uncertain convex inequalities. *Available on Optimization Online*, 2018.

A. Takeda, S. Taguchi, and R. Tütüncü. Adjustable robust optimization models for a nonlinear two-period system. *Journal of Optimization Theory and Applications*, 136(2):275–295, 2008.

G. Xu and S. Burer. A copositive approach for two-stage adjustable robust optimization with uncertain right-hand sides. *Computational Optimization and Applications*, 70(1):33–59, 2018.

B. Zeng and L. Zhao. Solving two-stage robust optimization problems using a column-and-constraint generation method. *Operations Research Letters*, 41:457–461, 2013.

J. Zhen, D. den Hertog, and M. Sim. Adjustable robust optimization via Fourier-Motzkin elimination. *Operations Research*, 66(4):1086–1100, 2017.

Dual approach for two-stage robust nonlinear optimization E-Companion

Frans J.C.T. de Ruiter
Operations Research and Logistics Group, Wageningen University and Research, Wageningen, The Netherlands
Jianzhe Zhen
Risk Analytics and Optimization Chair, École Polytechnique Fédérale de Lausanne, Switzerland
Dick den Hertog
Amsterdam Business School, University of Amsterdam, Amsterdam, The Netherlands

1. Proof for Theorem 1

Proof of Theorem 1. We adopt a similar proof strategy as for Theorem 1 in Bertsimas and de Ruiter (2016). We consider the inner infimum of (1) over y for a given $x \in \mathcal{X}$ and $\zeta \in \mathcal{U}$. Since Assumption 1 holds we can apply the Lagrangian principle to (1), and obtain the following equivalent reformulation:

$$\inf_{x\in\mathcal{X}} \sup_{\substack{\zeta\in\mathcal{U}\\ v\geq 0, w}} \inf_{y} \; f_0(x) + g_0(y) + \sum_{i=1}^{m_1} v_i \left(\zeta^\top F_{i\cdot}(x) + f_i(x) + g_i(y)\right) + w^\top \left(A(x)\zeta + By - b(x)\right).$$

We then use the definition of the conjugate functions and calculus rules for conjugate functions (specifically Rule 5 in Table 2 of Roos et al. (2020) for the conjugate of the sum of convex functions) to obtain the following inf-sup-sup reformulation:

$$\inf_{x\in\mathcal{X}} \sup_{\zeta\in\mathcal{U}} \sup_{(u,v,w)\in\mathcal{W}} \; f_0(x) + \sum_{i=1}^{m_1} v_i \left(\zeta^\top F_{i\cdot}(x) + f_i(x)\right) + w^\top \left(A(x)\zeta - b(x)\right) - \sum_{i=0}^{m_1} v_i \left(g_i\right)^* \left(\tfrac{u_i}{v_i}\right),$$

where $\mathcal{W} = \left\{(u,v,w):\; v \geq 0,\; v_0 = 1,\; \sum_{i=0}^{m_1} u_i = -B^\top w\right\}$. We can then switch the order of supremum such that the inner supremum is over $\zeta \in \mathcal{U}$. Since the inner supremum model is linear in ζ, we can apply strong duality for linear optimization to obtain the inf-sup-inf reformulation:

$$\inf_{x\in\mathcal{X}} \sup_{(u,v,w)\in\mathcal{W}} \inf_{\lambda} \left\{ \sum_{i=0}^{m_1} v_i f_i(x) + d^\top \lambda - w^\top b(x) - \sum_{i=0}^{m_1} v_i \left(g_i\right)^* \left(\tfrac{u_i}{v_i}\right) \;\middle|\right.$$

$$\sum_{k=1}^{p} D_{kj}\lambda_k \geq w^\top A_{\cdot j}(x) + \sum_{i=1}^{m_1} v_i F_{ij}(x), \quad j=1,\dots,n_\zeta \Bigg\}.$$

We then introduce epigraph variables z_i for every $v_i\,(g_i)^*\left(\frac{u_i}{v_i}\right)$, $i=0,...,m_1$, and finally obtain (6).

□

2. Example when conditions in Theorem 3 are violated

Consider the following problem:

$$\min_{(x_1,x_2)\in\mathcal{X}} \max_{(\zeta_1,\zeta_2)\in\mathcal{U}} \min_{y} \left\{-y \mid -1 + x_1\zeta_1 + x_2\zeta_2 + y^2 \leq 0\right\},$$

where $\mathcal{X} = \{(x_1,x_2) \geq 0 \mid x_1 + x_2 = 1\}$ and $\mathcal{U} = \{(\zeta_1,\zeta_2) \geq 0 \mid \zeta_1 + \zeta_2 \leq 1\}$. This problem satisfies the strong relatively complete recourse condition, because for all $(x_1,x_2)\in\mathcal{X}$ and $(\zeta_2,\zeta_2)\in\mathcal{U}$ the wait-and-see decision $y=0$ is feasible. Using Theorem 1 we can obtain the dual version of this problem:

$$\min_{(x_1,x_2)\in\mathcal{X}} \max_{v\geq 0} \min_{\lambda\geq 0} \left\{-\tfrac{1}{4v} - v + \lambda \mid \lambda \geq v x_1,\ \lambda \geq v x_2\right\}.$$

For this small problem the optimal solution can be determined without heavy computations. The optimal objective value for these problems is $-\frac{1}{\sqrt{2}}$ and is obtained for here-and-now decision $x_1^* = x_2^* = \frac{1}{2}$, and wait-and-see decision $y^* = \frac{1}{\sqrt{2}}$ in the primal formulation and $\lambda^* = \frac{1}{2}v$ in the dual formulation. The worst-case objective value in the dual formulation is achieved for $v^* = \frac{1}{\sqrt{2}}$. Suppose we solve the sampled version of the problem for only this worst-case scenario v^*. In that case, the sampled model looks like:

$$\min_{(x_1,x_2)\in\mathcal{X},\bar{\lambda}\geq 0} \left\{-\tfrac{2}{2\sqrt{2}} - \tfrac{1}{\sqrt{2}} + \bar{\lambda} \mid \bar{\lambda} \geq \tfrac{1}{\sqrt{2}}x_1,\ \bar{\lambda} \geq \tfrac{1}{\sqrt{2}}x_2\right\},$$

which has optimal objective value of $-\frac{1}{\sqrt{2}}$. Hence, the lower bound that follows from the sampled version of the dual formulation is tight. If we now want to match a critical scenario ζ^* using (12) we get

$$\begin{aligned}
\zeta^* &\in \arg\max_{\zeta\in\mathcal{U}} \{v^*(x_1^*\zeta_1 + x_2^*\zeta_2)\} \\
&= \arg\max_{\zeta\in\mathcal{U}} \left\{v^*(\tfrac{1}{2}\zeta_1 + \tfrac{1}{2}\zeta_2)\right\} \\
&= \{\zeta_1,\zeta_2 \geq 0 \mid \zeta_1 + \zeta_2 = 1\}.
\end{aligned}$$

Notice that there is no unique maximizer to (12) for this problem. If we take extreme point $(\zeta_1^*,\zeta_2^*) = (1,0)$, then the sampled version of the primal formulation is

$$\min_{(x_1,x_2)\in\mathcal{X},\bar{y}} \left\{-\bar{y} \mid -1 + x_1 + \bar{y}^2 \leq 0\right\},$$

which has optimal objective value of -1, which is strictly lower than the optimal solution to the original problem of $-\frac{1}{\sqrt{2}}$. Hence, in contrast to instances with only right-hand-side uncertainty, the sampled version of the dual formulation can give tighter lower bounds with left-hand-side uncertainty.

3. Example 1: distribution on a network with commitments

3.1. Problem formulation

This problem is adapted from Bertsimas and de Ruiter (2016). For the distribution on a network we determine the stock allocation x_i for location i, and the contracted transporting units z_{ij} from location i to location j, $i,j=1,...,N$, prior to knowing the realization of the demand at each location. The demand ζ is uncertain and assumed to be in a budget uncertainty set:

$$\mathcal{U}=\left\{\zeta\geq 0\ :\ \zeta\leq\hat{\zeta},\ e^\top\zeta\leq\Gamma\right\},$$

where $\hat{\zeta}_i\in\mathbb{R}_+$ denotes the maximum demand at location i, $i=1,...,N$, and $\Gamma\in\mathbb{R}_+$ denotes the maximum total demand. After we observe the realization of the demand we can transport stock y_{ij} from location i to location j at cost t_{ij} in order to meet all demand, $i,j=1,...,N$. The aim is to minimize the worst case total costs, which includes the storage costs (with unit costs c_i), the cost arising from shifting the products from one location to another (after the demands are realized), and the cost from violating the committed contract. A contract is violated if the transporting units y_{ij} differentiate from the committed units z_{ij}, $i,j=1,...N$. This distribution model can now be written as a specific instance of the primal problem as follows:

$$\inf_{\substack{x\in\mathcal{X}\\ z,\tau}}\sup_{\zeta\in\mathcal{U}}\inf_{y\geq 0}\left\{\sum_{i=1}^N c_ix_i+\tau\ \middle|\ \begin{array}{l}\sum_{i,j=1}^N t_{ij}y_{ij}+\frac{1}{2}\sum_{i,j=1}^N t_{ij}(y_{ij}-z_{ij})^2\leq\tau\\ \sum_{j=1}^N y_{ji}-\sum_{j=1}^N y_{ij}\geq\zeta_i-x_i\quad i=1,...,N\end{array}\right\},\tag{1}$$

where the quadratic terms in the first constraint captures the cost of contract violation, and $\mathcal{X}=\left\{x\in\mathbb{R}_+^N\mid e^\top x\geq\Gamma,\ x_i\leq K_i\quad i=1,...,N\right\}$. The set of linear constraints in (1) are the balance equations: we have to shift stock to and from location i such that the initial storage plus the net shift in stock still exceeds the demand at i. The constraints in $\mathcal{X}$ restrict the capacity of the stock at location i to at most K_i, $i=1,...,N$, as well as the total stock to be at least the maximum demand. The dualized formulation we obtain after consecutive dualization over the adjustable variables y and the uncertain parameters ζ is given below:

$$\inf_{\substack{x\in\mathcal{X}\\ z,\tau}}\sup_{(u,v,w)\in\mathcal{V}}\inf_{\lambda\geq 0}\left\{\sum_{i=1}^N c_ix_i+\tau\ \middle|\ \begin{array}{l}\sum_{i=1}^N(\hat{\zeta}\lambda_i-u_ix_i)+\Gamma\lambda_0\ \ldots\\ \qquad\qquad-\sum_{i,j=1}^N\left[(u_j-u_i-t_{ij}-v_{ij})z_{ij}+\frac{1}{2}w_{ij}\right]\leq\tau\\ \lambda_0+\lambda_i\geq u_i\qquad i=1,...,N\end{array}\right\},\tag{2}$$

where $\mathcal{V}=\{(u,v,w)\geq 0:\ (u_i-u_j+v_{ij}-t_{ij})^2\leq w_{ij}t_{ij}\quad\forall i,j=1,...N\}$. Note that in both problem formulations (1) and (2), the epigraphical auxiliary variable τ can be eliminated, then it can be verified that the resulting formulations satisfy (strongly relative) complete recourse.

3.2. Numerical setting

We choose $N \in \{5, 10, 20, 30, 40, 50, 60\}$ locations uniformly at random from $[0, 10]^2$. Let t_{ij}, the cost to transport one unit of demand from location i to j, be the Euclidean distance. The unit storage cost c_i are equal to 6 for $i = 1, ..., \lceil N/10 \rceil + 1$ warehouses and 10 for $i = \lceil N/10 \rceil + 1, ..., N$ stores. The individual maximum demand $\hat{\zeta}$ and the capacity K_i, $i = 1, ..., N$, of each location is set to 30 units. The total demand in the network is set to $20\sqrt{N}$. As an illustration, Figure 1 depicts a distribution on a network obtained from solving (2) with linear decision rules, which takes around 100s. All computations were carried out with MOSEK 8.0 (MOSEK ApS, 2017) on an Intel Core(TM) i5-4590 Windows computer running at 3.30GHz with 8GB of RAM. All modeling was done using the modeling package XProg (`http://xprog.weebly.com`). All the reported numbers in the tables are the average of 10 randomly generated instances.

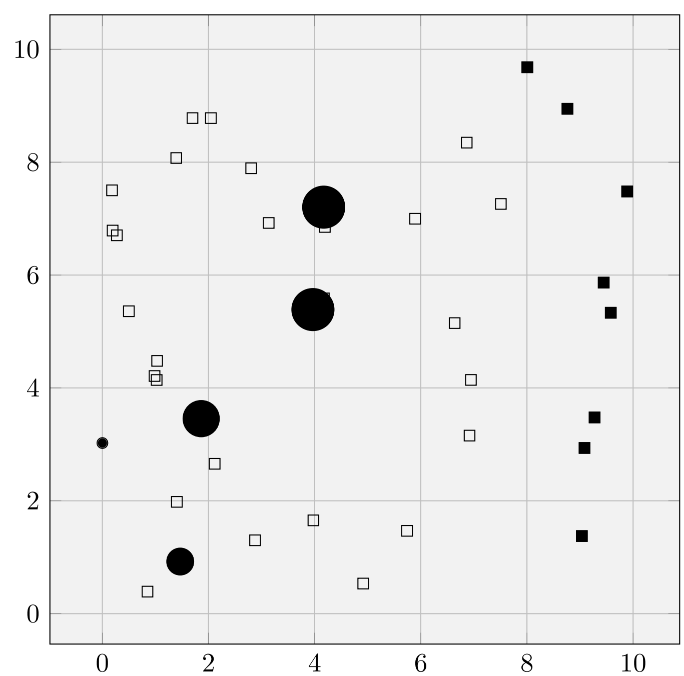

Figure 1 Stock allocation for $N = 40$ with 35 stores (squares) and 5 warehouses (circles) for one random instance. The filled dots have stock and the larger the dots are, the more stock is allocated.

3.3. Results on linear decision rules and bounds

For all cases we use linear decision rules to find solutions to the dualized model (2). For the smaller instance sizes, we also use the Fourier-Motzkin elimination of up to 10 adjustable variables. Since the model only has right-hand-side uncertainty, Theorem 3 states that we only have to use primal sampled scenarios to obtain a strong lower bound. The results are depicted in Table 1.

Table 1 Numerical results for distribution on a network with commitments. Objective values for the linear decision rule solution, Fourier-Motzkin up to 10 variable elimination and the sampled lower bound model using primal scenarios are depicted. The time refers to the computing time for the linear decision rule solution. All the numbers are the average of 10 randomly generated instances.

N	5	10	20	30	40	50	60
Linear decision rules	703	1029	1377	1606	1790	1962	2115
Fourier-Motzkin	635	944	1350	$*$	$*$	$*$	$*$
Lower bound	632	935	1272	1495	1681	1856	2004
Time linear decision rules (s)	<0.1	0.3	14	31	118	337	665

We observe that in all cases the linear decision rules give good performance, with objective values within 10% of the lower bound for the smaller cases and within 5% for the larger cases. Furthermore, the nonlinear models with linear decision rules can all be solved within seconds for the smaller cases. For a larger number of locations the number of variables grows quadratically, which explains the computation time increases to several minutes for $N=60$. The Fourier-Motzkin elimination only find a solution within one hour of computation time for $N \leq 20$, but has the potential to get the solution closer to the lower bound.

Remark 1. It follows from Theorem 2 that the infimum of (1) *with primal linear decision rules lower bounds that of* (2) *with dual linear decision rules. We propose to numerically evaluate the difference between the obtained infima from solving* (1) *and* (2) *with linear decision rules for* $N=5$*. From Table 1 we observe that for* $N=5$*, the average optimal value of* (2) *with dual linear decision rules is* 703*. By a vertex enumeration, we obtain the average optimal value of* (1) *with primal linear decision rules, i.e.,* 665*. The obtained average optimal values from solving* (1) *and* (2) *with linear decision rules are suboptimal because they are higher than that from Fourier-Motzkin elimination (i.e.,* 635*, see Table 1).*

4. Example 2: worst-case energy configuration of system with piecewise-linear springs

4.1. Problem formulation

The problem described in this section is adopted from Lobo et al. (1998). We consider a mechanical system that consists of N nodes at positions $x_1, \ldots, x_N \in \mathbb{R}^2$, with node i connected to node $i+1$, for $i = 1, \ldots, N-1$, by a nonlinear spring. The nodes x_1 and x_N are fixed at given values a and b, respectively. The tension in spring i is a nonlinear function of the distance between its endpoints, i.e., $\|x_i - x_{i+1}\|_2$:

$$s\left(\|x_i - x_{i+1}\|_2 - l_i^0\right)_+ ,$$

where $z_+ = \sup\{z, 0\}$, $s \in \mathbb{R}_+$ is the stiffness of the springs, and $l_i^0 \in \mathbb{R}_+$ is the natural (no tension) length of spring i. In this model the springs can only produce positive tension (which would be the

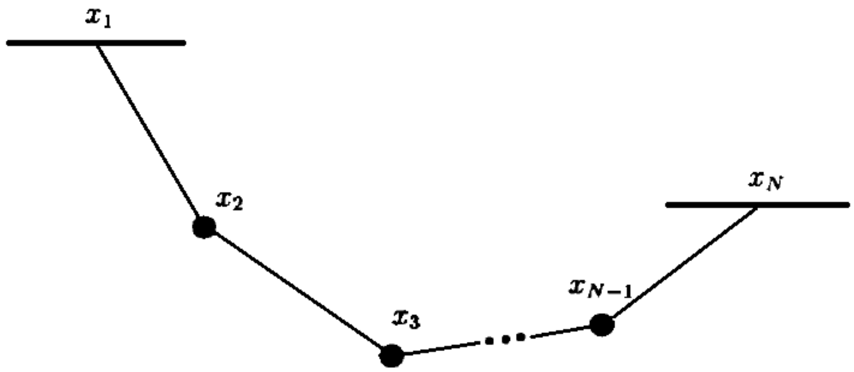

Figure 2 System of nodes (weights) connected by springs from Lobo et al. (1998). The first and last node positions, *i.e.*, x_1 and x_N, are fixed.

case if they buckled under compression). Each node has a mass of weight w attached to it. This is shown in Figure 2. The problem is to compute the equilibrium configuration of the system, *i.e.*, values of $x_1, \ldots, x_N$ such that the net force on each node is zero. This can be done by finding the minimum energy configuration, *i.e.*, solving a second-order cone optimization problem:

$$\begin{aligned} &\inf_{x\geq 0} \; w\sum_{i=1}^{N} x_{i2} + \frac{s}{2}\sum_{i=1}^{N-1}\left[\left(\|x_i - x_{i+1}\|_2 - l_i^0\right)_+\right]^2 \\ &\text{s.t.} \quad x_1 = a, \quad x_N = b, \end{aligned} \tag{3}$$

where x_{i2} is the second element of the vector x_i. For more detailed description of this problem, we refer to the original paper Lobo et al. (1998). Suppose the length of the springs are uncertain. The uncertainty may arise due to variations in the production process. Of course other parameters, *e.g.*, weight(w), stiffness(s), initial location of x_1 and x_N, may also be uncertain. Here we focus on uncertainty in the length of the springs, *i.e.*, $l(\zeta) = l^0 - \zeta$ (because only positive tension is considered), and the uncertain parameter $\zeta \in \mathbb{R}^{N-1}$ resides in a budget uncertainty set:

$$\mathcal{U} = \left\{\zeta \geq 0 \;:\; \zeta \leq \hat{\zeta},\; e^\top \zeta \leq \Gamma\right\},$$

where $\hat{\zeta}_i \in \mathbb{R}_+$ denotes the maximum deviation from the nominal length l^0 of spring i, $i = 1, ..., N-1$, and $\Gamma \in \mathbb{R}_+$ denotes the maximum total deviation of the springs. The minimum energy configuration model becomes a robust optimization model:

$$\inf_{x\geq 0}\sup_{\zeta\in\mathcal{U}} \left\{ w\sum_{i=1}^{N} x_{i2} + \frac{s}{2}\sum_{i=1}^{N-1}\left[\left(\|x_i - x_{i+1}\|_2 - l_i(\zeta)\right)_+\right]^2 \;\middle|\; x_1 = a, \quad x_N = b \right\}, \tag{4}$$

which can be rewritten as a two-stage robust optimization problem:

$$\inf_{x\geq 0}\sup_{\zeta\in\mathcal{U}}\inf_{y\geq 0} \left\{ w\sum_{i=1}^{N} x_{i2} + \frac{s}{2}\sum_{i=1}^{N-1} y_i^2 \;\middle|\; \begin{array}{l} \|x_i - x_{i+1}\|_2 - l_i(\zeta) \leq y_i \;\; i = 1, ..., N-1 \\ x_1 = a, \quad x_N = b \end{array} \right\}. \tag{5}$$

It can be verified that models (4) and (5) are equivalent, that is, eliminating all the y_i's for $i = 1, \dots, N-1$ in (5) we obtain (4). We solve the dualized formulation of (5) via linear decision rules. Note that here the strong relatively complete recourse assumption is satisfied.

4.2. Numerical setting

We consider $N \in \{15, 20, 30, 45, 60, 100\}$ nodes that are connecting $N-1$ springs. The nodes x_1 and x_N are fixed at given values $a = (0, 90)$ and $b = (100, 50)$, respectively. The natural (no tension) nominal length is $l_i^0 = 1 + \epsilon_i$, where ϵ_i is a random number drawn from a uniform distribution $U(0, 4)$, $i = 1, \dots, N-1$, and the stiffness of the springs is $s = 2$. Each node has a mass of weight $w = \frac{1}{10}$ attached to it. The upper-bound $\hat{\zeta}_i$ is set at $15\% l_i^0$ for $i = 1, \dots, N-1$, and $\Gamma = \frac{1}{2} e^\top \hat{\zeta}$. The computations is carried out with MOSEK 8.1 (MOSEK ApS, 2017) on an Intel(R) Xeon(R) E3-1241 v3 Windows computer running at 3.50GHz with 16GB of RAM. All modeling was done using the modeling package XProg (`http://xprog.weebly.com`).

4.3. Results

Figure 3 illustrates the static and robust locations of the nodes for $N = 45$, which shows that in order to minimize energy configuration under length uncertainty, in the solution from linear decision rules consecutive nodes are placed closer to each other than in the solution from static decision rules. Figure 4 depicts the robust locations obtained from solving the dualized model of (5) with linear decision rules. It shows that as N increases, the curvature of the connection between x_1 and x_N becomes severer; if N is large enough, *i.e.*, $N = 100$, then there are too many nodes with positive weights, all the useless nodes will simply be closely placed on the ground.

Since the infimum obtained from linear decision rule coincides with its lower bound (LB-P) in Table 2, which implies that the approximations from solving the dualized model of (5) via linear decision rules are tight. For small N, we observe that the infimum from primal static decision rule values are larger than that from solving the dualized model of (5) via linear decision rules, which means that the approximated solutions obtained via static decision rules are suboptimal. Since the robust problem (5) becomes easier to solve as N becomes larger, the infimum from primal static decision rule values becomes closer to that from solving the dualized model of (5) with linear decision rules.

5. Distribution on a network without commitments

Consider the linear variant of Problem (1) considered in §3, that is, distribution on a network *without* commitments, which can be written as the following two-stage robust linear optimization problem:

$$\inf_{x \in \mathcal{X}} \sup_{\zeta \in \mathcal{U}} \inf_{y \geq 0} \left\{ \sum_{i=1}^{N} c_i x_i + \sum_{i,j=1}^{N} t_{ij} y_{ij} \;\middle|\; \sum_{j=1}^{N} y_{ji} - \sum_{j=1}^{N} y_{ij} \geq \zeta_i - x_i \quad i = 1, ..., N \right\}. \tag{6}$$

Table 2 Equilibrium of system with $N-1$ piecewise-linear springs for $N \in \{15, 20, 30, 45, 65, 100\}$. Objective values for the primal static decision rule solution, the linear decision rule solution and the sampled lower bound model using primal scenarios are depicted. The time refers to the computing time for the linear decision rule solution.

N	15	20	30	45	65	100
Static decision rule	535.8	344.5	254.1	213.7	189.1	180.7
Linear decision rule	507.6	320.2	239.0	202.8	183.5	180.7
Lower bound	507.6	320.2	239.0	202.8	183.5	180.7
Time(s)	0.05	0.06	0.13	0.42	1.50	4.97

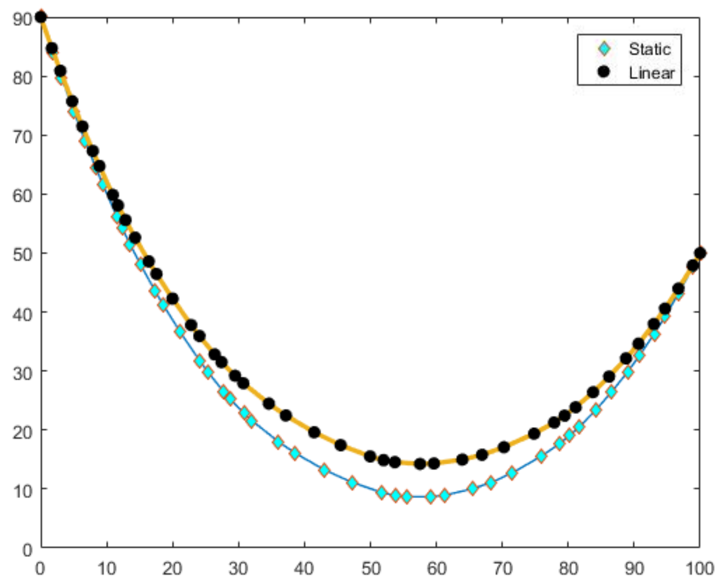

Figure 3 System of nodes (weights) connected by 44 springs for $N = 45$. The diamonds and dots represent the robust locations of the nodes from solving (5) and its dualized formulation with static decision rules and linear decision rules, respectively.

We choose $N \in \{10, 20, 30, 40, 50\}$ locations uniformly at random from $[0, 10]^2$. Let t_{ij}, the cost to transport one unit of demand from location i to j, be the Euclidean distance. The unit storage cost c_i are equal to 10 for $i = 1, ..., N$ stores. The individual maximum demand $\hat{\zeta}$ and the capacity K_i, $i = 1, ..., N$, of each location is set to 20 units. The total demand in the network is set to $20\sqrt{N}$. Note that here we consider the exact same problem setting as in Bertsimas and de Ruiter (2016), and compare the numerical performance of the lower bounding scheme proposed in §4 with the primal-dual lower bounding scheme proposed in Bertsimas and de Ruiter (2016). Table 3 reports the numerical results. One can observe that the optimality gaps obtained from our method almost halve the ones obtained from using the technique of Bertsimas and de Ruiter (2016), where the

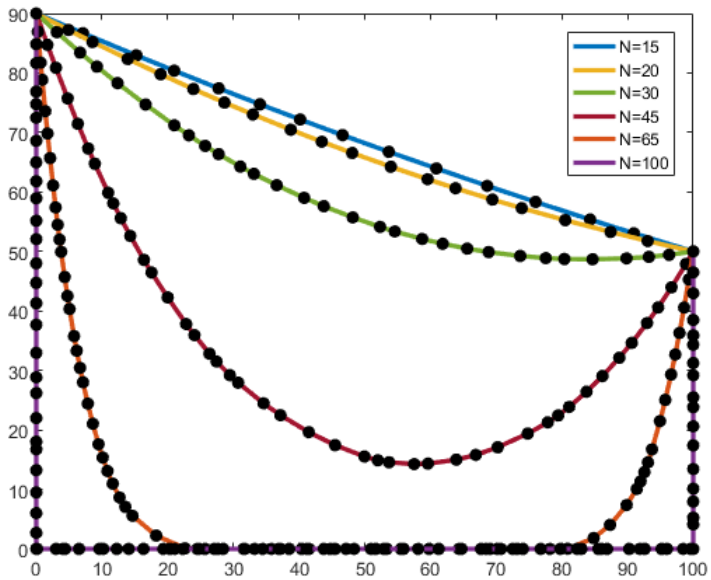

Figure 4 System of nodes (weights) connected by $N-1$ springs for $N \in \{15, 20, 30, 45, 65, 100\}$. The dots represent the robust location of the nodes.

optimality gap is computed via:

$$p\% = \frac{v(\mathrm{DL}) - v(\mathrm{LB})}{v(\mathrm{DL})} \times 100\%,$$

where $v(\cdot)$ denotes the optimal value of the corresponding problems, *e.g.*, $v(\mathrm{DL})$ is the optimal value obtained from solving the dualized model of (6) with linear decision rules, while LB$\in\{$LB-P, LB-BR$\}$.

Table 3 Lot-sizing problem with $N \in \{10, 20, 30, 40, 50\}$. LB-P and LB-BR denotes the approximated optimality gap using the primal scenarios (see §4) and using the primal and dual scenarios in Bertsimas and de Ruiter (2016). All the numbers are the average of 10 randomly generated instances.

N	10	20	30	40	50
LB-P	6.0%	5.7%	6.0%	5.0%	5.2%
LB-BR	13.9%	12.9%	10.4%	11.2%	10.7%

References

D. Bertsimas and F. de Ruiter. Duality in two-stage adaptive linear optimization: faster computation and stronger bounds. *INFORMS Journal on Computing*, 28(3):500–511, 2016.

M. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret. Applications of second-order cone programming. *Linear Algebra and its Applications*, 284(1-3):193–228, 1998.

MOSEK ApS. *The MOSEK optimization toolbox for MATLAB manual. Version 8.1*, 2017. URL `http://docs.mosek.com/8.1/toolbox.pdf`.

K. Roos, M. Balvert, B. Gorissen, and D. den Hertog. A universal and structured way to derive dual optimization problem formulations. *INFORMS Journal on Optimization*, 2(4):229–255, 2020.