Globalized Robust Optimization for Nonlinear Uncertain Inequalities

Ben-Tal, A.; Brekelmans, Ruud; den Hertog, Dick; Vial, J.P.

CentER Discussion Paper

No. 2015-031

GLOBALIZED ROBUST OPTIMIZATION FOR NONLINEAR UNCERTAIN INEQUALITIES

By

Aharon Ben-Tal, Ruud Brekelmans,
Dick den Hertog, Jean-Philippe Vial

15 June 2015

ISSN 0924-7815
ISSN 2213-9532

Tilburg University

Globalized Robust Optimization For Nonlinear Uncertain Inequalities

Aharon Ben-Tal
Department of Industrial Engineering and Management
Technion – Israel Institute of Technology, Haifa 32000, Israel
abental@ie.technion.ac.il

Ruud Brekelmans
Department of Econometrics and Operations Research,
Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
r.c.m.brekelmans@tilburguniversity.edu

Dick den Hertog
Department of Econometrics and Operations Research,
Tilburg University, P.O. Box 90153, 5000 LE Tilburg, The Netherlands
d.denhertog@tilburguniversity.edu

Jean-Philippe Vial
Ordecsys, Place de l'Etrier 4, CH-1224, Chêne-Bougeries, Switzerland
jphvial@gmail.com

June 10, 2015

Abstract

Robust optimization is a methodology that can be applied to problems that are affected by uncertainty in the problem's parameters. The classical robust counterpart (RC) of the problem requires the solution to be feasible for all uncertain parameter values in a so-called uncertainty set, and offers no guarantees for parameter values outside this uncertainty set. The globalized robust counterpart (GRC) extends this idea by allowing controlled constraint violations in a larger uncertainty set. The constraint violations are controlled by the distance of the parameter to the original uncertainty set. We derive tractable GRCs that extend the initial GRCs in the literature: our GRC is applicable to nonlinear constraints instead of only linear or conic constraints, and the GRC is more flexible with respect to both the uncertainty set and distance measure function, which are used to control the constraint violations. In addition, we present a GRC approach that can be used to provide an extended trade-off overview between the objective value and several robustness measures.

Keywords: robust optimization, globalized robust counterpart, constraint violations

JEL-classification: C61, C63, M11

1 Introduction

The robust optimization (RO) methodology supports decision makers in handling optimization problems affected by uncertainty in the problem's parameters. Many optimization problems depend on parameters that are the outcome of measurements, estimation procedures, or other methods that cannot guarantee absolute knowledge about the underlying parameters. Therefore, decision makers are interested in solutions that are robust to situations where the problem's parameters deviate from their nominal values, i.e., the solution should be feasible and attain a satisfactory objective value. The RO methodology attains this goal by formulating a *robust counterpart* (RC) of the original optimization problem, where the solution should be feasible for all uncertain parameter values in a so-called *uncertainty set*. For several optimization problems and choices of the uncertainty set, the resulting RO problem can be formulated as a computationally tractable optimization problem. For an overview about the theory and applications of RO we refer to Ben-Tal et al. (2009a), Bertsimas et al. (2011), and Gabrel et al. (2014).

The *globalized robust counterpart* (GRC) is an extension of the traditional RC, which was introduced by Ben-Tal et al. (2006) and originally called the comprehensive robust counterpart. The main idea of the GRC is to give the decision maker more control by alleviating the feasibility requirement in some parts of the uncertainty set in a controlled way. Hereto, the GRC approach uses two uncertainty sets: an inner uncertainty set and an outer uncertainty set, which contains the former. The GRC requires full feasibility for all parameter values in the inner uncertainty set analogous to the traditional RC. However, infeasibilities are allowed for parameter values in the outer uncertainty set, where the violation is controlled by the distance of the parameter value from the inner uncertainty set. The rationale behind this idea is that parameter values that are further away from the inner uncertainty set are less likely to occur in practice, but not so unlikely that they can be neglected altogether.

Since the introduction in 2006, several papers have appeared that apply GRC. Ben-Tal et al. (2009b) apply the GRC approach to an inventory model, and Babonneau et al. (2010) to an environmental problem. Furthermore, GRC has been applied to orienteering problems (Evers et al., 2011), facility location problems (Naseraldin and Baron, 2011), support vector machines (El Ghaoui, 2012) and porfolio optimization problems (Zymler et al., 2011). Xu et al. (2012) show that a probabilistic envelope constraint can be reformulated as a GRC.

In this paper, we present an extension of the initial GRC formulation that applies to larger classes of uncertain constraints and permits richer classes of controls on the infeasibilities. Using results in Ben-Tal et al. (2015), we are able to derive a tractable equivalent of the GRC in the following cases.

- In the initial GRC formulation, the outer uncertainty set is restricted to a special form: the inner uncertainty plus a cone. In our formulation, the only requirement is that the outer uncertainty be convex and that a tractable representation of its support function exists.

- In the initial GRC formulation, the distance measure function, which is used to control the constraint violations, is a linear function of the distance to the inner uncertainty set measured by a norm in the space of the cone used to define the outer uncertainty set. Our formulation allows more general distance measure functions.

- Our GRC is applicable to nonlinear constraint that are convex in the decision variables and concave in the uncertain parameters. The initial GRC formulation is only applicable to constraints that are linear in both decision variables and uncertain parameters.

Besides the above mentioned theoretical contributions, we show how GRCs can be used to offer a decision maker a more complete overview of the trade-off between a problem's objective value and the robustness of the solution. This is especially relevant for an optimization problem for which the RC is infeasible for the uncertainty set that was initially constructed to obtain a robust solution. We know of at least two methods based on the original RC aimed at handling the feasibility issue, but both suffer from drawbacks.

In the first method, the size of the uncertainty set is shrinked until the RC problem becomes feasible. This usually results in solutions that yield good performance with respect to the objective value, but constraint violations may be unacceptably large in the original uncertainty set. The second method simply minimizes the infeasibilities in the uncertainty set, thereby completely ignoring the objective value. Moreover, the average behavior of the constraint violations in the uncertainty set can become extremely bad, even though the worst-case constraint violations are minimized.

Our GRC approach uses the original uncertainty set as the outer uncertainty set and restricts constraint violations outside the inner uncertainty set. Only the general form of the distance measure functions for all constraints has to be specified, but the weights of these functions, which control the constraint violations, are determined implicitly in the approach. The results indicate that solutions can be obtained that perform better with respect to many robustness measures than solutions obtained by RC-based approaches. This extends the trade-off overview between objective value and robustness that can already be obtained by other approaches.

The remainder of this paper is organized as follows. In Section 2, we present the globalized robust optimization methodology. In Section 3, we show the theoretical foundation for deriving tractable GRCs. In Section 4, tractable GRCs are derived for several distance measures, and a comparison with other GRCs is given. Section 5 presents the GRC approach for dealing with infeasible robust optimization problems, and in Section 6 a numerical example is presented that illustrates this approach. Section 7 concludes.

Notation

Throughout this paper we use the following notation.

For any function $f : \mathbb{R}^n \to \mathbb{R}$ we let $\mathrm{dom}(f) = \{x \in \mathbb{R}^n \mid f(x) < \infty\}$. The *convex conjugate* of f is defined as

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

The *concave conjugate* of f is defined as

$$f_*(y) = \inf_{x \in \mathrm{dom}(-f)} \{y^T x - f(x)\}.$$

For a function $g(.,.)$ of two vector variables, $g^*(.,.)$ and $g_*(.,.)$ will denote the partial convex and partial concave conjugate with respect to the *first* variable, respectively. On the other hand, $g^*(.;.)$ and $g_*(.;.)$ will denote the convex and concave conjugate function with respect to *both* variables, respectively.

The *indicator function* on the set S is defined as

$$\delta(x \mid S) = \begin{cases} 0 & \text{if } x \in S \\ \infty & \text{otherwise.} \end{cases}$$

Then the conjugate function of $\delta(x \mid S)$, i.e.,

$$\delta^*(x \mid S) = \sup_{y \in S} \; y^T x,$$

is the so-called *support function* of the set S.

The relative interior of a set S is denoted by $\mathrm{ri}(S)$.

2 Globalized Robust Optimization Methodology

The Robust Optimization (RO) methodology considers problems that are affected by parameter uncertainty. Here we specifically focus on parameter uncertainty in a nonlinear constraint

$$f(a, x) \leq 0, \tag{1}$$

where $x \in \mathbb{R}^n$ is the optimization variable and $a \in \mathbb{R}^m$ is the vector of uncertain parameters. The function $f(., x)$ is assumed to be concave in the uncertain parameters, i.e., in its first argument, for all $x \in \mathbb{R}^n$.

The standard approach in RO is to require that the uncertain constraint (1) is satisfied for all vectors a in a so-called uncertainty set U, i.e.,

$$f(a, x) \leq 0, \qquad \forall a \in U, \tag{2}$$

which is known as the robust counterpart (RC) of the original uncertain constraint (1). In Ben-Tal et al. (2015) computationally tractable representations of the RC are derived using Fenchel duality.

This is a sensible approach in many applications. For instance, when designing a building you may require that the building must be able to withstand wind speeds for a whole range of forces and directions, without incurring any damage to the building.

The fact that in (2) the decision maker has full responsibility for the feasibility of the solution x for all $a \in U$ may lead to overly pessimistic solutions in practice. In the example of designing a building, this might happen when the uncertainty set U includes all wind forces up to Beaufort scale 12. A straightforward, but naive, way to deal with this undesired behavior is to reduce the original uncertainty set to, say wind forces up to Beaufort scale 10. The resulting solution will be indeed less pessimistic, however, it does not offer any guarantee about the feasibility of the solution for wind forces above Beaufort scale 10. For instance, the building could collapse entirely in the case of Beaufort scale 11. This is clearly undesirable, however, modest damage to the building might be acceptable in the case of Beaufort scale 11. In the case of a hurricane (Beaufort scale 12) even larger damages might be acceptable. The GRC tries to accomplish this goal by controlling the deterioration of infeasibilities outside the smaller shrinked set.

The GRC considers two uncertainty sets U_1 and U_2, both assumed to be convex, with $U_1 \subset U_2$. The inner uncertainty set U_1, which is assumed to be compact as well, is dealt with exactly as in the standard RC. However, for realizations of the uncertain vector $a \in U_2 \setminus U_1$ constraint violations of the original constraint (1) are allowed, where the magnitude of the allowed violation depends on the distance of a to the smaller set U_1. Formally, the GRC of the uncertain constraint (1) is defined as

$$f(a, x) \leq \min_{a' \in U_1} \phi(a, a'), \qquad \forall a \in U_2, \tag{3}$$

where $\phi(a, a')$ measures the distance between the parameters a and a'. The distance measure function $\phi(.,.)$ is assumed to be nonnegative and jointly convex in both arguments and $\phi(a, a) = 0$ for all $a \in \mathbb{R}^m$.

Now, if $a \in U_1$, then obviously $\min_{a' \in U_1} \phi(a, a') = 0$ and (3) simplifies to $f(a, x) \leq 0$. For $a \in U_2 \setminus U_1$, violation of the original inequality is allowed, which is controlled by the distance of a to U_1. For more extreme values of a the distance will be larger and consequently the allowed constraint violation will be larger as well.

One obvious choice for the distance measure function is to take a function of the norm of the difference between a and a': $\phi(a, a') = \alpha(\|a - a'\|)$, where α is a convex and nonnegative function with $\alpha(0) = 0$. In Section 4 we consider another choice of distance measures based on the so-called phi-divergence functions. An example is the χ^2-distance: $\phi(a, a') = \sum_{i=1}^m (a_i - a_i')^2 / a_i$ for $a, a' \in \mathbb{R}_+^m$.

For the uncertain parameter $a \in \mathbb{R}^m$, we consider uncertainty sets U_i defined by

$$U_i = \{a = a_0 + A\zeta \mid \zeta \in Z_i\}, \qquad i = 1, 2, \tag{4}$$

where $a_0 \in \mathbb{R}^m$ is the "nominal value", $A = [\, a_1 \ \dots \ a_L \,] \in \mathbb{R}^{m \times L}$ is the "perturbation set", $\zeta \in \mathbb{R}^L$ is the vector of "primitive uncertainties", and $Z_1 \subset \mathbb{R}^L$ is a given nonempty, convex and compact set, with $0 \in \mathrm{ri}(Z_1)$. The set Z_2 is also convex but not necessarily compact, and we assume that $Z_1 \subset Z_2$, implying that $U_1 \subset U_2$.

3 Deriving computationally tractable GRCs

Constraint (3) is a semi-infinite constraint, because it has to hold for all $a \in U_2$, which makes it difficult to use for optimization purposes. In the following theorem we derive a generic formula for the GRC (3), which will be the main tool to derive tractable GRCs for a host of functions f and ϕ, and for diverse sets Z_1 and Z_2.

Theorem 1. *Let $f(., x)$ be a concave function in $\mathbb{R}^m$ for all $x \in \mathbb{R}^n$, and $\phi : \mathbb{R}^m \times \mathbb{R}^m \to \mathbb{R}$ a convex and nonnegative function for which $\phi(a, a) = 0$ for all $a \in \mathbb{R}^m$. Let the set $Z_1 \subset \mathbb{R}^L$ be nonempty, convex, and compact with $0 \in \mathrm{ri}(Z_1)$, let Z_2 be a convex set such that $Z_1 \subset Z_2$, and let the sets U_1 and U_2 be defined by* (4) *for fixed $a_0 \in \mathbb{R}^m$ and $A \in \mathbb{R}^{m \times L}$.*

Then, vector $x \in \mathbb{R}^n$ satisfies (3) *if and only if there exist $v, w \in \mathbb{R}^m$ that satisfy the single inequality*

$$a_0^T(v + w) + \delta^*(A^T v \mid Z_1) + \delta^*(A^T w \mid Z_2) - f_*(v + w, x) + \phi^*(v; -v) \leq 0. \tag{5}$$

Proof. Note that (3) is satisfied if and only if $F(x) \leq 0$ with

$$\begin{aligned}
F(x) &\equiv \max_{a \in U_2} \{f(a, x) - \min_{a' \in U_1} \phi(a, a')\} \\
&= \max_{a \in U_2, a' \in U_1} \{f(a, x) - \phi(a, a')\} \\
&= \max_{a \in U_2, a' \in U_1, t, u} \{f(a, x) - \phi(t, u) \mid t = a, \ u = a'\} \\
&= \min_{v,z} \max_{a \in U_2, a' \in U_1, t, u} \{f(a, x) - \phi(t, u) - v^T(u - a') - z^T(t - a)\} \qquad (6) \\
&= \min_{v,z} \Big\{ \underbrace{\max_{a \in U_2} \{f(a, x) + z^T a\}}_{h_1(z,x)} + \underbrace{\max_{t,u} \{-\phi(t, u) - z^T t - v^T u\}}_{h_2(v,z)} + \underbrace{\max_{a' \in U_1} \{v^T a'\}}_{h_3(v)} \Big\}, \qquad (7)
\end{aligned}$$

where (6) follows by duality. Let us now analyze all parts of (7). Rewriting the first part yields

$$\begin{aligned} h_1(z,x) &\equiv \max_{a\in U_2}\{f(a,x)+z^T a\} \\ &= \max_{a}\{f(a,x)+z^T a-\delta(a\mid U_2)\} \\ &= \min_{w}\{\delta^*(w\mid U_2)-\left[f(w,x)+z^T w\right]_*\}, \end{aligned} \tag{8}$$

where the last equality follows by Fenchel duality (see Theorem A.4 in the Appendix). Using the relationship between U_2 and Z_2 (see (4)), it can be verified that $\delta^*(w|U_2) = a_0^T w+\delta^*(A^T w|Z_2)$. Furthermore, it is straightforward to derive the concave conjugate of the second term in (8) with respect to the variable w, which gives

$$h_1(z,x) = \min_{w}\{a_0^T w+\delta^*(A^T w\mid Z_2)-f_*(w-z,x)\}.$$

The second term in (7) simplifies to

$$\begin{aligned} h_2(v,z) &\equiv \max_{t,u}\{-\phi(t,u)-z^T t-v^T u\} \\ &= \phi^*(-z;-v), \end{aligned}$$

and finally,

$$h_3(v) \equiv \max_{a'\in U_1}\{v^T a'\} = \delta^*(v\mid U_1) = a_0^T v+\delta^*(A^T v\mid Z_1).$$

Hence, if we substitute these expressions in (7), then we get that (3) is equivalent to

$$\begin{aligned} F(x)\le 0 \Leftrightarrow \min_{v,z}\bigg\{ &\min_{w}\Big\{a_0^T w+\delta^*(A^T w\mid Z_2)-f_*(w-z,x)\Big\} \\ &+\phi^*(-z;-v)+a_0^T v+\delta^*(A^T v\mid Z_1)\bigg\}\le 0. \end{aligned}$$

Therefore, $F(x)\le 0$ if and only if x together with variables v,w and z satisfy

$$a_0^T w+\delta^*(A^T w\mid Z_2)-f_*(w-z,x)+\phi^*(-z;-v)+a_0^T v+\delta^*(A^T v\mid Z_1)\le 0.$$

Next, note that

$$\begin{aligned} \phi^*(-z;-v) &= \max_{t,u}\{-z^T t-v^T u-\phi(t,u)\} \\ &\ge \max_{t}\{-(z+v)^T t-\phi(t,t)\} = \max_{t}\{-(z+v)^T t\} \\ &= \begin{cases} 0 & \text{if } z=-v, \\ \infty & \text{if } z\ne -v. \end{cases} \end{aligned}$$

It follows that $z=-v$ and the result (5) follows. □

We make the following observations with respect to Theorem 1:

- The computations involving f, Z_1, Z_2 and ϕ are all separated. Later we shall show that (5) can be rewritten in a computationally tractable way for several choices of f, Z_i $(i=1,2)$ and ϕ.

- Suppose that $U_2 = \mathbb{R}^m$. The relationship between U_2 and Z_2 defined by (4) now implies that $L \geq \text{rank}(A) \geq m$. Then it follows from

$$\delta^*(A^T w \mid Z_2) = \max_{\zeta \in Z_2} w^T A\zeta = \max_{y \in \mathbb{R}^m} w^T y = \begin{cases} 0 & \text{if } w = 0, \\ \infty & \text{if } w \neq 0, \end{cases}$$

  that $w = 0$ and the GRC reduces to

$$a_0^T v + \delta^*(A^T v \mid Z_1) - f_*(v, x) + \phi^*(v; -v) \leq 0.$$

- Suppose that $f(a, .)$ is linear, i.e., $f(a, x) = a^T x - b$, with b a fixed parameter. We find that

$$f_*(v + w, x) = \min_a \{a^T(v + w) - a^T x + b\} = \begin{cases} b & \text{if } v + w = x, \\ -\infty & \text{if } v + w \neq x. \end{cases}$$

  Hence, $v + w = x$ and the GRC reduces to

$$a_0^T x + \delta^*(A^T v \mid Z_1) + \delta^*(A^T(x - v) \mid Z_2) + \phi^*(v; -v) \leq b.$$

- We can rewrite the GRC (3) as the standard RC of the uncertain constraint $f(a, x) \leq d$, where the right-hand side d is also uncertain. The uncertainty set that we need to consider in this case is

$$\tilde{U} = \left\{(a, d) \,\middle|\, a \in U_2, d \in \mathbb{R} : \exists a' \in U_1 : \phi(a, a') \leq d\right\}$$

  Theorem 1 then also follows by applying the main result about robust counterparts for nonlinear inequalities in Ben-Tal et al. (2015) to this new uncertain constraint and uncertainty set $\tilde{U}$.

An alternative to the GRC in (3) is to have the allowed constraint violation depend on the distance measured in the space of the primitive uncertain parameter ζ instead of in the space of a. Hence, we consider the GRC

$$f(a(\zeta), x) \leq \min_{\zeta' \in Z_1} \phi(\zeta, \zeta'), \qquad \forall \zeta \in Z_2, \tag{9}$$

where $a(\zeta) \equiv a_0 + A\zeta$ and ϕ now measures the distance in the space of ζ. Following a derivation analogous to Theorem 1, we find that (9) is satisfied if and only if there exist $v \in \mathbb{R}^L$ and $w \in \mathbb{R}^m$ such that

$$a_0^T w + \delta^*(v \mid Z_1) + \delta^*(A^T w - v \mid Z_2) - f_*(w, x) + \phi^*(v; -v) \leq 0.$$

We conclude this section by presenting two extensions of Theorem 1. First, we analyze the case that there are more than two uncertainty sets for the uncertain parameter a. Suppose that there are K uncertainty sets $U_1 \subset U_2 \subset \cdots \subset U_K \subset \mathbb{R}^m$, which all have the same structure as before (see (4)), thus we also have the uncertainty sets $Z_1 \subset Z_2 \subset \cdots \subset Z_K \subset \mathbb{R}^L$ in the space of the primitive uncertainties. The following GRC formulation allows the decision maker to control the infeasibilities in the sets $U_1, \ldots, U_{K-1}$ using different distance measure functions ϕ_k:

$$f(a, x) \leq \sum_{k=1}^{K-1} \min_{a_k \in U_k} \phi_k(a, a_k), \qquad \forall a \in U_K. \tag{10}$$

Analogous to Theorem 1, we find that (10) is satisfied if and only if there exist $v_1, \ldots, v_{K-1} \in \mathbb{R}^m$ and $w \in \mathbb{R}^m$ such that

$$a_0^T\Big(\sum_{k=1}^{K-1} v_k + w\Big) + \sum_{k=1}^{K-1} \delta^*(A^T v_k \mid Z_k) + \delta^*(A^T w \mid Z_k) \\ - f_*\Big(\sum_{k=1}^{K-1} v_k + w, x\Big) + \sum_{k=1}^{K-1} \phi_k^*(v_k; -v_k) \leq 0. \quad (11)$$

Finally, we consider the extension of Theorem 1 to multiple and weighted distance measure functions, i.e., the GRC that we consider is

$$f(a, x) \leq \sum_{k=1}^{K} \theta_k \min_{a' \in U_1} \phi_k(a, a'), \qquad \forall a \in U_2, \quad (12)$$

where θ_k are the nonnegative weights of the distance measures $\phi_k(., .)$ for $k = 1, \ldots, K$. Analogous to (11) we find that (12) is satisfied if and only if there exist $v_1, \ldots, v_K \in \mathbb{R}^m$ and $w \in \mathbb{R}^m$ such that

$$a_0^T\Big(w + \sum_{k=1}^{K} v_k\Big) + \sum_{k=1}^{K} \delta^*(A^T v_k \mid Z_1) + \delta^*(A^T w \mid Z_2) \\ - f_*\Big(w + \sum_{k=1}^{K} v_k, x\Big) + \sum_{k=1}^{K} \theta_k \phi_k^*\Big(\frac{v_k}{\theta_k}; \frac{-v_k}{\theta_k}\Big) \leq 0, \quad (13)$$

where $0\phi_k^*(v_k/0; -v_k/0) = \lim_{\theta_k \to \infty} \theta_k \phi_k(v_k/\theta_k; -v_k/\theta_k)$ is the recession function of ϕ_k^* (see Rockafellar, 1970). In Section 5, this GRC formulation is used, where the weights θ_k are included as decision variables in the optimization problem.

4 GRCs for special distance measures

4.1 Distance measures based on the difference between parameters

In this section we discuss a special class of functions that can be used for the general distance measure function ϕ introduced in Section 2. In many cases it makes sense to let the distance measure for two parameter values a and a' depend on the difference $a - a'$ or even only on the norm $\|a - a'\|$. The convex conjugate ϕ^* needed in Theorem 1 can be simplified in those cases:

- If $\phi(a, a') = \beta(a - a')$ for some convex nonnegative function $\beta : \mathbb{R}^m \to \mathbb{R}$ with $\beta(0) = 0$, then it easily follows that $\phi^*(v; -v) = \beta^*(v)$.

- Let $\phi(a, a') = \alpha(\|a - a'\|)$, where $\|.\|$ can be any norm and $\alpha(.)$ is a convex, nonnegative function such that $\alpha(0) = 0$. This gives

$$\begin{aligned} \phi^*(v; -v) &= \max_{s,t}\{(s - t)^T v - \alpha(\|s - t\|)\} \\ &= \max_{s}\{s^T v - \alpha(\|s\|)\} \\ &= \max_{\lambda \geq 0} \max_{s:\|s\|=\lambda} \{s^T v - \alpha(\lambda)\} \\ &= \max_{\lambda \geq 0}\{\lambda\|v\|^* - \alpha(\lambda)\} \\ &= \alpha^*(\|v\|^*), \end{aligned} \quad (14)$$

where $\|.\|^*$ denotes the dual norm of $\|.\|$.

Z_1	Z_2	$\|.\|$	$\alpha(.)$	GRC
ℓ_1/ℓ_∞/ polyh.	ℓ_1/ℓ_∞/ polyh.	ℓ_1/ ℓ_∞	linear	linear
ℓ_1/ℓ_∞/ polyh.	ℓ_1/ℓ_∞/ polyh.	ℓ_2	linear	quadratic
ℓ_2	ℓ_2	ℓ_2	linear	conic quadratic
ℓ_1/ℓ_∞/ polyh.	ℓ_1/ℓ_∞/ polyh.	ℓ_1/ℓ_∞	quadratic	quadratic
ℓ_2	ℓ_2	ℓ_2	quadratic	conic quadratic

Table 1: Tractability of GRC of a linear constraint for different choices for the uncertainty sets and distance measure function $\phi(a, a') = \alpha(\|a - a'\|)$.

Now, we illustrate how Theorem 1 can be used to derive tractable GRCs for these type of distance measures.

Example 1. Let

$$\begin{aligned} f(a,x) &= a^T x - b \\ Z_i &= \{\zeta \in \mathbb{R}^L \mid \|\zeta\|_{p_i} \le \rho_i\}, \qquad i = 1,2, \\ \phi(a,a') &= \alpha(\|a - a'\|_{p_0}) \qquad \text{with} \qquad \alpha(t) = \theta t, \qquad t \ge 0, \end{aligned}$$

where $\theta \ge 0$, $0 \le \rho_1 \le \rho_2$, and $p_i \ge 1$ $(i = 0,1,2)$.

The support functions for Z_i are

$$\delta^*(u \mid Z_i) = \max_\zeta \{\zeta^T u \mid \|\zeta\|_{p_i} \le \rho_i\} = \rho_i \|u\|_{q_i},$$

where $\|.\|_{q_i}$, with $1/p_i + 1/q_i = 1$, are the $\|.\|_{p_i}$ dual norms $(i = 0,1,2)$. Furthermore, the convex conjugate of $\phi(.;.)$ can be derived from (14) where

$$\alpha^*(s) = \begin{cases} 0 & \text{if } s \le \theta, \\ \infty & \text{if } s > \theta. \end{cases} \tag{15}$$

Using these results, (5) reduces to the system of inequalities

$$\begin{cases} a_0^T x + \rho_1 \|A^T v\|_{q_1} + \rho_2 \|A^T (x - v)\|_{q_2} \le b \\ \|v\|_{q_0} \le \theta. \end{cases}$$

If all norms are ℓ_1 or ℓ_∞ norms, then both inequalities above can be represented by a linear system of inequalities. For other norms, the inequalities can also be reformulated in tractable versions in many situations. Table 1 summarizes the tractability of the resulting GRC for different choices for Z_1, Z_2, $\|.\|$ and $\alpha(.)$. □

Example 2. Let

$$\begin{aligned} f(a,x) &= a^T x - b, \\ Z_i &= \{\zeta \in \mathbb{R}^L \mid C_i \zeta \le d_i\}, \qquad i = 1,2, \\ \phi(a,a') &= \alpha(\|a - a'\|_2) \qquad \text{with} \qquad \alpha(t) = \tfrac{1}{2}\theta t^2, \qquad t \ge 0, \end{aligned}$$

where $\theta > 0$ is fixed and C_i and d_i are given matrices and vectors that define the polyhedrons Z_i $(i = 1,2)$. For these choices, we obtain

$$\delta^*(u \mid Z_i) = \max_\zeta \{\zeta^T u \mid C_i \zeta \le d_i\} = \min_{w \ge 0} \{d_i^T w \mid C_i^T w = u\},$$

and $\alpha^*(s) = \frac{1}{2\theta}s^2$ for $s \geq 0$. Hence, GRC (5) now reduces to the system of (in)equalities

$$\begin{cases} a_0^T x + d_1^T w + d_2^T z + \frac{1}{2\theta}\|v\|_2^2 \leq b \\ C_1^T w = A^T v \\ C_2^T z = A^T(x - v) \\ w, z \geq 0, \end{cases}$$

where w and z are additional analysis variables. □

Example 3. Let the uncertainty in (1) be linear in a but nonlinear in x:

$$f(a, x) = \sum_{i=1}^{m} a_i f_i(x) - b,$$

with $a_i \geq 0$ and $f_i(.)$ nonnegative and convex. Let Z_1 be the intersection of ℓ_1 and ℓ_∞ balls, i.e.,

$$Z_1 = \{\zeta \mid \|\zeta\|_1 \leq B,\ \|\zeta\|_\infty \leq \rho_1\},$$

and let Z_2 be the box

$$Z_2 = \{\zeta \mid \|\zeta\|_\infty \leq \rho_2\},$$

with $0 < \rho_1 \leq \rho_2$. Along with these choices of Z_1 and Z_2, we make the additional assumption that the nominal value a_0 and the matrix A guarantee that U_1 and U_2 are nonnegative, i.e., $U_1 \subset U_2 \subset \mathbb{R}^m_+$.

In view of the nonnegativity assumptions on a_i, f_i and U_1, we observe that

$$\sum_{i=1}^{m} (a_i - \Delta_i) f_i(x) \leq \sum_{i=1}^{m} a_i f_i(x) \qquad \forall a \in U_1, \Delta \in \mathbb{R}^m_+.$$

Hence, any deviation from $a \in U_1$ by only negative components relaxes the constraint $f(a, x) \leq 0$. Therefore, the distance measure should assign a zero distance to such parameters, effectively ensuring feasibility for those parameters even if they are outside the inner set U_1. For instance this goal can be attained by choosing the distance measure function

$$\phi(a, a') = \beta(a - a') \qquad \text{with} \qquad \beta(v) = \theta \sum_{i=1}^{m} \max\{v_i, 0\},$$

with θ a positive constant.

A tractable GRC can now be obtained from Theorem 1 by deriving the respective terms in (5). We derive

$$f_*(u, x) = \min_{a \geq 0} \left\{ \sum_{i=1}^{m} a_i (u_i - f_i(x)) + b \right\} = \begin{cases} b & \text{if } f_i(x) \leq u_i \ \forall i \\ -\infty & \text{otherwise.} \end{cases}$$

Furthermore, we find

$$\delta^*(A^T v \mid Z_1) = \min_{s,t} \{B\|s\|_\infty + \rho_1 \|t\|_1 \mid s + t = A^T v\},$$

by Lemma A.1, and $\delta^*(A^T w \mid Z_2) = \rho_2 \|A^T w\|_1$. Finally, we obtain

$$\phi^*(v; -v) = \beta^*(v) = \begin{cases} 0 & \text{if } 0 \leq v_i \leq 1 \quad \forall i, \\ \infty & \text{otherwise.} \end{cases}$$

If we combine all these results, then for this particular example, GRC (3) reduces to

$$\begin{cases} \sum_{i=1}^m (a_0)_i f_i(x) + B\|s\|_\infty + \rho_1\|t\|_1 + \rho_2\|A^T w\|_1 \le b \\ A^T v = s + t \\ f_i(x) \le v_i + w_i \qquad \forall i \\ 0 \le v_i \le 1 \qquad \forall i. \end{cases}$$

□

4.2 Distance measures based on phi-divergence functions

Another way to measure the distance between two parameters is by a phi-divergence distance defined by

$$\phi(a, a') = \sum_{i=1}^m a'_i \varphi\left(\frac{a_i}{a'_i}\right), \qquad a, a' \in \mathbb{R}_+^m, \tag{16}$$

where $\varphi(t)$ is convex for $t \ge 0$, $\varphi(1) = 0$, $0\varphi(a/0) \equiv a \lim_{t\to\infty} \varphi(t)/t$ for $a > 0$ and $0\varphi(0/0) \equiv 0$. The function φ is referred to as the phi-divergence function. The phi-divergence measure is often applied to, but not restricted to, probability vectors. It is, however, required that the arguments are nonnegative vectors.

If we let $\eta(a, b) \equiv b\varphi(a/b)$, then $\phi(a, a') = \sum_{i=1}^m \eta(a_i, a'_i)$. Because $\phi(.,.)$ is separable it follows that $\phi^*(s; t) = \sum_{i=1}^m \eta^*(s_i; t_i)$. The convex conjugate of $\eta(a, b)$ is

$$\begin{aligned} \eta^*(s; t) &= \max_{a,b\ge 0} \left\{ as + bt - b\varphi\left(\frac{a}{b}\right) \right\} \\ &= \max_{a,b\ge 0} b\left\{ t + \frac{a}{b}s - \varphi\left(\frac{a}{b}\right) \right\} \\ &= \max_{b\ge 0} b\{t + \varphi^*(s)\} \\ &= \begin{cases} 0 & \text{if } \varphi^*(s) \le -t, \\ \infty & \text{otherwise.} \end{cases} \end{aligned}$$

Hence, by Theorem 1, we find that, for the ϕ-divergence distance (16), x satisfies GRC (3) if and only if there exist $v, w \in \mathbb{R}^m$ that satisfy

$$\begin{cases} a_0^T(v + w) + \delta^*(A^T v \mid Z_1) + \delta^*(A^T w \mid Z_2) - f_*(v + w, x) \le 0, \\ \varphi^*(v_i) \le v_i, \qquad i = 1, \dots, m. \end{cases} \tag{17}$$

Example 4. Suppose that it is required that the expectation of a function of the decision variable $x \in \mathbb{R}^n$ and a random variable $Y \in \mathbb{R}$ is less or equal to b, i.e.,

$$\mathrm{E}[g(x, Y)] \le b. \tag{18}$$

Let Y be a discrete random variable:

$$\Pr(Y = y_i) = a_i, \qquad i = 1, \dots, m,$$

with $a = (a_1, \dots, a_m)$ being a probability vector: $a \ge 0$ and $a^T e = 1$. Constraint (18) can be rewritten as

$$f(a, x) = \sum_{i=1}^m a_i g(x, y_i) - b \le 0. \tag{19}$$

Suppose that a is an uncertain parameter with associated uncertainty sets defined by (4) where

$$Z_1 = \{\zeta \in \mathbb{R}^L \mid a_0 + A\zeta \geq 0,\ \|\zeta\|_\infty \leq \rho\},$$
$$Z_2 = \{\zeta \in \mathbb{R}^L \mid a_0 + A\zeta \geq 0\},$$

with $a_0^T e = 1$ and $A \in \mathbb{R}^{m\times L}$ a matrix such that $A^T e = 0$. Note that these assumptions guarantee that all vectors in U_1 and U_2 are probability vectors.

Suppose that we want to handle the constraint in a globalized manner as in GRC (3) by using the phi-divergence distance (16) with

$$\varphi(t) = \tfrac{1}{t}(t-1)^2,$$

which is known as the χ^2-distance. The convex conjugate of this function is (see Ben-Tal et al., 2013)

$$\varphi^*(s) = 2 - 2\sqrt{1-s}, \qquad s < 1.$$

Note that the restriction $\varphi^*(v_i) \leq v_i$ now simplifies to $0 \leq v_i < 1$. The function $f(a,x)$ is similar to the one in Example 3, hence

$$f_*(v+w,x) = \begin{cases} b & \text{if } g(x,y_i) \leq v_i + w_i \quad \forall i, \\ -\infty & \text{otherwise.} \end{cases}$$

Note that Z_2 is polyhedral and Z_1 the intersection of a polyhedral set and ℓ_∞ norm ball. This yields

$$\delta^*(A^T w \mid Z_2) = \min_s \{a_0^T s \mid A^T s + A^T w = 0,\ s \geq 0\}$$

and

$$\delta^*(A^T v \mid Z_1) = \min_{u_1,u_2} \left\{ \min_t \{a_0^T t \mid A^T t + u_1 = 0,\ t \geq 0\} + \rho\|u_2\|_1 \,\middle|\, u_1 + u_2 = A^T v \right\}.$$

Inserting the above results in (17), then after some simplifications the GRC for this example reduces to

$$\begin{cases} a_0^T(v+w+s+t) + \rho\|A^T(v+t)\| \leq b & \\ A^T(s+w) = 0 & \\ g(x,y_i) \leq v_i + w_i & \forall i \\ 0 \leq v_i < 1 & \forall i \\ s,t \geq 0 & \\ v,w,s,t \in \mathbb{R}^m,\ x \in \mathbb{R}^n. & \end{cases}$$

□

4.3 Relation to initial GRC

Consider GRC (9) in the case of a linear uncertain constraint, i.e., $f(a,x) = a^T x - b$. In this case, the GRC can be rewritten as

$$(a_0 + A\zeta)^T x - b \leq \min_{\zeta' \in Z_1} \phi(\zeta,\zeta'), \qquad \forall \zeta \in Z_2. \tag{20}$$

This GRC is very similar to the GRC previously considered by Ben-Tal et al. (2009a, see Chapter 3). If we focus on the uncertain parameter a, and thus ignore any possible uncertainty in b, then their GRC translates to

$$(a_0 + A\zeta)^T x - b \leq \theta \min_{\zeta'} \{\|\zeta - \zeta'\| \mid \zeta' \in Z_1,\ \zeta - \zeta' \in C\}, \qquad \forall \zeta \in Z_2, \tag{21}$$

where $\theta \geq 0$ is a fixed constant, C is a closed convex cone, and $Z_2 = Z_1 + C$. Ben-Tal et al. (2009a) show that x satisfies (21) if and only if x satisfies

$$\begin{cases} (a_0 + A\zeta)^T x \leq b, & \forall \zeta \in Z_1, \\ (A\Delta)^T x \leq \theta, & \forall \Delta \in C : \|\Delta\| \leq 1. \end{cases} \tag{22}$$

Each of these semi-infinite constraints can be dealt with in the usual robust optimization approach for RCs.

We observe a number of differences between GRCs (20) and (21). First, the controlled violation in (21) is linear in the norm of $\zeta - \zeta'$, whereas this is not required in (20). Second, the largest uncertainty set Z_2 in (21) needs the special structure that it can be written as the sum of Z_1 and a closed convex cone C, which is not required in (20). Third, in some situations, the distance between the uncertain parameter ζ and Z_1 is measured in a counter-intuitive manner in (21), because the difference between ζ and ζ' has to be in the cone C.

If we assume $\phi(\zeta, \zeta') = \theta\|\zeta - \zeta'\|$, then, in light of the differences mentioned above, the GRCs (20) and (21) could turn out different or not, depending on the choices Z_1, Z_2 and C. This is illustrated by the following two examples.

Example 5. Let $Z_1 = \{\zeta \in \mathbb{R}^L \mid \|\zeta\|_\infty \leq \rho\}$, $C = \mathbb{R}^L$, and let $\|.\|$ be any fixed norm. Using the standard results in RO yields that (22) is equivalent to

$$\begin{cases} a_0^T x + \rho\|A^T x\|_1 \leq b, \\ \|A^T x\|^* \leq \theta. \end{cases}$$

On the other hand, using the result derived in (15), (20) is equivalent to

$$\begin{cases} a_0^T x + \delta^*(v \mid Z_1) + \delta^*(A^T x - v \mid Z_2) \leq b, \\ \|v\|^* \leq \theta. \end{cases}$$

Because $Z_2 = Z_1 + C = \mathbb{R}^L$, it readily follows that $A^T x = v$, and, since $\delta^*(v \mid Z_1) = \rho\|v\|_1$, it follows that GRCs (20) and (21) are identical for this example. $\square$

Example 6. Let $Z_1 = \{\zeta \in \mathbb{R}^L \mid \|\zeta\|_\infty \leq \rho\}$, $C = \{\gamma e \mid \gamma \in \mathbb{R}\}$, $Z_2 = Z_1 + C$, $\tilde{Z} = \{\Delta \in C \mid \|\Delta\| \leq 1\}$, with e the all-one vector in $\mathbb{R}^L$ and $\|.\|$ any fixed norm. So the cone C represents deviations from Z_1 that are of equal magnitude in all dimensions.

We derive the following support functions:

$$\begin{aligned} \delta^*(v \mid Z_1) &= \rho\|v\|_1, \\ \delta^*(v \mid C) &= \begin{cases} 0 & \text{if } v^T e = 0, \\ \infty & \text{if } v^T e \neq 0, \end{cases} \\ \delta^*(v \mid Z_1 + C) &= \delta^*(v \mid Z_1) + \delta^*(v \mid C), \\ \delta^*(v \mid \tilde{Z}) &= \max_{\gamma \in \mathbb{R}} \{\gamma v^T e \mid \|\gamma e\| \leq 1\} = |v^T e| / \|e\|. \end{aligned}$$

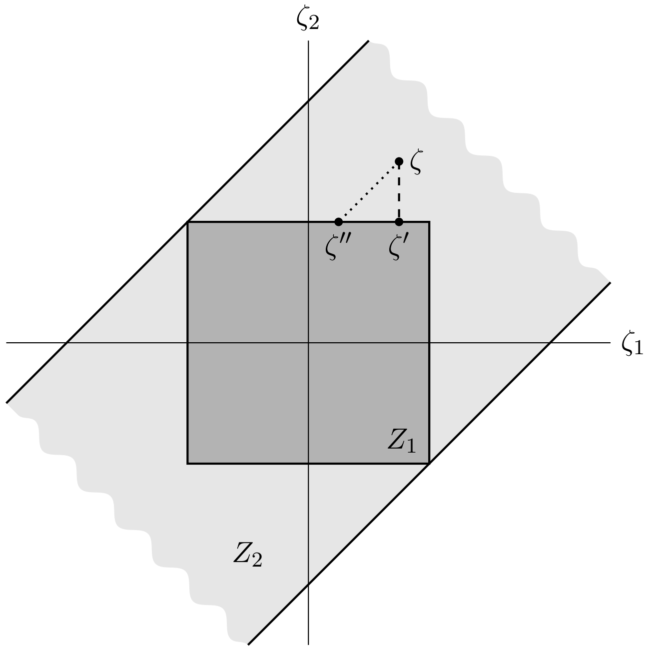


Figure 1: Different measurement of distance of ζ to Z_1: regular norm distance (dashed line) and norm distance limited to vectors in the cone C (dotted line).

Now, we find that (22) reduces to

$$\begin{cases} a_0^T x + \rho \|A^T x\|_1 \le b \\ \delta^*(A^T x \mid \tilde{Z}) \le \theta \end{cases} \quad \Leftrightarrow \quad \begin{cases} a_0^T x + \rho \|A^T x\|_1 \le b \\ |e^T A^T x| / \|e\| \le \theta. \end{cases}$$

On the other hand, we find that (20) reduces to

$$\begin{cases} a_0^T x + \rho \|v\|_1 \le b \\ e^T (A^T x - v) = 0 \\ \|v\|^* \le \theta. \end{cases}$$

Hence, in this example we see that the two GRCs are not identical. The reason for this difference is illustrated in Figure 1, which shows the two uncertainty sets Z_1 and Z_2 in $\mathbb{R}^2$. The usual distance of the point ζ to Z_1 is illustrated by the dashed line. The distance measure used in (21) is illustrated by the dotted line, $\zeta - \zeta''$ which has to be in the cone C. □

5 GRC approach to deal with infeasibilities

5.1 Trade-off between primary objective and constraint violations

If the robustness of a solution in an uncertain environment is an issue, then this implies that we are dealing with a multi-objective problem: both the primary objective and robustness are important. The traditional RO approach handles these two objectives by requiring full feasibility of the solution in an uncertainty set. A drawback of this approach is that it obfuscates the trade-off between the two objectives that is actually reflected by the choice of the uncertainty set. A larger uncertainty set puts more emphasis on the robustness than a smaller uncertainty set would. In the end, an uncertainty set has to be chosen, and for a decision maker it should be clear what this choice means in terms of the trade-off between the primary objective(s) and the robustness of the resulting solution.

Before we can actually make the trade-off, it has to be determined how the objectives, i.e., the primary objective and robustness, are actually measured in an uncertain environment. For example, the objectives can be measured by a single scenario criterion (e.g., the nominal scenario), a worst-case criterion, but also by an average behavior criterion. Moreover, the robustness can be measured by absolute feasibility, i.e., the solution is feasible or not, but also by the sum or maximum of constraint violations.

A straightforward approach to determine the trade-off between the primary objective and feasibility of a problem is to evaluate solutions of the problem's RC solved for different uncertainty sets. However, the RC can become infeasible, when the uncertainty set gets too large. So, when robustness is the main concern, then the question becomes whether this approach can offer a complete trade-off overview, or that other approaches are required to yield additional insight.

In the next section, several approaches are presented that can be used to gain more insight in the trade-off between the primary objective and the robustness of a solution.

5.2 Approaches to control constraint violations

Instead of focusing on an individual uncertain constraint, we now look at an entire optimization problem. We start by defining the uncertain optimization problem:

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(a,x) \leq 0 \qquad i = 1,\ldots,p \\ & x \in X, \end{aligned} \tag{P}$$

with an uncertain parameter $a \in \mathbb{R}^m$. For simplicity, we assume that all constraints depend on the same uncertain parameter. The restriction $x \in X$ is assumed to represent the certain constraints of the problem, where $X \subset \mathbb{R}^m$.

Suppose that (P) is feasible for the nominal value $a = a_0$, but we are interested in finding a robust solution for the uncertainty set $U_2 = \{a = a_0 + A\zeta \mid \zeta \in Z_2\}$. Therefore, we consider the RC

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(a,x) \leq 0, \qquad \forall a \in U_2,\ i = 1,\ldots,p \\ & x \in X. \end{aligned} \tag{RC}$$

The size of the uncertainty set can be varied to obtain a different trade-off between objective and robustness. However, if U_2 is too large it may happen that there is no robust feasible solution. Hence, if the uncertainty set U_2 reflects the uncertainty set where robustness of the solution is desired, then the RC approach has no way of controlling constraint violations in this uncertainty set.

An alternative approach, which can be applied to larger uncertainty sets, is to minimize the constraint violations. Taking (RC) as a starting point, we can minimize the sum of the violations by solving

$$\begin{aligned} \min \quad & \sum_{i=1}^{p} y_i \\ \text{s.t.} \quad & f_i(a,x) \leq y_i, \qquad \forall a \in U_2,\ i = 1,\ldots,p \\ & x \in X,\ y \in \mathbb{R}^p_+. \end{aligned} \tag{RC-sum}$$

Note that as an alternative, it is also possible to minimize the maximum constraint violation instead of the sum. As explained in Iancu and Trichakis (2014), solutions to RO problems are not necessarily unique. This could be the situation for (RC-sum) as well, and because the original objective function is completely ignored in (RC-sum), a second optimization problem should be solved afterwards taking the optimal objective value Y^* of (RC-sum) as an input. This yields a variation of the (RC) problem where the original constraints are now relaxed:

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(a,x) \leq y_i, && \forall a \in U_2,\ i = 1, \dots, p \\ & \sum_{i=1}^{p} y_i \leq Y^* \\ & x \in X,\ y \in \mathbb{R}^p_+. \end{aligned} \tag{RC-rel}$$

Just as it is a legitimate goal to optimize the worst-case objective in the standard RO approach, it is also a legitimate goal to minimize the worst-case sum of constraint violation as formulated in (RC-sum). However, just as for standard RO, the average behavior of the solution should be considered as well. The same can be said about problem (RC-rel). It is not necessarily the case that for a fixed uncertainty set U_2 the optimal solution to (RC-rel) also results in acceptable average behavior with respect to the constraint violations. The worst-case sum of violations might be obtained for a solution that results in unacceptably high violations for scenarios close to the nominal value. Therefore, to improve average behavior, it might be worthwhile to consider problems (RC-sum) and (RC-rel) with uncertainty sets of different sizes.

Yet another alternative to deal with the infeasibility of (RC) is to take a GRC approach. Hereto, first an inner uncertainty set U_1 needs to be defined for which (RC) is feasible. Similar to (RC-sum), we can then minimize the sum of the weights of the distance measure functions ϕ_i. This leads to the problem

$$\begin{aligned} \min \quad & \sum_{i=1}^{p} \theta_i \\ \text{s.t.} \quad & f_i(a,x) \leq \theta_i \min_{a' \in U_1} \phi_i(a,a'), && \forall a \in U_2,\ i = 1, \dots, p \\ & x \in X,\ \theta \in \mathbb{R}^p_+. \end{aligned} \tag{GRC-sum}$$

Once the optimal objective value Θ^* has been found, we can find the solution with the best objective value for the original problem among all optimal solutions to (GRC-sum) by solving

$$\begin{aligned} \min \quad & f_0(x) \\ \text{s.t.} \quad & f_i(a,x) \leq \theta_i \min_{a' \in U_1} \phi_i(a,a'), && \forall a \in U_2,\ i = 1, \dots, p \\ & \sum_{i=1}^{p} \theta_i \leq \Theta^* \\ & x \in X,\ \theta \in \mathbb{R}^p_+. \end{aligned} \tag{GRC-rel}$$

The advantage of the GRC approach is that the average behavior of the constraint violations can be further improved compared to the RC approach. The reason is that the RC approach either takes full responsibility for the constraint violation, or no responsibility at all in the case of a shrinked uncertainty set. The GRC approach, on the other hand, considers all possible constraint violations but with varying importance.

Note that the bound on the weights θ_i in (GRC-rel) can be relaxed by

$$\sum_{i=1}^{p} \theta_i \leq (1+\kappa)\Theta^*, \tag{23}$$

where κ is a nonnegative parameter that gives the decision maker further control on the trade-off between the solution's objective value and robustness. As κ increases, constraint violations outside U_1 get more relaxed and the problem approaches the (RC) model applied to uncertainty set U_1 instead of U_2.

6 Numerical results

In this section we study an inventory model that has previously been studied by Ben-Tal et al. (2004). In this model production levels have to be chosen such that demand can be supplied and the inventory level is between given lower and upper bounds. The demands in all periods of the planning horizon are uncertain. Ben-Tal et al. (2004) solve the model using an affinely adjustable robust counterpart (AARC) approach, i.e., production decisions follow from decision rules that depend on the observed demands in the previous periods. This leads to much better results than the RC approach, where the production levels of all periods have to be made at the start of the planning horizon.

In the current example, we return to the more classical situation where all decisions are here-and-now decisions. This is not uncommon for short term or medium term planning periods. Often, decisions have to be communicated to suppliers, employees, and customers well in advance. Therefore, it is impossible or undesirable to change decisions when new information becomes available. Thus, our goal is to minimize the costs of the production plan, but the primary focus is to obtain a production plan that is robust, i.e., the plan is feasible for all demand scenarios where feasibility is a reasonable requirement, and the plan yields constraint violations that are few in number and small in magnitude when the demand scenario is such that feasibility is not possible or not a reasonable requirement.

6.1 The production-inventory model

We consider a single product inventory system, which is comprised of a warehouse and I factories. A planning horizon of T periods is used. In the model we use the following parameters and variables, using the same notation as in Ben-Tal et al. (2004).

Parameters.

d_t : Demand for the product in period t;

$P_i(t)$: Production capacity of factory i in period t;

$c_i(t)$: Cost of producing one product unit at factory i in period t;

$V_{\min}$: Minimal allowed level of inventory at the warehouse;

$V_{\max}$: Storage capacity of the warehouse;

Q_i : Cumulative production capacity of the ith factory throughout the planning horizon.

Variables.

$p_i(t):$ The amount of the product to be produced in factory i in period t;

$v(t):$ Inventory level at the beginning of period t ($v(1)$ is given).

Optimization model. We try to minimize the total production costs over all factories and the whole planning horizon. The restriction is that all demand in period t must be satisfied by units on stock in the warehouse or by the production in period t. If all the demand, and all other parameters, are certain in all periods $1, \dots, T$, then the problem is modeled by the following linear optimization model:

$$
\begin{aligned}
\min_{p_i(t), v(t), F} \quad & F \\
\text{s.t.} \quad & \sum_{t=1}^{T} \sum_{i=1}^{I} c_i(t) p_i(t) \leq F \\
& 0 \leq p_i(t) \leq P_i(t) && i = 1, \dots, I,\ t = 1, \dots, T \\
& \sum_{t=1}^{T} p_i(t) \leq Q_i && i = 1, \dots, I \\
& v(t+1) = v(t) + \sum_{i=1}^{I} p_i(t) - d_t && t = 1, \dots, T \\
& V_{\min} \leq v(t) \leq V_{\max} && t = 2, \dots, T+1.
\end{aligned}
$$

By eliminating the v-variables, we obtain the following linear optimization model:

$$
\begin{aligned}
\min_{p_i(t), F} \quad & F \\
\text{s.t.} \quad & \sum_{t=1}^{T} \sum_{i=1}^{I} c_i(t) p_i(t) \leq F \\
& 0 \leq p_i(t) \leq P_i(t) && i = 1, \dots, I,\ t = 1, \dots, T \\
& \sum_{t=1}^{T} p_i(t) \leq Q_i && i = 1, \dots, I \\
& V_{\min} \leq v(1) + \sum_{s=1}^{t} \sum_{i=1}^{I} p_i(s) - \sum_{s=1}^{t} d_t \leq V_{\max} && t = 1, \dots, T.
\end{aligned}
\tag{24}
$$

6.2 Illustrative data set

We use the data set from the illustrative example by Ben-Tal et al. (2004). Hence, there are $I = 3$ factories and the planning period is $T = 24$ periods, representing one full season of 48 weeks. The nominal demand $d_t^\star$ has a seasonal pattern given by

$$
d_t^\star = 1000 \left(1 + \tfrac{1}{2} \sin\left(\tfrac{\pi(t-1)}{12}\right)\right), \quad t = 1, \dots, T.
$$

The production costs depend on the factory and on time and follow the same seasonal pattern as the demand. The production cost per unit for factory i at period t is given by:

$$
c_i(t) = \alpha_i \left(1 + \tfrac{1}{2} \sin\left(\tfrac{\pi(t-1)}{12}\right)\right), \qquad t = 1, \dots, T,
$$

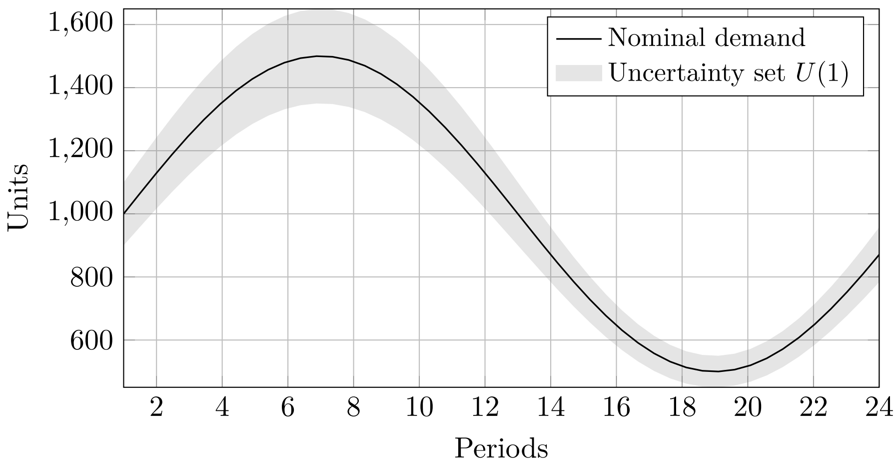


Figure 2: The nominal demand pattern and the uncertainty set $U(1)$.

with

$$\alpha_1 = 1, \quad \alpha_2 = 1.5, \quad \alpha_3 = 2.$$

The maximal production capacity of each factory during each period is $P_i(t) = 567$ units, and the integral production capacity of each one of the factories for the planning horizon is $Q_i = 13{,}600$. The inventory at the warehouse should be no less than 500 units, and cannot exceed 2,000 units.

It is assumed that the demand is the only uncertain data in the problem. We assume that $d_t \in [0.9d_t^\star, 1.1d_t^\star]$, i.e., the uncertainty level is 10%, which is lower than the 20% used in Ben-Tal et al. (2004). We use a lower uncertainty level to counterbalance the fact that we consider here-and-now decisions instead of wait-and-see decisions. Using the notation for uncertainty sets introduced in Section 2, this corresponds to the uncertainty set

$$U(\rho) = \left\{ d \in \mathbb{R}^T \;\middle|\; d = d^\star + A\zeta,\ \zeta \in Z(\rho) \right\},$$

where

$$Z(\rho) = \left\{ \zeta \in \mathbb{R}^T \;\middle|\; \|\zeta\|_\infty \leq \rho \right\}, \quad d^\star = (d_1^\star, \ldots, d_T^\star)^T, \quad A = 0.1\,\mathrm{diag}(d^\star).$$

The parameter ρ determines the size of the ℓ_∞-box for the primitive uncertainties ζ, and consequently also the size of the uncertainty set $U(\rho)$. The original uncertainty set, which corresponds to the uncertainty level 10%, is $U(1)$. This is illustrated in Figure 2. In the next section, also other uncertainty sets will be used by taking different values for ρ.

6.3 The experiments

In this section, several solutions for the production-inventory problem are obtained using the approaches from Section 5. The solutions are compared by the objective value and several robustness measures related to the average behavior of constraint violations in the optimization problem (24). The average behavior is evaluated by sampling 5,000 demand patterns uniformly from the uncertainty set $U(1)$. The reported performance measures related to the constraint violations of a solution are:

Model	ρ	Objective value	Deviation (%)	Average behavior (uniform distribution) Constraints violated (%)	Infeasible (%)	Max violation	Sum violation
P(nominal)	–	33,822	0	14.6	94.8	219.4	1,331.1
RC	0.31	35,758	5.7	1.8	34.2	27.7	67.3
RC-rel	0.41	35,778	5.8	1.0	22.1	16.9	34.1
RC-rel	1.00	33,221	−1.8	19.1	99.9	899.5	3,551.1
GRC-rel-L	0.31	35,918	6.2	0.8	8.8	9.1	34.1
GRC-rel-Q	0.31	36,529	8.0	0.2	1.7	1.7	8.5
GRC-rel-Q	0.01	36,426	7.7	0.2	2.1	2.0	8.9
GRC-rel-Q[a]	0.01	35,721	5.6	0.5	7.9	9.2	23.5

Table 2: Numerical results for the inventory problem. For the RC and RC-rel models, ρ denotes the radius of primitive uncertainties. For the GRC models, this refers to the inner uncertainty set. The performance of the solutions is given by the objective value, its deviation from the nominal model's objective value, and the average behavior with respect to the (48) uncertain constraints. The percentage of violated constraints, the percentage of infeasible solutions, the maximum violation and the sum of the constraint violations are reported for a simulation study of 5,000 demand scenarios sampled uniformly from the uncertainty region (radius 1).

[a]Obtained by using the relaxed constraint (23) on the weights with $\kappa = 0.03$.

Constraints violated (%) The percentage (averaged over the sampled demand patterns) of uncertain constraints that are violated by the solution. Recall that problem (24) has $2T = 48$ uncertain constraints, but at most $T = 24$ constraints can be violated simultaneously.

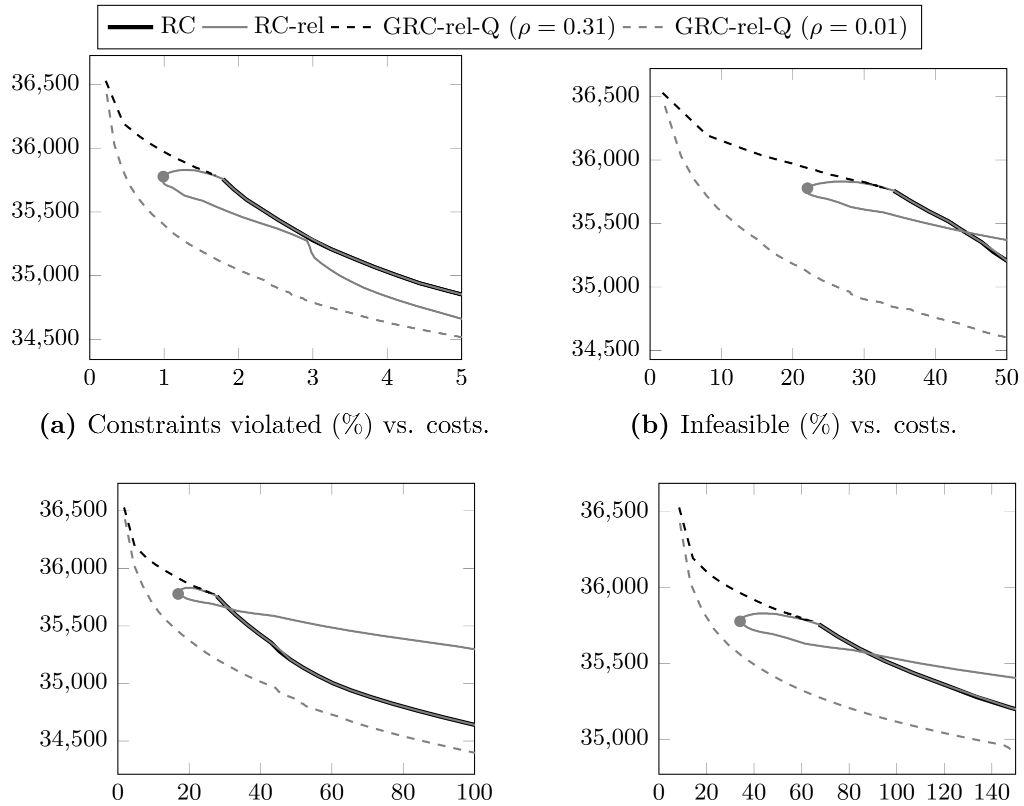
Infeasible (%) The percentage of demand patterns for which the solution violates at least one uncertain constraint.

Max violation The average of the solution's largest constraint violation for a demand pattern.

Sum violation The average of the sum of all constraint violations.

The optimization problem (24) can be solved for the nominal demand pattern $d^\star$. This model does not consider any uncertainty at all, and therefore, its solution cannot be expected to perform well on the average behavior for the constraint violations. This is indeed the case as can be seen in Table 2. For example, almost 95% of the sampled demand patterns result in at least one constraint violation, and the average sum of constraint violations is 1,331.1.

Next, we shall explore the trade-off between objective value (total production costs) and the robustness by considering the approaches from Section 5.2. First, consider the RC of problem (24), analogously to the RC of problem (P) we considered in Section 5. This problem is infeasible for uncertainty set $U(1)$. The largest uncertainty set for which the RC is feasible is $U(0.31)$. The optimal solution for this model is not necessarily unique, and therefore, we obtained the solution, among all optimal solutions, that minimizes the sum of the slacks of all uncertain constraints, which guarantees a Pareto robustly optimal (PRO) solution (see Iancu and Trichakis, 2014). The RC solution results in a deterioration of 5.7% in objective value compared to the nominal problem's objective value. On the other hand, a significant performance increase is obtained for the constraint violations. If we apply the RC model for



(a) Constraints violated (%) vs. costs.

(b) Infeasible (%) vs. costs.

(c) Maximum violation vs. costs.

(d) Sum of constraint violations vs. costs.

Figure 3: Trade-off between objective value (costs) and four robustness measures for solutions obtained by different approaches. The most robust points on the RC curves are obtained with $U(0.31)$. The most robust points on the RC-rel curves are marked by a dot and are obtained by using $U(0.41)$.

smaller uncertainty sets $U(\rho)$, with $\rho < 0.31$, then the objective value improves at the expense of the robustness. This is visualized in Figure 3.

Another class of solutions can be obtained by applying model (RC-rel) to the production-inventory problem (24). This model is feasible for any uncertainty set $U(\rho)$, so we can apply it for the original uncertainty set $U(1)$, but also for smaller sets. If we look at Figure 3, we can make some interesting observations. For $\rho \leq 0.31$, the (RC-rel) model yields similar results as the (RC) model. If we increase ρ, then we see that the robustness of the solutions can be improved until $\rho = 0.41$. For $\rho > 0.41$, the robustness starts deteriorating again. Apparently, protecting against the worst-case constraint violations for larger uncertainty sets is not necessarily good for the robustness criteria based on the average behavior.

The (GRC-rel) model is also applied to the production-inventory problem with inner uncertainty set $U_1 = U(0.31)$ and outer uncertainty set $U_2 = U(1)$. This inner uncertainty set guarantees feasibility for the same uncertainty set for which feasibility can also be guaranteed for the (RC) model, but later we shall also apply the model with a smaller inner uncertain set.

Two different types of distance measure functions are applied to the (GRC-rel) model:

$$\phi_i(a, a') = \|a - a'\|_1,$$

and

$$\phi_i(a, a') = \|a - a'\|_2^2,$$

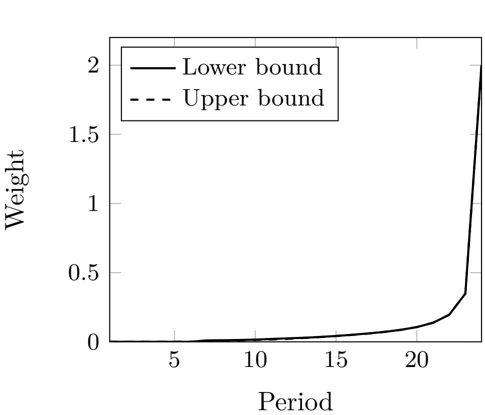
and the resulting problems are referred to by (GRC-rel-L) and (GRC-rel-Q), respectively. Note that problem (GRC-rel-L) can be formulated as a linear optimization problem, and problem (GRC-rel-Q) is a conic quadratic optimization problem.

It can be seen from Table 2 that with the (GRC-rel) models solutions can be obtained that improve the robustness beyond a level than is possible with the previously discussed approaches. In this respect, the (GRC-rel-Q) model yields better results than the (GRC-rel-L) model, which is also more sensitive to the choice of the inner uncertainty set. The (GRC-rel-L) model suffers from the same drawback that we observed for the (RC-rel) model albeit less severely: protecting against worst-case sum of violations is not necessarily good for the average behavior. The (GRC-rel-Q) model with the squared norm, on the other hand, allows relatively large constraint violations for parameters further away from the inner uncertainty set than the (GRC-rel-L) model, but it does not neglect them either such as the (RC-rel) model does for small ρ. The solutions obtained for the (GRC-rel-Q) model violated only 0.2% of the 48 uncertain constraints and the solutions are infeasible for only 1.7% of the demand patterns. Also the maximum and sum of violations are very small compared to the other approaches. We also solved the (GRC-rel) model with a distance measure function that is linear in the ℓ_2-norm, i.e., $\phi_i(a, a') = \|a - a'\|_2$, and the results were very similar to the (GRC-rel-L) model. Thus, the better results of the (GRC-rel-Q) model compared with the (GRC-rel-L) models are mainly caused by the fact that the (GRC-rel-Q) model uses a quadratic distance measure function instead of a linear one, and not by the underlying norm (2-norm vs. 1-norm).

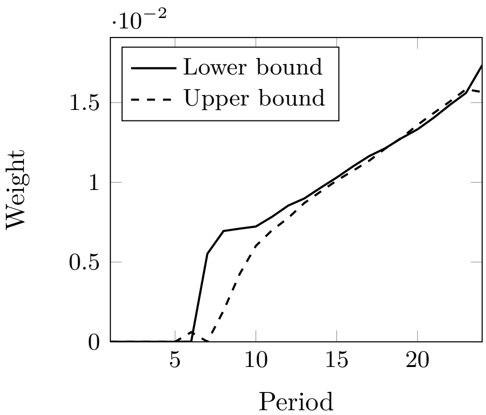
There is actually quite a gap between the most robust solutions for model (GRC-rel-Q) and the (RC) model. It turns out that we can close this gap and obtain intermediate solutions by relaxing the upper bound on the distance measure weights in (GRC-rel-Q) by using (23) instead. Figure 3 illustrates the trade-off between the objective value and robustness measures that can be obtained by this relaxation.

The (GRC-rel-Q) model can also applied with smaller inner uncertainty set than $U(0.31)$. An interesting suggestion could be to pick the uncertainty set for which the nominal solution remains feasible. Here we choose the uncertainty set $U(0.01)$, which is actually a very small region surrounding the nominal value. This inner uncertainty set does no longer guarantee feasibility in the largest possible uncertainty set where this is possible, which is $U(0.31)$. However, the performances of the solutions that we obtain for the (GRC-rel) model with inner uncertainty sets $U(0.01)$ and $U(0.31)$ are almost identical for both the objective value and average robustness measures. On the other hand, the trade-off curves for both uncertainty sets, which are obtained by applying the relaxation on the weights, are quite different. As we have already observed, the trade-off curve for inner uncertainty set $U(0.31)$ approaches the solution obtained by the (RC) model. The trade-off curve for inner uncertainty set $U(0.01)$ yields combinations of objective values and averge robustness measures that are well beyond the results obtained by the alternative models. The performance of one of these solutions, which has a comparable objective value to the solutions obtained by the (RC) model with $U(0.31)$ and (RC-rel) model with $U(0.41)$, is shown in Table 2.

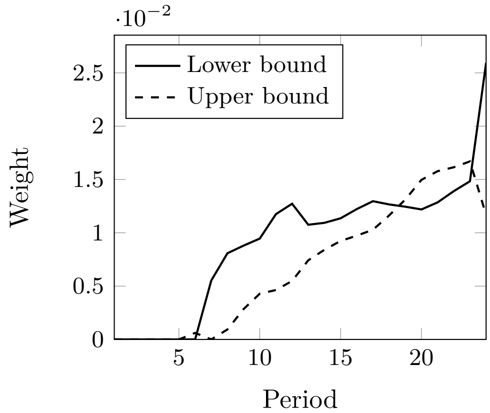
Solving the (GRC-rel) model also yields the optimal weight θ_i for the distance measure functions in the GRC of the ith constraint. This gives an indication what constraints are more likely to be violated or more likely to have larger violations



(a) (GRC-rel-Q) with $U(0.31)$.



(b) (GRC-rel-Q) with $U(0.01)$.



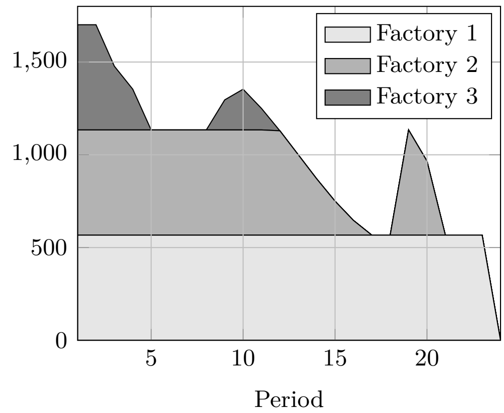
(c) (GRC-rel-Q) with $U(0.01)$ and relaxed weights with $\kappa = 0.03$.

Figure 4: Optimal weights for distance measure functions corresponding to the GRCs for the stock lower and upper bound in each period.

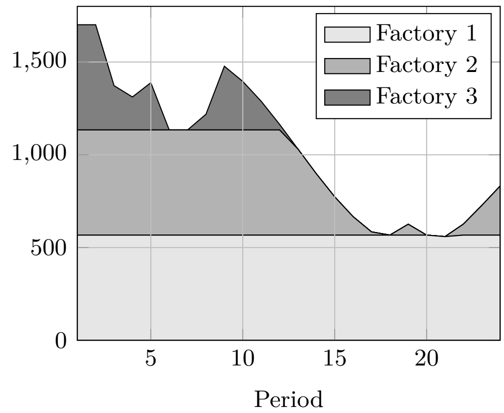
than other constraints. Figure 4 shows the weights for constraints (lower and upper bound) on the stock levels in each period of the planning horizon. Not surprisingly, the weights are larger towards the end of the planning horizon, because corresponding constraints are affected by more uncertain demands than stock level constraints near the start of planning horizon.

Figure 5 shows a graphical illustration of the actual production plans for four different models. Recall that the factory's production costs follow the seasonal pattern and that factory 1 has the lowest production costs, and factory 3 the largest. Near the end of the planning horizon, that is roughly from period 15 to 24, we observe different patterns of inventory building to match the demand in the last periods. Clearly, the nominal solution builds up inventory in the low-cost season (period 17–21) and stops production in in the last period. The (RC) and (RC-rel) solutions postpone this production more towards the end of the planning horizon. The (GRC-rel-Q) solution follows the nominal demand pattern more directly, and does not build up inventory at all, and produces less in the low-demand season and consequently more towards the end of the planning horizon.

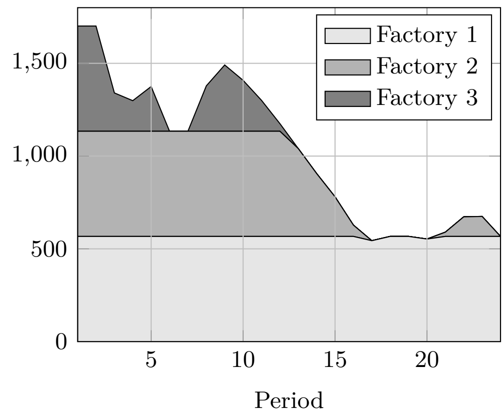
The sampling method used in Table 2 is the uniform distribution in $U(1)$, which corresponds to i.i.d. random samples taken from Unif$(-1, 1)$ for ζ_t, $t = 1, \dots, T$. Alternatively, we also consider sampling ζ_t from the triangular distribution, and two



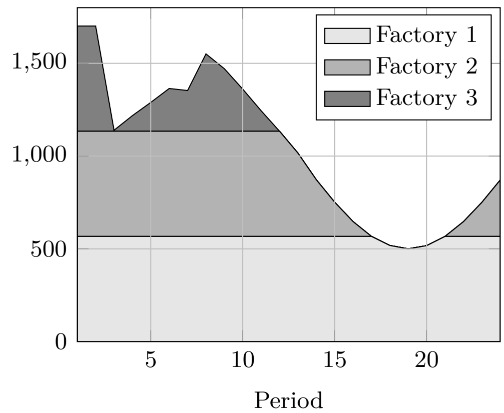
(a) Nominal solution



(b) RC solution for $\rho = 0.31$



(c) RC-rel solution for $\rho = 0.41$



(d) GRC-rel-Q solution for $\rho = 0.31$

Figure 5: Graphical illustration of production plans.

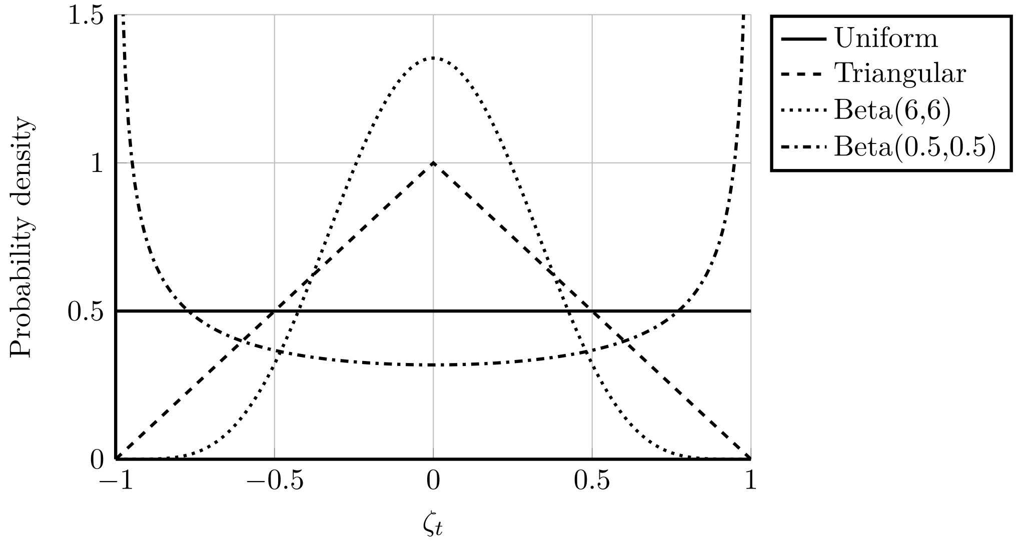


Figure 6: Distributions used for sampling ζ_t from $U(1)$, $t = 1, \ldots, T$.

different Beta distributions mapped on the $[-1, 1]$ interval. All four distributions are illustrated in Figure 6.

We have also evaluated the average behavior of a selection of the most interesting models in Table 2 for the alternative sampling methods. The results are shown in Table 3. The results are according to our expectations: compared to the uniform distribution, the average performance of all solutions improves for the triangular and Beta$(6, 6)$ distributions and deteriorates for the Beta$(0.5, 0.5)$ distribution. However, the relative performance differences of the solutions does not change much, i.e., the GRC-rel-Q solution gives still the best average performance for the constraint violations.

7 Concluding remarks

In this paper, we have derived tractable GRCs. Compared to the initial GRCs in the literature, our GRC is applicable to nonlinear constraints instead of only linear or conic constraints, and the GRC is more flexible with respect to both the uncertainty set and distance measure function, which are used to control the constraint violations. These extensions make the concept of globalized robust optimization more applicable.

We showed that GRCs are useful to provide an extended trade-off overview between the objective value and robustness measures of different solutions. This is particularly relevant when the RC model turns out to be infeasible for a chosen uncertainty set. However, also in case the RC model is feasible, the GRC approach could be used as a means to explore the trade-off between relaxation of the feasibility requirement in the uncertainty set against improvement of the problem's objective value.

The GRC approach presented in this paper offers the ability to find solutions that improve average robustness measures beyond the possibilities of RC-based approaches. This ability appears to be, at least partially, based on the newly derived tractable GRCs, which allow the use of nonlinear distance measure functions to control constraint violations, such as the quadratic distance measure used in Section 6.3.

Model	ρ	Objective value	Average behavior				
			Distribution	Constraints violated (%)	Infeasible (%)	Max violation	Sum violation
P(nominal)	–	33,822	Uniform	14.6	94.8	219.4	1,331.1
			Triangular	14.0	93.8	150.9	905.3
			Beta(6,6)	13.4	94.5	103.4	599.6
			Beta(0.5,0.5)	15.5	96.7	282.4	1,714.7
RC	0.31	35,758	Uniform	1.8	34.2	27.7	67.3
			Triangular	0.5	15.0	6.4	10.4
			Beta(6,6)	0.1	3.2	0.7	0.9
			Beta(0.5,0.5)	3.3	50.0	58.3	172.1
RC-rel	0.41	35,778	Uniform	1.0	22.1	16.9	34.1
			Triangular	0.2	6.4	2.4	3.1
			Beta(6,6)	0.0	0.4	0.1	0.1
			Beta(0.5,0.5)	2.2	37.1	41.9	105.9
GRC-rel-L	0.31	35,918	Uniform	0.8	8.8	9.1	34.1
			Triangular	0.1	1.7	0.8	1.8
			Beta(6,6)	0.0	0.1	0.0	0.0
			Beta(0.5,0.5)	1.8	17	24.5	107.1
GRC-rel-Q	0.31	36,529	Uniform	0.2	1.7	1.7	8.5
			Triangular	0.0	0.0	0.0	0.0
			Beta(6,6)	0.0	0.0	0.0	0.0
			Beta(0.5,0.5)	0.8	6.4	7.8	41.1

Table 3: Performance of solutions for various distributions for ζ_t, $t = 1, \ldots, T$. For the RC and RC-rel models, ρ denotes the radius of primitive uncertainties. For the GRC models, this refers to the inner uncertainty set. The performance of the solutions is given by the objective value and the average behavior with respect to the (48) uncertain constraints. The percentage of violated constraints, the percentage of infeasible solutions, the maximum violation and the sum of the constraint violations are reported by sampling from four different distribution on $U(1)$.

References

Babonneau, F., J.-Ph. Vial, and R. Apparigliato (2010). Robust Optimization for Environmental and Energy Planning. *Uncertainty and Environmental Decision Making*. Ed. by J. A. Filar and A. Haurie. Vol. 138. International Series in Operations Research & Management Science. Springer US, pp. 79–126.

Ben-Tal, A., D. den Hertog, and J.-Ph. Vial (2015). Deriving Robust Counterparts of Nonlinear Uncertain Inequalities. *Mathematical Programming* 149 (1-2), pp. 265–299.

Ben-Tal, A., A. Goryashko, E. Guslitzer, and A. Nemirovski (2004). Adjustable Robust Solutions of Uncertain Linear Programs. *Mathematical Programming* 99 (2), pp. 351–376.

Ben-Tal, A., S. Boyd, and A. Nemirovski (2006). Extending Scope of Robust Optimization: Comprehensive Robust Counterparts of Uncertain Problems. *Mathematical Programming* 107 (1), pp. 63–89.

Ben-Tal, A., D. den Hertog, A. De Waegenaere, B. Melenberg, and G. Rennen (2013). Robust Solutions of Optimization Problems Affected by Uncertain Probabilities. *Management Science* 59 (2), pp. 341–357.

Ben-Tal, A., L. El Ghaoui, and A. Nemirovski (2009a). *Robust Optimization*. Princeton Series in Applied Mathematics. Princeton University Press.

Ben-Tal, A., B. Golany, and S. Shtern (2009b). Robust Multi-echelon Multi-period Inventory Control. *European Journal of Operational Research* 199 (3), pp. 922–935.

Bertsimas, D., D. B. Brown, and C. Caramanis (2011). Theory and Applications of Robust Optimization. *SIAM Review* 53 (3), pp. 464–501.

El Ghaoui, L. (2012). *Robust Optimization and Machine Learning*. Lecture slides. TRANSP-OR Seminar, Zinal, Switserland. UC Berkely.

Evers, L., T. Dollevoet, A. I. Barros, and H. Monsuur (2011). *Robust UAV Mission Planning*. Econometric Institute Report EI 2011-07. The Netherlands: Econometric Institute, Erasmus University Rotterdam.

Gabrel, V., C. Murat, and A. Thiele (2014). Recent Advances in Robust Optimization: An Overview. *European Journal of Operational Research* 235 (3), pp. 471–483.

Iancu, D. A. and N. Trichakis (2014). Pareto Efficiency in Robust Optimization. *Management Science* 60 (1), pp. 130–147.

Naseraldin, H. and O. Baron (2011). *Robust Facility Location Under Service Constraints*. Working paper.

Rockafellar, R. T. (1970). *Convex Analysis*. Princeton Mathematical Series. Princeton University Press.

Xu, H., C. Caramanis, and S. Mannor (2012). Optimization Under Probabilistic Envelope Constraints. *Operations Research* 60 (3), pp. 682–699.

Zymler, S., B. Rustem, and D. Kuhn (2011). Robust Portfolio Optimization With Derivative Insurance Guarantees. *European Journal of Operational Research* 210 (2), pp. 410–424.

A Conjugate functions, support functions and Fenchel duality

In this section we give some basic results on conjugate functions, support functions and Fenchel duality. For a detailed treatment we refer to Rockafellar (1970).

Lemma A.1. *Let $S_1, ..., S_k$ be closed convex sets, such that $\cap_i \operatorname{ri}(S_i) \neq \emptyset$, and let $S = \cap_{i=1}^k S_i$. Then*

$$\delta^*(y \mid S) = \min\left\{ \sum_{i=1}^k \delta^*(v_i \mid S_i) \;\middle|\; \sum_{i=1}^k v_i = y \right\}.$$

Lemma A.2. *Assume that f_i, $i = 1, \dots, k$, are concave, and the intersection of the relative interiors of the domains of f_i, $i = 1, \dots, k$, is nonempty, i.e., $\cap_{i=1}^k \operatorname{ri}(\operatorname{dom}(f_i)) \neq \emptyset$. Then*

$$\left(\sum_{i=1}^k f_i \right)_* (s) = \sup_{v_1, \dots, v_k} \left\{ \sum_{i=1}^k (f_i)_*(v_i) \;\middle|\; \sum_{i=1}^k v_i = s \right\},$$

and the $\sup$ *is attained for some $v_1, \dots, v_k$.*

Lemma A.3. *Assume that f_i, $i = 1, \dots, k$, are convex, and the intersection of the relative interiors of the domains of f_i, $i = 1, \dots, k$, is nonempty, i.e., $\cap_{i=1}^k \operatorname{ri}(\operatorname{dom}(f_i)) \neq \emptyset$. Then*

$$\left(\sum_{i=1}^k f_i \right)^* (s) = \inf_{v_1, \dots, v_k} \left\{ \sum_{i=1}^k (f_i)^*(v_i) \;\middle|\; \sum_{i=1}^k v_i = s \right\},$$

and the $\inf$ *is attained for some $v_1, \dots, v_k$.*

Theorem A.4 (Fenchel duality theorem)**.** *Let f be a proper convex function and g a proper concave function and define the primal problem*

$$\inf\{f(x) - g(x) \mid x \in \operatorname{dom}(f) \cap \operatorname{dom}(-g)\} \tag{25}$$

and its Fenchel dual problem

$$\sup\{g_*(y) - f^*(y) \mid y \in \operatorname{dom}(-g_*) \cap \operatorname{dom}(f^*)\}. \tag{26}$$

1. *If $\operatorname{ri}(\operatorname{dom}(f)) \cap \operatorname{ri}(\operatorname{dom}(-g)) \neq \emptyset$, then the optimal values of* (25) *and* (26) *are equal and the minimum value of* (25) *is attained.*
2. *If $\operatorname{ri}(\operatorname{dom}(f^*) \cap \operatorname{ri}\operatorname{dom}(-g_*) \neq \emptyset$, then the optimal values of* (25) *and* (26) *are equal and the maximum value of* (26) *is attained.*

Note that since $f^{**} = f$ and $g_{**} = g$, we have that the dual of (26) is (25).