Tractable approximation of hard uncertain optimization problems

Ernst Roos
ORTEC, Zoetermeer, The Netherlands, ernst.roos@ortec.com

Aharon Ben-Tal
Faculty of Industrial Engineering and Management, Technion - Israel Institute of Technology & Shenkar College, Israel

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

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

Jianzhe Zhen
School of Economics and Management, University of Chinese Academy of Sciences, Beijing, China

In many optimization problems uncertain parameters appear in a convex way, which is problematic as common techniques can only handle concave uncertainty. In this paper, we provide a systematic way to construct conservative approximations to such problems. Specifically, we reformulate the original problem as an adjustable robust optimization problem in which the nonlinearity of the original problem is captured by the new uncertainty set. This adjustable robust optimization problem is linear both in the variables and the uncertain parameters whenever the original uncertainty set is polyhedral, which allows for the application of a multitude of techniques from adjustable robust optimization. Our approach can be applied to a wealth of constraints, including constraints that are convex quadratic, sum-of-max (with and without square), log-sum-exp, norms, and negative entropy.

Key words: robust optimization, nonlinear inequality, convex analysis

1. Introduction

1.1. Problem Formulation

We consider a general convex constraint given by

$$f\left(A(x)\zeta + b(x)\right) \leq 0 \qquad \forall \zeta \in U, \tag{1}$$

where $f : \mathbb{R}^p \to [-\infty, +\infty]$ is proper, closed and convex, $A : \mathbb{R}^n \to \mathbb{R}^{p\times L}$ and $b : \mathbb{R}^n \to \mathbb{R}^p$ are affine. For ease of exposure, we initially consider nonempty polyhedral uncertainty sets

$$U = \left\{ \zeta \in \mathbb{R}^L_+ \;\middle|\; D\zeta = d \right\}, \tag{2}$$

for some $D \in \mathbb{R}^{q\times L}$ and $d \in \mathbb{R}^q$. We also show that we can extend our results to general convex uncertainty sets, although results are more conservative in that case.

This formulation allows for many important classes of constraints, such as convex quadratic, sum-of-max (with and without square), log-sum-exp functions, norms and negative entropy. Furthermore, all functions $g(x, \zeta)$ that are jointly convex in x and ζ can be written as $f\left(A(x)\zeta + b(x)\right)$, by choosing

$$A(x) = \begin{bmatrix} O \\ I \end{bmatrix}, \qquad b(x) = \begin{bmatrix} x \\ 0 \end{bmatrix}.$$

Despite its generality, it's worth noting that we cannot handle all constraints convex in both x and ζ by choosing to consider the function $f(A(x)\zeta + b(x))$. For instance, functions of the form $b(x)^\top g(\zeta)$, where b is an affine function and g is a convex convex function, cannot be formulated as $f(A(x)\zeta + b(x))$. Examples of such functionscan be found in brachytherapy optimization (Gorissen et al. 2013). Other examples include capital budgeting problems and multinomial logit models (Alfandari and García 2018).

1.2. Literature

Robust models with constraints that are *convex* in the uncertain parameters are, in general, hard (Chassein and Goerigk 2019) but are common in models. They appear, for example, in inventory management problems, geometric programming and conic optimization. Therefore, several papers focus on specific type of robust convex constraints.

In Bertsimas and Sim (2006), a method for norm-based uncertainty sets and homogeneous constraint functions is proposed. We show in Section 3 that their method is a special case of our approach. Even more, for polyhedral uncertainty sets we show that our approach is tighter. Our

approach can also be seen as a generalization of the approach in Zhen et al. (2022). There, second-order cone and semidefinite programming constraints are reformulated to an adjustable robust optimzation model. However, that approach does not handle general convex constraints mentioned above, nor can it be extended to non-polyhedral uncertainty sets. When the function f is a sum-of-max-of-linear functions, then the method proposed by Ben-Tal et al. (2005) and Gorissen and Den Hertog (2013) coincides with our method. There, they propose to use linear decision rules. It can be shown that our method together with linear decision rules coincides with their approximation.

Also the seminal papers on robust optimization (El Ghaoui and Lebret (1997), El Ghaoui et al. (1998), Ben-Tal et al. (2002)) pay attention to specific cases of constraints with convex uncertainty. These papers consider the special combination of (conic) quadratic constraint functions and ellipsoidal types of uncertainty sets. For this specific case, their approach is exact. However, the final problem that needs to be solved with their approach is a semidefinite optimization program. Our approach is a conservative approximation for these cases, but is computationally far less demanding as it yields a conic quadratic optimization problem.

1.3. Contributions

This paper presents a general and extended approach to general uncertain convex constraints from the literature. The major contributions of this paper are:

- We show that an uncertain convex constraint (1) can be written as a set of linear adjustable robust constraints for polyhedral uncertainty sets.
- Using linear decision rules we obtain a tractable conservative approximation for the adjustable robust model. Furthermore, we show that this single approach is either equivalent to or tighter than the approach proposed in Bertsimas and Sim (2006).
- We extend our approach to convex constraints with non-polyhedral uncertainty sets and show that the resulting nonlinear adjustable robust constraints can be solved with static policies.

1.4. Paper organization and notation

The paper is organized as follows. Section 2 treats the approach for (1) for polyhedral uncertainty sets. In Section 3 we extend the results from Section 2 to general convex uncertainty sets and provides additional theory that compares our approach with the existing literature. The Electronic Companion contains numerical results on problems with these hard convex constraints: a robust geometric programming problem and a practical radiotherapy optimization example.

Notation. Throughout this paper we use the following notation.

Let $f:\mathbb{R}^n \to \mathbb{R}$ be a closed convex function with domain $\mathrm{dom}(f) = \{x \mid f(x) < \infty\}$. The *convex conjugate*, which we refer to as conjugate, of f is defined as

$$f^*(y) = \sup_{x \in \mathrm{dom}(f)} \left\{ y^\top x - f(x) \right\}.$$

The *support function* of a set U is the conjugate of that set's *indicator function*. This indicator function is defined as:

$$\delta\left(x \mid U\right) = \begin{cases} 0 & \text{if } x \in U \\ \infty & \text{otherwise,} \end{cases}$$

and thus the support function is given by

$$\delta^*\left(y \mid U\right) = \sup_{x \in U} y^\top x.$$

The *perspective function* of f is defined by

$$\left(f\lambda\right)(x) = \begin{cases} \lambda f\left(\frac{x}{\lambda}\right) & \lambda > 0 \\ f_\infty(x) & \lambda = 0, \end{cases}$$

where f_∞ is the *recession function* of f, defined by (Rockafellar 1970, Theorem 13.3):

$$f_\infty(y) = \delta^*\left(y \mid \mathrm{dom}\left(f^*\right)\right). \tag{3}$$

We remark that the alternative definition of the recession function in (Rockafellar 1970, Corollary 8.5.2), i.e.,

$$f_\infty(x) = \lim_{\lambda \downarrow 0} \lambda f\left(\frac{x}{\lambda}\right), \tag{4}$$

may not be closed. For the ease of exhibition, we write $\lambda f\left(\frac{x}{\lambda}\right)$ for the perspective function throughout the rest of the paper, implicitly assuming that for $\lambda = 0$, we consider the definition in (3). A table of recession functions of some well-known functions, as well as composition rules, can be found in Appendix A.

2. The Robust Counterpart

2.1. Reformulation to ARO

In order to find a conservative approximation to (1), we first transform the problem into an equivalent linear adjustable robust optimization problem.

THEOREM 1. *Let $f : \mathbb{R}^p \to [-\infty, +\infty]$ be a proper, closed and convex function and let $A : \mathbb{R}^n \to \mathbb{R}^{p\times L}$ and $b : \mathbb{R}^n \to \mathbb{R}^p$ be affine functions. Let $U \subseteq \mathbb{R}^L$ be a polyhedron as defined in* (2)*. Then, $x \in \mathbb{R}^n$ satisfies*

$$f\left(A(x)\zeta + b(x)\right) \leq 0 \qquad \forall \zeta \in U,$$

if and only if it satisfies the following set of adjustable robust optimization constraints:

$$\forall w \in \operatorname{dom}\left(f^*\right),\ \exists \lambda \in \mathbb{R}^q : \begin{cases} d^\top \lambda + w^\top b(x) - f^*(w) \leq 0 & \text{(5a)} \\ D^\top \lambda \geq A(x)^\top w. & \text{(5b)} \end{cases}$$

Proof. Because f is a closed convex function we have that

$$f(z) = f^{**}(z) = \sup_{w \in \operatorname{dom} f^*} \left\{ z^\top w - f^*(w) \right\}.$$

Substituting this into (1) yields

$$\begin{aligned}
& & \forall \zeta \in U \ :\ f\left(A(x)\zeta + b(x)\right) \leq 0 \\
\iff & & \forall \zeta \in U \ :\ \sup_{w \in \operatorname{dom} f^*} \left\{ \left(A(x)\zeta + b(x)\right)^\top w - f^*(w) \right\} \leq 0 \\
\iff & & \sup_{\zeta \in U} \left\{ \sup_{w \in \operatorname{dom} f^*} \left\{ \left(A(x)^\top w\right)^\top \zeta + b(x)^\top w - f^*(w) \right\} \right\} \leq 0 \\
\iff & & \sup_{w \in \operatorname{dom} f^*} \left\{ \sup_{\zeta \in U} \left\{ \left(A(x)^\top w\right)^\top \zeta \right\} + b(x)^\top w - f^*(w) \right\} \leq 0 \qquad (6)
\end{aligned}$$

$$\iff \qquad \sup_{w \in \operatorname{dom} f^*} \left\{ \inf_{\lambda \in \mathbb{R}^q} \left\{ d^\top \lambda \mid D^\top \lambda \geq A(x)^\top w \right\} + b(x)^\top w - f^*(w) \right\} \leq 0 \tag{7}$$

$$\iff \qquad \forall w \in \operatorname{dom} f^*,\ \exists \lambda \in \mathbb{R}^q \ : \ \begin{cases} d^\top \lambda + b(x)^\top w - f^*(w) \leq 0 \\ D^\top \lambda \geq A(x)^\top w, \end{cases}$$

where (6) and (7) are equivalent because of strong LP duality. □

In the field of Robust Optimization the variable λ in (5) is referred to as an adjustable variable, as its value can be adjusted after the value of the uncertain parameter w is known. We note that a similar result holds if the nonnegativity constraint in U is omitted. Then the inequality in (5b) becomes an equality constraint, which can be used to eliminate some adjustable variables λ. Eliminating variables in this way is equivalent to imposing linear decision rules for those variables (Zhen and den Hertog 2017, Lemma 2). We also remark that Theorem 1 does not rely on A and b being affine in x.

The robust constraint (1) allows for many related robust optimization constraints. One particularly interesting setting is where the original constraint contain adjustable variables, that is:

$$g\left(x, y(\zeta)\right) \leq 0 \qquad \forall \zeta \in U, \tag{8}$$

where g is jointly convex in x and y for adjustable variables y, that is, jointly convex constraints in adjustable robust optimization problems. Specifically, such constraints can be treated by substituting a linear decision rule $y = s + S\zeta$, such that $g\left(x, s + S\zeta\right)$ can be written as $f\left(A(x,S)\zeta + b(x,s)\right)$. We do remark that substituting a linear decision rule for y yields a conservative approximation to (8), and thus our approach yields a conservative approximation to this conservative approximation of (8).

2.2. Conservative Approximation

Tractable conservative solutions to (5) can be constructed using affine decision rules as the resulting problem is linear in w. To overcome some of the conservativeness, we can lift the nonlinearity of $f^*(w)$ to the uncertainty set and using a slightly more involved decision rule. Lifting a set has been applied in adaptive *distributionally* robust settings by Bertsimas et al. (2019) in a similar spirit.

THEOREM 2. *If there exist $u \in \mathbb{R}^q$, $V \in \mathbb{R}^{q\times p}$ and $r \in \mathbb{R}^q$ for a given $x \in \mathbb{R}^n$ such that*

$$\begin{cases} d^\top u + (1+d^\top r)\, f\left(\frac{V^\top d + b(x)}{1+d^\top r}\right) \le 0 & \text{(9a)} \\ 1 + d^\top r \ge 0 & \\ -D_i^\top u + (-D_i^\top r)\, f\left(\frac{A_i(x) - V^\top D_i}{-D_i^\top r}\right) \le 0 & \\ -D_i^\top r \ge 0 & \end{cases} \quad i = 1,\dots,L, \qquad \text{(9b)}$$

holds, then x also satisfies (1). *Here, D_i and $A_i(x)$ denote the i-th column of D and $A(x)$, respectively.*

The resulting system of inequalities (9) is convex because the perspective function of convex functions are convex, and it includes $2q + qp + n$ variables compared to the original n in (1). The proof of Theorem 2 using these lifted decision rules can be found in Appendix B.

We would like to note that for certain special classes of function f, the set of inequalities in (9) remains convex even if A and b are not affine in x. For instance, if f is also non-decreasing or piece-wise affine, then the constraints in (9) are convex as long as A and b are convex in x.

The conservative approximation derived in Theorem 2 is generally tighter than that in Theorem 1. The constraints in (9) are represented by the perspective of the original constraint function f and its tractability is thus highly reliant on the tractability of this perspective. A disadvantage of perspective functions is that they can lead to numerical issues in practice (Jung et al. 2013). One way to overcome this, is to set $r = 0$ and end up with a problem without perspective functions and the recession function of f in (9). However, for conically representable f, we know that the perspective is conically representable as well, and hence these numerical issues can be circumvented by using this conic reformulation. We refer to Appendix C for the mathematical proof of this statement.

3. Extension to general convex uncertainty sets

In this section, we consider (1) with a general convex uncertainty set that is given by

$$U = \left\{ \zeta \in \mathbb{R}^L_+ \,\middle|\, h_\ell(\zeta) \le 0 \quad \ell = 1,\dots,q \right\}, \qquad (10)$$

where $h_\ell : \mathbb{R}^p \to [-\infty, +\infty]$ is proper, closed and convex. Restricting the uncertainty set U to the nonnegative orthant $\mathbb{R}^L_+$ can be done without loss of generality because we can always lift the set U into $\mathbb{R}^L_+$ by setting $\zeta = \zeta^+ - \zeta^-$ in (1), where $\zeta^+, \zeta^- \in \mathbb{R}^L_+$, and incorporate the non-convex projection of Bertsimas and Sim (2006) if U is norm-based.

In case the set U is conic quadratic representable, we can approximate the set U by a polyhedron using the work by Ben-Tal and Nemirovski (2001). After having the polyhedral description, all techniques from the previous section can be applied. We note that the polyhedral approximation is polynomial in the dimension of the conic quadratic representation of the set, as well as $\frac{1}{\epsilon}$, where ϵ is the accuracy of approximation. Furthermore, a large value of ϵ, and therefore a crude approximation of the uncertainty set, is often acceptable as the uncertainty set is a modelers' choice and not a hard constraint. There is another option, which can also be applied for sets that are not conic quadratic representable, and is outlined in the theorem below. The proof of Theorem 3 can be found in Appendix B.

THEOREM 3. *Let $f : \mathbb{R}^p \to [-\infty, +\infty]$ be a proper, closed and convex function and let $A : \mathbb{R}^n \to \mathbb{R}^{p \times L}$ and $b : \mathbb{R}^n \to \mathbb{R}^p$ be affine functions. Let $U \subseteq \mathbb{R}^L_+$ be a convex set as defined in* (10) *and* $\mathrm{ri}(U) \neq \emptyset$. *Then, $x \in \mathbb{R}^n$ satisfies*

$$f\left(A(x)\zeta + b(x)\right) \leq 0 \qquad \forall \zeta \in U,$$

if and only if it satisfies the following set of adjustable robust optimization constraints:

$$\forall (w_0, w) \in W,\ \exists \lambda \in \mathbb{R}^q_+, \{u_\ell\}_\ell \subset \mathbb{R}^p : \begin{cases} \displaystyle\sum_{\ell=1}^{q} \lambda_\ell h^*_\ell(u_\ell/\lambda_\ell) + b(x)^\top w + w_0 \leq 0 & \text{(11a)} \\ A(x)^\top w \leq \displaystyle\sum_{\ell=1}^{q} u_\ell, & \text{(11b)} \end{cases}$$

where the uncertainty set W is defined by

$$W = \left\{ (w_0, w) \in \mathbb{R}^{p+1} \,\middle|\, w_0 + f^*(w) \leq 0 \right\}.$$

The adjustable variables $(\lambda, \{u_\ell\}_\ell)$ in (11) may appear in a nonlinear way, and imposing linear decision rules would again lead to robust constraints with convex uncertainties. In order to obtain

the tractable conservative approximation of (11), we treat the adjustable variables $(\lambda, \{u_\ell\}_\ell)$ in (11) as static instead of the more flexible affine decision rules. The derivation of this conservative approximation can be found in Appendix B.

THEOREM 4. *If there exist $\lambda \in \mathbb{R}^q_+$ and $u_\ell \in \mathbb{R}^p$, $\ell = 1, ..., q$, for a given $x \in \mathbb{R}^n$ such that*

$$\begin{cases} \sum_{\ell=1}^{q} \lambda_\ell h^*_\ell(u_\ell/\lambda_\ell) + f(b(x)) \leq 0 & \text{(12a)} \\ f_\infty(A_i(x)) \leq \sum_{\ell=1}^{q} u_{i\ell} \qquad i = 1, \ldots, L, & \text{(12b)} \end{cases}$$

holds, then x also satisfies (1) *with $U \subseteq \mathbb{R}^L$ as defined in* (10).

For a special case when $f(\cdot)$ is positively homogeneous, the set of constraints in (12) is equivalent to

$$\begin{cases} \sum_{\ell=1}^{q} \lambda_\ell h^*_\ell(u_\ell/\lambda_\ell) + f(b(x)) \leq 0 \\ f(A_i(x)) \leq \sum_{\ell=1}^{q} u_{i\ell} \qquad i = 1, ..., L \end{cases} \tag{13}$$

as $f_\infty = f$ in this case. The obtained set of finite convex constraints (13) is in fact the tractable reformulation of the following robust convex constraint

$$\sum_{i=1}^{L} \zeta_i f\left(A(x)_i\right) + f\left(b(x)\right) \leq 0 \qquad \forall \zeta \in U,$$

which coincides with the conservative approximation of Bertsimas and Sim (2006) for a robust convex constraint (1) where f is positively homogeneous. Since our approach does not require f to be homogeneous in (1), our approach generalizes the approach of Bertsimas and Sim (2006) to non-homogeneous functions. For (1) with a polyhedral uncertainty set, we propose to impose linear decision rules to adjustable variables in the adjustable robust *linear* reformulation (5), and the obtained approximation (9) is tighter than the one from (12) using static decision rules. Hence, in this case our approximations via linear decision rules are tighter than the ones from Bertsimas and Sim (2006). Additionally, our approach allows for a progressive approximation using a finite sample from W as described in Appendix D, while their method does not. Note that this progressive approximation works similarly for a general convex uncertainty set. We also remark that in case the

uncertainty set consists of both linear and nonlinear constraints, the approaches can be combined. In those situations, linear decision rules are used for the adjustable variables corresponding to the linear constraints and static rules for the other constraints.

Appendix A: Recession Functions

The recession function can be defined in multiple ways. In this paper, we mainly use it to concisely denote the support function of the domain of a function's conjugate. An advantage of the recession function besides concise notation is the relative ease of computing a recession function. Let $f^1, \ldots, f^m$ be convex, proper and lower semicontinuous functions. Then, the following composition rules for recession functions are valid (Auslender and Teboulle 2006, Proposition 2.6.1, 2.6.2):

1. Let f be defined by $f(x) = \sum_{i=1}^m f^i(x)$. Then $f_\infty(y) = \sum_{i=1}^m f^i_\infty(y)$;
2. Let f be defined by $f(x) = \sup_{i\in\{1,\ldots,m\}} f^i$. Then $f_\infty(y) = \sup_{i\in\{1,\ldots,m\}} f^i_\infty(y)$.

Moreover, if $g : \mathbb{R}^m \to [-\infty, +\infty]$ is a proper convex function, A is a linear map from $\mathbb{R}^n$ to $\mathbb{R}^m$ and $\psi : (-\infty, b) \to \mathbb{R}$ for $0 \le b \le +\infty$ is convex and nondecreasing with $\psi_\infty(1) > 0$ it holds that (Auslender and Teboulle 2006, Proposition 2.6.3, 2.6.4):

3. Let f be defined by $f(x) = g(Ax)$. Then $f_\infty(y) = g_\infty(Ay)$;
4. Let f be defined by
$$f(x) = \begin{cases} \psi(g(x)) & \text{if } x \in \operatorname{dom}(g) \\ +\infty & \text{otherwise.} \end{cases}$$
Then
$$f_\infty(y) = \begin{cases} \psi_\infty(f_\infty(y)) & \text{if } y \in \operatorname{dom}(f_\infty) \\ +\infty & \text{otherwise.} \end{cases}$$

Using the above composition rules as well as the recession functions of some often encountered basic functions f, one can directly find the recession function of the function of interest. An overview of some common recession functions is given in Table 1. It should be noted that the recession function is always conically representable, as its epigraph is the recession cone of the epigraph of f and thus is a cone by definition (Rockafellar 1970, p. 66). We additionally remark that for all

$f(x)$	$f_\infty(y)$
$\sqrt{1 + x^\top Q x}$ $(Q \succeq 0)$	$\sqrt{y^\top Q y}$
$x^\top Q x + q^\top x + c$ $(Q \succeq 0)$	$\begin{cases} q^\top y & \text{if } Qy = 0 \\ +\infty & \text{if } Qy \neq 0 \end{cases}$
$\log \sum_{i=1}^n e^{x_i}$ $(n > 1)$	$\max\{y_i \mid i = 1, \ldots, n\}$
$\sum_{i=1}^n \sqrt{1 + x_i^2}$	$\lVert y \rVert_2$
$\sum_{i=1}^m \max_{k\in K_i}\{x_k\}$	$\sum_{i=1}^m \max_{k\in K_i}\{y_k\}$
$\lVert x \rVert_2$	$\lVert y \rVert_2$

Table 1 Some examples of functions f with recession functions f_∞.

positively homogeneous functions of order one, or equivalently all functions such that $f^*(y)=0$ on its domain, it holds that $f_\infty(x)=f(x)$.

Appendix B: Proofs for Conservative Approximations

Proof of Theorem 2. From Theorem 1 we know (1) is equivalent to (5). We lift the nonlinear term $f^*(w)$ to the uncertainty set, that is, we introduce an auxiliary uncertain parameter w_0 such that (5) is equivalent to

$$\forall \begin{pmatrix} w_0 \\ w \end{pmatrix} \in W,\ \exists \lambda \in \mathbb{R}^q : \begin{cases} d^\top \lambda + w^\top b(x) + w_0 \le 0 & \text{(14a)} \\ D^\top \lambda \ge A(x)^\top w, & \text{(14b)} \end{cases}$$

where the new uncertainty set W is defined by

$$W = \left\{ \begin{pmatrix} w_0 \\ w \end{pmatrix} \in \mathbb{R}^{p+1} \,\middle|\, w_0 + f^*(w) \le 0 \right\}.$$

The support function of this new uncertainty set is essential for deriving a tractable robust counterpart and is equal to:

$$\begin{aligned} \delta^*\left(\begin{pmatrix} z_0 \\ z \end{pmatrix} \middle| W \right) &= \sup_{(w_0\ w)^\top \in W} \left\{ z_0 w_0 + z^\top w \right\} \\ &= \begin{cases} \sup_{w\in\mathbb{R}^p} \left\{ z^\top w - z_0 f^*(w) \right\} & \text{if } z_0 > 0 \\ \sup_{w \in \operatorname{dom} f^*} \left\{ z^\top w \right\} & \text{if } z_0 = 0 \\ +\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} z_0 \sup_{w\in\mathbb{R}^p} \left\{ w^\top \frac{z}{z_0} - f^*(w) \right\} & \text{if } z_0 > 0 \\ f_\infty(z) & \text{if } z_0 = 0 \\ +\infty & \text{otherwise} \end{cases} \\ &= \begin{cases} z_0 f\left(\frac{z}{z_0}\right) & \text{if } z_0 \ge 0 \\ +\infty & \text{otherwise.} \end{cases} \end{aligned} \tag{15}$$

Now, we once again use a linear decision rule for λ of the form

$$\lambda = u + Vw + r w_0, \tag{16}$$

where $u \in \mathbb{R}^q$, $V \in \mathbb{R}^{q\times p}$ and $r \in \mathbb{R}^q$, and thus we obtain a conservative approximation for (1). Substituting this decision rule in (14a) yields

$$\begin{aligned} & & d^\top \lambda + b(x)^\top w + w_0 \le 0 & & \forall \begin{pmatrix} w_0 \\ w \end{pmatrix} \in W \\ &\Longrightarrow & d^\top (u + Vw + r w_0) + b(x)^\top w + w_0 \le 0 & & \forall \begin{pmatrix} w_0 \\ w \end{pmatrix} \in W \\ &\Longleftrightarrow & d^\top u + \begin{pmatrix} w_0 \\ w \end{pmatrix}^\top \begin{pmatrix} 1 + d^\top r \\ V^\top d + b(x) \end{pmatrix} \le 0 & & \forall \begin{pmatrix} w_0 \\ w \end{pmatrix} \in W \\ &\Longleftrightarrow & d^\top u + \delta^*\left(\begin{pmatrix} 1 + d^\top r \\ V^\top d + b(x) \end{pmatrix} \middle| W \right) \le 0 & & \\ &\Longleftrightarrow & \begin{cases} d^\top u + (1 + d^\top r)^\top f\left(\frac{V^\top d + b(x)}{1 + d^\top r} \right) \le 0 \\ 1 + d^\top r \ge 0, \end{cases} & & \end{aligned} \tag{17}$$

where the last equivalence holds because of the definition of the support of W in (15). Note that (17) is exactly (9a). Similarly, substituting the linear decision rule for λ in (14b) we find

$$\begin{aligned}
& & D_i^\top \lambda \geq A_i(x)^\top w & \qquad \forall \begin{pmatrix} w_0 \\ w \end{pmatrix} \in W, \ i = 1, \ldots, L \\
& \Longrightarrow & -D_i^\top u + \begin{pmatrix} w_0 \\ w \end{pmatrix}^\top \begin{pmatrix} -D_i^\top r \\ A_i(x) - V^\top D_i \end{pmatrix} \leq 0 & \qquad \forall \begin{pmatrix} w_0 \\ w \end{pmatrix} \in W, \ i = 1, \ldots, L \\
& \Longleftrightarrow & -D_i^\top u + \delta^* \left(\begin{pmatrix} -D_i^\top r \\ A_i(x) - V^\top D_i \end{pmatrix} \middle| W \right) \leq 0 & \qquad i = 1, \ldots, L \\
& \Longleftrightarrow & \begin{cases} -D_i^\top u + (-D_i^\top r) \, f\left(\frac{A_i(x) - V^\top D_i}{-D_i^\top r} \right) \leq 0 \\ -D_i^\top r \geq 0 \end{cases} & \qquad i = 1, \ldots, L,
\end{aligned}$$

which is exactly (9b). □

Proof of Theorem 3. Because f is a closed convex function we have that

$$f(z) = f^{**}(z) = \sup_{w \in \operatorname{dom} f^*} \left\{ z^\top w - f^*(w) \right\}.$$

Substituting this into (1) yields

$$\begin{aligned}
& & \forall \zeta \in U \, : \, f\left(A(x)\zeta + b(x)\right) \leq 0 \\
& \Longleftrightarrow & \forall \zeta \in U \, : \, \sup_{w \in \operatorname{dom} f^*} \left\{ \left(A(x)\zeta + b(x)\right)^\top w - f^*(w) \right\} \leq 0 \\
& \Longleftrightarrow & \sup_{\zeta \in U} \left\{ \sup_{w \in \operatorname{dom} f^*} \left\{ \left(A(x)^\top w\right)^\top \zeta + b(x)^\top w - f^*(w) \right\} \right\} \leq 0 \\
& \Longleftrightarrow & \sup_{(w_0, w) \in W} \left\{ \sup_{\zeta \in U} \left\{ \left(A(x)^\top w\right)^\top \zeta \right\} + b(x)^\top w + w_0 \right\} \leq 0 \qquad (18) \\
& \Longleftrightarrow & \sup_{(w_0, w) \in W} \left\{ \inf_{\substack{\lambda \in \mathbb{R}^q_+ \\ \{u_\ell\}_\ell}} \left\{ \sum_{\ell=1}^q \lambda_\ell h_\ell^*(u_\ell/\lambda_\ell) + b(x)^\top w + w_0 \mid A(x)^\top w \leq \sum_{\ell=1}^q u_\ell \right\} \right\} \leq 0 \qquad (19) \\
& \Longleftrightarrow & \forall (w_0, w) \in W, \, \exists \lambda \in \mathbb{R}^q_+, \{u_\ell\}_\ell \subset \mathbb{R}^p : \begin{cases} \sum_{\ell=1}^q \lambda_\ell h_\ell^*(u_\ell/\lambda_\ell) + b(x)^\top w + w_0 \leq 0 \\ A(x)^\top w \leq \sum_{\ell=1}^q u_\ell, \end{cases}
\end{aligned}$$

where (18) and (19) are equivalent because of strong duality for convex optimization problems, which applies because $\operatorname{ri}(U) \neq \emptyset$. □

Proof of Theorem 4. From Theorem 3, we know that (1) is equivalent to (11). To obtain the conservative solution, we restrict the adjustable variables $(\lambda, \{u_\ell\}_\ell)$ to static decision rules, which yields

$$\sum_{\ell=1}^q \lambda_\ell h_\ell^*(u_\ell/\lambda_\ell) + b(x)^\top w + w_0 \leq 0 \qquad \forall (w_0, w) \in W$$

$$\begin{aligned}
\Longrightarrow \quad & \sum_{\ell=1}^{q} \lambda_\ell h_\ell^*(u_\ell/\lambda_\ell) + \sup_{(w_0,w)\in W} \left\{ b(x)^\top w + w_0 \right\} \le 0 \\
\Longleftrightarrow \quad & \sum_{\ell=1}^{q} \lambda_\ell h_\ell^*(u_\ell/\lambda_\ell) + f\left(b(x)\right) \le 0,
\end{aligned}$$

where the last equivalence follows from the definition of the conjugate. For the second constraint we find

$$\begin{aligned}
& A_i(x)^\top w \le \sum_{\ell=1}^{q} u_{i\ell} && \forall w \in \operatorname{dom} f^*,\ i=1,\dots,L \\
\Longrightarrow \quad & \delta^*\left(A_i(x) \,|\, \operatorname{dom} f^*\right) \le \sum_{\ell=1}^{q} u_{i\ell} && i=1,\dots,L \\
\Longleftrightarrow \quad & f_\infty\left(A_i(x)\right) \le \sum_{\ell=1}^{q} u_{i\ell} && i=1,\dots,L,
\end{aligned}$$

where the last equivalence follows from (3). □

Appendix C: Proof of Conically Representable Perspective

We use the definition of conically representable from Serrano (2015), that is, a function $f:\mathbb{R}^n \to [-\infty,+\infty]$ is proper and conically representable if its epigraph can be written as

$$\begin{aligned}
\operatorname{Epi} f = & \ \{(x,t) \ | \ f(x) \le t\} \\
= & \ \{(x,t) \ | \ \exists u \in \mathbb{R}^m,\ S(x,u,t)=0,\ T(x,u,t) \in \mathcal{K}\},
\end{aligned}$$

where $S:\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{k_1}$ and $T:\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \to \mathbb{R}^{k_2}$ are affine mappings and $\mathcal{K}$ is a cone.

THEOREM 5. *If f is conically representable, so is its perspective (fv).*

Proof. Let S,T be the affine mappings that define the conic representation of f and let $\mathcal{K}$ be the corresponding cone. Define $S^{per}:\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}^{k_1}$ and $T^{per}:\mathbb{R}^n \times \mathbb{R}^m \times \mathbb{R} \times \mathbb{R}_{++} \to \mathbb{R}^{k_2}$ by

$$S^{per}(x,u,t,v) = vS\left(\frac{x}{v},\frac{u}{v},\frac{t}{v}\right), \qquad T^{per}(x,u,t,v) = vT\left(\frac{x}{v},\frac{u}{v},\frac{t}{v}\right).$$

Clearly, S^{per} and T^{per} are affine mappings. Moreover we find

$$\begin{aligned}
\operatorname{Epi}(fv) = & \left\{(x,v,t) \ \middle| \ vf\left(\frac{x}{v}\right) \le t\right\} \\
= & \left\{(x,v,t) \ \middle| \ \left(\frac{x}{v},\frac{t}{v}\right) \in \operatorname{Epi} f\right\} \\
= & \left\{(x,v,t) \ \middle| \ \exists u \in \mathbb{R}^m,\ S\left(\frac{x}{v},u,\frac{t}{v}\right)=0,\ T\left(\frac{x}{v},u,\frac{t}{v}\right) \in \mathcal{K}\right\} \\
= & \left\{(x,v,t) \ \middle| \ \exists u \in \mathbb{R}^m,\ S\left(\frac{x}{v},\frac{u}{v},\frac{t}{v}\right)=0,\ T\left(\frac{x}{v},\frac{u}{v},\frac{t}{v}\right) \in \mathcal{K}\right\} \\
= & \ \{(x,v,t) \ | \ \exists u \in \mathbb{R}^m,\ S^{per}(x,u,t,v)=0,\ T^{per}(x,u,t,v) \in \mathcal{K}\},
\end{aligned}$$

which concludes the proof. □

Appendix D: Progressive Approximation

As all sets of constraints described in Section 2.2 are conservative approximations to our original constraint (1), they can yield suboptimal solutions. In particular, we propose linear decision rules to solve (5), which is equivalent to (1), of which we know they generally do not guarantee to solve adjustable robust optimization problems to optimality (Ben-Tal et al. 2004). Moreover, as our adjustable formulation (5) exhibits left-hand side uncertainty, that is, the uncertain parameter w directly interacts with decision variables x, little is known with regard to the approximative power of linear decision rules.

In this section, therefore, we focus on finding a good progressive approximation to (1) such that we can gauge the quality of the conservative approximations we propose. A simple method detailed by Zhen et al. (2022) to obtain such approximation is to only require (1) to hold for a finite subset of scenarios from the uncertainty set U. The approximation is then given by

$$f\left(A(x)\zeta^{(k)} + b(x)\right) \leq 0 \qquad k = 1, \ldots, K, \tag{20}$$

where $\left\{\zeta^{(1)}, \ldots, \zeta^{(K)}\right\} \subseteq U$. We note that these constraints are exactly as computationally tractable as the original constraint without uncertainty, given that K is not too large. In fact, because we assume a polyhedral set U and f is convex, (20) is equivalent to (1) if $\left\{\zeta^{(1)}, \ldots, \zeta^{(K)}\right\}$ contains all extreme points of U. Generally, U has prohibitively many extreme points though, and we must resort to some other way of finding scenarios $\zeta^{(1)}, \ldots, \zeta^{(K)}$.

We can apply the same reasoning as above to (5) to find an approximation:

$$\text{For } k = 1, \ldots, K, \quad \exists \lambda^{(k)} \in \mathbb{R}^q : \begin{cases} w_0^{(k)} + b(x)^\top w^{(k)} + d^\top \lambda^{(k)} \leq 0 \\ D^\top \lambda^{(k)} \geq A(x)^\top w^{(k)}, \end{cases}$$

where $\left\{\begin{pmatrix} w_0^{(1)} \\ w^{(1)} \end{pmatrix}, \ldots, \begin{pmatrix} w_0^{(K)} \\ w^{(K)} \end{pmatrix}\right\} \subset W$ and $\lambda^{(k)} \in \mathbb{R}^q$ is a non-adjustable variable. Recall that

$$W = \left\{\begin{pmatrix} w_0 \\ w \end{pmatrix} : w_0 + f^*(w) \leq 0\right\},$$

which generally has infinitely many extreme points.

An approach to find a small and efficient set of scenarios for two-stage fixed-recourse robust constraints is suggested by Hadjiyiannis et al. (2011). For any feasible solution $\hat{x}$ and linear decision rule $\hat{\lambda} = \hat{u} + \hat{V}w + \hat{r}w_0$, we find scenarios that are worst-case for the constraints in (5). We then hope that these scenarios are also worst-case for the actual optimal solution x^*, λ^* of (5). For our problem, this means that we obtain scenarios

$$\begin{pmatrix} \bar{w}_0 \\ \bar{w} \end{pmatrix} = \underset{\begin{pmatrix} w_0 \\ w \end{pmatrix} \in W}{\arg\max} \left\{d^\top \left(\hat{u} + \hat{V}w + \hat{r}w_0\right) + b(\hat{x})^\top w + w_0\right\},$$

as well as the worst-case scenarios from (5b). An extension proposed by Zhen et al. (2022) is to use these $L+1$ scenarios to also obtain scenarios $\zeta^{(1)}, \ldots, \zeta^{(L+1)}$ by solving

$$\bar{\zeta}^{(k)} = \arg\max_{\zeta \in U} \left\{ \left(A(\hat{x})\zeta + b(\hat{x})\right)^\top \bar{w}^{(k)} \right\}.$$

Acknowledgments

We thank Erick Delage from HEC Montréal for sharing the proof to Theorem 1, which is more concise than our original proof. The research of the first author was funded by the Netherlands Organisation for Scientific Research (NWO) Research Talent [Grant 406.17.511].

References

Alfandari L, García JCE (2018) Robust optimization for non-linear impact of data variation. *Computers & Operations Research* 99:38–47.

Auslender A, Teboulle M (2006) *Asymptotic cones and functions in optimization and variational inequalities* (Springer Science & Business Media).

Balvert M, van Hoof SJ, Granton PV, Trani D, Hertog DD, Hoffmann AL, Verhaegen F (2015) A framework for inverse planning of beam-on times for 3d small animal radiotherapy using interactive multi-objective optimisation. *Physics in Medicine & Biology* 60(14):5681–5698.

Ben-Tal A, Golany B, Nemirovski A, Vial JP (2005) Retailer-supplier flexible commitments contracts: A robust optimization approach. *Manufacturing & Service Operations Management* 7(3):248–271.

Ben-Tal A, Goryashko A, Guslitzer E, Nemirovski A (2004) Adjustable robust solutions of uncertain linear programs. *Mathematical Programming* 99(2):351–376.

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

Ben-Tal A, Nemirovski A, Roos C (2002) Robust solutions of uncertain quadratic and conic-quadratic problems. *SIAM Journal on Optimization* 13(2):535–560.

Bertsimas D, Sim M (2006) Tractable approximations to robust conic optimization problems. *Mathematical Programming* 107(1-2):5–36.

Bertsimas D, Sim M, Zhang M (2019) Adaptive distributionally robust optimization. *Management Science* 65(2):604–618.

Chassein A, Goerigk M (2019) On the complexity of robust geometric programming with polyhedral uncertainty. *Operations Research Letters* 47(1):21–24.

Dunning I, Huchette J, Lubin M (2017) Jump: A modeling language for mathematical optimization. *SIAM Review* 59(2):295–320, URL `http://dx.doi.org/10.1137/15M1020575`.

El Ghaoui L, Lebret H (1997) Robust solutions to least-squares problems with uncertain data. *SIAM Journal on Matrix Analysis and Applications* 18(4):1035–1064.

El Ghaoui L, Oustry F, Lebret H (1998) Robust solutions to uncertain semidefinite programs. *SIAM Journal on Optimization* 9(1):33–52.

Fredriksson A (2013) *Robust optimization of radiation therapy accounting for geometric uncertainty.* Ph.D. thesis, KTH Royal Institute of Technology.

Gorissen BL, Den Hertog D (2013) Robust counterparts of inequalities containing sums of maxima of linear functions. *European Journal of Operational Research* 227(1):30–43.

Gorissen BL, Den Hertog D, Hoffmann AL (2013) Mixed integer programming improves comprehensibility and plan quality in inverse optimization of prostate hdr brachytherapy. *Physics in Medicine & Biology* 58(4):1041.

Hadjiyiannis MJ, Goulart PJ, Kuhn D (2011) A scenario approach for estimating the suboptimality of linear decision rules in two-stage robust optimization. *50th IEEE Conference on Decision and Control and European Control Conference (CDC-ECC)*, 7386–7391 (IEEE).

Hsiung KL, Kim SJ, Boyd S (2008) Tractable approximate robust geometric programming. *Optimization and Engineering* 9(2):95–118.

Jung MN, Kirches C, Sager S (2013) On perspective functions and vanishing constraints in mixed-integer nonlinear optimal control. *Facets of Combinatorial Optimization*, 387–417 (Springer).

MOSEK ApS (2019) *The MOSEK Optimizer API for C manual. Version 9.2.5.* URL `https://docs.mosek.com/9.2/capi/index.html`.

Rockafellar RT (1970) *Convex analysis* (Princeton University Press, Princeton, NJ).

Serrano SA (2015) *Algorithms for unsymmetric cone optimization and an implementation for problems with the exponential cone*. Ph.D. thesis, Stanford University.

Van der Merwe D, Van Dyk J, Healy B, Zubizarreta E, Izewska J, Mijnheer B, Meghzifene A (2017) Accuracy requirements and uncertainties in radiotherapy: a report of the international atomic energy agency. *Acta Oncologica* 56(1):1–6.

Van Dye J, Batista J, Bauman GS (2013) Accuracy and uncertainty considerations in modern radiation oncology. *The Modern Technology of Radiation Oncology* 3:361–412.

Zhen J, de Ruiter FJ, Roos E, Den Hertog D (2022) Robust optimization for models with uncertain second-order cone and semidefinite programming constraints. *INFORMS Journal on Computing* 34(1):196–210.

Zhen J, den Hertog D (2017) Computing the maximum volume inscribed ellipsoid of a polytopic projection. *INFORMS Journal on Computing* 30(1):31–42.

Additional Proofs

Appendix EC.1: Numerical Results

EC.1.1. Geometric Programming

For our first numerical experiment we test our approach on several randomly generated geometric programming instances, identically structured to the instances used by Hsiung et al. (2008). In particular, this means we treat geometric programming problems with a linear objective, and a number of two-term log-sum-exp inequality constraints with uncertainty:

$$\begin{aligned} \min_{x} \quad & c^\top x \\ \text{s.t.} \quad & \log\left(e^{\left(-\mathbf{1}+B_i^{(1)}\zeta\right)^\top x} + e^{\left(-\mathbf{1}+B_i^{(2)}\zeta\right)^\top x}\right) \leq 0 \qquad \forall \zeta \in U,\, i=1,\ldots,m, \end{aligned} \tag{EC.1}$$

where $c = \mathbf{1} \in \mathbb{R}^n$ is the all ones vector, and $B_i^{(1)}, B_i^{(2)} \in \mathbb{R}^{n\times L}$ are randomly generated sparse matrices with sparsity density 0.1 whose nonzero elements are uniformly distributed on the interval $[0,1]$. The uncertainty set U is assumed to be a box, that is,

$$U = \left\{\zeta \in \mathbb{R}^L \,\middle|\, \|\zeta\|_\infty \leq 1\right\}. \tag{EC.2}$$

Note that since U is symmetric around 0, we can restrict $B_i^{(1)}, B_i^{(2)}$ to be nonnegative.

We first consider a set of 20 small examples with $n = m = 100$ and $L = 5$. Since L, the number of uncertain parameters, is small, (EC.1) can be solved exactly by enumerating the 2^L vertices of U. For larger L, however, we need to resort to comparing our solutions' objective value to a lower bound. To this end, we use a lower bound based on the work of Hadjiyiannis et al. (2011) and Zhen et al. (2022) that uses the optimal solution to a conservative approximation to find potentially critical scenarios in the uncertainty set. The lower bound is then constructed by solving a model that only safeguards for this finite set of critical scenarios. For more details we refer the reader to Appendix D.

Table EC.1 lists the approximation error with respect to the exact solution and computation time of the solutions to the conservative approximations resulting from Theorem 2. Moreover, to

	Approximation Error	Computation Time (s)
Lower bound	-0.00%	3.3
Theorem 2	0.02%	1.3

Table EC.1 Approximation error with respect to the exact solution and computation time for 20 randomly generated instances of type (EC.1) with $n = m = 100$ and $L = 5$.

evaluate the quality of the obtained lower bound, we have included the approximation error with respect to the exact solution and computation time of the proposed lower bound as well. We define the approximation error (in percentage) with respect to a solution x^* equally to (Hsiung et al. 2008):

$$100\left(\frac{e^{c^\top \hat{x}}}{e^{c^\top x^*}} - 1\right),$$

where $\hat{x}$ is the solution to our approximation. In other words, we compare the objective value of different solutions to the robust geometric programming problem in posynomial form. We note that the -0.00 we report for the lower bound means we are unable to differentiate the objective value from the optimal objective value within a reasonable numerical precision. We remark that the lower bound does not necessarily yield a feasible solution to the original problem, but it serves us well in evaluating the approximations in higher dimensions, where we are unable to obtain the exact objective value.

Clearly, for instances of this size the lower bound is particularly good. Moreover, it is an order of magnitude closer to the exact robust objective value compared to the solutions we find using our conservative approximation. Therefore, we expect that using the lower bound instead of the exact robust solution for larger instances has hardly any effect on the approximation error we report.

To analyze how our approach scales with more uncertain parameters, Figure EC.1 shows the average approximation error with respect to the lower bound and computation time of solutions using Theorem 2 for several values of L over 20 random instances. The affine decision rule used in Theorem 2 performs very well, having an approximation error below 0.5% for all sizes except $L = 18$. The resulting model of Theorem 2 is highly tractable as in our experiments, the computation

time increases linearly with respect to the number of uncertain parameters L. We remark that no results for $L = 19$ or $L = 20$ are included, as a large proportion of the randomly generated instances of this size were infeasible. Note that not only the approximations we propose are infeasible, but the robust instances itself are as well. This can be verified through noting that the optimization problem used to obtain the lower bound is infeasible when enough scenarios are included. We note that Hsiung et al. (2008) report approximation errors between 30% and 0.1% dependent on the quality of approximation used, for $L = 5$ and $n = m = 500$.

Besides box uncertainty, we also consider a budget uncertainty set given by

$$U = \left\{ \zeta \in \mathbb{R}^L \;\middle|\; \|\zeta\|_\infty \leq 1, \;\; \|\zeta\|_1 \leq \Gamma \cdot L \right\}. \tag{EC.3}$$

In this uncertainty set, the parameter $\Gamma \in [0, 1]$ controls the level of uncertainty. It can be interpreted as the maximum fraction of uncertain parameters that is allowed to deviate maximally at the same time. Figure EC.2 depicts the numerical results for a budget uncertainty set with $\Gamma = \frac{1}{2}$. We first note that for this budget uncertainty set, all instances with $L = 19$ and $L = 20$ are feasible. The approximation error follows a very similar trend to the one observed for box uncertainty in

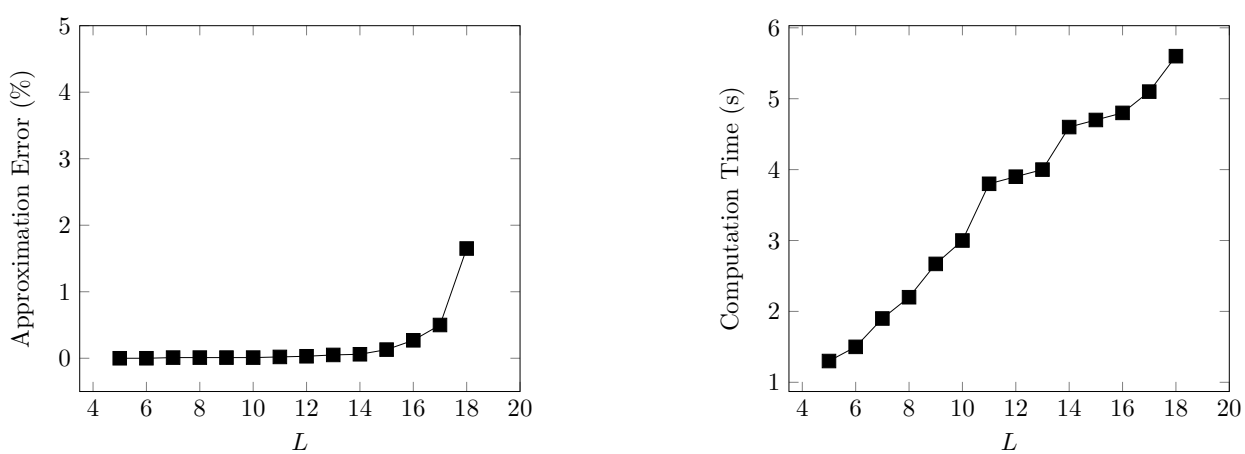

Figure EC.1 Average results of solving conservative approximations from Theorem 2 applied to (EC.1) over twenty randomly generated instances for a box uncertainty set (EC.2). The approximation error is reported with respect to the lower bound obtained from the approach proposed by Hadjiyiannis et al. (2011).

Figure EC.1 for smaller L, but there is a clear difference for larger L, where the approximation error for the budget uncertainty set is smaller. This causes us to suspect that the (extreme) increase in approximation error that was observed for the box uncertainty is related to the problems getting close to infeasibility for $L \geq 17$. The computation time follows a slightly more erratic pattern for the budget uncertainty set, and is slightly higher than for the box uncertainty set, but is largely comparable in magnitude. For all these uncertain geometric programming problems we find a solution very close to the optimal solution using this approximation.

The solutions to the geometric programming problems have been obtained using Julia with the JuMP interface (Dunning et al. 2017) and the Mosek solver for exponential cones (MOSEK ApS 2019). The experiments were conducted on a desktop with 8 GB RAM and a 3.4 GHz Intel Core i7 processor.

EC.1.2. Radiotherapy Optimization

Our second numerical experiment concerns a specific problem from radiotherapy optimization: inverse treatment planning of beam-on times for 3D small animal radiotherapy (Balvert et al. 2015). The core problem in treatment planning is ensuring a sufficient dose γ of radiation to the

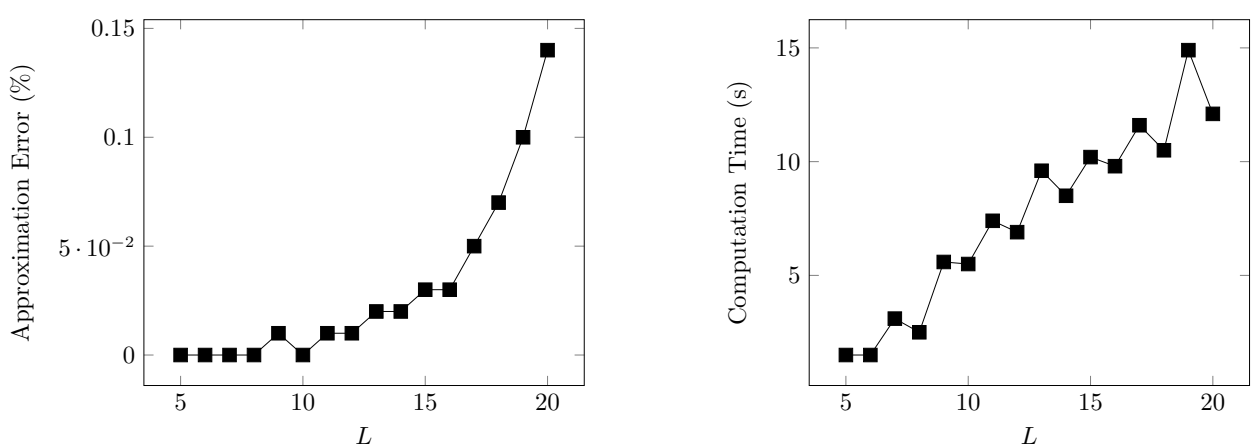

Figure EC.2 Average results of solving approximations to (EC.1) over twenty randomly generated instances for a budget uncertainty set (EC.3) with $\Gamma = \frac{1}{2}$. The approximation error is reported with respect to the lower bound obtained from the approach proposed by Hadjiyiannis et al. (2011).

planning target volume (PTV) while minimizing the dose to the tissue around that target volume, also known as the organs at risk ($\mathcal{OAR}$). To this end, we are interested in minimizing a weighted combination of the dose 'shortage' in the PTV and the dose delivered to the $\mathcal{OAR}$. The decision variables in this problem are the locations and beam-on times for all beams used. In this specific application, we assume the beam locations are given and we attempt to find optimal beam-on times t.

It is customary in radiotherapy optimization to discretize each tissue structure into voxels. Sets of these voxels are denoted by $\mathcal{I}_{PTV}$ and $\mathcal{I}_s$ for all $s \in \mathcal{OAR}$, respectively. The dose delivered to a tissue structure is then computed as the average dose delivered to its voxels. Given these voxels, one can compute the dose rates from all beams to all voxels, referred to as the matrix Γ. The i-th row of this matrix, Γ_i, then corresponds to the dose rate of all beams to voxel i. We specifically consider the following mathematical optimization problem:

$$\min_{t,\tau} \ \tau \tag{EC.4a}$$

$$\text{s.t.} \ w_{PTV} \frac{1}{|\mathcal{I}_{PTV}|} \sum_{i \in \mathcal{I}_{PTV}} \max\left\{\gamma - \Gamma_i^\top t, 0\right\}^2 + (1 - w_{PTV}) \sum_{s \in \mathcal{OAR}} w_s \frac{1}{|\mathcal{I}_s|} \sum_{i \in \mathcal{I}_s} \Gamma_i^\top t \leq \tau \tag{EC.4b}$$

$$t \geq 0, \tag{EC.4c}$$

which is a slight adaptation of the problem described by Balvert et al. (2015). Here, w_{PTV} and w_s for all $s \in \mathcal{OAR}$ represent predefined weights. In particular, we choose to use a squared penalty function for undelivered dose to the PTV, similar to Fredriksson (2013, Eq. 1). Irregardless of whether the regular or squared penalty function is used, little research has been done on robust or uncertain versions of (EC.4). An important reason for this is the general convex nature of constraint (EC.4b), along with the fact that a natural type of uncertainty in this problem is implementation error (Van Dye et al. 2013, Van der Merwe et al. 2017), which always leads to constraints that are convex in the uncertain parameters.

Maximum Implementation Error	1.00%	5.00%	10.00%
Approximation Error	1.13%	5.57%	10.91%

Table EC.2 Approximation error of Theorem 2 with respect to the lower bound for Case 3 discussed by Balvert et al. (2015) for different sizes of the uncertainty set.

In this numerical example, we therefore focus on implementation error. In particular, we consider multiplicative implementation error, that is, we replace t by $t \circ (\mathbf{1} + \epsilon)$, where $\circ$ denotes the element-wise multiplication of two vectors, and ϵ is the uncertain vector that models the implementation error. We note that, at least in this context, additive implementation error of the form $t + \Delta t$ would make little sense, as this would presume that there would also potentially be some implementation error if one chooses not to use a certain beam ($t_b = 0$).

We construct the conservative solutions derived in this paper for Case 3 discussed by Balvert et al. (2015). In this case, there are 6 different beam angles, that is, $t \in \mathbb{R}^6$, and the PTV consists of 112,738 voxels, while the four organs at risk consist of 207,974, 2,261,739, 177,165 and 212,864 voxels. We consider box uncertainty for ϵ, with three different maximum values: 0.01, 0.05 and 0.1. Note that there is a need for conservative solutions, as we cannot obtain exact robust solutions by enumeration due to the size of the problem. We are, however, able to obtain lower bounds using the technique described in Appendix D and report the approximation error with respect to that lower bound in Table EC.2. Furthermore, we find that the nominal solution performs 4.4% worse in the worst-case than the robust solution we find, which in turn performs 4.9% worse than the nominal solution when no uncertainty is present. It should be noted that the conservative solution was found in a matter of seconds, much like the nominal model that disregards uncertainty.

All results in this section have been obtained using Julia with the JuMP interface (Dunning et al. 2017) and the Gurobi solver. The experiments were conducted on a desktop with 8 GB RAM and a 3.4 GHz Intel Core i7 processor.