ROBUST CONVEX OPTIMIZATION

A. BEN-TAL AND A. NEMIROVSKI

We study convex optimization problems for which the data is not specified exactly and it is only known to belong to a given uncertainty set $\mathcal{U}$, yet the constraints must hold for all possible values of the data from $\mathcal{U}$. The ensuing optimization problem is called robust optimization. In this paper we lay the foundation of robust convex optimization. In the main part of the paper we show that if $\mathcal{U}$ is an ellipsoidal uncertainty set, then for some of the most important generic convex optimization problems (linear programming, quadratically constrained programming, semidefinite programming and others) the corresponding robust convex program is either exactly, or approximately, a tractable problem which lends itself to efficient algorithms such as polynomial time interior point methods.

AMS 1991 subject classification. Primary: 90C05, 90C25, 90C30.
OR/MS subject classification. Primary: Programming/Convex.
Key words. Convex optimization, data uncertainty, robustness, linear programming, quadratic programming, semidefinite programming, geometric programming.

1 Introduction

Robust counterpart approach to uncertainty. Consider an optimization problem of the form

$$\begin{array}{ll} (p_\zeta) & \min\limits_{x\in\mathbf{R}^n} f(x,\zeta) \\ \text{s.t.} & \\ & F(x,\zeta)\in\mathbf{K}\subset\mathbf{R}^m, \end{array} \tag{1}$$

where

- $\zeta \in \mathbf{R}^M$ is the data element of the problem;
- $x \in \mathbf{R}^n$ is the decision vector;
- the dimensions n, m, M, the mappings $f(\cdot,\cdot)$, $F(\cdot,\cdot)$ and the convex cone $\mathbf{K}$ are structural elements of the problem.

In this paper we deal with a "decision environment" which is characterized by

> (i) A crude knowledge of the data: it may be partly or fully "uncertain", and all that is known about the data vector ζ is that it belongs to a given *uncertainty set* $\mathcal{U} \subset \mathbf{R}^M$
>
> (ii) The constraints $F(x,\zeta) \in \mathbf{K}$ *must* be satisfied, whatever the actual realization of $\zeta \in \mathcal{U}$ is.

In view of (i) and (ii) we call a vector x *feasible solution* to the *uncertain optimization problem* $(P) = \{(p_\zeta)\}_{\zeta\in\mathcal{U}}$, if x satisfies all possible realizations of the constraints:

$$F(x,\zeta) \in \mathbf{K} \quad \forall \zeta \in \mathcal{U}. \tag{2}$$

Note that this is the same notion of feasibility as in Robust Control (see, e.g., Zhou, Doyle and Glover (1995)).

We can define in the same manner the notion of an *optimal* solution to the uncertain optimization problem (P): such a solution must give the best possible *guaranteed value*

$$\sup_{\zeta \in \mathcal{U}} f(x, \zeta)$$

of the original objective under constraints (2), i.e., it should be an optimal solution to the following "certain" optimization problem

$$(P^*) \quad \min \left\{ \sup_{\zeta \in \mathcal{U}} f(x, \zeta) : F(x, \zeta) \in \mathbf{K} \quad \forall \zeta \in \mathcal{U} \right\}.$$

From now on we call feasible/optimal solutions to (P^*) *robust feasible/optimal* solutions to the uncertain optimization problem (P), and the optimal value in (P^*) is called the *robust optimal value* of uncertain problem (P); the problem (P^*) itself will be called the *robust counterpart* of (P).

Motivation. The "decision environment" we deal with is typical for many applications. The reasons for a crude knowledge of the data may be:

- ζ is unknown at the time the values of x should be determined and will be realized in the future, e.g.: x is a vector of production variables and ζ is comprised of future demands, market prices, etc., or: x represents the cross-sections of bars in a truss construction like a bridge, and ζ represents locations and magnitudes of future loads the bridge will have to carry, etc.
- ζ is realizable at the time x is determined, but it cannot be measured/estimated/computed exactly, e.g.: material properties like Young's modulus, pressures and temperatures in remote places,...
- Even if the data is certain and an optimal solution x^* can be *computed* exactly, it cannot be *implemented* exactly, which is in a sense equivalent to uncertainty in the data. Consider, e.g., a linear programming problem $\min\{c^T x : Ax + b \geq 0, x \geq 0\}$ of an engineering origin, where the components of the decision vector x correspond to "physical characteristics" of the construction to be manufactured (sizes, weights, etc.). Normally it is known in advance that what will be actually produced will be not exactly the computed vector x, but a vector x' with components close (say, within 5% margin) to those of x. This situation is equivalent to the case where there are no inaccuracies in producing x, but there is uncertainty in the constraint matrix A and the objective c – they are known "up to multiplication by a diagonal matrix with diagonal entries varying from 0.95 to 1.05".

Situations when the constraints are "hard", so that (ii) is a "must", are quite common in reality, especially in engineering applications. Consider, e.g., a typical technological process in the chemical industry. Such a process consists of several decomposition – recombination stages, the yields of the previous stages being the "raw materials" for the next one. As a result, the model of the process includes many balance constraints indicating that the amount of materials of different types used at a given stage cannot exceed the amount of these materials produced at the preceding stages. On one hand, the data coefficients in these constraints are normally inexact – the contents of the input raw materials, the parameters of the devices, etc., typically are "floating". On the other hand, violation of physical balance constraints can "kill" the process completely and is therefore inadmissible. Similar situation arises in other engineering applications, where violating physical constraints may cause an explosion or a crush of the construction. As a concrete example of this type, consider the *Truss Topology Design* (TTD) problem (for more details, see Ben-Tal and Nemirovski (1997)).

A *truss* is a construction comprised of thin elastic *bars* linked with each other at *nodes* – points from a given finite (planar or spatial) set. When subjected to a given *load* – a collection of external forces acting at some specific nodes – the construction deformates, until the tensions caused by the deformation compensate the external load. The deformated truss capacitates certain potential energy, and this energy – the *compliance* – measures stiffness of the truss (its ability to withstand the load); the less is compliance, the more rigid is the truss.

In the usual TTD problem we are given the initial nodal set, the external "nominal" load and the total volume of the bars. The goal is to allocate this resource to the bars in order to minimize the compliance of the resulting truss.

Fig. 1.a) shows a cantilever arm which withstands optimally the nominal load – the unit force f^* acting down at the most right node. The corresponding optimal compliance is 1.

It turns out that the construction in question is highly unstable: a small force f (10 times smaller than f^*) depicted by small arrow on Fig. 1.a) results in a compliance which is more than 3000 times larger than the nominal one.

In order to improve design's stability, one may use the robust counterpart methodology. Namely, let us treat the external load as the uncertain data of the associated optimization problem. Note that in our example a load is, mathematically, a collection of ten 2D vectors representing (planar) external forces acting at the ten non-fixed nodes of the cantilever arm; in other words, the data in our problem is a 20-dimensional vector. Assume that the construction may be subjected not to the nominal load only, but also to a "small occasional load" represented by an arbitrary 20-dimensional vector of the Euclidean norm ≤ 0.1 (i.e., of the norm 10 times smaller than the norm of the nominal load). Thus, we allow the uncertain data of our problem to take the nominal value f^*, as well as any value from the 20-dimensional Euclidean ball $B_{0.1}$ centered at the origin. In order to get an efficiently solvable robust counterpart, we extend the resulting data set $B_{0.1} \cup \{f^*\}$ to its "ellipsoidal envelope" $\mathcal{U}$, i.e., the smallest volume 20-dimensional ellipsoid centered at the origin and containing f^* and $B_{0.1}$. This ellipsoid is then taken as our uncertainty set $\mathcal{U}$.

Solving the robust counterpart of the resulting uncertain problem, we get the cantilever arm shown on Fig. 1.b).

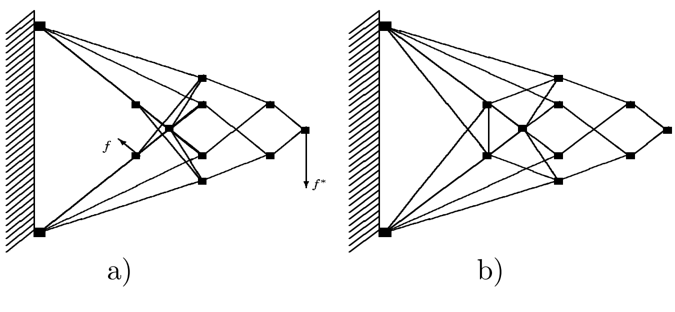

Figure 1. Cantilever arm: nominal design (left) and robust design (right)

The compliances of the original and the new constructions with respect to the nominal load and their worst-case compliances with respect to the "occasional loads" from $B_{0.1}$ are as follows:

Design	Compliance w.r.t. f^*	Compliance w.r.t. $B_{0.1}$
single-load	1	> 3360
robust	1.0024	1.003

We see that in this example the robust counterpart approach improves dramatically the stability of the resulting construction, and that the improvement is in fact "costless" – the robust optimal solution is nearly optimal for the nominal problem as well.

There are of course situations in reality when "occasional violations" of the constraints are not that crucial, and in these situations the robust counterpart approach may seem to be too conservative. We believe, however, that even in these situations, the robust optimization methodology should be at least tried – it may well happen that ensuring robust feasibility is "nearly costless", as it is in the above TTD example and other examples from Ben-Tal and Nemirovski (1997). Problems arising in applications often are highly degenerate, and as a result possess a quite "massive" set of nearly optimal solutions; ignoring the uncertainty and obtaining the decision by solving the problem with the "nominal" data, we normally end up with a point on the boundary of the above set, and such a point may be a very bad solution to a problem with a slightly perturbed data. In contrast to this, the robust counterpart tries to choose the "most inner" solution from the above massive set, and such a solution is normally much more stable w.r.t. data perturbations than a boundary one.

Robust counterpart approach vs. traditional ways to treat uncertainty. Uncertainty[1)] is treated in the Operations Research/Engineering literature in several ways.

- In many cases the uncertainty is simply ignored – at the stage of building the model and finding an optimal solution uncertain data are replaced with some nominal values, and the only care of uncertainty (if any) is taken by *sensitivity analysis*. This is a *post-optimization* tool allowing just to analyze the stability properties of the already generated solution. Moreover, this analysis typically can deal only with "infinitesimal" perturbations of the nominal data.

- There exist modeling approaches which handle uncertainty directly and from the very beginning, notably Stochastic Programming. This approach is limited to those cases where the uncertainty is of stochastic nature (which is not always the case) and we are able to identify the underlying probability distributions (which is even more problematic, especially in the large-scale case). Another essential difference between the SP and the robust counterpart approaches is that the majority of SP models allow a solution to violate the constraints affected by uncertainty and thus do not meet the requirement (ii).

  Another approach to handle uncertainty is "Robust Mathematical Programming" (RMP) proposed recently by Mulvey, Vanderbei and Zenios (1995). Here instead of fixed "nominal" data several scenarios are considered, and a candidate solution is allowed to violate the "scenario realizations" of the constraints. These violations are included via penalty terms in the objective, which allows to take care of the stability of the resulting solution. The RMP approach again may produce solutions which are not feasible for realizations of the constraints, even for the scenario ones. When treating the scenario realizations of the constraints as "obligatory", the RMP approach, under mild structural assumptions, becomes a particular case of the robust counterpart scheme, namely, the case of an uncertainty set given as the convex hull of the scenarios.

[1)] We refer here to uncertainty in the constraints. For uncertainty in the objective function, there is a vast literature in economics, statistics and decision theory.

Robust counterpart approach: the goal. As we have already mentioned, our approach is motivated by the need to comply with the difficulties imposed by the environment described in (i), (ii) above. But there is another equally important consideration: for robust optimization to be an *applicable* methodology for real life large scale problems, it is essential that

(iii) The robust counterpart must be "computationally tractable".

At a first glance this requirement looks unachievable, as the robust counterpart (P^*) of an uncertain problem (P) is a *semi-infinite* optimization problem. It is well-known that semi-infinite programs, even convex, often cannot be efficiently solved numerically, and in particular are not amenable to the use of advanced optimization tools like interior point methods. This in practice imposes severe restrictions on the sizes of problems which can be actually solved. Therefore a goal of primary importance for applications of the robust optimization approach is converting the robust counterparts of generic convex problems to "explicit" convex optimization programs, accessible for high-performance optimization techniques. Possibilities of such a conversion depend not only on the analytical structure of the generic program in question, but also on the geometry of the uncertainty set $\mathcal{U}$. In the case of "trivial geometry" of the uncertainty set, when $\mathcal{U}$ is given as a convex hull of a finite set of "scenarios" $\{\zeta^1, ..., \zeta^N\}$, the "conversion problem" typically does not arise at all. Indeed, if $f(x, \zeta)$ and $F(x, \zeta)$ "depend properly" on ζ, e.g., are affine in the data (which is the case in many of applications), the robust counterpart of $(P) = \{(p_\zeta)\}_{\zeta \in \mathcal{U}}$ is simply the problem

$$\min \left\{ \tau : \tau \geq f(x, \zeta^i), F(x, \zeta^i) \geq 0, i = 1, ..., N \right\}.$$

The analytical structure of this problem is exactly the same as the one of the instances. In our opinion, this simple geometry of the uncertainty set is not that interesting mathematically and is rather restricted in its expressive abilities. Indeed, a "scenario" description/approximation of typical uncertainty sets (even very simple ones, given by element-wise bounds on the data, or more general standard "inequality-represented" polytopes) requires an unrealistically huge number of scenarios. We believe that more reasonable types of uncertainty sets are ellipsoids and intersections of finitely many ellipsoids (the latter geometry covers also the case of a polytope uncertainty set given by a list of inequalities – a half-space is a "very large" ellipsoid). Indeed:

- An ellipsoid is a very convenient entity from the mathematical point of view, it has a simple parametric representation and can be easily handled numerically.

- In many cases of *stochastic* uncertain data there are probabilistic arguments allowing to replace stochastic uncertainty by an *ellipsoidal* deterministic uncertainty. Consider, e.g., an uncertain Linear Programming (LP) problem with random entries in the constraint matrix (what follows can be straightforwardly extended to cover the case of random uncertainty also in the objective and the right hand side of the constraints). For a given x, the left hand side $l_i(x) = a_i^T x + b_i$ of the i-th constraint in the system $A^T x + b \geq 0$ is a random variable with expectation $e_i(x) = (a_i^*)^T x + b_i$, and standard deviation $v_i(x) = \sqrt{x^T V_i x}$, a_i^* being the expectation and V_i being the covariance matrix of the random vector a_i. A "typical" value of the random variable $l_i(x)$ will therefore be $e_i(x) \pm O(v_i(x))$, and for a "light tail" distribution of the random data a "likely" lower bound on this random variable is $\hat{l}_i(x) = e_i(x) - \theta v_i(x)$ with "safety parameter" θ of order of one (cf. the engineers' "3σ-rule" for Gaussian random variables). This bound leads to the "likely reliable" version

$$e_i(x) - \theta v_i(x) \geq 0$$

of the constraint in question. Now note that the latter constraint is exactly the robust counterpart

$$a_i^T x + b_i \geq 0 \quad \forall a_i \in \mathcal{U}_i$$

of the original uncertain constraint if $\mathcal{U}_i$ is specified as the ellipsoid

$$\mathcal{U}_i = \left\{ a : (a - a_i^*)^T V_i^{-1} (a - a_i^*) \leq \theta^2. \right\}$$

- Finally, properly chosen ellipsoids and especially intersections of ellipsoids can be used as reasonable approximations to more complicated uncertainty sets.

As it is shown later in the paper, in several important cases (linear programming, convex quadratically constrained quadratic programming and some others) the use of ellipsoidal uncertainties leads to "explicit" robust counterparts which can be solved both theoretically and in practice (e.g., by interior-point methods). The derivation of these "explicit forms" of several generic uncertain optimization problems is our main goal.

Robust counterpart approach: previous work. As it was already indicated, the approach in question is directly related to Robust Control. For Mathematical Programming, this approach seems to be new. The only setting and results of this type known to us are those of Singh (1982) and Falk (1976) originating from the 1973 note of A.L. Soyster; this setting deals with a very specific and in fact the most conservative version of the approach (see discussion in Ben-Tal and Nemirovski (1995a)).

The general robust counterpart scheme as outlined below was announced in recent paper Ben-Tal and Nemirovski (1997) on robust Truss Topology Design. Implementation of the scheme in the particular case of uncertain linear programming problems is considered in Ben-Tal and Nemirovski (1995a). Many of the results of the current paper, and some others, were announced in Ben-Tal and Nemirovski (1995b). Finally, when working on the present paper, we became aware of the papers of Oustry, El Ghaoui and Lebret (1996) and El-Ghaoui and Lebret (1996) which also develop the robust counterpart approach, mainly as applied to uncertain semidefinite programming.

For recent progress in robust settings of *discrete* optimization problems, see Kouvelis and Yu (1997).

Summary of results. We mainly focus on (a) general development of the robust counterpart approach, and (b) applications of the scheme to generic families of nonlinear convex optimization problems – linear, quadratically constrained quadratic, conic quadratic and semidefinite ones. As far as (a) is concerned, we start with a formal exposition of the approach (Section 2.1) and investigate general qualitative (Section 2.2) and quantitative (Section 2.3) relations between solvability properties of the instances of an uncertain problem and those of its robust counterpart. The questions we are interested in are: assume that all instances of an uncertain program are solvable. Under what condition is it true that robust counterpart also is solvable? When is there no "gap" between the optimal value of the robust counterpart and the worst of the optimal values of instances? What can be said about the proximity of the robust optimal value and the optimal value in a "nominal" instance?

The main part of our efforts is devoted to (b). We demonstrate that

- The robust counterpart of an uncertain linear programming problem with ellipsoidal or $\bigcap$-ellipsoidal (intersection of ellipsoids) uncertainty is an explicit conic quadratic program (Section 3.1).

- The robust counterpart of an uncertain convex quadratically constrained quadratic programming problem with ellipsoidal uncertainty is an explicit semidefinite program, while a general-type $\bigcap$-ellipsoidal uncertainty leads to an NP-hard robust counterpart (Section 3.2).
- The robust counterpart of an uncertain conic quadratic programming problem with ellipsoidal uncertainty, under some minor restrictions, is an explicit semidefinite program, while a general-type $\bigcap$-ellipsoidal uncertainty leads to an NP-hard robust counterpart (Section 3.3).
- In the case of uncertain semidefinite program with a general-type ellipsoidal uncertainty set the robust counterpart is NP-hard. We present a generic example of a "well-structured" ellipsoidal uncertainty which results in a tractable robust counterpart (an explicit semidefinite program). Moreover, we propose a "tractable" *approximate* robust counterpart of a general uncertain semidefinite problem (Section 3.4).
- We derive an explicit form of the robust counterpart of an *affinely parameterized* uncertain problem (e.g., geometric programming problem with uncertain coefficients of the monomials) (Section 3.5) and develop a specific saddle point form of the robust counterpart for these problems (Section 4).

2 Robust counterparts of uncertain convex programs

In what follows we deal with *parametric convex programs* of the type (1). Technically, it is more convenient to "standardize" the objective – to make it linear and data-independent. To this end it suffices to add a new variable t to be minimized and to add to the list of original constraints the constraint $t - f(x,\zeta) \geq 0$. In what follows we assume that such a transformation has been already done, so that the optimization problems in question are of the type

$$(p): \quad \min \left\{ c^T x : F(x,\zeta) \in \mathbf{K},\ x \in \mathcal{X} \right\} \tag{3}$$

where

- $x \in \mathbf{R}^n$ is the design vector;
- $c \in \mathbf{R}^n$ is *certain* objective;
- $\zeta \in \mathcal{A} \subset \mathbf{R}^M$ is a parameter vector.

From now on we make the following assumption (which in particular ensures that (p) is a convex program whenever $\zeta \in \mathcal{A}$):

Assumption A:

- $\mathbf{K} \subset \mathbf{R}^N$ *is a closed convex cone with a nonempty interior;*
- $\mathcal{X}$ *is a closed convex subset in* $\mathbf{R}^n$ *with a nonempty interior;*
- $\mathcal{A}$ *is closed convex subset in* $\mathbf{R}^M$*; with a nonempty interior;*
- $F(x,\zeta) : \mathcal{X} \times \mathcal{A} \to \mathbf{R}^N$ *is assumed to be continuously differentiable mapping on* $\mathcal{X} \times \mathcal{A}$ *and* $\mathbf{K}$*-concave in* x*, i.e.*

$$\begin{array}{l} \forall (x', x'' \in \mathcal{X}, \zeta \in \mathcal{A})\ \forall (\lambda \in [0,1]) : \\ F(\lambda x' + (1-\lambda)x'', \zeta) \geq_{\mathbf{K}} \lambda F(x',\zeta) + (1-\lambda) F(x'',\zeta), \end{array}$$

where "$b \geq_{\mathbf{K}} a$" stands for $b - a \in \mathbf{K}$.

2.1 Uncertain convex programs and their robust counterparts

We define an *uncertain convex program* (P) as a family of "instances" – programs (p) of the type (3) with common *structure* $[n, c, \mathbf{K}, F(\cdot,\cdot)]$ and the *data vector* ζ running through a given *uncertainty set* $\mathcal{U} \subset \mathcal{A}$:

$$(P) = \left\{(p) : \quad \min\left\{c^T x : F(x,\zeta) \in \mathbf{K}, x \in \mathcal{X}\right\}\right\}_{\zeta\in\mathcal{U}}. \tag{4}$$

We associate with an uncertain program (P) its *robust counterpart* – the following certain optimization program

$$(P^*) : \quad \min\left\{c^T x : x \in \mathcal{X},\ F(x,\zeta) \in \mathbf{K} \quad \forall \zeta \in \mathcal{U}\right\}.$$

Feasible solutions to (P^*) are called *robust feasible solutions to* (P), the infimum of the values of the objective at robust feasible solutions to (P) is called the *robust optimal value* of (P), and, finally, an optimal solution to (P^*) is called *a robust optimal* solution to (P). Note that (P^*) is in fact a semi-infinite optimization problem.

It makes sense (see Propositions 2.1 and 2.2 below) to restrict ourselves to a particular case of uncertainty – the one which is concave in the data:

Definition 2.1 *The uncertain optimization problem (P) has concave uncertainty, if the underlying mapping $F(x,\zeta)$ is $\mathbf{K}$-concave in ζ:*

$$\begin{array}{l}\forall (x \in \mathcal{X}, \zeta', \zeta'' \in \mathcal{A})\ \forall (\lambda \in [0,1]) : \\ F(x, \lambda\zeta' + (1-\lambda)\zeta'') \geq_{\mathbf{K}} \lambda F(x,\zeta') + (1-\lambda)F(x,\zeta'').\end{array}$$

In some cases we shall speak also about *affine uncertainty*:

Definition 2.2 *The uncertain optimization problem (P) has affine uncertainty, if the underlying mapping $F(x,\zeta)$ is affine in ζ for every $x \in \mathcal{X}$.*

As we shall see in a while, a number of important generic convex programs (linear, quadratic, semidefinite,...) are indeed affine in the data.

In the case of concave uncertainty we can assume without loss of generality that the uncertainty set $\mathcal{U}$ is a closed convex set:

Proposition 2.1 *Let (P) be an uncertain convex program with concave uncertainty, and let (P') be the uncertain convex program obtained from (P) by replacing the original uncertainty set $\mathcal{U}$ by its closed convex hull. Then the robust counterparts of (P) and (P') are identical to each other.*

Proof. By assumption **A** the robust counterpart (P^*) clearly remains unchanged when we replace $\mathcal{U}$ by its closure. Thus, all we need is to verify that in the case of concave uncertainty the robust counterpart remains unchanged when the uncertainty set is extended to its convex hull. To this end, in turn, it suffices to verify that if x is robust feasible for (P) and $\zeta \in \mathrm{Conv}\,\mathcal{U}$, then $F(x,\zeta) \geq_{\mathbf{K}} 0$, which is evident: ζ is a convex combination $\sum_{i=1}^k \lambda_i \zeta_i$ of data vectors from $\mathcal{U}$, and due to concavity of F w.r.t. data,

$$[0 \leq_{\mathbf{K}}] \quad \sum_{i=1}^{k} \lambda_i F(x,\zeta_i) \leq_{\mathbf{K}} F(x,\zeta). \qquad \blacksquare$$

In view of the above considerations, from now on we make the following

> **Default Assumption:** *All uncertain problems under consideration satisfy Assumption* **A**, *and the uncertainty is concave. The uncertainty sets $\mathcal{U}$ underlying the problems are closed and convex.*

Let us also make the following

> **Default Notational Convention:** *Indices of all matrices and vectors are always superscripts; $A_i^{\cdots}$ denotes the i-th row of a matrix $A^{\cdots}$ regarded as a row vector.*

2.2 Solvability properties of the robust counterpart

We start with the following direct consequence of our Default Assumption:

Proposition 2.2 *The feasible set of the robust counterpart (P^*) is a closed convex set, so that the robust counterpart of an uncertain convex program is a convex program. If the robust counterpart is feasible, then all instances of the uncertain program are feasible, and the robust optimal value of the program is greater or equal to the optimal values of all the instances.*

It is easily seen that in the most general case there could be a substantial "gap" between the solvability properties of the instances of an uncertain program and those of the robust counterpart of the program; e.g., it may happen that all instances of a given uncertain program are feasible, while the robust counterpart is not, or that the optimal value of the robust counterpart is "essentially worse" than the optimal values of all instances (see e.g. the example in Ben-Tal and Nemirovski (1995b)). There is, however, a rather general case when the above "gap" vanishes – this is the case of *constraint-wise affine* uncertainty.

The notion of a constraint-wise uncertainty relates to the particular case of an uncertain convex program (P) associated with $\mathbf{K} = \mathbf{R}^m_+$, where it indeed makes sense to speak about separate scalar constraints

$$f_i(x,\zeta) \geq 0,\ i = 1, ..., m \qquad \left[F(x,\zeta) = \begin{pmatrix} f_1(x,\zeta) \\ ... \\ f_m(x,\zeta) \end{pmatrix} \right]$$

Note that the robust counterpart of (P) in this case is the program

$$(P^*): \quad \min \left\{ c^T x : x \in \mathcal{X}, f_i(x,\zeta) \geq 0 \ \forall i = 1, ..., m \ \forall \zeta \in \mathcal{U} \right\}.$$

By construction, the robust counterpart remains unchanged if instead thinking of a common data vector ζ for all constraints we think about a separate data vector for every constraint, i.e., instead of the original mapping

$$F(x,\zeta) = \begin{pmatrix} f_1(x,\zeta) \\ f_2(x,\zeta) \\ ... \\ f_m(x,\zeta) \end{pmatrix} : \mathcal{X} \times \mathcal{A} \to \mathbf{R}^m$$

and the original uncertainty set $\mathcal{U}$, we consider the mapping

$$\hat{F}(x,\zeta_1,...,\zeta_m) = \begin{pmatrix} f_1(x,\zeta_1) \\ f_2(x,\zeta_2) \\ ... \\ f_m(x,\zeta_m) \end{pmatrix} : \mathcal{X} \times \mathcal{A} \times ... \times \mathcal{A} \to \mathbf{R}^m$$

and the uncertainty set

$$\hat{\mathcal{U}} = \underbrace{\mathcal{U} \times ... \times \mathcal{U}}_{m}.$$

The constraint-wise uncertainty is exactly the one of the latter type: the i-th component of F depends on x and on the i-th portion ζ_i of the data, and the uncertainty set $\mathcal{U}$ is a direct product of closed and convex uncertainty sets $\mathcal{U}_i$ in the spaces of ζ_i's, $i = 1, ..., m$.

Theorem 2.1 *Assume that (P) is an uncertain problem with constraint-wise affine uncertainty and that the set $\mathcal{X}$ is compact. Then the robust counterpart (P^*) of the problem is feasible if and only if all instances of the problem are feasible, and in this case the robust optimal value is the supremum of those of the instances.*

Proof. In view of Proposition 2.2 it suffices to prove that if all instances of (P) are feasible, then (i) (P^*) is feasible, and (ii) the optimal value of (P^*) is the supremum of optimal values of the instances.

To prove (i), assume, on the contrary, that (P^*) is infeasible. Then the family of sets

$$\{X_i(\zeta_i) \equiv \{x \in X \mid f_i(x, \zeta_i) \geq 0\}\}_{i=1,...,m,\zeta_i \in \mathcal{U}_i}$$

has empty intersection. Since all these sets are closed subsets of a compact set $\mathcal{X}$, it follows that some finite subfamily of the family has empty intersection, so that there exists a collection $\zeta_{i,j} \in \mathcal{U}_i$, $i = 1, ..., m$, $j = 1, ..., M$ such that

$$\bigcap_{i \leq m, j \leq M} X_i(\zeta_{i,j}) = \emptyset,$$

or, which is the same,

$$\max_{x \in \mathcal{X}} \min_{i \leq m, j \leq M} f_i(x, \zeta_{i,j}) < 0.$$

Since all the functions $f_i(\cdot, \zeta_{i,j})$ are concave and continuous on convex compact set $\mathcal{X}$, it follows that there exists a convex combination

$$f(x) = \sum_{i \leq m, j \leq M} \lambda_{i,j} f_i(x, \zeta_{i,j})$$

of the functions $f_i(\cdot, \zeta_{i,j})$ which is strictly negative on $\mathcal{X}$. Let us set

$$\begin{array}{rcl} \lambda_i & = & \sum_{j=1}^{M} \lambda_{i,j}, \\ \zeta_i & = & \lambda_i^{-1} \sum_{j=1}^{m} \lambda_{i,j} \zeta_{i,j} \end{array}$$

(in the case of $\lambda_i = 0$, ζ_i is a fixed in advance point from $\mathcal{U}_i$). Since $f_i(x, \cdot)$ is affine in ζ_i, we clearly have

$$\lambda_i f_i(x, \zeta_i) = \sum_{j=1}^{M} \lambda_{i,j} f_i(x, \zeta_{i,j}),$$

so that

$$\sum_{i=1}^{m} \lambda_i f_i(x, \zeta_i) < 0 \quad \forall x \in \mathcal{X}. \tag{5}$$

Since the uncertainty is constraint-wise and the sets $\mathcal{U}_i$ are convex, the point $\zeta = (\zeta_1, ..., \zeta_m)$ belongs to $\mathcal{U}$; the instance corresponding to the data ζ is infeasible in view of (5), which is a contradiction.

To prove (ii), note that since $\mathcal{X}$ is bounded, the supremum (denoted by c^*) of optimal values of the instances is finite. Now let us add to all instances of the original uncertain problem (P) the additional "certain" constraint $c^T x \leq c^*$. Applying (i) to the resulting problem, we conclude that its robust counterpart is feasible, and hence the optimal value in (P^*) is at most (and in view of Proposition 2.2 – also at least) c^*, as required in (ii). ■

2.3 Elementary sensitivity analysis for uncertain programs

In this section we address the following natural question: let

$$(p^0): \quad \min\left\{c^T x : x \in \mathcal{X}, F(x,\zeta^0) \in \mathbf{K}\right\} \tag{6}$$

be a "nominal" instance of an uncertain convex program (P) with uncertainty set

$$\mathcal{U} = \zeta^0 + \mathcal{V}:$$

$$(P) = \left\{\min\left\{c^T x : x \in \mathcal{X},\ F(x,\zeta) \in \mathbf{K}\right\}\right\}_{\zeta\in\zeta^0+\mathcal{V}}$$

Assume that (p^0) is solvable, and let G^0 be the feasible set, x_0^* be an optimal solution, and c_0^* be the optimal value of this instance.

Let

$$(P^*): \quad \min\left\{c^T x : x \in \mathcal{X},\ F(x,\zeta) \in \mathbf{K}\ \ \forall \zeta \in \zeta^0 + \mathcal{V}\right\}.$$

be the robust counterpart of (P). What can be said about the proximity of the optimal value of the nominal instance to that of the robust counterpart (P^*)?

We are going to give a partial answer in the case of the mapping $F(x,\zeta)$ being affine in x (and, as always, being $\mathbf{K}$-concave in ζ). Even in this particular case the answer clearly should impose appropriate restrictions on the nominal instance – in the general case it may well happen that the robust counterpart is infeasible. We now define a quantity – the *feasibility margin* of the nominal instance – in terms of which we will express the closeness of the nominal and the robust optimal values.

Definition 2.3 *For $x \in G^0$, the feasibility margin $\lambda_{\mathcal{V}}(x)$ of x with respect to $\mathcal{V}$ is*

$$\lambda_{\mathcal{V}}(x) = \sup\{\lambda \geq 0 : \forall(\delta \in \mathcal{V}) : [\zeta^0 \pm \lambda\delta \in \mathcal{A}]\ \ \&\ \ [F(x,\zeta^0 \pm \lambda\delta) \in \mathbf{K}]\}.$$

We start with the following

Lemma 2.1 *Let $\lambda > 1$, $\rho \in (0,1)$ and $\widehat{x} \in G^0$ be such that every point x from the set*

$$G^0_{\rho,\widehat{x}} = (1-\rho)\widehat{x} + \rho G^0$$

satisfies the relation

$$\lambda_{\mathcal{V}}(x) \geq \lambda. \tag{7}$$

Let also the mapping $F(x,\zeta)$ be affine in x. Then the optimal value $c^*_{\mathcal{U}}$ of (P^*) satisfies the inequality

$$c^*_{\mathcal{U}} \leq c_0^* + \frac{2(1-\rho)}{\rho\lambda + 2 - \rho}\left[c^T\widehat{x} - c_0^*\right], \tag{8}$$

Proof. It suffices to prove that the point

$$\begin{array}{rcl} y & = & \widehat{x} + \omega u, \\ u & = & x_0^* - \widehat{x}, \\ \omega & = & 1 - \frac{2(1-\rho)}{\rho\lambda+2-\rho} \\ & = & \frac{\rho(\lambda+1)}{\rho\lambda+2-\rho} \end{array}$$

is robust feasible (note that the right hand side in (8) is exactly the value of the objective at y).

Let $\delta \in \mathcal{V}$. We have the following inclusions:

$$F(\widehat{x}, \zeta^0 - \lambda\delta) \in \mathbf{K}, \tag{9}$$
$$F(\widehat{x} + \rho u, \zeta^0 + \lambda\delta) \in \mathbf{K}, \tag{10}$$
$$F(\widehat{x} + u, \zeta^0) \in \mathbf{K}. \tag{11}$$

Indeed, (9) comes from $\lambda_{\mathcal{V}}(\widehat{x}) \geq \lambda$, (10) comes from $\lambda_{\mathcal{V}}(\widehat{x}+\rho u) \geq \lambda$, and (11) comes from $F(\widehat{x}+u, \zeta^0) = F(x_0^*, \zeta^0) \in \mathbf{K}$.

Note that $\omega > \rho$ due to $\rho \in (0,1), \lambda > 1$. Now let

$$\begin{array}{rcl} \mu & = & \frac{1-\omega}{2\omega} > 0, \\ \nu & = & \frac{1-\omega}{2(\omega-\rho)} > 0. \end{array}$$

In view of (9) − (11), for every $\delta \in \mathcal{V}$ we have

$$\begin{array}{rcl}
0 & \leq_{\mathbf{K}} & \nu F(\widehat{x} + \rho u, \zeta^0 + \lambda\delta) + \mu F(\widehat{x}, \zeta^0 - \lambda\delta) + F(\widehat{x} + u, \zeta^0) \\
& = & \left[\nu \frac{\rho}{\omega} F(\widehat{x} + \omega u, \zeta^0 + \lambda\delta) + \nu \frac{\omega-\rho}{\omega} F(\widehat{x}, \zeta^0 + \lambda\delta)\right] \\
& & +\mu F(\widehat{x}, \zeta^0 - \lambda\delta) + F(\widehat{x} + u, \zeta^0) \\
& & [\text{since } F(x,\zeta) \text{ is affine in } x] \\
& = & \nu \frac{\rho}{\omega} F(y, \zeta^0 + \lambda\delta) \\
& & +\mu \left[F(\widehat{x}, \zeta^0 + \lambda\delta) + F(\widehat{x}, \zeta^0 - \lambda\delta)\right] \\
& & +F(\widehat{x} + u, \zeta^0) \\
& & [\text{since } \nu \frac{\omega-\rho}{\omega} = \mu] \\
& \leq_{\mathbf{K}} & \nu \frac{\rho}{\omega} F(y, \zeta^0 + \lambda\delta) \\
& & +2\mu F(\widehat{x}, \zeta^0) + F(\widehat{x} + u, \zeta^0) \\
& & [\text{since } F(x,\zeta)\ \mathbf{K}\text{-concave in } \zeta] \\
& = & (2\mu+1) F(\widehat{x} + \frac{1}{2\mu+1} u, \zeta^0) \\
& & +\nu \frac{\rho}{\omega} F(y, \zeta^0 + \lambda\delta) \\
& & [\text{since } F(x,\zeta) \text{ is affine in } x] \\
& = & (2\mu+1) F(y, \zeta^0) + \nu \frac{\rho}{\omega} F(y, \zeta^0 + \lambda\delta) \\
& & [\text{since } \frac{1}{2\mu+1} = \omega] \\
& \leq_{\mathbf{K}} & \left(2\mu + 1 + \nu \frac{\rho}{\omega}\right) F(y, \zeta^0 + \gamma\delta),
\end{array}$$

where

$$\gamma = \lambda\nu \frac{\rho}{\omega} \frac{1}{2\mu + 1 + \nu\frac{\rho}{\omega}}$$

(the last inequality in the chain is given by the fact that $F(x,\zeta)$ is $\mathbf{K}$-concave in ζ).

From the definition of ω it immediately follows that $\gamma = 1$, and we conclude that $F(y, \zeta^0 + \delta) \in \mathbf{K}$ whenever $\delta \in \mathcal{V}$. Thus, y is robust feasible. ∎

The result of Lemma 2.1 motivates the following

Definition 2.4 *We say that a pair (λ, ρ) with $\lambda > 1$, $\rho > 0$, is admissible, if there exists $\widehat{x}$ satisfying, for these λ and ρ, the premise in Lemma 2.1. The upper bound of the quantities $\lambda\rho/(1-\rho)$ taken over all admissible pairs is called the* feasibility margin *$\kappa_{\mathcal{V}}(p^0)$ of the nominal instance (p^0) with respect to the perturbation set $\mathcal{V}$.*

Lemma 2.1 clearly implies the following

Proposition 2.3 *Let (p^0) possess positive feasibility margin with respect to the perturbation set $\mathcal{V}$, let the function $F(x,\zeta)$ be affine in x and let the feasible set G^0 of (p^0) be bounded. Then the robust counterpart (P^*) of (P) is solvable, and the robust optimal value $c^*_{\mathcal{V}}$ satisfies the inequality*

$$c^*_{\mathcal{V}} \le c^*_0 + \frac{2}{1+\kappa_{\mathcal{V}}(p^0)}\mathrm{Var}_{G_0}(c), \quad \mathrm{Var}_G(c) = \max_{x\in G} c^T x - \min_{x\in G} c^T x. \tag{12}$$

Perhaps a better interpretation of the result stated by Proposition 2.3 is as follows. Let $\mathcal{W}$ be a symmetric set (with respect to the origin) of "unit" perturbations of the data. Consider the situation where the actual perturbations in the data are "at most of level π", so that the associated perturbation set is

$$\mathcal{V}_\pi = \pi\mathcal{W}.$$

What happens with the robust counterpart as $\pi \to +0$? Proposition 2.3 says the following. Assume that the nominal instance (p^0) has a bounded feasible set and is "strictly feasible" with respect to the set of perturbations $\mathcal{W}$: there exist $\pi^* > 0$ and $\rho^* > 0$ such that the set

$$\{x \mid [\zeta^0 + \pi^*\delta \in \mathcal{A}] \ \& \ [F(x,\zeta^0+\pi^*\delta) \in \mathbf{K}] \quad \forall \delta \in \mathcal{W}\}$$

contains a ρ^*-dilatation of the nominal feasible set G^0. Then the robust counterpart (P^*_π) of the uncertain version of (p^0) associated with perturbations of the level $\pi < \pi^*$ – i.e., with the uncertainty set $\mathcal{U} = \zeta^0 + \pi\mathcal{W}$ – is solvable, and the corresponding robust optimal value c^*_π differs from the nominal one c^*_0 by at most $O(\pi)$:

$$c^*_\pi \le c^*_0 + \frac{2\pi(1-\rho^*)}{\pi(1-\rho^*)+\pi^*\rho^*}\mathrm{Var}_{G^0}(c). \tag{13}$$

The required pair π^*, ρ^* surely exists if G^0 and the set $\mathcal{W}$ of "unit perturbations" are bounded and the problem is strictly feasible (i.e., $F(\widehat{x},\zeta^0) \in \mathrm{int}\,\mathbf{K}$ for some $\widehat{x} \in \mathcal{X}$).

In the case where the nominal program is not strictly feasible, c^*_π may happen to be a very bad function of π, as it is seen from the following one-dimensional example where the nominal problem is

$$\begin{array}{lrcl} & \min\ x & & \\ \text{s.t.} & & & \\ (*): & \begin{pmatrix}0\\1\end{pmatrix} x + \begin{pmatrix}0\\1\end{pmatrix} & \ge & 0. \end{array}$$

Here the nominal optimal value is -1. Now, if our "unit perturbations" affect only the coefficients of x in the left hand side of $(*)$, and the corresponding set $\mathcal{W}$ is a central-symmetric convex set in $\mathbf{R}^2$ with a nonempty interior, then the robust counterpart (P^*_π) has, for every positive π, the singleton feasible set $\{0\}$, so that $c^*_\pi = 0$ whenever $\pi > 0$.

3 Robust counterparts of some generic convex programs

In this section we derive the robust counterparts of some important generic uncertain convex programs. Of course, the possibility of obtaining an explicit form of the robust counterpart depends not only on the structural elements of the uncertain programs in question, but also on the choice of the uncertainty sets. We intend to focus on the *ellipsoidal uncertainties* – those represented by ellipsoids or intersections of finitely many ellipsoids; this choice is motivated primarily by the desire to deal with "mathematically convenient" uncertainty sets.

Ellipsoids and ellipsoidal uncertainties. In geometry, a K-dimensional ellipsoid in $\mathbf{R}^K$ can be defined as an image of a K-dimensional Euclidean ball under a one-to-one affine mapping from $\mathbf{R}^K$ to $\mathbf{R}^K$. For our purposes this definition is not general enough. On one hand, we would like to consider "flat" ellipsoids in the space $E = \mathbf{R}^N$ of data, i.e., the usual ellipsoids in proper affine subspaces of E (such an ellipsoid corresponds to the case when we deal with "partial uncertainty": some of the entries in the data vector satisfy a number of given linear equations, e.g., some of the entries in ζ are "certain"). On the other hand, we want to incorporate also "ellipsoidal cylinders", i.e., the sets of the type "sum of a flat ellipsoid and a linear subspace". These sets occur when we impose on ζ several ellipsoidal restrictions, each of them dealing with part of the entries. For example, an "interval" uncertainty set ($\mathcal{U}$ is given by upper and lower bounds on the entries of the data vector) clearly is an intersection of M ($= \dim \zeta$) ellipsoidal cylinders. In order to cover all these cases, we define in this paper an ellipsoid in $\mathbf{R}^K$ as a set of the form

$$U \equiv U(\Lambda, \Pi) = \Pi(B) + \Lambda, \tag{14}$$

where $u \to \Pi(u)$ is an affine embedding of $\mathbf{R}^L$ into $\mathbf{R}^K$, $B = \{u \in \mathbf{R}^L \mid ||u||_2 \le 1\}$ is the unit Euclidean ball in $\mathbf{R}^L$ and Λ is a linear subspace in $\mathbf{R}^K$; here and in what follows $||\cdot||_2$ is the standard Euclidean norm. This definition covers all cases mentioned above: when $\Lambda = \{0\}$ and $L = K$, we obtain the standard K-dimensional ellipsoids in $\mathbf{R}^K$; "flat" ellipsoids correspond to the case when $\Lambda = 0$ and $L < K$; and the case of nontrivial Λ corresponds to ellipsoidal cylinders of different types.

From now on we say that $\mathcal{U} \in \mathbf{R}^N$ is an *ellipsoidal uncertainty*, if $\mathcal{U}$ is an explicitly given ellipsoid in the above sense, and we say that $\mathcal{U}$ is an $\bigcap$-*ellipsoidal uncertainty*, if

- A. $\mathcal{U}$ is given as an intersection of finitely many ellipsoids:

$$\mathcal{U} = \bigcap_{l=0}^{k} U(\Lambda_l, \Pi_l) \tag{15}$$

  with explicitly given Λ_l and Π_l;

- B. $\mathcal{U}$ is bounded;

- C. ["Slater condition"] there is at least one vector $\zeta \in \mathcal{U}$ which belongs to the relative interior of all ellipsoids U_l:

$$\forall l \le k \quad \exists u_l : \quad \zeta \in \Pi_l(u^l) + \Lambda_l \;\; \& \;\; ||u_l||_2 < 1.$$

3.1 Robust linear programming

Uncertain linear programming problems are of primary interest in the methodology we are developing. This is so because of the wide-spread practical use of linear programming and the quality of implementable results we are able to obtain.

An uncertain LP problem is

$$(P) = \left\{ \min \left\{ c^T x : Ax + b \ge 0 \right\} \right\}_{[A;b] \in \mathcal{U}};$$

this is a particular case of (4) corresponding to

$$\mathbf{K} = \mathbf{R}^m_+, \zeta = [A; b] \in \mathcal{A} = \mathbf{R}^{m \times (n+1)}, \mathcal{X} = \mathbf{R}^n, F(x, \zeta) = Ax + b.$$

Robust linear programming was already considered in Ben-Tal and Nemirovski (1995a); the main result of that paper is

Theorem 3.1 [Ben-Tal and Nemirovski (1995a), Theorem 3.1] *The robust counterpart* (P^*) *of an uncertain linear programming problem with ellipsoidal or* $\bigcap$*-ellipsoidal uncertainty is a conic quadratic program with the data given explicitly in terms of structural parameters of LP and the parameters defining the uncertainty ellipsoids.*

The conic quadratic program mentioned in Theorem 3.1 is as follows.

Case of ellipsoidal uncertainty set:

$$\mathcal{U} = \{[A; b] = [A^0; b^0] + \sum_{j=1}^{k} u_j [A^j; b^j] + \sum_{p=1}^{q} v_p [C^p; d^p] \mid u^T u \leq 1\}.$$

For this case the robust counterpart of (P) is the program

$$\begin{array}{lcl}
(P^*) & & \min_x \ c^T x \\
\text{w.r.t. } x \in \mathbf{R}^n \text{ subject to} & & \\
\quad C_i^p x + d_i^p & = & 0, \ p = 1, ..., q \\
\quad A_i^0 x + b_i^0 & \geq & \sqrt{\sum_{j=1}^{k} (A_i^j x + b_i^j)^2}, \\
& & i = 1, 2, ..., m.
\end{array}$$

Case of $\bigcap$-ellipsoidal uncertainty:

$$\begin{array}{rcl}
\mathcal{U} & = & \bigcap_{s=0}^{t} \mathcal{U}_s, \\
\mathcal{U}_s & = & \{[A; b] = [A^{s0}; b^{s0}] + \sum_{j=1}^{k_s} u_j [A^{sj}; b^{sj}] + \sum_{p=1}^{q_s} v_p [C^{sp}; d^{sp}] \mid \\
& & u^T u \leq 1\}, \\
& & s = 0, 1, ..., t.
\end{array}$$

For this case the robust counterpart of (P) is the program

$$(\mathcal{P}) \qquad \min \ c^T x$$

w.r.t. vectors $x \in \mathbf{R}^n$, $\mu^{is} \in \mathbf{R}^m$ and matrices $\lambda^{is} \in \mathbf{R}^{m\times n}$, $i = 1, ..., m, s = 1, ..., t$ subject to

$$
\begin{array}{lrcl}
(a): & \mathrm{Tr}([C^{sp}]^T\lambda^{is}) + [d^{sp}]^T\mu^{is} & = & 0,\ 1 \le i \le m, 1 \le s \le t, 1 \le p \le q_s;\\[2ex]
(b): & \sum\limits_{s=1}^{t}\Big(\mathrm{Tr}([C^{0p}]^T\lambda^{is}) + [d^{0p}]^T\mu^{is}\Big) & = & C_i^{0p}x + d_i^{0p},\ 1 \le i \le m,\ 1 \le p \le q_0;\\[2ex]
(c): & \multicolumn{3}{l}{\sum\limits_{s=1}^{t}\Big\{\mathrm{Tr}([A^{s0} - A^{00}]^T)\lambda^{is} + [b^{s0} - b^{00}]^T\mu^{is}\Big\} + A_i^{00}x + b_i^{00}}
\end{array}
$$

$$
\ge \left\| \begin{pmatrix} A_i^{01}x + b_i^{01} - \sum\limits_{s=1}^{t}\Big\{\mathrm{Tr}([A^{01}]^T\lambda^{is}) + [b^{01}]^T\mu^{is}\Big\} \\ A_i^{02}x + b_i^{02} - \sum\limits_{s=1}^{t}\Big\{\mathrm{Tr}([A^{02}]^T\lambda^{is}) + [b^{02}]^T\mu^{is}\Big\} \\ \cdots \\ A_i^{0k_0}x + b_i^{0k_0} - \sum\limits_{s=1}^{t}\Big\{\mathrm{Tr}([A^{0k_0}]^T\lambda^{is}) + [b^{0k_0}]^T\mu^{is}\Big\} \end{pmatrix} \right\|_2
$$

$$
+ \sum_{s=1}^{t} \left\| \begin{pmatrix} \mathrm{Tr}([A^{s1}]^T\lambda^{is}) + [b^{01}]^T\mu^{is} \\ \mathrm{Tr}([A^{s2}]^T\lambda^{is}) + [b^{02}]^T\mu^{is} \\ \cdots \\ \mathrm{Tr}([A^{sk_s}]^T\lambda^{is}) + [b^{0k_s}]^T\mu^{is} \end{pmatrix} \right\|_2, \quad i = 1, ..., m.
$$

Note that (a) and (b) are linear equality constraints, while the constraints (c) can be easily converted to conic quadratic constraints.

3.2 Robust quadratic programming

In this section we are interested in the analytical structure of the robust counterpart (P^*) of an uncertain Quadratically Constrained Convex Quadratic (QCQP) program.

An uncertain QCQP problem is

$$
\begin{array}{rcl}
(P) & = & \{\min\{c^Tx : -x^T[A^i]^TA^ix + 2[b^i]^Tx + \gamma^i \ge 0, \\
& & i = 1, 2, ..., m\}\}_{(A^1,b^1,\gamma^1;A^2,b^2,\gamma^2;...;A^m,b^m,\gamma^m)\in\mathcal{U}},
\end{array}
$$

where A^i are $l_i \times n$ matrices. This is the particular case of (4) corresponding to

$$
\begin{array}{c}
\mathbf{K} = \mathbf{R}^m_+, \zeta = (A^1, b^1, \gamma^1; A^2, b^2, \gamma^2; ...; A^m, b^m, \gamma^m) \in \mathcal{A} = \{A^i \in \mathbf{R}^{l_i\times n}, b^i \in \mathbf{R}^n, \gamma^i \in \mathbf{R}\},\\
\mathcal{X} = \mathbf{R}^n, F(x,\zeta) = \begin{pmatrix} -x^T[A^1]^TA^1x + 2[b^1]^Tx + \gamma^1 \\ \cdots \\ -x^T[A^m]^TA^mx + 2[b^m]^Tx + \gamma^m \end{pmatrix}.
\end{array}
$$

3.2.1 Case of ellipsoidal uncertainty

Here we demonstrate that in the case of *bounded* ellipsoidal uncertainty the robust counterpart to (P) is an explicit SDP (Semidefinite Programming) problem.

If the uncertainty set is a bounded ellipsoid, so is the projection $\mathcal{U}_i$ of $\mathcal{U}$ on the space of data of i-th constraint,

$$\mathcal{U}_i = \left\{ (A^i, b^i, \gamma^i) = (A^{i0}, b^{i0}, \gamma^{i0}) + \sum_{j=1}^{k} u_j (A^{ij}, b^{ij}, \gamma^{ij}) \mid u^T u \leq 1 \right\}, \ i = 1, ..., m; \tag{16}$$

note that the robust counterpart of the problem is the same as in the case of constraint-wise ellipsoidal uncertainty

$$\mathcal{U} = \mathcal{U}_1 \times ... \times \mathcal{U}_m. \tag{17}$$

Theorem 3.2 *The robust counterpart of* (P) *associated with the uncertainty set* $\mathcal{U}$ *given by* (16) – (17) *is equivalent to the SDP problem*

$$(\mathcal{P}) \qquad \min \ c^T x$$

w.r.t. $x \in \mathbf{R}^n, \lambda^1, ..., \lambda^m \in \mathbf{R}$ *subject to*

$$(a_i) \quad \left(\begin{array}{c|cccc|c} \gamma^{i0} + 2x^T b^{i0} - \lambda^i & \frac{\gamma^{i1}}{2} + x^T b^{i1} & \frac{\gamma^{i2}}{2} + x^T b^{i2} & \cdots & \frac{\gamma^{ik}}{2} + x^T b^{ik} & [A^{i0}x]^T \\ \hline \frac{\gamma^{i1}}{2} + x^T b^{i1} & \lambda^i & & & & [A^{i1}x]^T \\ \frac{\gamma^{i2}}{2} + x^T b^{i2} & & \lambda^i & & & [A^{i2}x]^T \\ \vdots & & & \ddots & & \cdots \\ \frac{\gamma^{ik}}{2} + x^T b^{ik} & & & & \lambda^i & [A^{ik}x]^T \\ \hline A^{i0}x & A^{i1}x & A^{i2}x & \cdots & A^{ik}x & I_{l_i} \end{array} \right) \geq 0,$$

$$i = 1, 2, ..., m,$$

I_l *being the unit* $l \times l$ *matrix.*

From now on, inequality $A \geq B$ *for matrices* A, B *means that* A, B *are symmetric and* $A - B$ *is positive semidefinite.*

Proof. It is clear that the robust counterpart of (P) is the problem of minimizing $c^T x$ under the constraints

$$(C_i): \quad -x^T A^T A x + 2b^T x + \gamma \leq 0 \quad \forall (A, b, \gamma) \in \mathcal{U}_i, \ i = 1, ..., m.$$

To establish the theorem, it suffices to prove that (C_i) is equivalent to existence of a real λ^i such that (x, λ^i) satisfies the constraint (a_i).

Let us fix i. We have

$$\begin{array}{rrl} (C_i): & -x^T [A^{i0} + \sum_{j=1}^k u_j A^{ij}]^T [A^{i0} + \sum_{j=1}^k u_j A^{ij}] x & \\ & +2[b^{i0} + \sum_{j=1}^k u_j b^{ij}]^T x + [\gamma^{i0} + \sum_{j=1}^k u_j \gamma^{ij}] & \geq \ 0 \\ & \forall (u : \|u\|_2 \leq 1). & \end{array}$$

$$\Updownarrow$$

$$\begin{array}{rrl} (C_i'): & -x^T [A^{i0}\tau + \sum_{j=1}^k u_j A^{ij}]^T [A^{i0}\tau + \sum_{j=1}^k u_j A^{ij}] x & \\ & +2\tau[b^{i0}\tau + \sum_{j=1}^k u_j b^{ij}]^T x + \tau[\gamma^{i0}\tau + \sum_{j=1}^k u_j \gamma^{ij}] & \geq \ 0 \\ & \forall ((\tau, u) : \|u\|_2 \leq \tau). & \end{array}$$

$$\Updownarrow$$

$$\begin{array}{rrl} (C_i''): & -x^T [A^{i0}\tau + \sum_{j=1}^k u_j A^{ij}]^T [A^{i0}\tau + \sum_{j=1}^k u_j A^{ij}] x & \\ & +2\tau[b^{i0}\tau + \sum_{j=1}^k u_j b^{ij}]^T x + \tau[\gamma^{i0}\tau + \sum_{j=1}^k u_j \gamma^{ij}] & \geq \ 0 \\ & \forall ((\tau, u) : \|u\|_2^2 \leq \tau^2), & \end{array}$$

the last equivalence is a consequence of the left hand side in (C_i') being an even function of (τ, u). We see that (C_i) is equivalent to the following implication:

$$(D_i): \quad P(\tau, u) \equiv \tau^2 - u^T u \geq 0 \Rightarrow Q_i^x(\tau, u) \geq 0,$$

$Q_i^x(\tau, u)$ being the homogeneous quadratic form of (τ, u) in the left hand side of (C_i'').

Now let us make use of the following well-known fact (see, e.g., Boyd, El Ghaoui, Feron and Balakrishnan (1994)):

Lemma 3.1 *Let P, Q be two symmetric matrices such that there exists z_0 satisfying $z_0^T P z_0 > 0$. Then the implication*

$$z^T P z \geq 0 \Rightarrow z^T Q z \geq 0$$

holds true if and only if

$$\exists \lambda \geq 0 : Q \geq \lambda P.$$

In view of Lemma 3.1, (D_i) is equivalent to the existence of $\lambda^i \geq 0$ such that the quadratic form of (τ, u)

$$Q_i^x(\tau, u) - \lambda^i(\tau^2 - u^T u)$$

is positive semidefinite. It is immediately seen that this quadratic form is of the type

$$R_{\lambda^i}(\tau, u) = \begin{pmatrix} \tau & u^T \end{pmatrix} [E^i(x) - [F^i]^T(x) F^i(x)] \begin{pmatrix} \tau \\ u \end{pmatrix} - \lambda^i(\tau^2 - u^T u),$$

with symmetric $(k+1) \times (k+1)$ matrix $E^i(x)$ and rectangular $l_i \times (k+1)$-matrix $F^i(x)$ affinely depending on x. By standard reasoning via Schur complement such a quadratic form is positive semidefinite if and only if the matrix

$$\begin{pmatrix} E^i(x) + \begin{pmatrix} -\lambda^i & \\ & \lambda^i I_k \end{pmatrix} & [F^i]^T(x) \\ F^i(x) & I_{l_i} \end{pmatrix}$$

is positive semidefinite. The latter property of (x, λ^i), along with the nonnegativity of λ^i is exactly the property expressed by the constraint (a_i). ■

3.2.2 Intractability of (P^*) in the case of general $\bigcap$-ellipsoidal uncertainty

It turns out that an uncertainty set which is the intersection of general-type ellipsoids leads to a computationally intractable robust counterpart. To see it, consider the particular case of (P) where $m = 1$ and where the uncertainty set is

$$\mathcal{U} = \{\zeta = (A^1, b^1, \gamma^1) \mid A^1 = \mathrm{Diag}(a); b^1 = 0; \gamma^1 = 1\}_{a \in B},$$

B being a parallelotope in $\mathbf{R}^n$ centered at the origin. Note that since a parallelotope in $\mathbf{R}^n$ is intersection of n stripes, i.e., n simplest "elliptic cylinders", the uncertainty set $\mathcal{U}$ indeed is $\bigcap$-ellipsoidal .

Now, for the case in question, to check the robust feasibility of the particular point $x = (1, ..., 1)^T \in \mathbf{R}^n$ is the same as to verify that $||a||_2 \leq 1$ for all $a \in B$, i.e., to verify that the parallelotope B is contained in the unit Euclidean ball. It is known that the latter problem is NP-hard.

3.3 Robust conic quadratic programming

Now consider the case of uncertain Conic Quadratic Programming (ConeQP) problem.

An uncertain ConeQP problem is

$$\begin{array}{rcl}(P) & = & \{\min\{c^T x : \|A^i x + b^i\|_2 \le [d^i]^T x + \gamma^i, \\ & & i = 1,2,...,m\}\}_{(A^1,b^1,d^1,\gamma^1;...;A^m,b^m,d^m,\gamma^m)\in\mathcal{U}},\end{array}$$

A^i being $l_i \times n$ matrices. This is a particular case of (4) corresponding to

$$\mathbf{K} = \prod_{i=1}^{m} \mathbf{K}_i,\ \mathbf{K}_i = \left\{(t,u) \in \mathbf{R} \times \mathbf{R}^{l_i} \mid t \ge \|u\|_2\right\},$$
$$\zeta = (A^1, b^1, d^1, \gamma^1; ...; A^m, b^m, d^m, \gamma^m) \in \mathcal{A} = \left\{A^i \in \mathbf{R}^{l_i \times n}, b^i \in \mathbf{R}^{l_i}, d^i \in \mathbf{R}^n, \gamma^i \in \mathbf{R}, i = 1,...,m\right\},$$
$$\mathcal{X} = \mathbf{R}^n, F(x,\zeta) = \left(\begin{array}{c} \begin{pmatrix} [d^1]^T x + \gamma^1 \\ A^1 x + b^1 \end{pmatrix} \\ \cdots \\ \begin{pmatrix} [d^m]^T x + \gamma^m \\ A^m x + b^m \end{pmatrix} \end{array}\right).$$

Here, as in the case of QCQP, a general-type $\bigcap$-ellipsoidal uncertainty leads to computationally intractable robust counterpart. Let us then focus on the following particular type of uncertainty:

"Simple" ellipsoidal uncertainty. Assume that

- (I) The uncertainty is "constraint-wise":

  $$\mathcal{U} = \mathcal{U}_1 \times ... \times \mathcal{U}_m,$$

  where $\mathcal{U}_i$ is an uncertainty set associated with the i-th conic quadratic constraint; specifically,

- (II) For every i, $\mathcal{U}_i$ is the direct product of two ellipsoids in the spaces of (A^i, b^i)- and (d^i, γ^i)-components of the data:

  $$\begin{array}{rcl}\mathcal{U}_i & = & \mathcal{V}_i \times \mathcal{W}_i, \\ \mathcal{V}_i & = & \{[A^i; b^i] = [A^{i0}; b^{i0}] + \sum_{j=1}^{k_i} u_j [A^{ij}; b^{ij}] \\ & & + \sum_{p=1}^{q_i} v_p [E^{ip}; f^{ip}] \mid u^T u \le 1\}, \\ \mathcal{W}_i & = & \{(d^i, \gamma^i) = (d^{i0}, \gamma^{i0}) + \sum_{j=1}^{k'_i} u_j (d^{ij}, \gamma^{ij}) \\ & & + \sum_{p=1}^{q'_i} v_p (g^{ip}, h^{ip}) \mid u^T u \le 1\}.\end{array}$$

Theorem 3.3 *The robust counterpart of the uncertain conic quadratic problem* (P) *corresponding to* *the uncertainty set* $\mathcal{U}$ *described in* (I), (II) *is equivalent to the following SDP problem:*

$$
\begin{array}{ll}
(\mathcal{P}) & \min\ c^T x \\
\textit{w.r.t.} & x \in \mathbf{R}^n, \lambda^1, ..., \lambda^m \in \mathbf{R}, \mu^1, ..., \mu^m \in \mathbf{R} \textit{ subject to} \\
(\mathrm{I}_{ip}) & E^{ip}x + f^{ip} = 0,\ i = 1, ..., m, p = 1, ..., q_i; \\
(\mathrm{II}_{ip}) & [g^{ip}]^T x + h^{ip} = 0,\ i = 1, ..., m, p = 1, ..., q_i'; \\
(\mathrm{III}_i) & \begin{pmatrix} [d^{i0}]^T x + \gamma^{i0} - \lambda^i & [d^{i1}]^T x + \gamma^{i1} & [d^{i2}]^T x + \gamma^{i2} & \cdots & [d^{ik_i'}]^T x + \gamma^{ik_i'} \\ [d^{i1}]^T x + \gamma^{i1} & [d^{i0}]^T x + \gamma^{i0} - \lambda^i & & & \\ [d^{i2}]^T x + \gamma^{i2} & & [d^{i0}]^T x + \gamma^{i0} - \lambda^i & & \\ \vdots & & & \ddots & \cdots \\ [d^{ik_i'}]^T x + \gamma^{ik_i'} & & & \cdots & [d^{i0}]^T x + \gamma^{i0} - \lambda^i \end{pmatrix} \geq 0, \\
& i = 1, ..., m; \\
(\mathrm{IV}_i) & \left(\begin{array}{c|cccc|c} \lambda^i - \mu^i & & & \cdots & & [A^{i0}x + b^{i0}]^T \\ \hline & \mu^i & & & & [A^{i1}x + b^{i1}]^T \\ & & \mu^i & & & [A^{i2}x + b^{i2}]^T \\ \vdots & & & \ddots & & \cdots \\ & & & & \mu^i & [A^{ik_i}x + b^{ik_i}]^T \\ \hline A^{i0}x + b^{i0} & A^{i1}x + b^{i1} & A^{i2}x + b^{i2} & \cdots & A^{ik_i}x + b^{ik_i} & \lambda^i I_{l_i} \end{array}\right) \geq 0, \\
& i = 1, ..., m.
\end{array}
$$

Proof. In the robust counterpart of (P) every uncertain quadratic constraint

$$\left|\left|A^i x + b^i\right|\right|_2 \leq [d^i]^T x + \gamma^i$$

is replaced by its robust version

$$\left|\left|A^i x + b^i\right|\right|_2 \leq [d^i]^T x + \gamma^i \quad \forall (A^i, b^i, d^i, \gamma^i) \in \mathcal{U}_i.$$

It follows that in order to understand what is, analytically, the robust counterpart of (P) it suffices to understand what is the robust version of a single quadratic constraint of (P); to simplify notation, we drop the index i so that the robust constraint in question becomes

$$
\begin{array}{rcl}
||Ax + b||_2 & \leq & d^T x + \gamma \quad \forall (A, b) \in \mathcal{V} \quad \forall (d, \gamma) \in \mathcal{W}, \\
\mathcal{V} & = & \{[A; b] = [A^0; b^0] + \sum_{j=1}^{k} u_j [A^j; b^j] \\
& & + \sum_{p=1}^{q} v_p [E^p; f^p] \mid u^T u \leq 1\}, \\
\mathcal{W} & = & \{(d, \gamma) = (d^0, \gamma^0) + \sum_{j=1}^{k'} u_j (d^j, \gamma^j) \\
& & + \sum_{p=1}^{q'} v_p (g^p, h^p) \mid u^T u \leq 1\}.
\end{array}
\tag{18}
$$

Let us set

$$
\begin{array}{rcl}
\phi(x) & = & x^T d^0 + \gamma^0 : \mathbf{R}^n \to \mathbf{R}, \\
\Phi(x) & = & \begin{pmatrix} x^T d^1 + \gamma^1 \\ x^T d^2 + \gamma^2 \\ \cdots \\ x^T d^{k'} + \gamma^{k'} \end{pmatrix} : \mathbf{R}^n \to \mathbf{R}^{k'}, \\
\psi(x) & = & A^0 x + b^0 : \mathbf{R}^n \to \mathbf{R}^l, \\
\Psi(x) & = & [A^1 x + b^1; A^2 x + b^2; ...; A^k x + b^k] : \mathbf{R}^n \to \mathbf{R}^{l \times k}.
\end{array}
\tag{19}
$$

In the above notation the robust constraint (18) becomes the constraint

$$
\begin{array}{rcl}
\multicolumn{3}{l}{\forall (u : u^T u \leq 1, w : w^T w \leq 1,)} \\
\multicolumn{3}{l}{\forall ((g,h) \in \mathrm{Span}\{(g^p, h^p), 1 \leq p \leq q'\})} \\
\multicolumn{3}{l}{\forall ((E,f) \in \mathrm{Span}\{(E^p, f^p), 1 \leq p \leq q\})} \\
\phi(x) + \Phi^T(x)u + [g^T x + h] & \geq & ||\psi(x) + \Psi(x)w + [Ex+f]||_2\,;
\end{array}
$$

The latter constraint is equivalent to the system of constraints (I_{ip}), $p = 1,...,q$, (II_{ip}), $p = 1,...,p'$ and the following constraint:

$$
\begin{array}{rcl}
\exists \lambda : & & \\
\phi(x) + \Phi^T(x)u & \geq & \lambda \quad \forall u, u^T u \leq 1, \\
\lambda & \geq & ||\psi(x) + \Psi(x)w||_2 \quad \forall w, w^T w \leq 1.
\end{array}
\tag{20}
$$

A pair (x, λ) satisfies (20) if and only if it satisfies (III_i) along with the constraints

$$\lambda \geq 0 \tag{21}$$

and

$$\lambda^2 \geq ||\psi(x) + \Psi(x)w||_2^2 \quad \forall w, w^T w \leq 1. \tag{22}$$

The constraint (22) is equivalent to the constraint

$$\forall ((t,w), w^T w \leq t^2) : \quad \lambda^2 t^2 \geq ||\psi(x)t + \Psi(x)w||_2^2\,. \tag{23}$$

In other words, a pair (x, λ) satisfies (21), (22) if and only if λ is nonnegative, and nonnegativity of the quadratic form $(t^2 - w^T w)$ of the variables t, w implies nonnegativity of the quadratic form

$$\lambda^2 t^2 - ||\psi(x)t + \Psi(x)w||_2^2$$

of the same variables. By Lemma 3.1, the indicated property is equivalent to the existence of a nonnegative ν such that the quadratic form

$$W(t,w) = \lambda^2 t^2 - ||\psi(x)t + \Psi(x)w||_2^2 - \nu(t^2 - w^T w)$$

is positive semidefinite. We claim that ν can be represented as $\mu\lambda$ with some nonnegative μ, $\mu = 0$ in the case of $\lambda = 0$. Indeed, our claim is evident if $\lambda > 0$. In the case of $\lambda = 0$ the form $W(t,w)$ clearly can be positive semidefinite only when $\nu = 0$ (look what happens when $w = 0$), and we indeed have $\nu = \mu\lambda$ with $\mu = 0$.

We have demonstrated that a pair (x, λ) satisfies (21), (22) if and only if there exists a μ such that the triple (x, λ, μ) possesses the following property

$(\pi) : \lambda, \mu \geq 0$; $\mu = 0$ when $\lambda = 0$; the quadratic form

$$
\begin{array}{rcl}
W(t,w) & = & \lambda(\lambda - \mu)t^2 + \lambda\mu w^T w - (\,t \quad w^T\,)\, R^T(x) R(x) \begin{pmatrix} t \\ w \end{pmatrix}, \\
R(x) & = & [\psi(x); \Psi(x)],
\end{array}
$$

of t, w is positive semidefinite.

Now let us prove that the property (π) of (x, λ, μ) is equivalent to positive semidefiniteness of the matrix $S = S(x, \lambda, \mu)$ in the left hand side of (IV_i).

Indeed, if $\lambda > 0$, positive semidefiniteness of W is equivalent to positive semidefiniteness of the quadratic form

$$(\lambda - \mu)t^2 + \mu w^T w - \begin{pmatrix} t & w^T \end{pmatrix} R^T(x)(\lambda I_l)^{-1} R(x) \begin{pmatrix} t \\ w \end{pmatrix},$$

which, via Schur complement, is exactly the same as positive semidefiniteness of $S(x, \lambda, \mu)$. Of course, the matrix in the left hand side of (IV_i) can be positive semidefinite only when $\lambda \equiv \lambda^i, \mu \equiv \mu^i$ are nonnegative. Thus, for triples (x, λ, μ) with $\lambda > 0$ the property (π) indeed is equivalent to positive semidefiniteness of $S(x, \lambda, \mu)$. Now consider the case of $\lambda = 0$, and let $(x, \lambda = 0, \mu)$ possesses property (π). Due to (π), $\mu = 0$ and W is positive semidefinite, which for $\lambda = 0$ is possible if and only if $R(x) = 0$; of course, in this case $S(x, 0, 0)$ is positive semidefinite. Vice versa, if $\lambda = 0$ and $S(x, \lambda, \mu)$ is positive semidefinite, then, of course, $R(x) = 0$ and $\mu = 0$, and the triple (x, λ, μ) possesses property (π).

The summary of our equivalences is that x satisfies (18) if and only if it satisfies (I_{ip}), (II_{ip}) for all possible p and there exist $\lambda = \lambda^i, \mu = \mu^i$ such that the pair (x, λ) satisfies (III_i) and the triple (x, λ, μ) satisfies (IV_i). This is exactly the assertion of the theorem. ■

3.4 Robust semidefinite programming

Next we consider uncertain Semidefinite Programming (SDP) problems.

An uncertain SDP problem is

$$(P) = \left\{ \min \left\{ c^T x : A(x) \equiv A^0 + \sum_{i=1}^{n} x_i A^i \in \mathbf{S}^l_+ \right\} \right\}_{(A,b) \in \mathcal{U}}, \tag{24}$$

where $\mathbf{S}^l_+$ is the cone of positive semidefinite $l \times l$ matrices and $A^0, A^1, ..., A^n$ belong to the space $\mathbf{S}^l$ of symmetric $l \times l$ matrices. This is a particular case of (4) corresponding to

$$\begin{array}{c} \mathbf{K} = \mathbf{S}^l_+, \zeta = (A^0, A^1, ..., A^n) \in \mathcal{A} = (\mathbf{S}^l)^n, \\ \mathcal{X} = \mathbf{R}^n, F(x, \zeta) = A^0 + \sum_{i=1}^n x_i A^i. \end{array}$$

As before, we are interested in ellipsoidal uncertainties, but a general-type uncertainty of this type turns out to be too complicated.

3.4.1 Intractability of the robust counterpart for a general-type ellipsoidal uncertainty

We start with demonstrating that a general-type ellipsoidal uncertainty affecting even b only leads to computationally intractable (NP-hard) robust counterpart. To see it, note that in this case already the problem of verifying robust feasibility of a given candidate solution x is at least as complicated as the following problem:

> (*) *Given k $l \times l$ symmetric matrices $A^1, ..., A^k$, (k is the dimension of the uncertainty ellipsoid), check whether $\sum_{i=1}^k u_i A^i \le I_l \quad \forall (u \in \mathbf{R}^k : u^T u \le 1)$.*

In turn, (*) is the same as the problem

(**) *Given k $l \times l$ symmetric matrices $A^1, ..., A^k$, and a real r, check whether*

$$r \geq \max_{\xi \in \mathbf{R}^l, ||\xi||_2^2 = l} f(\xi), \quad f(\xi) = \sum_{i=1}^{k} (\xi^T A^i \xi)^2.$$

We will show that in the case of $k = \frac{l(l-1)}{2} + 1$ the problem (**) is computationally intractable. Indeed, given an l-dimensional integral vector a, let us specify the data of (**) as follows: the first $k - 1 = \frac{l(l-1)}{2}$ matrices A^i are appropriately enumerated symmetric matrices B^{pq}, $1 \leq p < q \leq l$, given by the quadratic forms

$$\xi^T B^{pq} \xi = \sqrt{2} \xi_p \xi_q,$$

and the last matrix is

$$A^k = I - \frac{aa^T}{1 + a^T a}.$$

With this choice of $A^1, ..., A^k$, the function $f(\xi)$ (see (**)) restricted to the sphere $S = \{\xi \in \mathbf{R}^l \mid \xi^T \xi = l\}$ is

$$\begin{array}{rcl} f(\xi) & = & f_1(\xi) + f_2(\xi), \\ f_1(\xi) & = & \sum_{i=1}^{k-1} (\xi^T A^i \xi)^2 \\ & = & \sum_{1 \leq p < q \leq l} 2 \xi_p^2 \xi_q^2 \\ & = & l^2 - \sum_{p=1}^{l} \xi_p^4, \\ f_2(\xi) & = & \left(l - \frac{(a^T \xi)^2}{1 + a^T a} \right)^2, \end{array}$$

The maximum value of f_1 on S is $l^2 - l$, and the set F^* of maximizers of f_1 on the sphere is exactly the set of vertices of the cube $C = \{\xi \mid \|\xi\|_\infty \leq 1\}$. The maximum value of f_2 on S is l^2, and the set of maximizers of f_2 on S intersects F^* if and only if the equation

$$a^T z = 0 \tag{25}$$

has a solution with all coordinates being ± 1. In this latter case the maximum value of f on S is the sum of those of f_1 and f_2, i.e., it is $r^* = 2l^2 - l$. In the opposite case the maximum value of f on S is strictly less than r^*, and from the integrality of a it follows that this maximum value is less than $r^* - 2\pi^{-1}(l \, ||a||_2)$, π being a properly chosen polynomial[2]. We see that if r in (**) is specified as $r^* - \pi^{-1}(l \, ||a||_2)$, then the possibility to solve (**) implies the ability to say whether the equation (25) has a solution with entries ± 1. The latter problem is known to be NP-complete. ■

In spite of the last intractability result there are interesting "well-structured" uncertainty ellipsoids which do lead to computationally tractable robust counterparts. An example of this type is discussed next.

3.4.2 "Rank 2" ellipsoidal uncertainty and robust truss topology design

Consider a "nominal" (certain) SDP problem

$$(p^0): \quad \min \left\{ c^T x : \mathcal{A}^0(x) \in \mathbf{S}_+^l \right\}; \tag{26}$$

[2] For details, see Margelit, T. (1997). *Robust Mathematical Programming with Applications to Portfolio Selection*, M.Sc. Thesis, Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology.

here x is n-dimensional design vector and $\mathcal{A}^0(x)$ is $l \times l$ matrix affinely depending on x. Let d be a fixed nonzero l-dimensional vector; let us call *rank 2 perturbation of* $\mathcal{A}^0(\cdot)$ *associated with* d a perturbation of the form

$$\mathcal{A}^0(x) \mapsto \mathcal{A}^0(x) + b(x)d^T + db^T(x),$$

where $b(x)$ is an affine function of x taking values in $\mathbf{R}^l$. Now consider the uncertain SDP problem (P) obtained from (p^0) by all possible perturbations associated with a fixed vector $d \neq 0$ and with $b(\cdot)$ varying in a (bounded) ellipsoid:

$$(P) = \left\{ \min \left\{ c^T x : \mathcal{A}^0(x) + [\sum_{j=1}^{k} u_j b^j(x)]d^T + d[\sum_{j=1}^{k} u_j b^j(x)]^T \geq 0 \right\} \right\}_{u \in \mathbf{R}^k : u^T u \leq 1} ; \qquad (27)$$

here $b^j(x)$ are given affine functions of x taking values in $\mathbf{R}^l$.

Proposition 3.1 *The robust counterpart (P^*) of the uncertain SDP problem (P) is equivalent to the following SDP problem:*

$$\begin{array}{lcc} (\mathcal{P}) & \min\ c^T x & \\ \textit{w.r.t. } x \in \mathbf{R}^n, \lambda \in \mathbf{R} \textit{ subject to} & & \\ (a) & \begin{pmatrix} \lambda I_k & [b^1(x); b^2(x); ...; b^k(x)]^T \\ [b^1(x); b^2(x); ...; b^k(x)] & \mathcal{A}^0(x) - \lambda dd^T \end{pmatrix} & \geq \ 0. \end{array} \qquad (28)$$

Proof. Let

$$\beta(x) = [b^1(x); b^2(x); ...; b^k(x)];$$

a vector x is feasible for (P^*) if and only if

$$\forall (\xi \in \mathbf{R}^l)\ \forall (u \in \mathbf{R}^k, u^T u \leq 1) : \quad \xi^T \mathcal{A}^0(x)\xi + 2(d^T\xi)(u^T \beta^T(x)\xi) \geq 0, \qquad (29)$$

or equivalently, if and only if

$$\forall (\xi \in \mathbf{R}^l) : \quad \xi^T \mathcal{A}^0(x)\xi - 2|d^T\xi| \left|\left|\beta^T(x)\xi\right|\right|_2 \geq 0. \qquad (30)$$

Condition (30), in turn, is equivalent to the following one:

$$\begin{array}{l} \forall (\xi \in \mathbf{R}^l, \eta \in \mathbf{R}^k) : \\ P(\xi,\eta) \equiv (d^T\xi)^2 - \eta^T\eta \geq 0 \Rightarrow Q(\xi,\eta) \equiv \xi^T \mathcal{A}^0(x)\xi + 2\eta^T\beta^T(x)\xi \geq 0. \end{array} \qquad (31)$$

According to Lemma 3.1, (31) is equivalent to the existence of a nonnegative λ such that the quadratic form

$$Q(\xi,\eta) - \lambda P(\xi,\eta)$$

of ξ, η is positive semidefinite. Summarizing our observations, we see that x is feasible for (P^*) if and only if there exists $\lambda \geq 0$ such that the matrix

$$\begin{pmatrix} \lambda I_k & \beta^T(x) \\ \beta(x) & \mathcal{A}^0(x) - \lambda dd^T \end{pmatrix}$$

is positive semidefinite. ■

The situation described in Proposition 3.1 arises, e.g., in the truss topology design problem described in Introduction (see Ben-Tal and Nemirovski (1997) for more details). In the latter problem, the "nominal" program (26) is of the form

$$\begin{array}{rl} & \min\ \tau \\ \text{s.t.} & \\ & \left(\begin{array}{c|c|c} \tau & f^T & \\ \hline f & \sum_{i=1}^p t_i A^i & \\ \hline & & \mathrm{Diag}(t) \end{array}\right) \ \geq\ 0, \\ & \sum_{i=1}^p t_i = 1, \end{array}$$

the design vector being $(t_1, ..., t_p, \tau)$; t_i are bar volumes, τ represents the compliance, and f is the "nominal" external load. In the robust setting of the TTD problem, we replace the nominal load f with an ellipsoid of loads centered at f; this is nothing but rank 2 perturbation of the nominal problem associated with $d = e_1$ and a properly chosen $b^i(x)$ (in fact, independent of x).

3.4.3 "Approximate" robust counterpart of an uncertain SDP problem

As we have seen, an uncertain SDP problem with a general-type ellipsoid $\mathcal{U}$ may result in a computationally intractable robust counterpart. This can occur in other situations as well, and whenever this is the case, a natural way to overcome the difficulty is to replace the robust counterpart by a "tractable approximation". Let us start with introducing this latter concept.

Approximation of the robust counterpart. Consider an uncertain optimization problem:

$$(P) = \left\{ \min \left\{ c^T x : F(x, \zeta) \in \mathbf{K},\ x \in \mathcal{X} \right\} \right\}_{\zeta \in \zeta^0 + \mathcal{V}},$$

where ζ^0 is the "nominal" data and $\mathcal{V}$ is a convex "perturbation set" containing 0. Let

$$(P^*_\rho) \quad \min \left\{ c^T x : x \in \mathcal{X}, F(x, \zeta) \in \mathbf{K}\ \ \forall \zeta \in \zeta^0 + \rho \mathcal{V} \right\}$$

be the robust counterpart of the uncertain problem obtained from (P) by replacing the original perturbation set $\mathcal{V}$ by its ρ-enlargement, and let $G_*(\rho)$ denote the feasible set of (P^*_ρ).

Definition 3.1 *We say that a "certain" optimization problem*

$$(\Pi) : \quad \min \left\{ c^T x : (x, \lambda) \in \mathcal{X}^+ \right\}$$

with design variables x, λ and feasible domain $\mathcal{X}^+$ is an approximation of the robust counterpart $(P^*) \equiv (P^*_1)$ of (P), if the projection $G^+(\Pi)$ of the feasible set of (Π) on the plane of x-variables is contained in the feasible set $G_*(1)$ of the robust counterpart (P^*), i.e., if (Π) is "more conservative" than the robust counterpart of (P).

A natural "measure of conservatism" of (Π) as an approximation to (P^*) is as follows:

$$\mathrm{cons}(\Pi) = \inf\{\rho \geq 1 : G_*(\rho) \subset G^+(\Pi)\}.$$

Definition 3.2 *We say that (Π) is an α-conservative approximation of (P^*), if $\mathrm{cons}(\Pi) \leq \alpha$, i.e., if*

$$\left\{ \begin{array}{lll} x \in G^+(\Pi) & \Rightarrow & x \in G^*(1), \\ x \notin G^+(\Pi) & \Rightarrow & x \notin G_*(\rho) \quad \forall \rho > \alpha. \end{array} \right.$$

Note that from the viewpoint of practical modeling it is nearly the same to use the exact robust counterpart or its approximation with "moderate" level of conservativeness. Therefore, if the exact counterpart turns out to be a "computationally bad" problem (as it is the case for uncertain SDP problems with ellipsoidal uncertainties), it makes sense to look for an approximate counterpart which is on one hand with a "reasonable" level of conservativeness and on the other hand is computationally tractable. We now present a result in this direction.

A "universal" approximate robust counterpart to an uncertain SDP problem. Consider an uncertain SDP problem in the form

$$(P) = \left\{ \min \left\{ c^T x : \mathcal{A}^i(x) \in \mathbf{S}^{l_i}_+, i = 1, ..., m \right\} \right\}_{\mathcal{A}^i(\cdot) \in \mathcal{A}^{i0}(\cdot) + \mathcal{V}_i, i=1,...,m},$$

$\mathcal{A}^{i0}(\cdot)$ being an affine mapping from $\mathbf{R}^n$ to $\mathbf{S}^{l_i}$ and $\mathcal{V}_i$, $0 \in \mathcal{V}_i$, being convex perturbation sets in the spaces of mappings of this type.

Assume that

(#). For every $i \le m$ the set $\mathcal{V}_i$ can be approximated "up to factor γ_i" by an ellipsoid, i.e., we can find out k_i affine mappings $\mathcal{A}^{ij}(\cdot) : \mathbf{R}^n \to \mathbf{S}^{l_i}$, $j = 1, ..., k_i$, in such a way that $\mathcal{V}_i^- \subseteq \mathcal{V}_i \subseteq \gamma_i \mathcal{V}_i^-$, where

$$\mathcal{V}_i^- = \left\{ \sum_{j=1}^{k_i} u_j \mathcal{A}^{ij}(\cdot) \mid u^T u \le 1 \right\}.$$

Theorem 3.4 *Under assumption* **(#)** *the SDP program*

$$\begin{array}{ll} (\Pi): & \min_x \ c^T x \\ s.t. & \\ & \mathcal{B}_i(x) \equiv \begin{pmatrix} \mathcal{A}^{i0}(x) & \gamma_i \mathcal{A}^{i1}(x) & \gamma_i \mathcal{A}^{i2}(x) & \cdots & \gamma_i \mathcal{A}^{ik_i}(x) \\ \gamma_i \mathcal{A}^{i1}(x) & \mathcal{A}^{i0}(x) & & & \\ \gamma_i \mathcal{A}^{i2}(x) & & \mathcal{A}^{i0}(x) & & \\ \vdots & & & \ddots & \cdots \\ \gamma_i \mathcal{A}^{ik_i} & & & \cdots & \mathcal{A}^{i0}(x) \end{pmatrix} \ge 0, \\ & \qquad i = 1, ..., m. \end{array}$$

is an α-conservative approximate robust counterpart of (P) with

$$\alpha = \max_{i=1,...,m} \gamma_i \min[\sqrt{k_i}, \sqrt{l_i}]. \tag{32}$$

Note that the strength of the result is in the fact that the guaranteed level of conservatism α is proportional to the *minimum* of the *square roots* of the dimensions $\max_i k_i$ and $\max_i l_i$.

Proof of the theorem. We should prove that (i) (Π) indeed is an approximate robust counterpart of (P), and that (ii) the level of conservativeness of (Π) is as given in (32).

(i): Assume that x is feasible for (Π), and let us prove that x is robust feasible for (P). To this end it suffices to verify that for every $i = 1, ..., m$ one has

$$\mathcal{B}_i(x) \ge 0 \Rightarrow \mathcal{A}^{i0}(x) + \mathcal{A}(x) \ge 0 \quad \forall \mathcal{A} \in \mathcal{V}_i. \tag{33}$$

To simplify notation, let us drop the index i and write $A^0 = \mathcal{A}^{i0}(x)$, $A^j = \mathcal{A}^{ij}(x)$, $\gamma = \gamma_i$, etc.

Assume that $\mathcal{B}_i(x) \geq 0$. The matrix A^0 clearly is positive semidefinite. Let B be a positive definite symmetric matrix such that $B \geq A^0$; then the matrix

$$\mathcal{B} = \begin{pmatrix} B & \gamma A^1 & \gamma A^2 & \cdots & \gamma A^k \\ \gamma A^1 & B & & & \\ \gamma A^2 & & B & & \\ \vdots & & & \ddots & \cdots \\ \gamma A^k & & & \cdots & B \end{pmatrix}$$

is positive semidefinite, i.e., for every collection of vectors $\xi, \eta_1, \eta_2, ..., \eta_k$ one has

$$\xi^T B \xi + \sum_{j=1}^{k} \eta_j^T B \eta_j + 2\gamma \sum_{j=1}^{k} \xi^T A^j \eta_j \geq 0.$$

Minimizing the left hand side of this equation in $\eta_1, ..., \eta_k$, we get

$$\xi^T B \xi - \gamma^2 \sum_{j=1}^{k} \xi^T A^j B^{-1} A^j \xi \geq 0 \quad \forall \xi,$$

or, denoting $D_j = B^{-1/2} A^j B^{-1/2}$,

$$\sum_{j=1}^{k} \gamma^2 D_j^2 \leq I_l. \tag{34}$$

Now let $\mathcal{A}(\cdot) \in \mathcal{V}_i$ and let $A = \mathcal{A}(x)$; note that in view of **(#)**

$$\exists (u, u^T u \leq \gamma^2): \quad A = \sum_{j=1}^{k} u_j A^j.$$

For an arbitrary $\eta \in \mathbf{R}^l$, we have

$$\begin{array}{rcl} |\eta^T A \eta| & \leq & \sum_{j=1}^k |u_j| |\eta^T A^j \eta| \\ & \leq & \sqrt{\sum_{j=1}^k u_j^2} \sqrt{\sum_{j=1}^k (\eta^T A^j \eta)^2} \\ & \leq & \gamma \sqrt{\sum_{j=1}^k [(B^{1/2}\eta)^T (B^{-1/2} A^j \eta)]^2} \\ & \leq & \gamma \sqrt{\sum_{j=1}^k [\eta^T B \eta][\eta^T A^j B^{-1} A^j \eta]} \\ & = & \gamma \sqrt{\sum_{j=1}^k [\eta^T B \eta][\eta^T B^{1/2} D_j^2 B^{1/2} \eta]} \\ & \leq & \eta^T B \eta \\ & & \text{[we have used (34)].} \end{array}$$

Thus,

$$|\eta^T A \eta| \leq \eta^T B \eta \quad \forall \eta.$$

This relation is valid for all positive definite matrices B which are $\geq A^0$, so that $A^0 + A \geq 0$. Since the latter inequality is valid for all $A = \mathcal{A}(x)$, $\mathcal{A}(\cdot) \in \mathcal{V}_i$, x indeed satisfies the conclusion in (33). (i) is proved.

(ii): Assume that $x \in G_*(\alpha)$ with α given by (32); we should prove that x is feasible for (Π). Arguing by contradiction, assume that x is not feasible for (Π). First of all, we note that all $\mathcal{A}^{i0}(x)$ are positive semidefinite – otherwise x would not be feasible even for the nominal instance and therefore would not belong to $G_*(\alpha)$. Since x is not feasible for (Π), there exists i such that the matrix $\mathcal{B}_i(x)$ is not positive semidefinite. Choosing as B a positive definite matrix which is $\geq A^{i0}(x)$ and is close enough to $A^{i0}(x)$, we see that the matrix

$$\mathcal{B} = \begin{pmatrix} B & \gamma_i\mathcal{A}^{i1}(x) & \gamma_i\mathcal{A}^{i2}(x) & \cdots & \gamma_i\mathcal{A}^{ik_i}(x) \\ \gamma_i\mathcal{A}^{i1}(x) & B & & & \\ \gamma_i\mathcal{A}^{i2}(x) & & B & & \\ \cdots & & & \ddots & \cdots \\ \gamma_i\mathcal{A}^{ik_i} & & & & B \end{pmatrix}$$

is not positive semidefinite.

In view of the same reasoning as in the proof of (i), the latter fact means exactly that the matrix

$$I_l - \gamma^2\sum_{j=1}^{k} D_j^2, \qquad D_j = B^{-1/2}A^jB^{-1/2},$$

is not positive semidefinite (we use the same shortened notation as in the proof of (i)). In other words, there exists a vector h, $h^Th = 1$, such that

$$\begin{array}{rcl} \sum_{j=1}^k h_j^Th_j & > & 1, \\ h_j & = & \gamma D_jh, \; j = 1, 2, ..., k. \end{array} \tag{35}$$

We claim that

$$\exists u: \quad \{u^Tu = \alpha^2\} \;\&\; \{\|\Delta\| > 1\}, \quad \Delta = \sum_{j=1}^{k} u_jD_j, \tag{36}$$

where $\|\cdot\|$ is the operator norm of a symmetric matrix.

Indeed, consider two possible cases: (a) $k \equiv k_i \leq l \equiv l_i$, and (b) $k > l$.

Assume that (a) is the case. It is evident that one can choose numbers $\epsilon_j = \pm 1$ in such a way that

$$\left\|\sum_{j=1}^{k}\epsilon_jh_j\right\|_2^2 \geq \sum_{j=1}^{k} h_j^Th_j \quad [> 1],$$

whence

$$h^T[\sum_{j=1}^{k}\epsilon_jD_j]^2h > \gamma^{-2}.$$

In view of this inequality, setting $u_j = \alpha\epsilon_jk^{-1/2}$, $j = 1, ..., k$, we satisfy the first relation in the conclusion of (36), and the operator norm of the corresponding matrix Δ^2 is $> \alpha^2\gamma^{-2}k^{-1}$; the latter quantity is ≥ 1, since in case (a) one clearly has $\alpha \geq \gamma k^{1/2}$, see (32). Thus, $\|\Delta\|^2 \geq \|\Delta^2\| > 1$, as required in (36).

Now consider case (b). Denoting by d^p the p-th coordinate of an l-dimensional vector d and taking into account that in view of (35) one has

$$\sum_{j=1}^{k}\sum_{p=1}^{l}(h_j^p)^2 > 1,$$

we immediately conclude that there exists an index $q \in \{1, 2, ..., l\}$ such that

$$\sum_{j=1}^{k} (h_j^q)^2 > l^{-1}.$$

Consequently, there exists a vector $u \in \mathbf{R}^k$, $u^T u = \alpha^2$, such that

$$\sum_{j=1}^{k} u_j h_j^q > \alpha l^{-1/2},$$

or, which is the same,

$$\sum_{j=1}^{k} u_j (D_j h)^q > \alpha \gamma^{-1} l^{-1/2} \geq 1$$

(we have taken into account that in case (b), due to the origin of α, the quantity $\alpha\gamma^{-1}l^{-1/2}$ is ≥ 1). In other words, u satisfies the conclusion in (36) (recall that $h^T h = 1$).

Now let u, Δ be given by (36) (we just have seen that this relation is always true), and let $A = \sum_{j=1}^{k} u_j A^j$, so that $\Delta = B^{-1/2} A B^{-1/2}$. Since $\|\Delta\| > 1$, at least one of the matrices $I_l \pm \Delta$ is not positive semidefinite, or, which is the same, at least one of the matrices $B \pm A$ is not positive semidefinite. On the other hand, $u^T u \leq \alpha^2$, and therefore (see **(#)**) $\pm(\sum_{j=1}^{k} u_j \mathcal{A}^{ij}(\cdot)) \in \alpha \mathcal{V}_i$; since x was assumed to belong to $G_*(\alpha)$, the matrices $A^{i0} \pm A$ should be positive semidefinite; due to $B \geq A^{i0}$, it implies also positive semidefiniteness of the matrices $B \pm A$, and we come to the desired contradiction. ■

Remark 3.1 *In connection with Assumption* **(#)** *it is worth noting that, in the case when $\mathcal{V}$ is the intersection of k ellipsoids centered at the origin, this assumption is satisfied with $\gamma = \sqrt{k}$. Indeed, we may assume that the ellipsoids in question are $E^i = \{\zeta \mid \zeta Q_i^T Q_i \zeta \leq 1\}$; their intersection clearly contains the ellipsoid $\{\zeta \mid \zeta^T \sum_{i=1}^{k} Q_i^T Q_i \zeta \leq 1\}$ and is contained in a $\sqrt{k}$ times larger ellipsoid.*

3.5 Affinely parameterized uncertain programs

Consider the case of an uncertain problem with *affine* uncertainty and $\mathbf{K} = \mathbf{R}_+^m$:

$$(P) \quad = \quad \{\min\{c^T x : x \in \mathcal{X}, f_i(x, \zeta) \geq 0, \ i = 1, ..., m\}\}_{\zeta = \{\zeta_i^j\}_{i,j} \in \mathcal{U}},$$

where

$$f_i(x, \zeta) = f_i^0(x) + \sum_{j=1}^{k} \zeta_i^j f_i^j(x).$$

The robust counterpart of (P) is the problem

$$(P^*) \qquad \begin{array}{rcl} \min \ c^T x & & \\ F_i(x) & \geq & 0, \ i = 1, ..., m, \end{array}$$

where

$$F_i(x) = \inf_{\zeta = \{\zeta_i^j\} \in \mathcal{U}} [f_i^0(x) + \sum_{j=1}^{k} \zeta_i^j f_i^j(x)].$$

Assume that the uncertainty set $\mathcal{U}$ is a bounded ellipsoid, i.e., the image of the unit Euclidean ball under affine mapping:

$$\mathcal{U} = \{\zeta = \zeta^0 + Pu \mid u \in \mathbf{R}^l, u^T u \leq 1\}. \tag{37}$$

Our goal is to demonstrate that in the case in question the robust counterpart is basically "as computationally tractable" as the instances of our uncertain program.

Indeed, let us look what is the robust version of a single uncertain constraint. Dropping the index i, we write the uncertain constraint and its robust version, respectively, as

$$\begin{array}{rcrcl} 0 & \leq & f(x,\zeta) & \equiv & f^0(x) + \zeta^T f(x), \; f(x) = \begin{pmatrix} f^1(x) \\ f^2(x) \\ \cdots \\ f^k(x) \end{pmatrix}, \\ 0 & \leq & F(x) & \equiv & \min\limits_{\zeta \in \mathcal{U}} f(x,\zeta). \end{array} \tag{38}$$

We immediately see that

$$F(x) = f^0(x) + (\zeta^0)^T f(x) - \sqrt{f^T(x) P P^T f(x)}.$$

Note also that F is concave on $\mathcal{X}$ (recall that we always assume that the mapping $F(x,\zeta)$ is $\mathbf{K}$-concave in $x \in \mathcal{X}$). Thus, in the case under consideration the robust counterpart is an "explicitly posed" convex optimization program with constraints of the same "level of computability" as the one of the constraints of the instances. At the same time, the constraints of the robust counterpart may become worse than those of the instances, e.g., they may lose smoothness.

There are, however, cases when the robust counterpart turns out to be of the same analytical nature as the instances of the uncertain program in question.

Example: A simple uncertain geometric programming problem. Consider the case of an uncertain geometric programming program in the exponential form with uncertain coefficients of the exponential monomials, so that the original constraints are of the form

$$0 \leq f(x,\zeta) = 1 - \sum_{j=1}^{k} \zeta_j \exp\{[\beta^j]^T x\} \tag{39}$$

(β^j are "certain") and the uncertainty ellipsoid, for every constraint, is contained in the nonnegative orthant and is given by (37). Here the robust version of (39) is the constraint

$$\begin{array}{rcrcl} 0 & \leq & F(x) & \equiv & 1 - \sum_{j=1}^{k} \zeta_j^0 \exp\{[\beta^j]^T x\} - \sqrt{\phi^T(x) P P^T \phi(x)}, \\ & & \phi(x) & = & (\, \exp\{[\beta^1]^T x\} \quad \exp\{[\beta^2]^T x\} \quad \cdots \quad \exp\{[\beta^k]^T x\} \,)^T. \end{array} \tag{40}$$

Under the additional assumption that the matrix PP^T is a matrix with nonnegative entries, we can immediately convert (40) into a pair of constraints of the same form as (39); to this end it suffices to introduce for each constraint an additional variable t and to represent (40) equivalently by the constraints

$$\begin{array}{rcl} 0 & \leq & 1 - \sum_{j=1}^{k} \zeta_j^0 \exp\{[\beta^j]^T x\} - \exp\{t\}, \\ 0 & \leq & 1 - \sum_{p,q=1}^{k} (PP^T)_{pq} \exp\{[\beta^p + \beta^q]^T x - 2t\}. \end{array}$$

Thus, in the case under consideration the robust counterpart is itself a geometric programming program.

4 Saddle point form of programs with affine uncertainty

Consider once again an uncertain Mathematical Programming problem with *affine* uncertainty and $\mathbf{K} = \mathbf{R}^m_+$:

$$\begin{array}{rcl}(P) & = & \{\min\{c^T x : x \in \mathcal{X}, f_i(x, \zeta^i) \equiv f_i^0(x) + \sum_{j=1}^k \zeta^i_j f_i^j(x) \geq 0, \\ & & i = 1, 2, ..., m\}\}_{\zeta^i \in \mathcal{U}_i,\ i=1,...,m},\end{array}$$

each $\mathcal{U}_i$ being a closed convex set in some $\mathbf{R}^{k_i}$. Note that in the case in question Assumption **A** becomes

> **A.1:** *$\mathcal{X}$ is a closed convex set with a nonempty interior, the functions $f_i^j(x)$ are continuous on $\mathcal{X}$, $i = 1, ..., m, j = 1, ..., k$, and the functions $f_i(x, \zeta^i)$ are concave in $x \in \mathcal{X}$ for every $\zeta^i \in \mathcal{U}_i$, $i = 1, ..., m$.*

The robust counterpart of (P) is the problem

$$\begin{array}{lrcl}(P^*) & & & \\ & \min\ c^T x & & \\ & F_i(x) & \geq & 0,\ i = 1, ..., m, \\ & F_i(x) & = & \inf_{\zeta^i \in \mathcal{U}_i} [f_i^0(x) + \sum_{j=1}^k \zeta^i_j f_i^j(x)].\end{array}$$

which is equivalent to the problem

$$\min_{x \in \mathcal{X}} \sup_{\lambda \in \mathbf{R}^m_+} \left\{ c^T x - \sum_{i=1}^m \lambda_i F_i(x) \right\},$$

or explicitly the problem

$$\min_{x \in \mathcal{X}} \left\{ c^T x - \sum_{i=1}^m \inf_{\zeta^i \in \mathcal{U}_i, \lambda \in \mathbf{R}^m_+} \left[\lambda_i f_i^0(x) + \sum_{j=1}^k \lambda_i \zeta^i_j f_i^j(x) \right] \right\}. \tag{41}$$

Now let us set

$$\begin{array}{rcl}\hat{f}_i(x, \eta) & = & \sum_{j=0}^k \eta_j f_i^j(x) : \mathcal{X} \times \mathbf{R}^{k+1} \to \mathbf{R}, \\ \mathcal{W}_i & = & \mathrm{cl}\,\{(\mu, \mu\zeta^i_1, \mu\zeta^i_2, ..., \mu\zeta^i_k)^T \mid \mu \geq 0, \zeta^i \in \mathcal{U}_i\};\end{array}$$

note that $\mathcal{W}_i$ is a closed convex cone in $\mathbf{R}^{k+1}$. We clearly have

$$\inf_{\zeta^i \in \mathcal{U}_i, \lambda_i \geq 0} \left[\lambda_i f_i^0(x) + \sum_{j=1}^k \lambda_i \zeta^i_j f_i^j(x) \right] = \inf_{\eta^i \in \mathcal{W}_i} \hat{f}_i(x, \eta^i),$$

so that the problem (41) is exactly the problem

$$\min_{x \in \mathcal{X}} \left\{ c^T x - \sum_{i=1}^m \inf_{\eta^i \in \mathcal{W}_i} \hat{f}_i(x, \eta^i) \right\}. \tag{42}$$

Now let

$$\begin{array}{rcl}\mathcal{W} & = & \prod_{i=1}^m \mathcal{W}_i, \\ \eta & = & (\eta^1, ..., \eta^m), \\ L(x, \eta) & = & c^T x - \sum_{i=1}^m \hat{f}_i(x, \eta^i) : \mathcal{X} \times \mathcal{W} \to \mathbf{R}.\end{array}$$

Due to **A.1** and to the origin of $\mathcal{W}$, the function L is convex in $x \in \mathcal{X}$ for every $\eta \in \mathcal{W}$ and affine (and therefore concave) in η. We summarize the above facts in

Proposition 4.1 *Under assumption* **A.1**, *the robust counterpart* (P^*) *of uncertain problem* (P) *is equivalent to the saddle point problem*

$$\min_{x\in\mathcal{X}} \sup_{\eta\in\mathcal{W}} L(x,\eta) \tag{43}$$

with convex-concave and continuous on $\mathcal{X}\times\mathcal{W}$ *function* $L(x,\eta)$.

The latter result provides us with an opportunity to handle the uncertain problem under consideration via its dual problem

$$\max_{\eta\in\mathcal{W}} \inf_{x\in\mathcal{X}} L(x,\eta)$$

which in some cases may be better suited for a numerical solution. Another advantage of the saddle point formulation of the robust counterpart is seen from the following considerations. In simple enough cases (cf. Section 3.5) we can explicitly perform the inner maximization in η in (43) and end up with an explicit optimization form of the robust counterpart:

$$\min_{x\in\mathcal{X}} \hat{L}(x) \quad [\hat{L}(x) = \sup_{\eta\in\mathcal{W}} L(x,\eta)]. \tag{44}$$

We point out, however, that sometimes this policy, even when it can be carried out, is not necessarily the best one; there are cases when, computationally, it is preferable to work directly on the saddle point form of the robust counterpart than on its optimization form. The reason is that in the case when f_i^j possess nice analytical structure, the function $L(x,\eta)$ basically inherits this nice structure, and in favourable circumstances we can solve the saddle point problem by advanced interior point methods for variational inequalities (see Nesterov and Nemirovski (1994), Chapter 7). In contrast to this, the objective in the optimization program (44), even if it can be explicitly computed, typically is much worse analytically than the functions f_i^j; it may, e.g., loose smoothness even when all functions f_i^j are very smooth. As a result, it may happen that the only numerical methods which can be used to solve (44) directly are relatively low-performance "black box"-oriented tools of nonsmooth convex optimization, like the bundle methods, the Ellipsoid method, etc.

To illustrate the point, let us consider the following simple example:

Heat Dissipation problem. Consider an electric circuit (see Fig. 2) represented as a connected oriented graph Γ with N arcs, an arc (ij) possessing a given conductance $s_{ij}\in(0,\infty)$. Assume that the arcs of the circuit are charged with external constant voltages y_{ij}, and the vector y of these voltages is restricted to belong to a given convex and compact set $\mathcal{Y}$.

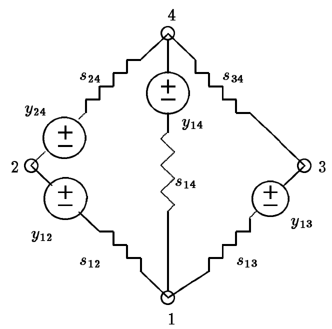

Figure 2. A simple circuit

The voltages induce currents in the circuit; as a result, the circuit dissipates heat. According to Kirchoff's laws, the heat dissipated by the circuit per unit time is

$$H = \min_{u\in\mathbf{R}^M} (Qu+Py)^T S(Qu+Py),$$

where Q and P are some matrices given by the topology of the circuit, M is the number of nodes of the circuit and S is $N \times N$ diagonal matrix with the diagonal entries indexed by the arcs of G and equal to s_{ij}. The problem is to choose $y \in \mathcal{Y}$ which minimizes the dissipated heat:

$$(p) \qquad \min_{x=(t,y)\in\mathcal{X}=\mathbf{R}\times\mathcal{Y}} \left\{ t : t - \min_{u\in\mathbf{R}^m} (Qu+Py)^T S(Qu+Py) \geq 0 \right\},$$

the data in the problem being the conductances s_{ij}, or, which is the same, the diagonal positive semidefinite matrix S. In fact, in what follows we treat the more general problem where S is symmetric positive semidefinite, but not necessarily diagonal.

Now assume that the data are uncertain: all we know is that $S \in \mathcal{U}$, $\mathcal{U}$ being a closed convex subset of the positive semidefinite cone $\mathbf{S}^N_+$; for the sake of simplicity, we assume that $\mathcal{U}$ is bounded. The saddle point problem (43) becomes here

$$\min_{(t,y)\in\mathcal{X}} \sup_{\lambda\geq 0, S\in\mathcal{U}} \left[t(1-\lambda) + \lambda \min_u (Qu+Py)^T S(Qu+Py) \right].$$

It is clearly seen that the latter problem is equivalent to the saddle point problem

$$\min_{y\in\mathcal{Y}} \max_{S\in\mathcal{U}} \left[\min_u (Qu+Py)^T S(Qu+Py) \right],$$

which in turn is equivalent to the saddle point problem

$$\min_{y\in\mathcal{Y},u} \max_{S\in\mathcal{U}} f(y,u;S), \quad f(y,u;S) = (Qu+Py)^T S(Qu+Py) \tag{45}$$

(we have used the standard min-max theorem (see Rockafellar (1970)) and have taken into account that $\mathcal{U} \subset \mathbf{S}^N_+$, so that the function $(Qu+Py)^T S(Qu+Py)$ is convex in (y,u) for every $S \in \mathcal{U}$ and is linear, and therefore concave, in S).

The nice analytical structure of f makes it possible to solve the saddle point problem by the polynomial time path-following interior point methods from (Nesterov and Nemirovski (1994), Chapter 7) provided that the convex sets $\mathcal{Y}, \mathcal{U}$ are "simple" – admit explicit self-concordant barriers. Thus, here the saddle point form of the robust counterpart yields a problem which can be solved by theoretically efficient interior point methods. This is not always the case for the optimization form (44) of the robust counterpart. Indeed, consider the simplest case where Q is the zero matrix (such a case has no actual sense in the Heat Dissipation problem, but it makes sense mathematically, which is sufficient to illustrate the point). In this case the objective $\hat{L}(x)$, $x = (t,y)$ in problem (44) becomes

$$\begin{cases} +\infty, & t < \max_{S\in\mathcal{U}} y^T P^T SPy \\ t, & \text{otherwise} \end{cases},$$

so that (44) is nothing but the problem

$$\min_{y\in\mathcal{Y}} F(y), \quad F(y) = \max_{S\in\mathcal{U}} y^T P^T SPy. \tag{46}$$

Now assume that $\mathcal{U} \subset \mathbf{S}^N_+$ is an ellipsoid:

$$\mathcal{U} = \{S^0 + \sum_{j=1}^k v_j S^j \mid v^T v \leq 1\}.$$

Then the objective F in (46) can be explicitly computed:

$$F(x) = x^T P^T S P x + \sqrt{\sum_{j=1}^{k} (x^T P^T S^j P x)^2}. \tag{47}$$

Note that since not all S^j necessarily are positive semidefinite, it is not even seen from (47) that F is convex. Moreover, direct minimization of convex functions of the type (47), where not all S^j's are positive semidefinite, is not as yet covered by the current interior point methodology.

Acknowledgment. The research was supported by the Israel Science Foundation grant # 306/94-3, by the Fund for the Promotion of Research at Technion, by the G.I.F. contract No. I-0455-214.06/95 and by the Israel Ministry of Science grant # 9636-1-96.

References

Ben-Tal, A., A. Nemirovski (1995a). *Robust solutions to uncertain linear problems*, Research Report # 6/95, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion – the Israel Institute of Technology, Technion City, Haifa 32000, Israel.

Ben-Tal, A., A. Nemirovski (1995b). *Robust Convex Programming*, Working paper, Faculty of Industrial Engineering and Management, Technion – the Israel Institute of Technology, Technion City, Haifa 32000, Israel.

Ben-Tal, A., A. Nemirovski (1997). Robust Truss Topology Design via Semidefinite Programming. *SIAM J. Optim.* **7** 991–1016.

Boyd, S., L. El Ghaoui, E. Feron, V. Balakrishnan (1994). *Linear Matrix Inequalities in System and Control Theory*, Society of Industrial and Applied Mathematics, Philadelphia.

El-Ghaoui, L., H. Lebret (1996). *Robust solutions to least-square problems with uncertain data matrices*, Working paper, Ecole Nationale Supérieure De Techniques Avancées, 32, Bd. Victor, 75739 Paris, France.

Kouvelis, P., G. Yu (1997). Robust Discrete Optimization and its Applications. Nonconvex Optimization and its Applications **14**, Kluwer Academic Publishers.

Falk, J.E. (1976). Exact Solutions to Inexact Linear Programs. *Oper. Res.* **34** 783–787.

Mulvey, J.M., R.J. Vanderbei, S.A. Zenios (1995). Robust optimization of large-scale systems. *Oper. Res.* **43** 264–281.

Nemirovski, A. (1997). *On self-concordant convex-concave functions*, Research Report # 3/97 (June 1997), Optimization Laboratory Faculty of Industrial Engineering and Management, Technion – the Israel Institute of Technology, Technion City, Haifa 32000, Israel.

Nesterov, Yu., A. Nemirovski (1994). *Interior point polynomial methods in Convex Programming*, Society of Industrial and Applied Mathematics, Philadelphia.

Oustry, H., L. El Ghaoui, H. Lebret (1996). *Robust solutions to uncertain semidefinite programs*, Working paper, Ecole Nationale Supérieure De Techniques Avancées, 32, Bd. Victor, 75739 Paris, France. *SIAM J. Optim.* (to appear).

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

Singh, C. (1982). Convex Programming with Set-Inclusive Constraints and its Applications to Generalized Linear and Fractional Programming. *JOTA* **38** 33–42.

Soyster, A.L. (1973). Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming. *Oper. Res.* **21** 1154–1157.

Zhou, K., J. Doyle, K. Glover (1995). *Robust and Optimal Control*, Prentice Hall, New Jersey.

A. Ben-Tal: Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Technion City, Haifa 32000, Israel; e-mail: morbt@ie.technion.ac.il

A. Nemirovski: Faculty of Industrial Engineering and Management, Technion – Israel Institute of Technology, Technion City, Haifa 32000, Israel; e-mail: nemirovs@ie.technion.ac.il