Math. Program., Ser. B 107, 63–89 (2006)

Digital Object Identifier (DOI) 10.1007/s10107-005-0679-z

Aharon Ben-Tal · Stephen Boyd · Arkadi Nemirovski

# Extending Scope of Robust Optimization: ⋆

## Comprehensive Robust Counterparts of Uncertain Problems

Received: 14 March 2005 / Accepted: 11 July 2005
Published online: February 2, 2006 – © Springer-Verlag 2006

**Abstract.** In this paper, we propose a new methodology for handling optimization problems with uncertain data. With the usual Robust Optimization paradigm, one looks for the decisions ensuring a required performance for all realizations of the data from a given bounded uncertainty set, whereas with the proposed approach, we require also a controlled deterioration in performance when the data is outside the uncertainty set.

The extension of Robust Optimization methodology developed in this paper opens up new possibilities to solve efficiently multi-stage finite-horizon uncertain optimization problems, in particular, to analyze and to synthesize linear controllers for discrete time dynamical systems.

## 1. Introduction

In this paper, our aim is to extend the scope of Robust Optimization (RO). To explain the new directions into which we take the basic RO, we start with brief overview of the main concepts of this methodology.

*Uncertain convex problems.* RO is a methodology for modelling *uncertain* optimization problems of the form

$$\min_{\chi} \{F_0(\chi, \zeta) : F_i(\chi, \zeta) \in K_i,\ i = 1, \dots, I\}, \qquad (1)$$

where

- $\chi \in \mathbf{R}^{n_\chi}$ is the vector of decision variables,
- $\zeta \in \mathbf{R}^{n_\zeta}$ is the vector of *problem's data*,
- $F_0(\chi, \zeta) : \mathbf{R}^{n_\chi} \times \mathbf{R}^{n_\zeta} \to \mathbf{R}$, $F_i(\chi, \zeta) : \mathbf{R}^{n_\chi} \times \mathbf{R}^{n_\zeta} \to \mathbf{R}^{k_i}$, $1 \le i \le I$, are given functions, and $K_i \subset \mathbf{R}^{k_i}$ are given nonempty sets.

A. Ben-Tal: Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Technion city, Haifa 32000, Israel

S. Boyd: Department of Electrical Engineering, Stanford University, Packard 264, Stanford, CA 94305, USA

A. Nemirovski: Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Technion city, Haifa 32000, Israel

*Mathematics Subject Classification (2000):* 90C05, 90C25, 90C34, 93C05

⋆ Research was supported by the Binational Science Foundation grant #2002038

Uncertainty means that *the data vector $\zeta$ is not known exactly at the time when the solution has to be determined. In RO, the aim is to choose a solution which is capable "to cope" best of all with various realizations of the data.* Needless to say, the latter sentence has no sense unless we specify what we mean by "to cope" and "best of all". At present, RO has focused on two specifications of this type: *non-adjustable* and *adjustable* ones.

*A. Non-adjustable Robust Optimization.* In hindsight, this approach is based on the following tacitly accepted assumptions:

**A.1.** All decision variables in (1) represent "here and now" decisions; they should get specific numerical values as a result of solving the problem and *before* the actual data "reveal itself";

**A.2.** The uncertain data are "unknown but bounded", so that one can specify an appropriate (typically, bounded) *uncertainty set* $\mathcal{U} \subset \mathbf{R}^{n_\zeta}$ of possible values of the data. The decision maker is fully responsible for consequences of his/her decisions when, and only when, the actual data is within this set;

**A.3.** The constraints in (1) are "hard", that is, we cannot tolerate violations of constraints, even small ones, when the data is in $\mathcal{U}$.

A natural conclusion from A.1 – A.3 is that the only admissible candidate solutions of problem (1) are fixed vectors $\chi \in \mathbf{R}^{n_\chi}$ which satisfy the semi-infinite constraints

$$F_i(\chi, \zeta) \in K_i \ \forall \zeta \in \mathcal{U}, \ 1 \le i \le I, \tag{2}$$

that is, remain feasible whatever is a realization of the data from $\mathcal{U}$. Applying the same worst-case-oriented approach to the objective function leads to a new "robust" objective function

$$\bar{F}_0(\chi) = \sup_{\zeta \in \mathcal{U}} F_0(\chi, \zeta). \tag{3}$$

Summarizing, with the above approach "to solve" uncertain optimization problem (2) means, *by definition*, to solve its *Robust Counterpart* (RC) – the semi-infinite optimization problem

$$\begin{array}{c} \min\limits_{\chi} \left\{ \sup\limits_{\zeta \in \mathcal{U}} F_0(\chi, \zeta) : F_i(\chi, \zeta) \in K_i \ \forall \zeta \in \mathcal{U} \right\} \\ \Updownarrow \\ \min\limits_{\chi, \sigma} \left\{ \sigma : \left. \begin{array}{l} F_0(\chi, \zeta) \le \sigma \\ F_i(\chi, \zeta) \in K_i, \ 1 \le i \le I \end{array} \right\} \forall \zeta \in \mathcal{U} \right\}. \end{array} \tag{4}$$

The idea of robust feasibility in Linear Programming was discussed as early as in 1973 by Soyster [15]. The in-depth developments of RO occurred since the mid-90's [1, 2, 12, 13, 3] and opened the way to extensive research on the subject in the recent years (see, e.g., [8–10] and references therein).

In perspective, the above developments dealt with *static* RO; recently, the RO methodology was enriched by introducing a novel concept of *Adjustable Robust Counterpart* [6], allowing to handle dynamical decision-making.

*B. Affinely Adjustable Robust Optimization.* The idea underlying the latter approach comes from revising Assumption A.1 "all decision variables in (1) represent "here and now" decisions and as such should be specified before the actual data become known". This assumption is, first, independent of two other assumptions and, second, it is unnatural in numerous models of real-life origin. For example, in dynamical decision-making only part of the decisions are "here and now" ones, while the remaining variables represent "wait and see" decisions. These latter decisions need to be fully specified (assigned numerical values) when part of the data is already known and thus are *adjustable* – can tune themselves to the actual data. Another typical source of adjustability are "analysis variables" which do not represent actual decisions and are introduced in order to convert constraints in (1) to a desired form (like slack variables which allow to represent a constraint $\sum_j |a_j \chi_j| \leq 1$ by a system of linear inequalities). Analysis variables usually merely certify certain property of actual decisions; even when the latter are "here and now" ones, there is absolutely no necessity for the certificates to be so as well.

A natural way to account for adjustability is as follows. We assign every decision variable $\chi_j$ in (1) with a "portion of data" $P_j\zeta$ on which $\chi_j$ can depend. Here $P_j$ are given matrices[1] which, in particular, could be zero (meaning that $\chi_j$ is a *non-adjustable* "here and now" decision). We now allow decision variable $\chi_j$ to depend on $P_j\zeta$ and seek for the dependencies $\chi_j = \chi_j(P_j\zeta)$ ("decision rules") which make the constraints feasible for all realizations of the data $\zeta \in \mathcal{U}$ and minimize, under this restriction, the guaranteed value of the objective. The resulting *Adjustable Robust Counterpart* of (1) is the *infinite-dimensional* optimization problem

$$\min_{\chi_j(\cdot),\sigma} \left\{ \sigma : \left. \begin{array}{l} F_0(\chi(\zeta),\zeta) \leq \sigma \\ F_i(\chi\zeta),\zeta) \in K_i,\ 1 \leq i \leq I \end{array} \right\} \forall \zeta \in \mathcal{U} \right\} \quad [\chi(\zeta) = \{\chi_j(P_j\zeta)\}]$$

Unfortunately, in contrast to the usual Robust Counterpart (4), the latter problem is almost always "completely computationally intractable" – in general, it is even unclear how to represent in a computationally tractable fashion candidate solutions (that is, multivariate functions of continuous variables), let alone how to optimize over these solutions. To recover computational tractability, it was proposed in [6] to restrict decision rules to be *affine* functions of their arguments. The resulting *Affinely Adjustable Robust Counterpart* (AARC) of (1) is still significantly more flexible than the non-adjustable RC, and at the same time it is computationally tractable for a wide class of uncertain problems satisfying appropriate (and not too restrictive) structural assumptions (for details and instructive application examples, see [6, 7]). Note that the RC (4) is a particular case of AARC corresponding to the situation when all matrices $P_j$ are zero; this allows us to focus in the sequel solely on AARC.

*The goal* of this paper is to extend the scope of RO methodology by modifying the "uncertain-but-bounded" model of data uncertainty postulated by A.2. The motivation is that in many applications, including into the uncertainty set all "physically possible" realizations of the data may lead to an overly "pessimistic" solution or even to an infeasible Robust Counterpart. At the same time it is unwise to neglect some of

[1] In principle, one can replace the affine functions $P_j\zeta$ of the data with arbitrary ones.

these realizations. To make the point clearer, consider for example a communication network where the uncertain data represent traffic between terminal nodes. It would be too expensive to build a network which could serve with the same small delays both a typical everyday traffic and a much larger traffic caused by an outstanding event. At the same time, it would be undesirable to loose completely part of the latter traffic. What is reasonable in this situation, is to allow for "normal" delays when serving "typical" traffic and for larger, but somehow controlled, delays when serving large traffic fluctuations.

Our goal in this paper is to propose and to investigate a novel "uncertainty-immunized" counterpart of uncertain problem (1) which adequately and comprehensively models the requirement for *normal performance in presence of "typical" uncertainty and controlled deterioration in performance for "large deviations" in the uncertain data.* A rough model, to be refined in the sequel, is as follows. Let us treat what was called earlier "uncertainty set" $\mathcal{U}$ as the "normal range" of the uncertain data, rather than the set of all data we want to take into account, and treat a candidate affinely adjustable solution $\chi(\cdot)$ as "admissible", if it meets two requirements:

**I.** When $\zeta \in \mathcal{U}$, the solution must satisfy the constraints $F_i(\chi(\zeta), \zeta) \in K_i, i \leq I$;

**II.** For *all* data $\zeta \in \mathbf{R}^{n_\zeta}$, the violations of the constraints should not exceed a prescribed multiple of the deviation of the data from its normal range:

$$\forall \zeta \in \mathbf{R}^{n_\zeta} : \mathrm{dist}(F_i(\chi(\zeta), \zeta), K_i) \leq \alpha_i \mathrm{dist}(\zeta, \mathcal{U}),\ i \leq I. \tag{5}$$

Here the distances from a vector $a$ to a set $A$ is measured in a prescribed norm on the corresponding space according to the standard rule $\mathrm{dist}(a, A) = \inf_{a' \in A} \|a - a'\|$. The constants $\alpha_i \geq 0$ in (5) serve as (upper bounds on) "global sensitivities" of the constraints, evaluated at the candidate solution in question, to deviations of the data from its normal range. Note that in the case when the sets $K_i$ are closed, requirement II dominates requirement I; indeed, (5) implies that when $\zeta \in \mathcal{U}$, we have $\mathrm{dist}(F_i(\chi(\zeta), \zeta), K_i) = 0$, that is, $F_i(\chi(\zeta), \zeta) \in K_i$. Thus, we loose nothing when focusing solely on II.

Similar logic can be applied to the objective function. Specifically, we call a real $\sigma$ an *achievable value of the objective, evaluated at a candidate solution* $\chi(\cdot)$, *with global sensitivity* $\alpha_0 \geq 0$, if

$$\forall \zeta \in \mathbf{R}^n : F_0(\chi(\zeta), \zeta) - \tau \leq \alpha_0 \mathrm{dist}(\zeta, \mathcal{U}). \tag{6}$$

Summarizing, we associate with uncertain problem (1) its *Comprehensive Robust Counterpart* (CRC)

$$\min_{\chi(\cdot), \sigma} \left\{ \sigma : \begin{array}{l} F_0(\chi(\zeta), \zeta) \leq \sigma + \alpha_0 \mathrm{dist}(\zeta, \mathcal{U}) \\ \mathrm{dist}(F_i(\chi(\zeta), \zeta), K_i) \leq \alpha_i \mathrm{dist}(\zeta, \mathcal{U}),\ 1 \leq i \leq I \end{array} \right\} \forall \zeta \right\}. \tag{7}$$

We could also treat the global sensitivities $\alpha_i$ as additional variables rather than given constants and replace the objective function with, say, $\sigma + g(\alpha)$ for an appropriate user-specified function $g$.

Note that problem (7) is "stronger" than AARC: every feasible solution to (7) is feasible for the AARC as well. At the same time, pushing in (7) the sensitivities to $+\infty$, we, essentially, recover the AARC. Thus, the modelling concept we intend to introduce and to investigate here is an extension of AARC and in particular encompasses also the usual RC (hence the term "comprehensive").

The rest of this paper is organized as follows. In Section 2, we refine the concept of CRC. Section 3 addresses the crucial issue of computational tractability of CRC. Specifically, we demonstrate that under appropriate structural assumptions on (1) ("bi-affine in $\chi$, $\zeta$ constraints and objective"), problem (7) (which by itself is semi-infinite and as such could be intractable) is computationally tractable, provided that $K_i$ are not too complicated convex sets. In Section 4, we apply the outlined approach to a problem which is important by its own right: that of specifying optimal finite-horizon linear control law in a linear discrete time dynamical system affected by uncertain inputs. In order to bring the latter problem into our "bi-affine" framework, we use an appropriate parameterization of the family of linear control laws, similar to what is called in Control "$Q$-parameterization".

## 2. Comprehensive Robust counterpart of an uncertain optimization problem

We are about to refine the concept of Comprehensive Robust Counterpart of uncertain problem (1), which was roughly outlined in the Introduction, with the purpose of adding more flexibility to this notion, so as to capture a wide spectrum of requirements which could arise in various potential applications.

### *2.1. The setup*

*2.1.1. Structure of normal ranges and sets of "physically possible" values of the uncertain data* For the time being, the normal range $\mathcal{U}$ of the data in (1) was defined as a predetermined nonempty subset of the "data universe" $\mathbf{R}^{n_\zeta}$. From now on, we assume that this set is *closed, bounded and convex*. Further, in many applications the uncertain data $\zeta$ is naturally partitioned into blocks $\zeta^\ell$, $\ell = 1, \dots, L$, representing "physically different" components. For example, in the communication network example mentioned in the Introduction, the uncertain data represent the traffic in the network, i.e., this is the vector comprised of communication demands of various origin-destination pairs of nodes. It could be natural to split this vector according to a given partition of the nodes such as domestic communications in various countries, communications between various pairs of countries, etc. To handle the case of "structured data" conveniently, we assume from now on that

1. The "data universe" $\mathbf{R}^{n_\zeta}$ is represented as the direct product $\mathbf{R}^{(1)} \times \dots \times \mathbf{R}^{(L)}$, and $\zeta^\ell$ is the projection of $\zeta$ on the corresponding direct factor.
2. The normal range $\mathcal{U}$ of $\zeta$ is the direct product of closed and bounded convex sets $\mathcal{U}_\ell \subset \mathbf{R}^{(\ell)}$ – normal ranges of the corresponding data components.

Further, whereas in the Introduction we considered *all* data vectors $\zeta$ to be "physically possible", from now on, in order to capture some applications (e.g. our "communication example", where the data definitely is nonnegative), we relax this assumption, allowing the set $\mathcal{Z}_\ell$ of "physically possible" values of $\zeta^\ell$ to be a proper subset of the corresponding "universe" $\mathbf{R}^{(\ell)}$. It, however, would be technically inconvenient to allow for arbitrary closed and convex sets $\mathcal{Z}_\ell$. From now on, we assume that the sets $\mathcal{Z}_\ell$ are of the form

$$\mathcal{Z}_\ell = \mathcal{U}_\ell + \mathcal{L}_\ell, \tag{8}$$

where $\mathcal{L}_\ell$ are closed convex cones in the corresponding spaces $\mathbf{R}^{(\ell)}$. For illustration, to model the case where $Z_\ell$ is the nonnegative orthant $\mathbf{R}_+^{(\ell)}$, we can choose as $\mathcal{U}_\ell$ an arbitrary closed, bounded, convex subset of $\mathbf{R}_+^{(\ell)}$ containing the origin and set $\mathcal{L}_\ell = \mathbf{R}_+^{(\ell)}$. With $\mathcal{L}_\ell = \mathbf{R}^{(\ell)}$, the set $\mathcal{Z}_\ell$, independently of what is the corresponding normal range $\mathcal{U}_\ell$, becomes the entire $\mathbf{R}^{(\ell)}$. Finally, the set $\mathcal{Z}$ of all physically possible values of the data is, by definition, the direct product $\mathcal{Z}_1 \times \ldots \times \mathcal{Z}_L$.

*2.1.2. Measuring global sensitivities* As outlined in the Introduction, global sensitivities are measured in terms of point-to-set distances, the latter being given by prescribed norms on the corresponding spaces. However, a norm imposes symmetry which is not necessary relevant in certain situations. Therefore we choose to work with a wider set of distances – those coming from *Minkowski functions.*

*Minkowski functions.* Recall that a *Minkowski function* $\phi(\cdot)$ on $\mathbf{R}^n$ is a real-valued function on $\mathbf{R}^n$ with the following properties:

1. [positivity] $\phi(w) > 0$ whenever $w \neq 0$;
2. [positive homogeneity] $\phi(\lambda w) = \lambda\phi(w)$ for all $w$ and all $\lambda \geq 0$;
3. [triangle inequality] $\phi(w' + w'') \leq \phi(w') + \phi(w'')$ for all $w'$, $w''$.

Note that a Minkowski function is a slight generalization of a norm; norms on $\mathbf{R}^n$ are exactly the *symmetric* Minkowski functions: $\phi(-w) \equiv \phi(w)$.

Given a Minkowski function $\phi(\cdot)$ on $\mathbf{R}^n$ and a nonempty closed set $K \subset \mathbf{R}^n$, we define the distance $\text{dist}_\phi(v, K)$ of a vector $v \in \mathbf{R}^n$ from $K$ by the natural relation

$$\text{dist}_\phi(v, K) = \min_{v' \in K} \phi(v - v'). \tag{9}$$

Similarly, given a nonempty closed and bounded set $Y \subset \mathbf{R}^n$ and a closed cone $L \subset \mathbf{R}^n$, we set

$$\text{dist}_\phi(v, Y|L) = \min_{v' \in Y: v - v' \in L} \phi(v - v'), \ \ v \in Y + L. \tag{10}$$

Note that $\text{dist}_\phi(v, K)$ is always nonnegative and is zero if $v \in K$, and similarly for $\text{dist}_\phi(v, Y|L)$. Note also that $\text{dist}_\phi(v, Y|\mathbf{R}^n) = \text{dist}_\phi(v, Y)$.

*Measuring sensitivities.* A definition of global sensitivity as outlined in the Introduction could be as follows:

> *Let $\chi(\cdot)$ be a candidate solution to an uncertain constraint $F(\chi, \zeta) \in K \subset \mathbf{R}^n$, let $\psi(\cdot)$ be a Minkowski function on $\mathbf{R}^n$, and let $\phi(\cdot)$ be a Minkowski function on $\mathbf{R}^{n_\zeta}$. We say that the global sensitivity of the constraint, evaluated at $\chi(\cdot)$, to the uncertain data $\zeta$ does not exceed a real $\alpha$, if*
>
> $$\forall \zeta \in \mathcal{Z} : \text{dist}_\psi(F(\chi(\zeta), \zeta), K) \leq \alpha \text{dist}_\phi(\zeta, \mathcal{U}).$$
>
> Note the major modification as compared with the Introduction: now we do not care what happens with the constraint when $\zeta$ is not "physically possible", that is, $\zeta \notin \mathcal{Z} \equiv \mathcal{Z}_1 \times \ldots \times \mathcal{Z}_L$.

It makes sense to refine the just outlined definition in order to respect data structure as defined in Section 2.1.1. A technically convenient way to do it is as follows:

**Definition 1.** *Let $\chi(\cdot)$ be a candidate solution to an uncertain constraint $F(\chi,\zeta)\in K\subset\mathbf{R}^n$, let $\psi(\cdot)$ be a Minkowski function on $\mathbf{R}^n$, and let $\phi_\ell(\cdot)$, $\ell=1,\dots,L$, be Minkowski functions on $\mathbf{R}^{(\ell)}$. We say that the global sensitivities of the constraint, evaluated at $\chi(\cdot)$, to the uncertain data $\zeta$ do not exceed reals $\alpha_\ell$, $\ell=1,\dots,L$, if*

$$\forall\zeta\in\mathcal{Z}\equiv\mathcal{Z}_1\times\ldots\times\mathcal{Z}_L:\mathrm{dist}_\psi(F(\chi(\zeta),\zeta),K)\le\sum_{\ell=1}^{L}\alpha_\ell\mathrm{dist}_{\phi_\ell}(\zeta^\ell,\mathcal{U}_\ell|\mathcal{L}_\ell).\qquad(11)$$

Note the major differences between this definition (which is the only one we use from now on) and the above "draft". First, in the right hand side we use a "separable" (w.r.t. data components) distance from the data to its normal range. Second, when measuring the distance from a data component $\zeta^\ell\in\mathcal{Z}_\ell$ to its normal range $\mathcal{U}_\ell$, we use the distance $\mathrm{dist}_{\phi_\ell}(\zeta^\ell,\mathcal{U}_\ell|\mathcal{L}_\ell)$ which respects the structure $\mathcal{Z}_\ell=\mathcal{U}_\ell+\mathcal{L}_\ell$ of $\mathcal{Z}_\ell$. Third, the sensitivity now becomes a vector rather than a scalar, so that we can speak about "partial sensitivities" $\alpha_\ell$ w.r.t. different data components.

### *2.2. Comprehensive Robust Counterpart of uncertain problem*

Following the approach outlined in the Introduction and incorporating the refinements presented in Section 2.1, we define the Comprehensive Robust Counterpart (CRC) of uncertain problem (1) as the optimization problem

$$\min_{\chi(\cdot)\in\mathrm{Aff},\sigma}\left\{\sigma:\begin{array}{r}F_0(\chi(\zeta),\zeta)-\sigma\le\sum_{\ell=1}^{L}\alpha_{0\ell}\mathrm{dist}_{\phi_{0\ell}}(\zeta^\ell,\mathcal{Z}_\ell|\mathcal{L}_\ell)\ \forall\zeta\in\mathcal{Z}\\ \mathrm{dist}_{\phi_i}(F_i(\chi(\zeta),\zeta),K_i)\le\sum_{\ell=1}^{L}\alpha_{i\ell}\mathrm{dist}_{\phi_{i\ell}}(\zeta^\ell,\mathcal{Z}_\ell|\mathcal{L}_\ell),\ \forall\zeta\in\mathcal{Z}\\ 1\le i\le I\end{array}\right\},\qquad(12)$$

where

- $\zeta\in\mathbf{R}^{n_\zeta}$ is the vector of uncertain data, and $\zeta^\ell$, $\ell=1,\dots,L$, are the blocks of $\zeta$ corresponding to a given decomposition $\mathbf{R}^{n_\zeta}=\mathbf{R}^{(1)}\times\ldots\times\mathbf{R}^{(L)}$,
- the "decision variable" $\chi(\cdot)$ runs through the set of affinely adjustable candidate solutions to (1), that is, through the set Aff of $n_\chi$-dimensional vector-functions of $\zeta$ with components of the form

$$\chi_j(\zeta)\equiv(\chi(\zeta))_j=\eta_j^0+\langle\eta_j,P_j\zeta\rangle,\ j=1,\dots,n_\chi.\qquad(13)$$

  Here $P_j$ are given matrices, while the reals $\eta_j^0$ and the vectors $\eta_j$ are the parameters specifying affinely adjustable candidate solutions to (1) (these parameters, along with $\sigma$, are the decision variables in (12)). From now on, $\langle\cdot,\cdot\rangle$ denotes the inner product in the Euclidean space in question;

- $F_0(\chi,\zeta): \mathbf{R}^{n_\chi}\times\mathbf{R}^{n_\zeta}\to\mathbf{R}$, $F_i(\chi,\zeta): \mathbf{R}^{n_\chi}\times\mathbf{R}^{n_\zeta}\to\mathbf{R}^{k_i}$, $1\le i\le I$, are the objective and the constraints of problem (1), and $K_i\subset\mathbf{R}^{k_i}$ are given nonempty closed sets;
- $\psi_i(\cdot)$ and $\phi_{i\ell}(\cdot)$ are given Minkowski functions on $\mathbf{R}^{k_i}$, $\mathbf{R}^{(\ell)}$, respectively;
- $\mathcal{U}_\ell\subset\mathbf{R}^{(\ell)}$ are given nonempty compact convex sets, $\mathcal{L}_\ell\subset\mathbf{R}^{(\ell)}$ are given closed convex cones, and $\mathcal{Z}=\underbrace{(\mathcal{U}_1+\mathcal{L}_1)}_{\mathcal{Z}_1}\times\ldots\times\underbrace{(\mathcal{U}_L+\mathcal{L}_L)}_{\mathcal{Z}_L}$;
- $\alpha_{i\ell}$ are given nonnegative constants.

Along with (12), we shall consider a more general problem, where the sensitivities $\alpha_{i\ell}$ are treated as variables rather than given constants, namely, the problem

$$\min_{\substack{\chi(\cdot)\in\mathrm{Aff},\sigma\\ \alpha=\{\alpha_{i\ell}\}}}\left\{\Phi(\sigma,\alpha):\begin{array}{l}\left.\begin{array}{r}F_0(\chi(\zeta),\zeta)-\sigma\le\sum_{\ell=1}^L\alpha_{0\ell}\mathrm{dist}_{\phi_{0\ell}}(\zeta^\ell,\mathcal{Z}_\ell|\mathcal{L}_\ell)\\ \forall\zeta\in\mathcal{Z}\\ \mathrm{dist}_{\psi_i}(F_i(\chi(\zeta),\zeta),K_i)\le\sum_{\ell=1}^L\alpha_{i\ell}\mathrm{dist}_{\phi_{i\ell}}(\zeta^\ell,\mathcal{Z}_\ell|\mathcal{L}_\ell),\\ \forall(i\le I,\zeta\in\mathcal{Z})\end{array}\right\}\\ \alpha\in\mathcal{A}\end{array}\right\}, \tag{14}$$

where $\Phi$ is a given objective and $\mathcal{A}$ is a given subset of the nonnegative orthant. Note that (12) is a particular case of the latter problem corresponding to a singleton set $\mathcal{A}$ and to $\Phi(\sigma,\alpha)\equiv\sigma$. In the sequel, unless explicitly stated otherwise, reference to the CRC of (1) points to (14) rather than (12).

## 3. Processing (14)

Problem (14) is semi-infinite – it has infinitely many constraints parameterized by the uncertain data $\zeta\in\mathcal{Z}$; as such, it can be computationally intractable even in the case where all instances of the underlying uncertain problem (1) are simple (e.g., are explicit convex programs). Our next goal is to present sufficient conditions for the CRC to be computationally tractable.

### *3.1. Bi-affinity and fixed recourse*

We start with imposing appropriate structural assumptions on the functions $F_i(\chi,\zeta)$ in (1). These assumptions by themselves do not necessarily guarantee tractability of the CRC, but they simplify dramatically the corresponding analysis. The assumptions are as follows:

*Bi-affinity: The functions $F_i(\chi,\zeta)$, $0\le i\le I$, are affine in $\chi\in\mathbf{R}^{n_\chi}$ for $\zeta$ fixed and are affine in $\zeta\in\mathbf{R}^{n_\zeta}$ for $\chi$ fixed.* In other words, we have

$$F_i(\chi,\zeta)=F_i^0[\zeta]+\sum_{j=1}^{n_\chi}\chi_jF_i^j[\zeta],\ 0\le i\le I \tag{15}$$

where $F_i^j[\zeta]$, $0\le i\le I$, $0\le j\le n_\chi$, are affine in $\zeta$.

Note that as far as uncertain *convex* programs are concerned, bi-affinity assumption is less restrictive than it might look. Indeed, a wide spectrum of convex programs (e.g., linear and semidefinite ones) can be represented in the form of (1) with functions $F_i$ affine in $\chi$. In this case, bi-affinity merely means that the uncertain data enter the coefficients of these affine functions in an affine fashion.

*Fixed recourse.* Let $F_i(\chi, \zeta)$ obey (15), and assume that we substitute into these functions, as the $\chi$-argument, an affinely adjustable candidate solution to (1), that is, an affine function $\chi(\zeta)$ of $\zeta$ with components of the form (13). The resulting functions of $\zeta$ are of the form

$$F_i(\chi(\zeta), \zeta) = \left[ F_i^0[\zeta] + \sum_{j=1}^{n_\chi} \eta_j^0 F_i^j[\zeta] \right] + \sum_{j=1}^{n_\chi} \langle \eta_j, P_j\zeta \rangle F_i^j[\zeta].$$

In general, these functions are quadratic in $\zeta$, except for the case of *fixed recourse*, where for all $j$ with adjustable $\chi_j$ (that is, with nonzero $P_j$), the corresponding "coefficient" $F_i^j[\zeta]$ is "certain" – independent of $\zeta$. In the latter case, the functions $F(\chi(\zeta), \zeta)$ are affine in $\zeta$. From now on, we make the following crucial

Assumption A: [fixed recourse] *For every $j$ such that $P_j \neq 0$, all functions $F_i^j$, $i = 0, 1, \dots, I$, are independent of $\zeta$.*

With this assumption, we clearly have

$$\begin{aligned} \chi_j(\zeta) = \eta_j^0 + \langle \eta_j, P_j\zeta, \rangle,\ j = 1, \dots, n_\chi \Rightarrow F_i(\chi(\zeta), \zeta) &\equiv \omega_i[\eta] + \Omega_i[\eta]\zeta \\ \equiv \omega_i[\eta] + \sum_{\ell=1}^{L} \Omega_i^\ell[\eta]\zeta^\ell, & \qquad (16) \end{aligned}$$

where $\eta = \{\eta_j^0, \eta_j\}_{j=1}^{n_\chi}$ is the vector of parameters specifying affine function $\chi(\zeta) = (\chi_1(\eta), \dots, \chi_{n_\chi}(\eta))'$ ($'$ stands for taking the transpose), and $\omega_i[\eta]$, $\Omega_i^\ell[\eta]$ are *affine in* $\eta$ vector- and matrix-valued functions readily given by the data of (1).

### 3.2. *Convexity assumption*

From now on, we assume that the sets $K_i$ in (1) are nonempty, closed and *convex*.

### 3.3. *Processing (14): decomposition*

Under the assumptions of bi-affinity, fixed recourse and convexity (which are our default assumptions from now on), all semi-infinite constraints in (14) are of the generic form

$$\operatorname{dist}_\psi\left(\omega[\gamma] + \sum_{\ell=1}^{L} \Omega^\ell[\gamma]\zeta^\ell, K\right) \le \sum_{\ell=1}^{L} \alpha_\ell \operatorname{dist}_{\phi_\ell}(\zeta^\ell, \mathcal{U}_\ell | \mathcal{L}_\ell) \quad \forall \begin{pmatrix} \zeta^\ell \in \mathcal{U}_\ell + \mathcal{L}_\ell, \\ 1 \le \ell \le L \end{pmatrix}, \qquad (17)$$

where

- $\omega[\gamma], \Omega^\ell[\gamma]$ are given *affine* functions of the design variables $\gamma = (\eta = \{\eta_j^0, \eta_j\}_{j=1}^{n_\chi}, \sigma)$ specifying, via (13), an affinely adjustable candidate solution to (1) and its quality,
- $K$ is a given nonempty closed convex set in $\mathbf{R}^k$, and
- $\psi(\cdot), \phi_\ell(\cdot)$ are given Minkowski functions on $\mathbf{R}^k$ and $\mathbf{R}^{(\ell)}$, respectively.

Now we show how to "decompose" the semi-infinite constraint (17) into a system of somewhat simpler semi-infinite constraints.

*Preliminaries.* Let $X$ be a nonempty closed convex set in $\mathbf{R}^q$. Recall that the *recessive cone* $\mathrm{Rec}(X)$ of $X$ is the set

$$\mathrm{Rec}(X) = \{e \in \mathbf{R}^q : x + te \in X \,\forall (x \in X, t \geq 0)\};$$

elements of $\mathrm{Rec}(X)$ are called *recessive directions* of $X$. It is well-known that $\mathrm{Rec}(X)$ is a closed convex cone which is trivial – the origin $\{0\}$ – if and only if $X$ is bounded. Finally, it is easily seen that $e \in \mathbf{R}^q$ is a recessive direction of $X$ if and only if $e$ can be represented as $\lim_{i\to\infty} t_i^{-1}(y_i - c)$, where $y_i, c \in X$ and $t_i \to \infty$ (see, e.g., [11], Chapter 1).

Given a Minkowski function $\pi(\cdot)$ on $\mathbf{R}^q$, let us set

$$\pi_X(y) = \mathrm{dist}_\pi(y, \mathrm{Rec}(X)) \equiv \min_{x\in \mathrm{Rec}(X)} \pi(y - x) : \mathbf{R}^q \to \mathbf{R}. \tag{18}$$

Note that $\pi_X(y)$ satisfies all requirements from the definition of a Minkowski function, except for positivity. Instead of the latter property we have $\pi_X(y) \geq 0$, and $\pi_X(y) = 0$ if and only if $y \in \mathrm{Rec}(X)$. Finally, $\pi_X(\cdot) \equiv \pi(\cdot)$ for bounded $X$.

The role of the outlined notions in our context stems from the following simple observation:

**Proposition 1.** *Let*

1. *$X \subset \mathbf{R}^q$ and $Y \subset \mathbf{R}^p$ be closed nonempty convex sets, with $Y$ bounded,*
2. *$L$ be a closed convex cone in $\mathbf{R}^p$,*
3. *$\pi(\cdot), \theta(\cdot)$ be Minkowski functions on $\mathbf{R}^q$ and $\mathbf{R}^p$, respectively,*
4. *$\alpha \geq 0$.*

*For an affine mapping $y \mapsto c + Sy : \mathbf{R}^p \to \mathbf{R}^q$, the condition*

$$\mathrm{dist}_\pi(c + Sy, X) \leq \alpha \mathrm{dist}_\theta(y, Y|L) \quad \forall y \in Y + L \tag{19}$$

*is equivalent to the pair of conditions*

$$\begin{array}{l} (a)\; c + Sy \in X \;\forall y \in Y \\ (b)\; \Gamma^X_{\theta,\pi}(S|L) \equiv \max\limits_y \{\pi_X(Sy) : y \in L, \theta(y) \leq 1\} \leq \alpha. \end{array} \tag{20}$$

*Proof.* Assume that (19) takes place. Then for $y \in Y$ we have $\mathrm{dist}_\pi(c + Sy, X) = 0$, or, which is the same, $c + Sy \in X$, so that (20.$a$) is valid. Further, let $y_0 \in Y$, and let $e \in L$ be such that $\theta(e) \leq 1$. For $\lambda \geq 0$ we should have $\mathrm{dist}_\pi(c + S(y_0 + \lambda e), x) \leq \alpha \mathrm{dist}_\theta(y_0 + \lambda e, Y|L) \leq \alpha\theta(\lambda e) = \alpha\lambda\theta(e) \leq \alpha\lambda$. Recalling the definition of $\mathrm{dist}_\pi(\cdot, X)$, we conclude that for every $\lambda > 0$ there exists $f_\lambda \in X$ such that $\pi(\underbrace{[c + Sy_0]}_{c_0} + \lambda Se - f_\lambda) \leq \alpha\lambda$, or, which is the same,

$$\pi(Se - \lambda^{-1}[f_\lambda - c_0]) \leq \alpha. \tag{21}$$

It follows that the sequence of vectors $i^{-1}[f_i - c_0]$ is bounded; passing to a subsequence, we may assume that the vectors $h_j = i_j^{-1}[f_{i_j} - c_0]$, where $i_j \to \infty$ as $j \to \infty$, converge to a vector $h$ as $j \to \infty$. Since $c_0 \in X$ by (20.$a$), $f_i \in X$ and $i_j \to \infty$ as $j \to \infty$, $h$ is a recessive direction of $X$, and (21) implies that $\pi(Se - h) \leq \alpha$, whence, by (18), $\pi_X(Se) \leq \alpha$ as well, as required in (20.$b$).

It remains to verify that (20) implies (19). Assuming that (20) is valid, let $y \in Y + L$; by definition of $\text{dist}_\theta(y, Y|L)$, there exist $\bar{y} \in Y$ and $e \in L$ such that $y = \bar{y} + e$ and $\theta(e) = \text{dist}_\theta(y, Y|L)$. By (20.$b$), there exists $f \in \text{Rec}(X)$ such that

$$\pi(Se - f) \leq \alpha\theta(e) = \alpha \text{dist}_\theta(y, Y|L).$$

Finally, the vector $c + S\bar{y}$ belongs to $X$ by (20.$a$), so that the vector $c + S\bar{y} + f$ also belongs to $X$ due to $f \in \text{Rec}(X)$. Therefore

$$\text{dist}_\pi(c + Sy, X) \leq \pi(c + Sy - [c + S\bar{y} + f]) = \pi(Se - f) \leq \alpha \text{dist}_\theta(y, Y|L).$$

The resulting inequality holds true for all $y \in Y + L$, that is, (19) is valid. □

*Decomposing (17).* The desired decomposition is given by the following

**Corollary 1.** *The semi-infinite constraint* (17) *in variables $\gamma$, $\alpha_\ell$, is equivalent to the system of constraints*

$$\begin{array}{ll}(a) & \omega[\gamma] + \sum_{\ell=1}^{L} \Omega^\ell[\gamma]\zeta^\ell \in K \quad \forall(\zeta^\ell \in \mathcal{U}_\ell, \ell = 1, \ldots, L) \\ (b_\ell) & \Gamma^K_{\phi_\ell,\psi}(\Omega^\ell[\gamma]|\mathcal{L}_\ell) \equiv \max\limits_{\zeta^\ell \in \mathcal{L}_\ell, \phi_\ell(\zeta^\ell) \leq 1} \min\limits_{v \in \text{Rec}(K)} \psi(\Omega^\ell[\gamma]\zeta^\ell - v) \leq \alpha_\ell, \\ & \ell = 1, \ldots, L\end{array} \tag{22}$$

*Proof.* To save notation, let us skip the pointer $[\gamma]$. If (17) holds true, then clearly $\omega + \sum_{\ell=1}^{L} \Omega^\ell\zeta^\ell \in K$ whenever $\zeta_\ell \in \mathcal{U}_\ell$, $\ell = 1, \ldots, L$, as required in (22.$a$). Now let $\bar{\zeta}^\ell \in \mathcal{U}_\ell$. $\ell = 1, \ldots, L - 1$. By (17), we have

$$\text{dist}_\psi\Bigg(\underbrace{\omega + \sum_{\ell=1}^{L-1} \Omega^\ell\bar{\zeta}^\ell}_{c} + \Omega^L\zeta^L, K\Bigg) \leq \alpha_L \text{dist}_{\phi_L}(\zeta^L, \mathcal{U}_L|\mathcal{L}_L) \quad \forall\zeta^L \in \mathcal{U}_L + \mathcal{L}_L,$$

whence, by Proposition 1, (22.$b_L$) holds true. By similar reasons, all remaining statements (22.$b_\ell$), $\ell = 1, \ldots, L - 1$, hold true as well. Thus, (17) implies (22).

Conversely, let (22) hold true. Given $\zeta$ with $\zeta^\ell \in \mathcal{U}_\ell + \mathcal{L}_\ell$, $\ell = 1, \dots, L$, let us find $\bar{\zeta}^\ell \in \mathcal{U}_\ell$ such that $\zeta^\ell - \bar{\zeta}^\ell \in \mathcal{L}_\ell$ and $\phi_\ell(\zeta^\ell - \bar{\zeta}^\ell) = \mathrm{dist}_{\phi_\ell}(\zeta^\ell, \mathcal{U}_\ell|\mathcal{L}_\ell)$, $\ell = 1, \dots, L$. By (22), for properly chosen $f_\ell \in \mathrm{Rec}(K)$ we have

$$\begin{array}{l} c \equiv \omega + \sum\limits_{\ell=1}^{L} \Omega^\ell \bar{\zeta}^\ell \in K, \\ \psi(\Omega^\ell[\zeta^\ell - \bar{\zeta}^\ell] - f_\ell) \le \Gamma^K_{\phi_\ell,\psi}(\Omega^\ell|\mathcal{L}_\ell)\phi_\ell(\zeta^\ell - \bar{\zeta}^\ell) \le \alpha_\ell \phi_\ell(\zeta^\ell - \bar{\zeta}^\ell) \\ = \alpha_\ell \mathrm{dist}_{\phi_\ell}\left(\zeta^\ell, \mathcal{U}_\ell|\mathcal{L}_\ell\right), \ \ell = 1, \dots, L. \end{array} \tag{23}$$

We now have

$$\omega + \sum_{\ell=1}^{L} \Omega^\ell \zeta^\ell - \underbrace{\left[\omega + \sum_{\ell=1}^{L}\left[\Omega^\ell \bar{\zeta}^\ell + f_\ell\right]\right]}_{\equiv e = c + \sum_\ell f_\ell} = \sum_{\ell=1}^{L}\left[\Omega^\ell[\zeta^\ell - \bar{\zeta}^\ell] - f_\ell\right]$$

whence, by (23) and the triangle inequality,

$$\psi\left(\omega + \sum_{\ell=1}^{L} \Omega^\ell \zeta^\ell - e\right) \le \sum_{\ell=1}^{L} \alpha_\ell \mathrm{dist}_{\phi_\ell}(\zeta^\ell, \mathcal{U}_\ell|\mathcal{L}_\ell).$$

Now, $e \in K$ due to $c \in K$, see (23), and $f_\ell \in \mathrm{Rec}(K)$, $\ell = 1, \dots, L$. We see that

$$\mathrm{dist}_\psi(\omega + \sum_{\ell=1}^{L} \Omega^\ell \zeta^\ell, K) \le \psi(\omega + \sum_{\ell=1}^{L} \Omega^\ell \zeta^\ell - e) \le \sum_{\ell=1}^{\ell} \alpha_\ell \mathrm{dist}_{\phi_\ell}(\zeta^\ell, \mathcal{U}_\ell|\mathcal{L}_\ell)$$

for all $\zeta \in \mathcal{Z}$, as required in (17). □

Corollary 1 implies that computationally tractable reformulation of (17) reduces to similar reformulations of the constraints in (22). These are the issues we are about to consider.

*3.3.1. Processing (22.a)* Tractable reformulation of semi-infinite inclusions like (22.$a$) is one of the major issues in Robust Optimization [1, 12, 13, 3, 2, 5, 4] and is reasonably well studied. A rough summary of these studies is as follows: aside of few rather special cases, there are just two generic situations where a tractable reformulation of (22.$a$) is possible:

**A.** $K$ is computationally tractable[2] and the normal range $\mathcal{U}$ of $\zeta$ is a polytope given as a convex hull of finite set: $\mathcal{U} = \mathrm{Conv}\{\zeta_\nu,\ \nu = 1, \dots, N\}$. In this (not too interesting) case the validity of (22.$a$) is clearly equivalent to the system of convex constraints in variables $\gamma$:

$$\omega[\gamma] + \sum_{\ell=1}^{L} \Omega^\ell[\gamma]\zeta^\ell_\nu \in K, \quad \nu = 1, \dots, N. \tag{24}$$

[2] A closed convex set $K$ is *computationally tractable*, if we are given in advance the affine hull of $K$, a point from the relative interior of $K$ and can check efficiently whether a given point belongs to $K$, see [14].

**B.** $K$ is a polyhedral set given by an explicit list of linear inequalities:

$$K = \{v : \langle p_\nu, v\rangle \geq r_\nu, \nu = 1, \ldots, N\} \subset \mathbf{R}^k. \tag{25}$$

In this case, the semi-infinite inclusion (17.$a$) admits a tractable reformulation, provided that $\mathcal{U}$ is computationally tractable. Here is a more explicit result in this direction, which covers a wide range of applications:

**Proposition 2.** [[3], Remark 4.1] *Let $K$ be given by* (25)*, and let the set $\mathcal{U}$ be of the form*

$$\mathcal{U} = \left\{\zeta : \exists u \in \mathbf{R}^s : \mathcal{A}(\zeta, u) - a \in \mathcal{K}\right\}, \tag{26}$$

*where $\mathcal{K}$ is either a nonnegative orthant $\mathbf{R}^M_+$ in the space $E = \mathbf{R}^M$, or the cone of positive semidefinite matrices $\mathbf{S}^M_+$ in the space $E = \mathbf{S}^M$ of $M \times M$ symmetric matrices equipped with the Frobenius inner product, and $(\zeta, u) \mapsto \mathcal{A}(\zeta, u)$ is a linear mapping from $\mathbf{R}^{n_\zeta} \times \mathbf{R}^s$ to $E$. In the case when $\mathcal{K}$ is the semidefinite cone, assume also that the image of the affine mapping $\mathcal{A}(\cdot, \cdot) - a$ intersects the interior of $\mathcal{K}$. Then the semi-infinite constraint in variables $\gamma$:*

$$\omega[\gamma] + \Omega[\gamma]\zeta \in K \quad \forall \zeta \in \mathcal{U} \tag{27}$$

*where $\omega[\gamma],\ \Omega[\gamma]$*

*Example 1: One-dimensional K and interval uncertainty.* Consider the semi-infinite constraint (27) and assume that $K$ in (17) is a proper subset of the real axis, while the normal range $\mathcal{U}$ is a box:

$$\mathcal{U} = \{\zeta \in \mathbf{R}^{n_\zeta} : \underline{z} \leq \zeta \leq \overline{z}\}.$$

Note that in this situation $\Omega[\gamma]$ is a row vector of dimension $n_\zeta$, and $\omega[\gamma]$ is a scalar. As for $K$, up to evident equivalent transformations (shift and scaling) there exist three possibilities:

E.1: $K = \{0\}$;
E.2: $K = [-1, 1]$;
E.3: $K = (-\infty, 0]$.

The corresponding semi-infinite inclusion (27) reduced to an explicit finite system of simple convex constraints in variables $\gamma$ (which can be further reduces to linear constraints):

*In the case of E.1:*

$$\omega[\gamma] + \Omega[\gamma]\frac{\underline{z} + \overline{z}}{2} = 0, \quad (\Omega[\gamma])_j = 0 \ \forall (j : \underline{z}_j \neq \overline{z}_j); \tag{29}$$

*In the case of E.2:*

$$\begin{aligned} \omega[\gamma] + \Omega[\gamma]\tfrac{\underline{z}+\overline{z}}{2} + \sum_{i=1}^{n_\zeta} \tfrac{\overline{z}_i - \underline{z}_i}{2} |(\Omega[\gamma])_i| \leq 1, \\ \omega[\gamma] + \Omega[\gamma]\tfrac{\underline{z}+\overline{z}}{2} - \sum_{i=1}^{n_\zeta} \tfrac{\overline{z}_i - \underline{z}_i}{2} |(\Omega[\gamma])_i| \geq -1; \end{aligned} \tag{30}$$

*In the case of E.3:*

$$\omega[\gamma] + \sum_{i=1}^{n_\zeta} \max\left[(\Omega[\gamma])_i \overline{z}_i, (\Omega[\gamma])_i \underline{z}_i\right] \leq 0. \tag{31}$$

*3.3.2. Processing (22.b)* All conditions (22.$b$) are of the same generic form

$$\psi_K(H[\zeta]e) \equiv \min_{f \in \mathrm{Rec}(K)} \psi(H[\gamma]e - f) \leq \alpha \ \forall (e \in L : \phi(e) \leq 1), \tag{32}$$

where $H[\gamma]$ is $k \times n$ matrix affinely depending on $\gamma$, $L$ is a closed convex cone in $\mathbf{R}^n$ and $\psi(\cdot)$, $\phi(\cdot)$ are Minkowski functions on $\mathbf{R}^k$, $\mathbf{R}^n$, respectively. Let us list several situations where (32) admits computationally tractable reformulation.

**A:** *The set* $\{e \in L : \phi(e) \leq 1\}$ *is a polytope given as a convex hull of finite set* $\{e^1, \ldots, e^M\}$ (e.g., $\phi(e) = \sum_i |e_i|$ is the $\|\cdot\|_1$-norm and $L$ is given by restrictions on signs of some or all coordinates of a vector) *and the function* $\psi_K(\cdot)$ *is efficiently computable* (in fact, the latter is the case when $K$ is computationally tractable, and $\psi(\cdot)$ is efficiently computable).

In this case (32) is equivalent to the following explicit system of convex constraints in the variables $\gamma, \alpha$:

$$\psi_K(H[\gamma]e^\ell) \leq \alpha, \ \ell = 1, \ldots, M \tag{33}$$

with efficiently computable left hand sides.

**B:** *$K$ is bounded, $L = \mathbf{R}^n$ and both $\psi(\cdot)$, $\phi(\cdot)$ are Euclidean norms.*
In this case, $\psi_K(\cdot) = \psi(\cdot)$ due to the boundedness of $K$, and (32) is equivalent to the efficiently computable convex constraint

$$\|H[\gamma]\| \leq \alpha$$

where $\|H[\gamma]\|$ is the standard matrix norm (maximal singular value).

**C:** *$K$ is bounded, $\psi(\cdot)$ is polyhedral:*

$$\psi(x) = \max_{1\leq \ell \leq M} \langle a_\ell, x\rangle \tag{34}$$

*$L$ is computationally tractable and $\phi(\cdot)$ is efficiently computable.* In this case $\psi_K(\cdot) \equiv \psi(\cdot)$, so that (32) reads

$$\underbrace{\max_{e\in L:\phi(e)\leq 1} \langle H^*[\gamma]a_\ell, e\rangle}_{\phi_*(H^*[\gamma]a_\ell|L)} \leq \alpha, \ \ \ell = 1, \dots, M. \tag{35}$$

The function $\phi_*(\cdot|L)$ is efficiently computable, since $L$ is computationally tractable and $\phi(\cdot)$ is efficiently computable, so that (35) is a finite system of efficiently computable convex constraints.

**D:** (D.1) $\mathrm{Rec}(K)$ *is comprised of all vectors $e \in \mathbf{R}^k$ with nonpositive coordinates $e_i$, $i \in J_- \subset \{1, \dots, k\}$ and zero coordinates $e_i$, $i \in J_0 \subset \{1, \dots, k\}$;*

(D.2) $\psi(\cdot)$ *is of the form*

$$\psi(v) = \max_{1\leq i\leq M} \sum_{s=1}^{k} a_{is} \max\left[\beta_{is}^+ v_s, -\beta_{is}^- v_s\right], \tag{36}$$

*where $a_{is}$, $\beta_{is}^\pm$ are nonnegative and the quantities $\mu_i = \mathrm{Card}(I_i)$, $I_i = \{s : a_{is} > 0\}$ do not exceed $\mu = O(1)\log_2(kM)$* (e.g., $\psi(v) = \|v\|_\infty \equiv \max_i |v_i|$, where $\mu = 2$);

(D.3) *$L$ is computationally tractable, and $\phi(\cdot)$ is efficiently computable.*
Processing (32) in the case of **D** is as follows. First, from (D.1) and (36) it clearly follows that

$$\begin{array}{l} \psi_K(v) = \max\limits_{1\leq i\leq M} \sum\limits_{s=1}^{k} a_{is} \max\left[\widehat{\beta}_{is}^+ v_s, -\widehat{\beta}_{is}^- v_i\right], \\ \widehat{\beta}_{is}^+ = \begin{cases} 0, & s \notin J_- \cup J_0 \\ \beta_{is}^+, & s \in J_- \cup J_0 \end{cases}, \widehat{\beta}_{is}^- = \begin{cases} 0, & s \notin J_0 \\ \beta_{is}^-, & s \in J_0 \end{cases}, \end{array}$$

or, which is the same,

$$\psi_K(v) = \max_{1\leq i\leq M} \max_{\epsilon^i = \{\epsilon_s^i = \pm 1\}_{s\in I_s}} \langle a^{i,\epsilon^i}, v\rangle, \ (a^{i,\epsilon^i})_s = \begin{cases} a_{is}\widehat{\beta}_{is}^+, & s \in I_i, \epsilon_s^i = 1 \\ a_{is}\widehat{\beta}_{is}^-, & s \in I_i, \epsilon_s^i = -1 \\ 0, & i \notin I_j \end{cases}. \tag{37}$$

Therefore (32) is equivalent to the system of at most $2^\mu M \leq \mathrm{Poly}(k, M)$ convex efficiently computable (by (D.3) constraints in variables $\gamma, \alpha$:

$$\phi_*(H^*[\gamma]a^{i,\epsilon^i}|L) \equiv \max_{e\in L:\phi(e)\leq 1} \langle H^*[\gamma]a^{i,\epsilon^i}, e\rangle \leq \alpha \ \forall \begin{pmatrix} \epsilon^i = \{\epsilon_s^i = \pm 1\}_{s\in I_i}, \\ i \leq M \end{pmatrix}. \tag{38}$$

*Example 2: One-dimensional K,*

$$\psi(s) = \max[r^+ s, -r^- s], \;\; \phi(e) = \max_{1 \le j \le \dim e} \max[p_j^+ e_j, -p_j^- e_j],$$
$$p_j^\pm, r^\pm > 0, \;\; L = \{e : e_j \ge 0, j \in J_+, \; e_j = 0, \; j \in J_0\}.$$

Up to the same equivalent transformations of $K$ as above, there are three possibilities: E.1: $K = \{0\}$, E.2: $K = [-1, 1]$, E.3: $K = (-\infty, 0]$. Let us set

$$\kappa_j^+ = \begin{cases} 1/p_j^+, & j \notin J_0 \\ 0, & j \in J_0 \end{cases}, \quad \kappa_j^- = \begin{cases} 1/p_j^-, & j \notin J_+ \cup J_0 \\ 0, & j \in J_+ \cup J_0 \end{cases}$$

Given a row vector $H[\gamma]$ affinely depending on variables $\gamma$, relation

$$\max_{e \in L: \phi(e) \le 1} \min_{f \in \mathrm{Rec}(K)} \psi(H[\gamma]e - f) \le \alpha$$

is equivalent to the following explicit finite system of convex constraints in variables $\gamma$, $\alpha$:

*In the cases of E.1, E.2* (see **C**):

$$\begin{aligned} r^+ \sum_{i=1}^{\dim e} \max\left[\kappa_j^+ (H[\gamma])_j, -\kappa_j^- (H[\gamma])_j\right] \le \alpha, \\ r^- \sum_{j=1}^{\dim e} \max\left[-\kappa_j^+ (H[\gamma])_j, \kappa_j^- (H[\gamma])_j\right] \le \alpha; \end{aligned} \tag{39}$$

*In the case of E.3* (see **D**):

$$r^+ \sum_{j=1}^{\dim e} \max\left[\kappa_j^+ (H[\gamma])_j, -\kappa_j^- (H[\gamma])_j\right] \le \alpha. \tag{40}$$

### *3.4. Summary*

An informal summary of our developments is as follows: *while the semi-infinite problem* (14) *can be difficult in general, there exists a reasonably wide spectrum of cases (see Sections 3.3.1, 3.3.2) where the problem admits "computationally tractable" reformulation.* The next statement illustrates this in the important particular case where all constraints in (1) are scalar (i.e., all $K_i$'s are one-dimensional) and the Minkowski functions $\phi_{i\ell}(\cdot)$ in (14) are of the form mentioned in Example 2, Section 3.3.2 (the latter assumption is made only for the sake of definiteness and in order to stay all the time within Linear Programming). For simplicity, we assume also that $\mathcal{U}$ is a box (this assumption also is non-critical, see Section 3.3.1). In the sequel, we denote $j$-th coordinate of $\ell$-th block $\zeta^\ell$ in $\zeta$ by $\zeta_j^\ell$; similarly, $\Omega_{ij}^\ell[\gamma]$ denotes $j$-th column in the matrix $\Omega_i^\ell[\gamma]$, see (16).

**Theorem 1.** *Consider semi-infinite problem* (14) *and assume that*

1. *The normal range $\mathcal{U}$ of the uncertain data is a box:*

$$\mathcal{U} = \{\zeta : \underline{z}^\ell \leq \zeta^\ell \leq \overline{z}^\ell,\ \ell = 1, \ldots, L\}. \tag{41}$$

*while the cones $\mathcal{L}_\ell$ are given by*

$$\mathcal{L}_\ell = \{\zeta^\ell : \zeta^\ell_j \geq 0,\ j \in J_\ell, \zeta^\ell_j = 0,\ j \in I_\ell\}; \tag{42}$$

2. *All the sets $K_i$, $i = 1, \ldots, I$, are one-dimensional, so that the set $\mathcal{I} = \{1, \ldots, I\}$ of constraint indices can be partitioned into three subsets $\mathcal{I}_0$, $\mathcal{I}_{[-1,1]}$ and $\mathcal{I}_{(-\infty,0]}$ in such a way that $K_i$ is*

   - *a singleton (w.l.o.g., $\{0\}$) for $i \in \mathcal{I}_0$,*
   - *a non-singleton bounded segment (w.l.o.g., $[-1, 1]$) for $i \in \mathcal{I}_{[-1,1]}$, and*
   - *a ray (w.l.o.g., $(-\infty, 0]$) for $i \in \mathcal{I}_{(-\infty,0]}$.*

   *Note that with one-dimensional $K_i$, $\psi_i(\cdot)$ are univariate Minkowski functions, so that*

$$\psi_i(s) = \max[r_i^+ s, -r_i^- s],\ \ r_i^\pm > 0,\ i = 1, \ldots, I; \tag{43}$$

3. *The set $\mathcal{A}$ of allowed sensitivities is a closed convex subset of the nonnegative orthant, and the objective function $\Phi(\sigma, \alpha)$ in (14) is convex.*
4. *Functions $\phi_{i\ell}(\cdot)$ are of the form*

$$\phi_{i\ell}(\zeta^\ell) = \max_{1 \leq j \leq \dim \zeta^\ell} \max[p^+_{i\ell j}\zeta^\ell_j, -p^-_{i\ell j}\zeta^\ell_j] \tag{44}$$

   *where all coefficients $p^\pm_\cdot$ are positive.*

*Under these assumptions, problem* (14) *is equivalent to the following explicit Convex Programming program:*

$$\min_{\substack{\eta=\{\eta^0_j,\eta_j\}_j,\\ \sigma,\alpha=\{\alpha_{i\ell}\}}} \Phi(\sigma, \alpha) \text{ subject to (46)} \tag{45}$$

*where*

$$\begin{array}{c} \omega_0[\eta] + \sum_{\ell,j} \max\left[\underline{z}^\ell_j \Omega^\ell_{0j}[\eta], \overline{z}^\ell_j \Omega^\ell_{0j}[\eta]\right] \leq \sigma \\ \sum_j \max\left[\kappa^+_{0\ell j}\Omega^\ell_{0j}[\eta], -\kappa^-_{0\ell j}\Omega^\ell_{0j}[\eta]\right] \leq \alpha_{0\ell},\ \ell = 1, ..., L \end{array} \tag{$46_a$}$$

$$\begin{array}{ll} \underline{\forall i \in \mathcal{I}_0 :} & \\ & \omega_i[\eta] + \sum_\ell \Omega^\ell_i[\eta]\frac{\underline{z}^\ell + \overline{z}^\ell}{2} = 0 \\ \forall(\ell, j : \underline{z}^\ell_j < \overline{z}^\ell_j) : & \Omega^\ell_{ij}[\eta] = 0 \\ \forall(\ell, j : \underline{z}^\ell_j = \overline{z}^\ell_j) : & \left\{ \begin{array}{l} r_i^+ \sum_j \max\left[\kappa^+_{i\ell j}\Omega^\ell_{ij}[\eta], -\kappa^-_{i\ell j}\Omega^\ell_{ij}[\eta]\right] \leq \alpha_{i\ell} \\ r_i^- \sum_j \max\left[-\kappa^+_{i\ell j}\Omega^\ell_{ij}[\eta], \kappa^-_{i\ell j}\Omega^\ell_{ij}[\eta]\right] \leq \alpha_{i\ell} \end{array}\right. \end{array} \tag{$46_b$}$$

$$
\begin{array}{l}
\underline{\forall i \in \mathcal{I}_{[-1,1]}}: \\
\omega_i[\eta] + \sum_\ell \Omega_i^\ell[\eta]\frac{\underline{z}^\ell+\overline{z}^\ell}{2} + \sum_{\ell,j}\frac{\overline{z}_j^\ell-\underline{z}_j^\ell}{2}\left|\Omega_{ij}^\ell[\eta]\right| \leq 1 \\
\omega_i[\eta] + \sum_\ell \Omega_i^\ell[\eta]\frac{\underline{z}^\ell+\overline{z}^\ell}{2} - \sum_{\ell,j}\frac{\overline{z}_j^\ell-\underline{z}_j^\ell}{2}\left|\Omega_{ij}^\ell[\eta]\right| \geq -1 \\
\quad r_i^+ \sum_j \max\left[\kappa_{i\ell j}^+\Omega_{ij}^\ell[\eta], -\kappa_{i\ell j}^-\Omega_{ij}^\ell[\eta]\right] \leq \alpha_{i\ell},\ \ell = 1, ..., L \\
\quad r_i^- \sum_j \max\left[-\kappa_{i\ell j}^+\Omega_{ij}^\ell[\eta], \kappa_{i\ell j}^-\Omega_{ij}^\ell[\eta]\right] \leq \alpha_{i\ell},\ \ell = 1, ..., L
\end{array}
\tag{46$_c$}
$$

$$
\begin{array}{l}
\underline{\forall i \in \mathcal{I}_{[-\infty,0]}}: \\
\omega_i[\eta] + \sum_{\ell,j} \max\left[\Omega_{ij}^\ell[\eta]\overline{z}_j, \Omega_{ij}^\ell[\eta]\underline{z}_j\right] \leq 0 \\
r_i^+ \sum_{\ell,j} \max\left[\kappa_{i\ell j}^+\Omega_{ij}^\ell[\eta], -\kappa_{i\ell j}^-\Omega_{ij}^\ell[\eta]\right] \leq \alpha_{i\ell},\ \ell = 1, ..., L
\end{array}
\tag{46$_d$}
$$

$$\alpha \in A.$$

*In the above relations,*

$$
\kappa_{i\ell j}^+ = \begin{cases} 1/p_{i\ell j}^+, & j \notin I_\ell \\ 0, & j \in I_\ell \end{cases}, \quad \kappa_{i\ell j}^- = \begin{cases} 1/p_{i\ell j}^-, & j \notin J_\ell \cup I_\ell \\ 0, & j \in J_\ell \cup I_\ell \end{cases}
\tag{47}
$$

*Proof.* The result is readily given by the constructions of Example 1, Section 3.3.1 expressing conditions ($22.a$) associated with the semi-infinite constraints of (14) and the constructions of Example 2, Section 3.3.2 expressing conditions ($22.b_\ell$) associated with these constraints. □

*Remark 1.* In fact, results of Sections 3.3.1, 3.3.2 allow to relax significantly the assumptions of Theorem 1, while keeping the conclusions intact. For example, we could allow for some of $K_i$ to be bounded multi-dimensional polyhedral sets given by explicit lists of linear inequalities, provided that the corresponding Minkowski functions $\psi_i(\cdot)$ are polyhedral (see Proposition 2 and item C in Section 3.3.2). Moreover, we could require no more than computational tractability of normal ranges $\mathcal{U}_\ell$, the cones $\mathcal{L}_\ell$ and Minkowski functions $\phi_{i\ell}$, etc.

## 4. Generic application: optimal finite-horizon linear control in linear dynamical system

In this section, we apply the CRC methodology to the problem of optimizing a finite-horizon linear control in a linear discrete time dynamical system affected by uncertain input.

### 4.1. The control problem

Consider a linear discrete time dynamical system given by

$$
\begin{array}{rl}
x_{t+1} &= A_t x_t + B_t u_t + R_t d_t,\ t = 0, 1, ... \\
x_0 &= z \\
y_t &= C_t x_t + D_t d_t,\ t = 0, 1, ...
\end{array}
\tag{48}
$$

where

- $x_t \in \mathbf{R}^{n_x}$ is the state at time $t$,
- $u_t \in \mathbf{R}^{n_u}$ is the endogeneous control at time $t$,
- $d_t \in \mathbf{R}^{n_d}$ is the exogeneous input at time $t$,
- $y_t \in \mathbf{R}^{n_y}$ is the observable output at time $t$,

and $A_t, B_t, C_t, D_t, t = 0, 1, ...,$ are given matrices of appropriate sizes.

Our goal is to optimize system's behaviour on a given finite time horizon $t = 0, 1, \dots, T$ by designing appropriate *non-anticipative linear control law*. The latter means that the controls at time $t$ should be affine functions of the outputs $y^t = (y_0, \dots, y_t)$ observed till time $t$:

$$u_t = g_t + \sum_{\tau=0}^{t} G_{t\tau} y_\tau,\ 0 \le t \le T; \tag{49}$$

here $g_t$, $G_{t\tau}$ are (in principle, arbitrary) vectors and matrices of appropriate sizes. We wish to choose the "parameter" $\gamma = \{g_t, G_{t\tau}\}_{0\le\tau\le t\le T}$ of the control law in a way which ensures a desired behaviour of the system and minimizes under this restriction a given loss function. We assume that both the desired behaviour of the system and the loss is expressed in terms of the resulting *state-control trajectory*

$$w^T = (x^{T+1} \equiv (x_0, \dots, x_{T+1}), u^T = (u_0, \dots, u_T)).$$

Specifically, the "desired behaviour" is modelled by a system of convex constraints

$$p_i + P_i w^T \in K_i,\ i = 1, \dots, I, \tag{50}$$

on the trajectory; here $p_i$ are given $k_i$-dimensional vectors, $P_i$ are given $k_i \times \dim w^T$ matrices and $K_i \subset \mathbf{R}^{k_i}$ are given nonempty closed convex sets. For the sake of simplicity, we restrict ourselves with a linear loss function

$$\langle c, w^T \rangle \tag{51}$$

Our goal is to minimize the loss function by choice of the parameters $\gamma$ of control law (49) restricted to ensure constraints (50).

*Uncertainty.* When specifying a control law, we do not fully specify the state-space trajectory – it depends on the control law *and* on the inputs $d^T = (d_0, \dots, d_T)$, as well as on the initial state $z$:

$$w^T = W^T(\gamma; z, d^T), \tag{52}$$

where $W^T(\cdot;\cdot)$ is a function readily given by the data in (48), that is, the matrices $A_t$, $B_t, C_t, R_t$. In typical applications, the inputs (and in many cases the initial state as well) are not fully known when building the control law, so that it is natural to treat them as the uncertain part $\zeta = (z, d^T)$ of the data. Consequently, when substituting the right hand side of (52) into (50) and (51), we get an *uncertain* optimization problem

$$\min_{\gamma} \Big\{ \underbrace{\langle c, W^T(\gamma;\zeta)\rangle}_{F_0(\gamma,\zeta)} : \underbrace{p_i + P_i W^T(\gamma;\zeta)}_{F_i(\gamma,\zeta)} \in K_i,\ i = 1, \dots, I \Big\} \tag{53}$$

which is in the form of (1). We could now apply to (53) the methodology we have developed. However, there is an obstacle: the problem "as it is" severely violates the assumption of bi-affinity of $F_i$, the assumption playing a crucial role in converting the CRC of (53) into a computationally tractable form. Indeed, due to the linearity of system (48) and affinity of control laws we intend to use, the state-space trajectory $w^T = W^T(\gamma;\zeta)$ does depend affinely on $\zeta = (z, d^T)$, but, in contrast, its dependence on the design variables $\gamma$ is highly non-linear. Consequently, the functions $F_i(\gamma, \zeta)$ are highly nonlinear functions of $\gamma$. Fortunately, we have a remedy: linear control laws can be re-parameterized in such a way that the state-space trajectory (and consequently functions $F_i(\cdot,\cdot)$) will become bi-affine in the new parameters of a control law and in $\zeta$. As a result, under reasonable structural restrictions on the constraints (50) the CRC of (53) turns out to be computationally tractable, specifically, an explicit Linear Programming problem of sizes polynomial in $T$, sizes of the matrices in (48) and the number $I$ of constraints in (50). We start processing (53) with developing the aforementioned re-parameterization of linear control laws (resembling what is called "$Q$-parameterization" in Control).

### 4.2. *Linear control revisited*

*Notational convention.* From now on, given a sequence $e_0, e_1, \ldots$ of vectors and an integer $t \geq 0$, we denote by $e^t$ the initial fragment $(e_0, \ldots, e_t)$ of this sequence; for $t < 0$, $e^t$ is, by definition, a zero vector.

*Purified outputs.* Assume that we "close" the open-loop system (48) with a (not necessary linear) control $u_t = U_t(y^t)$, and consider, along with the resulting closed loop system

$$\begin{array}{rl} x_{t+1} &= A_t x_t + B_t u_t + R_t d_t \\ x_0 &= z \\ y_t &= C_t x_t + D_t d_t \\ \hline u_t &= U_t(y^t) \end{array} \tag{54}$$

its *model*

$$\begin{aligned} \widehat{x}_{t+1} &= A_t \widehat{x}_t + B_t u_t \\ \widehat{x}_0 &= 0 \\ \widehat{y}_t &= C_t \widehat{x}_t \end{aligned} \tag{55}$$

where $u_t$ are the controls given by (54). Since we know the matrices $A_t, B_t, C_t$ (and of course know the controls we are generating), we can run the model in an "on-line" fashion, so that at time $t$, when the decision on $u_t$ should be made, we have in our disposal the model outputs $\widehat{y}_\tau$, $0 \leq \tau \leq t$. It follows that at this time we also know *the purified outputs* $v_t = y_t - \widehat{y}_t$.

*Re-parameterization of linear control laws.* Now let us equip (48) with a control law where controls $u_t$ are affine functions of the purified outputs observed till time $t$:

$$u_t = h_t + \sum_{\tau=0}^{t} H_{t\tau} v_\tau,$$

$h_t$, $H_{t\tau}$ being vectors and matrices of appropriate sizes. The resulting closed loop system is given by the relations

$$\begin{array}{|l|}\hline \text{plant:} \\ (a):\ \begin{cases} x_0 = z \\ x_{t+1} = A_t x_t + B_t u_t + R_t d_t \\ y_t = C_t x_t + D_t d_t \end{cases} \\ \hline \text{model:} \\ (b):\ \begin{cases} \widehat{x}_0 = 0 \\ \widehat{x}_{t+1} = A_t \widehat{x}_t + B_t u_t \\ \widehat{y}_t = C_t \widehat{x}_t \end{cases} \\ \hline \text{purified outputs:} \\ (c):\ v_t = y_t - \widehat{y}_t \\ \hline \text{control law:} \\ (d):\ u_t = h_t + \sum_{\tau=0}^{t} H_{t\tau} v_\tau \\ \hline \end{array} \tag{56}$$

We make the following simple observation:

**Proposition 3.** (i) *For every linear control law in the form of* (49)*, there exists a control law in the form of* (56.*d*) *which, whatever be the initial state and a sequence of inputs, results in exactly the same state-control trajectories of the closed loop system;*

(ii) *Vice versa, for every linear control law in the form of* (56.*d*)*, there exists a control law in the form of* (49) *which, whatever be the initial state and a sequence of inputs, results in exactly the same state-control trajectories of the closed loop system;*

(iii) [bi-affinity] *The state-control trajectory $w^T$ of closed loop system* (56) *is affine in $z$, $d^T$ when the parameters $\eta = \{h_t, H_{t\tau}\}_{0\le\tau\le t\le T}$ of the underlying control law are fixed, and is affine in $\eta$ when $z$, $d^T$ are fixed:*

$$w^T = \omega[\eta] + \Omega_z[\eta]z + \Omega_d[\eta]d^T \tag{57}$$

*for some vectors $\omega[\eta]$ and matrices $\Omega_z[\eta]$, $\Omega_d[\eta]$ affinely depending on $\eta$.*

*Proof.* (i): Let us fix a linear control law in the form of (49), and let $x_t = X_t(z, d^{t-1})$, $u_t = U_t(z, d^t)$, $y_t = Y_t(z, d^t)$, $v_t = V_t(z, d^t)$ be the corresponding states, controls, outputs and purified outputs. To prove (i) it suffices to show that for every $t \ge 0$ with properly chosen vectors $q_t$ and matrices $Q_{t\tau}$ one has

$$\forall (z, d^t) : Y_t(z, d^t) = q_t + \sum_{\tau=0}^{t} Q_{t\tau} V_\tau(z, d^\tau). \tag{I$_t$}$$

Indeed, given the validity of these relations and taking into account (49), we would have

$$U_t(z,d^t) \equiv h_t + \sum_{\tau=0}^{t} H_{t\tau} Y_\tau(z,d^\tau) \equiv \widetilde{h}_t + \sum_{\tau=0}^{t} \widetilde{H}_{t\tau} V(z,d^\tau) \qquad (\mathrm{II}_t)$$

with properly chosen $\widetilde{h}_t$, $\widetilde{H}_{t\tau}$, so that the control law in question can indeed be represented as a linear control law via purified outputs.

We shall prove ($\mathrm{I}_t$) by induction in $t$. The base $t = 0$ is evident, since by (56.*a*-*c*) we merely have $Y_0(z,d^0) \equiv V_0(z,d^0)$. Now let $s \geq 1$ and assume that relations ($\mathrm{I}_t$) are valid for $0 \leq t < s$. Let us prove the validity of ($\mathrm{I}_s$). From the validity of ($\mathrm{I}_t$), $t < s$, it follows that the relations ($\mathrm{II}_t$), $t < s$, take place, whence, by the description of the model system, $\widehat{x}_s = \widehat{X}_s(z,d^{s-1})$ is affine in the purified outputs, and consequently the same is true for the model outputs $\widehat{y}_s = \widehat{Y}_s(z,d^{s-1})$:

$$\widehat{Y}_s(z,d^{s-1}) = p_s + \sum_{\tau=0}^{s-1} P_{s\tau} V_\tau(z,d^\tau).$$

We conclude that with properly chosen $p_s$, $P_{s\tau}$ we have

$$\begin{aligned} Y_s(z,d^s) &\equiv \widehat{Y}_s(z,d^{s-1}) + V_s(z,d^s) \\ &= p_s + \sum_{\tau=0}^{s-1} P_{s\tau} V_\tau(z,d^\tau) + V_s(z,d^s), \end{aligned}$$

as required in ($\mathrm{I}_s$). Induction is completed, and (i) is proved.

(ii): Let us fix a linear control law in the form of (56.*d*), and let

$$x_t = X_t(z,d^{t-1}),\ \widehat{x}_t = \widehat{X}_t(z,d^{t-1}),\ u_t = U_t(z,d^t),\ y_t = Y_t(z,d^t),$$
$$v_t = V_t(z,d^t)$$

be the corresponding actual and model states, controls, and actual and purified outputs. We should verify that the state-control dynamics in question can be obtained from an appropriate control law in the form of (49). To this end, similarly to the proof of (i), it suffices to show that for every $t \geq 0$ one has

$$V_t(z,d^t) \equiv q_t + \sum_{\tau=0}^{t} Q_{t\tau} Y_\tau(z,d^\tau) \qquad (\mathrm{III}_t)$$

with properly chosen $q_t$, $Q_{t\tau}$. We again apply induction in $t$. The base $t = 0$ is again trivially true due to $V_0(z,d^0) \equiv Y_0(z,d^0)$. Now let $s \geq 1$, and assume that relations ($\mathrm{III}_t$) are valid for $0 \leq t < s$, and let us prove that ($\mathrm{III}_s$) is valid as well. From the validity of ($\mathrm{III}_t$), $t < s$, and from (56.*d*) it follows that

$$t < s \Rightarrow U_t(z,d^t) = h_t + \sum_{\tau=0}^{t} H_{t\tau} Y_\tau(z,d^\tau)$$

with properly chosen $h_t$ and $H_{t\tau}$. From these relations and the description of the model system it follows that its state $\widehat{X}_s(z, d^{s-1})$ at time $s$, and therefore the model output $\widehat{Y}_s(z, d^{s-1})$, are affine functions of $Y_0(z, d^0), \ldots ,Y_{s-1}(z, d^{s-1})$:

$$\widehat{Y}_s(z, d^{s-1}) = p_s + \sum_{\tau=0}^{s-1} P_{s\tau} Y_\tau(z, d^\tau)$$

with properly chosen $p_s$, $P_{s\tau}$. It follows that

$$\begin{aligned} V_s(z, d^s) &\equiv Y_s(z, d^s) - \widehat{Y}_s(z, d^{s-1}) \\ &= Y_s(z, d^s) - p_s - \sum_{\tau=0}^{s-1} P_{s\tau} Y_\tau(z, d^\tau), \end{aligned}$$

as required in (III$_s$). Induction is completed, and (ii) is proved.

(iii): For $0 \le s \le t$ let

$$A_s^t = \begin{cases} \prod_{r=s}^{t-1} A_r, & s < t \\ I, & s = t \end{cases}$$

Setting $\delta_t = x_t - \widehat{x}_t$, we have by (56.*a*-*b*)

$$\delta_{t+1} = A_t \delta_t + R_t d_t, \ \ \delta_0 = z \Rightarrow \delta_t = A_0^t z + \sum_{s=0}^{t-1} A_{s+1}^t R_s d_s$$

(from now on, sums over empty index sets are zero), whence

$$v_\tau = C_\tau \delta_\tau + D_\tau d_\tau = C_\tau A_0^\tau z + \sum_{s=0}^{\tau-1} C_\tau A_{s+1}^\tau R_s d_s + D_\tau d_\tau. \tag{58}$$

Therefore control law (56.*d*) implies that

$$\begin{aligned} u_t &= h_t + \sum_{\tau=0}^{t} H_{t\tau} v_\tau = \underbrace{h_t}_{\nu_t[\eta]} \\ &\quad + \underbrace{\left[\sum_{\tau=0}^{t} H_{t\tau} C_\tau A_0^\tau\right]}_{N_t[\eta)} z + \sum_{s=0}^{t-1} \underbrace{\left[H_{ts} D_s + \sum_{\tau=s+1}^{t} H_{t\tau} C_\tau A_{s+1}^\tau R_s\right]}_{N_{ts}[\eta]} d_s + \underbrace{H_{tt} D_t}_{N_{tt}[\eta]} d_t \\ &= \nu_t[\eta] + N_t[\eta] z + \sum_{s=0}^{t} N_{ts}[\eta] d_s \end{aligned} \tag{59}$$

whence, invoking (56.$a$),

$$\begin{aligned} x_t &= A_0^t z + \sum_{\tau=0}^{t-1} A_{\tau+1}^t [B_\tau u_\tau + R_\tau d_\tau] = \underbrace{\left[\sum_{\tau=0}^{t-1} A_{\tau+1}^t B_\tau \nu_t[\eta]\right]}_{\mu_t[\eta]} \\ &\quad + \underbrace{\left[A_0^t + \sum_{\tau=0}^{t-1} A_{\tau+1}^t B_\tau N_\tau[\eta]\right]}_{M_t[\eta]} z \\ &\quad + \sum_{s=0}^{t-1} \underbrace{\left[\sum_{\tau=s}^{t-1} A_{\tau+1}^t B_\tau N_{\tau s}[\eta] + A_{s+1}^t B_s R_s\right]}_{M_{ts}[\eta]} d_s \\ &= \mu_t[\eta] + M_t[\eta] z + \sum_{s=0}^{t-1} M_{ts}[\eta] d_s. \end{aligned} \tag{60}$$

We see that the states $x_t$, $0 \le t \le T+1$, and the controls $u_t$, $0 \le t \le T$, of the closed loop system (56) are affine functions of $z$, $d^T$, and the corresponding "coefficients" $\mu_t[\eta]$, ... ,$N_{ts}[\eta]$ are affine vector- and matrix-valued functions of the parameters $\eta = \{h_t, H_{t\tau}\}_{0 \le \tau \le t \le T}$ of the underlying control law (56.$d$). □

As we shall see in a while, the bi-affinity property proved in Proposition 3 allows to synthesize efficiently a linear control law, if any, which meets given finite-horizon control specifications when the latter are expressed by a system of linear (or convex nonlinear) constraints on the states and the controls. It should be stressed that this attractive option exists only in the case when we seek for a "general-type" linear control. Adding, along with linearity, other structural restrictions on the control law can make the synthesis problem difficult. For example, this is the case with the simple linear feedback control

$$u_t = K_t y_t \tag{61}$$

which is of primary interest in Control. Indeed, laws of the form (61) form a proper subset $\mathcal{C}$ in the set of all linear control laws. With parameterization (49), $\mathcal{C}$ looks very simple (it is just a linear subspace in the space of all linear control laws), but this does not help much in computationally efficient synthesis, since the parameterization itself is bad for this purpose. With the outlined re-parameterization which eliminates the latter difficulty, $\mathcal{C}$ is cut off the entire space of parameters by a system of high-order polynomial equations, and optimization over this highly nonlinear set seems to be computationally intractable.

### *4.3. Tractability of the CRC*

With Proposition 3 at hand, we can proceed to use the results of Section 3 in order to build and process the CRC of problem (53). Specifically, the state-space trajectory $w^T$

of (56) is a bi-affine function of the parameters $\eta = \{h_t, H_{t\tau}\}_{0 \le \tau \le t \le T}$ of the underlying control law and the uncertain data $\zeta = (z, d^T)$:

$$w^T = \omega[\eta] + \Omega[\eta]\zeta = \omega[\eta] + \Omega^z[\eta]z + \Omega^d[\eta]d^T \tag{62}$$

($\omega[\eta]$, $\Omega[\eta]$ are affine in $\eta$). Consequently, problem (53) can be equivalently rewritten as

$$\min_{\eta} \left\{ \underbrace{\langle c, \omega[\eta] + \Omega[\eta]\zeta \rangle}_{\widehat{F}_0(\eta,\zeta)} : \underbrace{p_i + P_i\,[\omega[\eta] + \Omega[\eta]\zeta]}_{\widehat{F}_i(\eta,\zeta)} \in K_i,\ i = 1, \ldots, I \right\}, \tag{63}$$

the functions $\widehat{F}_i(\eta, \zeta)$, $0 \le i \le I$ being bi-affine. All decision variables in (63) are non-adjustable, so that the Fixed Recourse assumption holds trivially true. Thus, we are in a situation which satisfies the assumptions of Section 3.1, and we can utilize all constructions and results of Section 3. In particular, consider the case when

- all $K_i$ in (50) are one-dimensional (or, equivalently, (50) is a system of *linear* equalities and inequalities on the states and the controls),
- the normal ranges of the components $\zeta^\ell$ of $\zeta$ are boxes, the cones $\mathcal{L}_\ell$ are of the form (42), and the Minkowski functions $\phi_{i\ell}(\cdot)$ are given by (44),

In this case, Theorem 1 ensures computational tractability of the CRC of (63) and, moreover, provides its equivalent reformulation as an explicit convex program[3]. As a result, we get a possibility to check efficiently whether given control specifications, expressed by a system of linear inequalities on states and controls over finite time horizon, can be satisfied by a linear control law, whatever be exogeneous inputs and initial states varying in their (bounded) normal ranges. We can further optimize the performance of the closed loop system, provided that the latter is quantified by a linear function of states and controls. Moreover, we can take care of global sensitivities of system's behaviour to deviations of the inputs/initial states from their normal ranges. These possibilities seem to be very attractive and, we believe, deserve extensive exploration. There is, however, a limitation of the approach we have outlined: by its nature, it is restricted to handle finite-horizon control problems only. Of course, "infinite time horizon" by itself is a mathematical abstraction we do not meet in real-life applications. The actual bottleneck in our approach is that the computational effort required to solve the control synthesis problem, grows nonlinearly (although polynomially) with the time horizon $T$, which makes the approach impractical when $T$ is large; for existing optimization techniques, already $T = 100$ is too much ... While time horizon of few tens could be appropriate for management applications, it may be too small for engineering ones. However, certain important "infinite horizon" control specifications are still amenable to the outlined CRC approach. An example is offered next.

*Example: stabilizing the closed-loop system.* Stability is one of the most typical specifications in infinite-horizon linear control; it requires from states and controls of the closed loop system to go to 0 as $t \to \infty$ whenever $d_t \to 0$, $t \to \infty$. At a first glance, the

[3] As explained in Remark 1, the latter conclusion remains valid under significantly milder assumptions on (50) and on the ingredients in the CRC setup.

requirement that a linear control law should make the closed loop system stable cannot be addressed by our finite-horizon-oriented synthesis approach, but in fact it can, at least in the time-invariant case (matrices $A_t$, $B_t$, $C_t$, $D_t$, $R_t$ in (48) are independent of $t$). In this case, when applying the CRC approach, we can specify the initial state $z$ of (48) as the first component $\zeta^1$ of the uncertain data $(z, d^T)$, $d^T$ as the second component of the data, choosing the corresponding cones $\mathcal{L}_\ell$, $\ell = 1, 2$ to be the entire spaces, thus making all initial states and all input sequences "physically possible". Let us also ensure that the normal range of inputs contains the origin and that system (50) includes bounds on states $x_t$ and controls $u_t$:

$$\|x_t\|_\infty \leq a_t,\ 0 \leq t \leq T+1, \quad \|u_t\|_\infty \leq b_t,\ 0 \leq t \leq T, \tag{64}$$

where we set $a_{T+1} = 0$. In the CRC of (63), the bound $\|x_{T+1}\|_\infty \leq a_{T+1} = 0$ will be represented by a system of semi-infinite constraints of the form

$$\begin{array}{l}\forall (z, d^T) : \mathrm{dist}_{|\cdot|}(\chi_j[\eta; z, d^T], \{0\}) \leq \alpha_j \mathrm{dist}_{\phi_{j1}}(z, \mathcal{U}_1) + \beta_j \mathrm{dist}_{\phi_{j2}}(d^T, \mathcal{U}_2 | \mathcal{L}_2),\\ 1 \leq j \leq \dim x,\end{array} \tag{65}$$

where $\chi_j[\eta; z, d^T]$ are the coordinates of $x_{T+1}$ expressed as bi-affine functions of the parameters $\eta$ of a control law and the uncertain data $\zeta = (z, d^T)$, while all other entities are ingredients of the CRC setup. Now let us specify the Minkowski functions $\phi_{j1}$ as the $\|\cdot\|_\infty$-norm on the state space and impose on the sensitivities $\alpha_j$ bounds

$$\alpha_j \leq \theta,$$

where $\theta \in (0, 1)$ is a given parameter. Note that with this setup, a linear control law which is feasible for the CRC of (63) possesses the property that the state $x_{T+1} = x_{T+1}(z, d^T)$ of the closed loop system at time $T+1$ satisfies the bound

$$\|x_{T+1}(z, d^T)\|_\infty \leq \theta \|z\|_\infty + \beta \|d^T\|$$

We can now use such a finite-horizon control law in cyclic fashion, that is, use "as it is" at the first $T+1$ time instants $0, 1, \ldots, T$, then shift by $T+1$ the origin on the time axis (thus making $x_{T+1}$ our new initial state) and use the same control law for $T+1$ instants more, then again shift our "instant 0" by $T+1$, use our control law for $T+1$ instants more, and so on. It is immediately seen that due to $\theta \in (0, 1)$, the resulting infinite-horizon control law stabilizes the closed loop system. Thus, in principle our finite-horizon approach allows to take care both of the "nearest" and the remote future.

## References

1. Ben-Tal, A., Nemirovski, A.: Stable Truss Topology Design via Semidefinite Programming. SIAM J. Optimization **7**, 991–1016 (1997)
2. Ben-Tal, A., Nemirovski, A.: Robust Convex Optimization. Math. Oper. Res. **23**, 769–805 (1998)
3. Ben-Tal, A., Nemirovski, A.: Robust solutions to uncertain linear programs. OR Letters **25**, 1–13 (1999)
4. Ben-Tal, A., El-Ghaoui, L., Nemirovski, A.: Robust semidefinite programming. In: Semidefinite Programming and Applications, R. Saigal, R., Vandenberghe, L., Wolkowicz, H. (eds), Kluwer Academic Publishers, 2000
5. Ben-Tal, A., Nemirovski, A.: Robust Optimization — Methodology and Applications. Math. Program. Series B **92**, 453–480 (2002)

6. Ben-Tal, A., Goryashko, A., Guslitzer, E., Nemirovski, A.: Adjustable Robust Solutions of Uncertain Linear Programs: Math. Program. **99**, 351–376 (2004)
7. Ben-Tal, A., Golany, B., Nemirovski, A., Vial, J.-Ph.: Supplier-Retailer Flexible Commitments Contracts: A Robust Optimization Approach. Accepted to Manufacturing & Service Operations Management
8. Bertsimas, D., Pachamanova, D., Sim, M.: Robust Linear Optimization under General Norms. Oper. Res. Lett. **32**, 510–516 (2004)
9. Bertsimas, D., Sim, M.: The price of Robustness. Oper. Res. **52**, 35–53 (2004)
10. Bertsimas, D., Sim, M.: Robust Discrete optimization and Network Flows. Math. Program. Series B **98**, 49–71 (2003)
11. Bertsekas, D.P., Nedic, A., Ozdaglar, A.E.: Convex Analysis and Optimization, Athena Scientific, 2003
12. El-Ghaoui, L., Lebret, H.: Robust solutions to least-square problems with uncertain data matrices. SIAM J. Matrix Anal. Appl. **18**, 1035–1064 (1997)
13. El-Ghaoui, L., Oustry, F., Lebret, H.: Robust solutions to uncertain semidefinite programs. SIAM J. Optimization **9**, 33–52 (1998)
14. Grötschel, M., Lovasz, L., Schrijver, A.: The Ellipsoid Method and Combinatorial Optimization, Springer, Heidelberg, 1988
15. Soyster, A.L.: Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. Oper. Res., 1973, pp. 1154–1157