Robust Semidefinite Programming*

Aharon Ben-Tal† Laurent El Ghaoui‡ Arkadi Nemirovski§

August 1st, 1998

Abstract

In this paper, we consider semidefinite programs where the data is only known to belong to some uncertainty set $\mathcal{U}$. Following recent work by the authors, we develop the notion of robust solution to such problems, which are required to satisfy the (uncertain) constraints whatever the value of the data in $\mathcal{U}$. Even when the decision variable is fixed, checking robust feasibility is in general NP-hard. For a number of uncertainty sets $\mathcal{U}$, we show how to compute robust solutions, based on a sufficient condition for robust feasibility, via SDP. We detail some cases when the sufficient condition is also necessary, such as linear programming or convex quadratic programming with ellipsoidal uncertainty. Finally, we provide examples, taken from interval computations and truss topology design.

1 Introduction

1.1 SDPs with uncertain data

We consider a semidefinite programming problem (SDP) of the form

$$\max\ b^T y \quad \text{subject to } F(y) = F_0 + \sum_{i=1}^{m} y_i F_i \succeq 0, \tag{1}$$

where $b \in \mathbf{R}^m$ is given, and F is an affine map from $y \in \mathbf{R}^m$ to $\mathcal{S}^n$.

In many practical applications, the "data" of the problem (the vector b and the coefficients matrices $F_0, \ldots, F_m$) is subject to possibly large uncertainties. Reasons for this include the following.

- *Uncertainty about the future.* The exact value of the data is unknown at the time the values of y should be determined and will be known only in the future. (In a risk management problem for example, the data may depend on future demands, market prices, etc, that are unknown when the decision has to be taken.)

*This research was partly supported by the Fund for the Promotion of Research at the Technion, by the G.I.F. contract No. I-0455-214.06/95, by the Israel Ministry of Science grant # 9636-1-96, and by Electricité de France, under contract P33L74/2M7570/EP777.

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

‡Ecole Nationale Suprieure de Techniques Avances, 32, Bvd. Victor, 75739 Paris, France; email: elghaoui@ensta.fr.

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

- *Errors in the data.* The data originates from a process (measurements, computational process) that is very hard to perform error-free (errors in the measurement of material properties in a truss optimization problem, sensor errors in a control system, floating-point errors made during computation, etc).

- *Implementation errors.* The computed optimal solution y^* cannot be implemented exactly, which results in uncertainty about the feasibility of the implemented solution. For example, the coefficients of an optimal finite-impulse response (FIR) filter can often be implemented in 8 bits only. As it turns out, this can be modeled as uncertainties in the coefficients matrices F_i.

- *Approximating nonlinearity by uncertainty.* The mapping $F(y)$ is completely known, but (slightly) non linear. The optimizer chooses to approximate the nonlinearity by uncertainty in the data.

- *Infinite number of constraints.* The problem is posed with an infinite number of constraints, indexed on a scalar parameter ω (for example, the constraints express a property of a FIR filter at each frequency ω). It may be convenient to view this parameter as a perturbation, and regard the semi-infinite problem as a robustness problem.

Depending on the assumptions on the nature of uncertainty, several methods have been proposed in the Operations Research/Engineering literature. Stochastic programming works with random perturbations and probabilities of satisfaction of constraints, and thus requires correct estimates of the distribution of uncertainties; sensitivity analysis assumes the perturbations is infinitesimal, and can be used only as a "post-optimization" tool; the "Robust Mathematical Programming" approach recently proposed by Mulvey, Vanderbei and Zenios [13] is based on the (sometimes very restrictive) assumption that the uncertainty takes a finite number of values (corresponding to "worst-case scenarios"). Interval arithmetic is one of the methods that have been proposed to deal with uncertainty in numerical computations; references to this large field of study include the book by one of its founders, Moore [12], the more recent book by Hansen [9], and also the very extensive web site developed by Kosheler and Kreinovich [11].

The approach proposed here assumes (following the philosophy of interval calculus) that the data of the problem is only known to belong to some "uncertainty set" $\mathcal{U}$; in this sense, the perturbation to the nominal problem is deterministic, unknown-but-bounded. A *robust solution* of the problem is one which satisfies the perturbed constraints for every value of the data within the admissible region $\mathcal{U}$. The *robust counterpart* of the SDP is to minimize the worst-case value of the objective, among all robust solutions. This approach was introduced by the authors independently in [1, 2, 3] and [15, 5]; although apparently new in mathematical programming, the notion of robustness is quite classical in control theory (and practice).

Even for simple uncertainty sets $\mathcal{U}$, the resulting robust SDP is NP-hard. Our main result is to show how to compute, via SDP, *upper bounds* for the robust counterpart. Contrarily to the other approaches to uncertainty, the robust method provides (in polynomial time) *guarantees* (of,*e.g.*, feasibility), at the expense of possible conservatism. Note that there is no real conflict between other approaches to uncertainty and ours; for example, it is possible to solve via robust SDP a stochastic programming problem with unknown-but-bounded distribution of the random parameters.

1.2 Problem definition

To make our problem mathematically precise, we assume that the perturbed constraint is of the form

$$\mathbf{F}(y, \delta) \succeq 0,$$

where $y \in \mathbf{R}^m$ is the decision vector, δ is a "perturbation vector" that is only known to belong to some "perturbation set" $\mathcal{D} \subseteq \mathbf{R}^l$, and $\mathbf{F}$ is a mapping from $\mathbf{R}^m \times \mathcal{D}$ to $\mathcal{S}^n$. We assume that $\mathbf{F}(y, \delta)$ is affine in y for each δ, and rational in δ for each y; also we assume that $\mathcal{D}$ contains 0, and that $\mathbf{F}(y, 0) = F(y)$ for every y. Without loss of generality, we assume that the objective vector b is independent of perturbation.

We consider the following problem, referred to as the *robust counterpart* to the "nominal" problem (1):

$$\max\ b^T y \text{ subject to } y \in \mathcal{X}_{\mathcal{D}} \tag{2}$$

where $\mathcal{X}_{\mathcal{D}}$ is the set of *robust feasible* solutions, that is,

$$\mathcal{X}_{\mathcal{D}} = \left\{ y \in \mathbf{R}^m \;\middle|\; \text{for every } \delta \in \mathcal{D},\ \mathbf{F}(y,\delta) \text{ is well-defined and } \mathbf{F}(y,\delta) \succeq 0 \right\}.$$

In this paper, we only consider *ellipsoidal uncertainty.* This means that the perturbation set $\mathcal{D}$ consists of block vectors, each block being subject to an Euclidean-norm bound. Precisely,

$$\mathcal{D} = \left\{ \delta \in \mathbf{R}^l \;\middle|\; \delta = \begin{bmatrix} \delta^1 \\ \vdots \\ \delta^N \end{bmatrix},\ \text{where } \delta^k \in \mathbf{R}^{n_k},\ \ \|\delta^k\|_2 \le \rho,\ \ k = 1, \ldots, N \right\}, \tag{3}$$

where $\rho \ge 0$ is a given parameter that determines the "size" of the uncertainty, and the integers n_k denote the lengths of each block vector δ_k (we have of course $n_1 + \ldots + n_N = L$).

There are many motivations for considering the above framework. It can be used when the perturbation is a vector with each component bounded in magnitude, in which case each block vector δ^k is actually a scalar ($n_1 = \ldots = n_N = 1$). It also can be used when the perturbation is bounded in Euclidean norm (which is often the case when the bounds on the parameters are obtained from statistics, and a Gaussian distribution is assumed). In some applications, there is a mixture of Euclidean-norm and maximum-norm bounds. (For example, we might have some parameters of the form $\delta_1 = \rho \cos\theta$, $\delta_2 = \rho \sin\theta$, where both ρ and θ are uncertain.)

It turns out that already in the case of affine perturbations ($\mathbf{F}(\cdot, \delta)$ is affine in δ) the robust counterpart (2), generally speaking, is NP-hard. This is why we are interested not only in the robust counterpart itself, but also in its *approximations* – "computationally tractable" problems with the same objective as in (2) and feasible sets contained in the set $\mathcal{X}_{\mathcal{D}}$ of robust solutions. We will obtain an upper bound (approximation) on the robust counterpart in the form of an SDP. The size of this SDP is linear in both the length n_k of each block and the number N of blocks.

The paper is organized as follows. In section 2, we consider the case when the perturbation vector affects the semidefinite constraint affinely, that is, the matrix $\mathbf{F}(y, \delta)$ is affine in δ; we provide not only an approximation of the robust counterpart, but also a result on the quality of the approximation, in an appropriately defined sense. Section 3 is devoted to the general case (the perturbation δ enters rationally in $\mathbf{F}(y, \delta)$). We provide interesting special cases (when the approximation is exact) in section 4, while section 5 describes several examples of application.

2 Affine perturbations

In this section, we assume that the matrix function $\mathbf{F}(y, \delta)$ is given by

$$\mathbf{F}(y, \delta) = F^0(y) + \sum_{i=1}^{l} \delta_i F^i(y) \tag{4}$$

where each $F^i(y)$ is a symmetric matrix, affine in y.

We have the following result.

Theorem 2.1 *Consider uncertain semidefinite program with affine perturbation* (4) *and ellipsoidal uncertainty* (3)*, and let* $\nu_0 = 0$*,* $\nu_k = \sum_{s=1}^{k} n_s$*. Then the semidefinite program*

$$\max\ b^T y$$

$$\begin{array}{lrcl}
s.t. & & & \\
(a) & \begin{bmatrix} S_k & \rho F_{\nu_{k-1}+1}(y) & \rho F_{\nu_{k-1}+2}(y) & \cdots & \rho F_{\nu_k}(y) \\ \rho F_{\nu_{k-1}+1}(y) & Q_k & & & \\ \rho F_{\nu_{k-1}+2}(y) & & Q_k & & \\ \vdots & & & \ddots & \vdots \\ \rho F_{\nu_k}(y) & & & & Q_k \end{bmatrix} & \succeq & 0,\ k = 1, 2, \ldots, N; \\
(b) & \sum_{k=1}^{N} (S_k + Q_k) & \preceq & 2F_0(y);
\end{array} \tag{5}$$

in variables $y, S_1, \ldots, S_N, Q_1, \ldots, Q_N$ *is an approximation of the robust counterpart* (2)*, i.e., the projection of the feasible set of* (5) *on the space of* y*-variables is contained in the set of robust feasible solutions.*

Proof. Let us fix a feasible solution $Y = (y, \{S_k\}, \{Q_k\})$ to (5), and let us set $F_i = F_i(y)$. We should prove that

> (*) For every $\delta = \{\delta_i\}_{i=1}^{l}$ such that
>
> $$\sum_{i=\nu_{k-1}+1}^{\nu_k} \delta_i^2 \le \rho^2,\ k = 1, \ldots, N, \tag{6}$$
>
> one has
>
> $$F_0 + \sum_i \delta_i F_i \succeq 0.$$

Since Y is feasible for (5), it follows that the matrices $F_0, S_1, \ldots, S_N, Q_1, \ldots, Q_N$ are positive semidefinite. By obvious regularization arguments, we may further assume these matrices to be positive definite. Finally, performing "scaling"

$$\begin{array}{rcl}
S_k & \mapsto & F_0^{-1/2} S_k F_0^{-1/2}, \\
Q_k & \mapsto & F_0^{-1/2} Q_k F_0^{-1/2}, \\
F_i & \mapsto & F_0^{-1/2} F_i F_0^{-1/2}, \\
\Phi_0 & \mapsto & I,
\end{array}$$

we reduce the situation to the one where $F_0 = I$, which we assume till the end of the proof.

Let $\mathcal{I}_k$ be the set of indices $\nu_{k-1}+1, \ldots, \nu_k$. Whenever δ satisfies (6) and $\xi \in \mathbf{R}^n$, n being the row size of

F_i's, we have

$$\begin{array}{rcl}\xi^T\left(I+\sum_i \delta_i F_i\right)\xi & = & \xi^T\xi + \sum_k \left[Q_k^{1/2}\xi\right]^T \left[\sum_{i\in\mathcal{I}_k} \delta_i Q_k^{-1/2} F_i S_k^{-1/2}\xi_k\right] \\ & & [\xi_k = S_k^{1/2}\xi,\ k=1,\dots,N] \\ & \geq & \xi^T\xi - \sum_k \|Q_k^{1/2}\xi\|_2 \left[\sum_{i\in\mathcal{I}_k} |\delta_i| \|Q_k^{-1/2}F_i S_k^{-1/2}\xi_k\|_2\right] \\ & \geq & \xi^T\xi - \sum_k \|Q_k^{1/2}\xi\|_2 \sqrt{\sum_{i\in\mathcal{I}_k} \rho^2 \|Q_k^{-1/2}F_i S_k^{-1/2}\xi_k\|_2^2} \\ & & \text{[we have used (6)]} \\ & = & \xi^T\xi - \sum_k \|Q_k^{1/2}\xi\|_2 \sqrt{\rho^2 \xi_k^T \left[\sum_{i\in\mathcal{I}_k} S_k^{-1/2} F_i Q_k^{-1} F_i S_k^{-1/2}\right]\xi_k} \\ & \geq & \xi^T\xi - \sum_k \|Q_k^{1/2}\xi\|_2 \sqrt{\xi_k^T\xi_k} \\ & & \text{[we have used (5.}a\text{)]} \\ & \geq & \xi^T\xi - \sqrt{\sum_k \|Q_k^{1/2}\xi\|_2^2}\sqrt{\sum_k \xi_k^T\xi_k} \\ & = & \xi^T\xi - \sqrt{\xi^T\left[\sum_k Q_k\right]\xi}\sqrt{\xi^T\left[\sum_k S_k\right]\xi} \end{array} \tag{7}$$

It remains to note that if $a = \xi^T\left[\sum_k Q_k\right]\xi$, $b = \xi^T\left[\sum_k S_k\right]\xi$, then $a+b \leq 2\xi^T\xi$ by (5.b) (recall that we are in the situation $F_0(y)=I$), so that $\sqrt{ab} \leq \xi^T\xi$. Thus, the concluding expression in (7) is nonnegative. ■

2.1 Quality of approximation

A general-type approximation of the robust counterpart (2) is an optimization problem

$$\max\ b^T y \quad \text{subject to } (y,z)\in\mathcal{Y} \tag{A}$$

(with variables y,z) such that the projection $\mathcal{X}(A)$ of its feasible set $\mathcal{Y}$ on the space of y-variables is contained in $\mathcal{X}_{\mathcal{D}}$, so that (A)-feasibility of (y,z) implies robust feasibility of y. For a particular approximation, a question of primary interest is how conservative the approximation is. A natural way to measure the "level of conservativeness" of an approximation (A) is as follows. Since (A) is an approximation of (2), we have $\mathcal{X}(A)\subset\mathcal{X}_{\mathcal{D}}$. Now let us increase the level of perturbations, i.e., let us replace the original set of perturbations $\mathcal{D}$ by its κ-enlargement $\kappa\mathcal{D}$, $\kappa\geq 1$. The set $\mathcal{X}_{\kappa\mathcal{D}}$ of robust feasible solutions associated with the enlarged set of perturbations shrinks as κ grows, and for large enough values of κ it may become a part of $\mathcal{X}(A)$. The lower bound of these "large enough" values of κ can be treated as the level of conservativeness $\lambda(A)$ of the approximation (A):

$$\lambda(A) = \inf\{\kappa\geq 1 : \mathcal{X}_{\kappa\mathcal{D}}\subset\mathcal{X}(A)\}.$$

Thus, we say that the level of conservativeness of an approximation (A) is $<\lambda$, if every y which is "rejected" by (A) (i.e., $y\notin\mathcal{X}(A)$) looses robust feasibility after the level ρ of perturbations in (3) is increased by factor λ.

The following theorem bounds the level of conservativeness of approximations we have derived so far:

Theorem 2.2 *Consider semidefinite program with affine perturbations* (4) *and ellipsoidal uncertainty* (3)*, the blocks δ^k of the perturbation vector being of dimensions n_k, $k=1,\dots,N$, and let $l=\sum_k n_k$ and n be the row size of $F(\cdot,\cdot)$. The level of conservativeness of the approximation* (5) *does not exceed* $\min\left[\sqrt{nN};\sqrt{l}\right]$.

Proof. Of course, it suffices to consider the case of $\rho=1$, which is assumed till the end of the proof.

Let $\mathcal{X}$ be the projection of the feasible set of (5) to the space of y-variables, and let $y\notin\mathcal{X}$. We should prove that $y\notin\mathcal{X}_{\lambda\mathcal{D}}$ at least in the following two cases:

(i.1): $\lambda > \sqrt{l}$; (i.2): $\lambda > \sqrt{nN}$.

To save notation, let us write F_i instead of $F_i(y)$. Note that we may assume that $F_0 \succeq 0$ – otherwise y is not robust feasible and there is nothing to prove. In fact we may assume even $F_0 \succ 0$, since from the structure of (5) it is clear that the relation $y \notin \mathcal{X}$, being valid for the original data $F_0, \dots, F_l$, remains valid when we replace a positive semidefinite matrix F_0 with a close positive definite matrix. Note that this regularization may only increase the robust feasible set, so that it suffices to prove the statement in question in the case of $F_0 \succ 0$. Finally, the same scaling as in the proof of Theorem 2.1 allows to assume that $F_0 = I$. Let also $\mathcal{I}_k$ be the same index sets as in the proof of Theorem 2.1.

1^0. Consider the case of (i.1). Let us set

$$\begin{array}{rcl} Q_k & = & \frac{n_k}{l} I, \\ S_k & = & \frac{l}{n_k} \sum_{i \in \mathcal{I}_k} F_i^2. \end{array}$$

The collection $(y, \{S_k, Q_k\})$ clearly satisfies $(5.a, c)$, and therefore it must violate $(5.b)$, since otherwise we would have $y \in \mathcal{X}$. Thus, there exists an n-dimensional vector ξ such that

$$\sum_{k=1}^{N} \frac{l}{n_k} \sum_{i \in \mathcal{I}_k} \|F_i \xi\|_2^2 > \xi^T \xi. \tag{8}$$

Setting

$$p_k = \max_{i \in \mathcal{I}_k} \|F_i \xi\|_2 = \|F_{i_k} \xi\|_2, \ i_k \in \mathcal{I}_k,$$

we come to

$$\sum_{k=1}^{N} l p_k^2 > \xi^T \xi. \tag{9}$$

Now let $\delta \in \lambda \mathcal{D}$ be a random vector with independent coordinates distributed as follows: a coordinate δ_i with index $i \in \mathcal{I}_k$ is zero, except the case of $i = i_k$, and the coordinate δ_{i_k} takes values $\pm\lambda$ with probabilities $1/2$. The expected squared Euclidean norm of the random vector $\sum_i \delta_i F_i \xi$ clearly is equal to $\lambda^2 \sum_k p_k^2$; thus, $\lambda \mathcal{D}$ contains a perturbation δ such that

$$\| \left[\sum_{i=1}^{l} \delta_i F_i \right] \xi \|_2^2 \geq \lambda^2 \sum_k p_k^2 > \lambda^2 l^{-1} \xi^T \xi;$$

since $\lambda^2 l^{-1} > 1$ by (i.1), we conclude that the spectral norm of the matrix $\sum_i \delta_i F_i$ is > 1, whence either this matrix, or its negation is not $\preceq F_0 = I$. Thus, $y \notin \mathcal{X}_{\lambda \mathcal{D}}$, as claimed.

2^0. Now let (i.2) be the case. Let us set

$$\begin{array}{rcl} Q_k & = & N^{-1} I, \\ S_k & = & N \sum_{i \in \mathcal{I}_k} F_i^2. \end{array}$$

By the same reasons as in 1^0, there exists an n-dimensional vector ξ such that

$$\sum_{k=1}^{N} N \sum_{i \in \mathcal{I}_k} \|F_i \xi\|_2^2 > \xi^T \xi. \tag{10}$$

Denoting $e_1, \dots, e_n$ the standard basic orths in $\mathbf{R}^n$, we conclude that there exists $p \in \{1, \dots n\}$ such that

$$\sum_{k=1}^{N} N \sum_{i \in \mathcal{I}_k} |e_p^T F_i \xi|^2 > \frac{1}{n} \xi^T \xi. \tag{11}$$

We clearly can choose $\delta \in \lambda \mathcal{D}$ in such a way that

$$\sum_{i \in \mathcal{I}_k} \delta_i e_p^T F_i \xi = \lambda \sqrt{\sum_{i \in \mathcal{I}_k} |e_p^T F_i \xi|^2}.$$

Setting

$$F = \sum_i \delta_i F_i,$$

we get

$$\begin{array}{rcl} e_p^T F \xi & = & \sum_k \sum_{i \in \mathcal{I}_k} e_p^T \delta_i F_i \xi \\ & \geq & \lambda \sum_{k=1}^N \sqrt{\sum_{i \in \mathcal{I}_k} |e_p^T F_i \xi|^2} \\ & \geq & \lambda \sqrt{\sum_{k=1}^N \sum_{i \in \mathcal{I}_k} |e_p^T F_i \xi|^2} \\ & > & \frac{\lambda}{\sqrt{mN}} \|\xi\|_2 \quad \text{[see (11)]} \\ & > & \|\xi\|_2 \quad \text{[by (i.2)]} \end{array}$$

We conclude that the spectral norm of F is > 1, so that either F or $-F$ is not $\succeq F_0 = I$. Thus, $y \notin \mathcal{X}_{\lambda \mathcal{D}}$, as claimed. ∎

3 Rational Dependence

In this section, we seek to handle cases when the matrix-valued function $\mathbf{F}(y, \delta)$ is rational in δ for every y. There are many practical situations when the perturbation parameters enter rationally, and not affinely, in a perturbed SDP. One important example arises with the problem of checking *robust singularity* of a square (non symmetric) matrix $\mathbf{A}$, which depends affinely on parameters. One has to check if $\mathbf{A}^T \mathbf{A} \succ 0$ for every $\mathbf{A}$ in the affine uncertainty set; this is a matrix inequality condition in which the parameters enter quadratically.

We will introduce a versatile framework, called linear-fractional representations, for describing rational dependence, and devise approximate robust counterparts that are based on this linear-fractional representation. Our framework will of course cover the cases when the perturbation enters affinely in the matrix $\mathbf{F}(y, \delta)$, which is a case already covered by theorem 2.1. At present we do not know if theorem 2.1 always yields more accurate results than those described next, except in the case $N = 1$ (Euclidean-norm bounds), where both results are actually equivalent.

In this section, we take $\rho = 1$.

3.1 Linear-fractional representations

We assume that the function $\mathbf{F}$ is given by a "linear-fractional representation" (LFR):

$$\begin{array}{c} \mathbf{F}(y, \Delta) = F(y) + L(y)\Delta(I - D\Delta)^{-1} R + R^T (I - \Delta^T D^T)^{-1} \Delta^T L(y)^T, (a) \\ \Delta = \mathrm{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}\right), \end{array}$$

where $F(y)$ is defined in (1), $L(\cdot)$ is an affine mapping taking values in $\mathbf{R}^{n \times p}$, $R \in \mathbf{R}^{q \times n}$ and $D \in \mathbf{R}^{q \times p}$ are given matrices, and $r_1, \ldots, r_l$ are given integers. We assume that the above LFR is well-posed over $\mathcal{D}$, meaning that $\det(I - D\Delta)$ for every Δ of the form above, with $\delta \in \mathcal{D}$; we return to the issue of well-posedness later.

The above class of models seem quite specialized. In fact, these models can be used in a wide variety of situations. For example, in the case of affine dependence:

$$\mathbf{F}(\delta) = F^0(y) + \sum_{i=1}^l \delta_i F^i(y),$$

we can construct a linear-fractional representation, for example

$$L(y) = \frac{1}{\sqrt{2}} \begin{bmatrix} F^1(y) \\ \vdots \\ F^l(y) \end{bmatrix}^T, \quad R = \frac{1}{\sqrt{2}} \begin{bmatrix} I \\ \vdots \\ I \end{bmatrix}, \quad D = 0, \quad r_1 = \ldots = r_l = n. \tag{12}$$

Our framework also covers the case when the matrix $\mathbf{F}$ is rational in δ. The representation lemma [5], given below, illustrates this point.

Lemma 3.1 *For any rational matrix function $\mathbf{M} : \mathbf{R}^l \to \mathbf{R}^{n\times c}$, with no singularities at the origin, there exist nonnegative integers $r_1, \ldots, r_l$, and matrices $M \in \mathbf{R}^{n\times c}$, $L \in \mathbf{R}^{n\times N}$, $R \in \mathbf{R}^{N\times c}$, $D \in \mathbf{R}^{N\times N}$, with $N = r_1 + \ldots + r_l$, such that $\mathbf{M}$ has the following Linear-Fractional Representation (LFR): For all δ where $\mathbf{M}$ is defined,*

$$\mathbf{M}(\delta) = M + L\Delta \left(I - D\Delta\right)^{-1} R, \text{ where } \Delta = \operatorname{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}\right). \tag{13}$$

In the above construction, the sizes of the matrices involved are polynomial in the number l of parameters.

In the sequel, we denote by $\boldsymbol{\Delta}$ the set of matrices defined by

$$\boldsymbol{\Delta} = \left\{\Delta = \operatorname{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}\right) \;\middle|\; \delta \in \mathcal{D}\right\}. \tag{14}$$

The following developments are valid for a large class of matrix sets $\boldsymbol{\Delta}$.

3.2 Robustness analysis via Lagrange relaxations

The basic idea behind linear-fractional representation is to convert a a robustness condition such as

$$\xi^T \mathbf{F}(\Delta)\xi \geq 0 \text{ for every } \xi, \quad \Delta \in \boldsymbol{\Delta} \tag{15}$$

into a *quadratic* condition involving ξ and some additional variables p, q. Then, using Lagrange relaxation, we can obtain an SDP that yields a sufficient condition for robustness.

Using the LFR of $\mathbf{F}(\Delta)$, and assuming the latter is well-posed, we rewrite (15) as

$$\xi^T \left(F\xi + 2Lp\right) \geq 0, \text{ for every } \xi, p, q, \Delta \text{ such that } q = R\xi + Dp, \quad p = \Delta q, \quad \Delta \in \boldsymbol{\Delta}. \tag{16}$$

where p, q are additional variables. In the above, the only non convex condition is $p = \Delta q$, $\Delta \in \boldsymbol{\Delta}$. It turns out that for the set $\mathcal{D}$ defined in (3) (as well as for many other sets), we can obtain a necessary and sufficient condition for $p = \Delta q$ for some $\Delta \in \boldsymbol{\Delta}$, in the form of a linear matrix inequality on the rank-one matrix zz^T, where

$$z = \begin{bmatrix} q \\ p \end{bmatrix} = \begin{bmatrix} R & D \\ 0 & I \end{bmatrix} \begin{bmatrix} \xi \\ p \end{bmatrix}.$$

Let us characterize this linear matrix inequality as

$$\Phi_{\boldsymbol{\Delta}}(zz^T) \succeq 0,$$

where $\Phi_{\boldsymbol{\Delta}}$ is a linear map from $\mathcal{S}^{2N}$ to $\mathcal{S}^N$ (recall N denotes the row size of matrix Δ, and is the size of vectors p, q). In table 1, we show the mappings $\Phi_{\boldsymbol{\Delta}}$ associated with various sets $\boldsymbol{\Delta}$.

set $\mathbf{\Delta}$	condition on p,q equivalent to $p=\Delta q,\ \Delta\in\mathbf{\Delta}$	$\Phi_{\mathbf{\Delta}}(Z),$ for $Z\in\mathcal{S}^{2N}$
$\{\Delta \mid \|\Delta\|\le 1\}$	$q^Tq-p^Tp\ge 0$	$\mathrm{Tr}(Z_{22}-Z_{11})$
$\{sI \mid s\in\mathbf{C},\ \Re(s)\ge 0\}$	$pq^H+qp^H\succeq 0$	$Z_{12}+Z_{21}^H$
$\{\delta I \mid \delta\in\mathbf{R},\ \lvert\delta\rvert\le 1\}$	$qq^T-pp^T\succeq 0$	$Z_{22}-Z_{11}$
$\{\mathrm{diag}(\delta_i I_{r_i})_{i=1}^l \mid \delta\in\mathbf{R}^l,\ \|\delta\|_\infty\le 1\}$	$\mathrm{diag}(q_i^Tq_i-p_i^Tp_i)_{i=1}^l\succeq 0$	$\mathrm{diag}(Z_{22}(i)-Z_{11}(i))_{i=1}^l$
$\{\mathrm{diag}(\delta_i I_{r_i})_{i=1}^l \mid \delta\in\mathbf{R}^l,\ \|\delta\|_2\le 1\}$	$\mathrm{diag}(q_i^Tq_i)_{i=1}^l-pp^T\succeq 0$	$\mathrm{diag}(Z_{22}(i))_{i=1}^l-Z_{11}$

Table 1: Linear mappings $\Phi_{\mathbf{\Delta}}$ associated with various sets $\mathbf{\Delta}\subseteq\mathbf{R}^{N\times N}$. The notation Z_{kj}, $k,j\in\{1,2\}$, refers to corresponding $N\times N$ blocks in matrix $Z\in\mathcal{S}^{2N}$, and $Z_{jj}(i)$ denotes the i-the diagonal block of Z_{jj}, of size $r_i\times r_i$.

Using this equivalent condition, we rewrite (16) as

$$\begin{array}{c}\mathrm{Tr}\begin{bmatrix} F & L \\ L^T & 0\end{bmatrix}\begin{bmatrix}\xi\\ p\end{bmatrix}\begin{bmatrix}\xi\\ p\end{bmatrix}^T\ge 0, \text{ for every } \xi,p \text{ such that}\\ \Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} R & D\\ 0 & I\end{bmatrix}\begin{bmatrix}\xi\\ p\end{bmatrix}\begin{bmatrix}\xi\\ p\end{bmatrix}^T\begin{bmatrix} R & D\\ 0 & I\end{bmatrix}\right)\succeq 0.\end{array}\tag{17}$$

The next step is to relax the non convex condition on p,ξ using Lagrange relaxation. Previous condition is true if there exist a positive semidefinite matrix S such that for every ξ,p, we have

$$\begin{bmatrix}\xi\\ p\end{bmatrix}^T\begin{bmatrix} F & L \\ L^T & 0\end{bmatrix}\begin{bmatrix}\xi\\ p\end{bmatrix}\ge \mathrm{Tr}\, S\Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} R & D\\ 0 & I\end{bmatrix}\begin{bmatrix}\xi\\ p\end{bmatrix}\begin{bmatrix}\xi\\ p\end{bmatrix}^T\begin{bmatrix} R & D\\ 0 & I\end{bmatrix}\right),$$

or, equivalently,

$$S\succeq 0,\quad \begin{bmatrix} F & L \\ L^T & 0\end{bmatrix}\succeq\begin{bmatrix} R & D\\ 0 & I\end{bmatrix}^T\Phi_{\mathbf{\Delta}}^*(S)\begin{bmatrix} R & D\\ 0 & I\end{bmatrix},\tag{18}$$

where $\Phi_{\mathbf{\Delta}}^*$ is the dual of $\Phi_{\mathbf{\Delta}}$.

Note that the dual of the above LMI condition amounts to enforce condition (17) not only on rank-one matrices, but for every positive semidefinite matrix.

Theorem 3.1 *Consider the uncertain semidefinite program with rational perturbation, described by the LFR (3.1), where the perturbation vector lies in an arbitrary set $\mathcal{D}\subseteq\mathbf{R}^l$. Let $\mathbf{\Delta}$ be defined by (14), and assume the LFR is well-posed over $\mathbf{\Delta}$. Assume that we can associate to this set a linear mapping $\Phi_{\mathbf{\Delta}}$ such that*

$$p=\Delta q \text{ for some } \Delta\in\mathbf{\Delta} \text{ if and only if } \Phi_{\mathbf{\Delta}}\left(\begin{bmatrix} q\\ p\end{bmatrix}\begin{bmatrix} q\\ p\end{bmatrix}^T\right)\succeq 0,$$

and denote by $\Phi_{\mathbf{\Delta}}^*$ the dual of this map.

Then the semidefinite program

$$\max\ b^T y \ \textit{subject to}$$
$$S \succeq 0, \quad \begin{bmatrix} F(y) & L(y) \\ L(y)^T & 0 \end{bmatrix} \succeq \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}^T \Phi^*_{\mathbf{\Delta}}(S) \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}$$

in variables y, S, is an approximation of the robust counterpart (2), i.e., the projection of the feasible set of (2) on the space of y-variables is contained in the set of robust feasible solutions.

It is now a simple matter to specialize the above result for the set $\mathcal{D}$ defined in (3). The condition on p, q that is equivalent to $p = \Delta q$, $\Delta \in \mathbf{\Delta}$ writes

$$\mathrm{diag}(q_i q_i^T)_{i \in I_k} - p^k (p^k)^T \succeq 0 \quad k = 1, \ldots, N, \tag{19}$$

where I_k be the set of indices $\nu_{k-1}+1, \ldots, \nu_k$, with $\nu_0 = 0$, $\nu_k = \sum_{s=1}^k n_s$, and p^k is the vector with elements $(p_i)_{i \in I_k}$.

The following result is then a corollary of theorem 3.1.

Corollary 3.1 *Consider the uncertain semidefinite program with rational perturbation, described by the LFR (3.1), where the perturbation vector lies in the set $\mathcal{D}$ defined in (3). Let $\mathbf{\Delta}$ be defined by (14), and assume the LFR is well-posed over $\mathbf{\Delta}$.*

Consider the semidefinite program

$$\max\ b^T y \ \text{ subject to } S \succeq 0,$$
$$\begin{bmatrix} F(y) & L(y) \\ L(y)^T & 0 \end{bmatrix} \succeq \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}^T \begin{bmatrix} T & 0 \\ 0 & -S \end{bmatrix} \begin{bmatrix} R & D \\ 0 & I \end{bmatrix} \tag{20}$$

where $S = \mathrm{diag}(S_1, \ldots, S_N)$, with each S_i of size $\sum_{i \in I_k} r_i$, and T is the block-diagonal matrix formed with the block-diagonal $r_i \times r_i$ blocks of S.

Then the above semidefinite program in variables y, S, is an approximation of the robust counterpart (2), i.e., the projection of the feasible set of (2) on the space of y-variables is contained in the set of robust feasible solutions.

Remark 3.1 *We note that the above condition, if it is strictly enforced, ensures well-posedness, meaning that $\det(I - D\Delta) \neq 0$ for every $\Delta \in \mathbf{\Delta}$. (To prove this, it suffices to apply the previous methodology to the matrix function $\mathbf{F}(\Delta) = (I - D\Delta)^T (I - D\Delta)$.)*

3.3 Comparison with earlier results

We do not have a general comparison theorem with the results of section 2, in the case of affine dependence. However, we can prove that, when $\mathcal{D}$ represents Euclidean-norm bounds (that is, there is only one block: $N = 1$), then both results are equivalent.

Indeed, assume that $\mathbf{F}(\delta, y)$ has the form (4); a linear-fractional representation of this dependence is (3.1), with $F = F_0$, $D = 0$, and L, R given by (12). The linear matrix inequality (20) then involves a full matrix S of row size n^2, and writes (dropping the dependence on y, and exchanging L^T and R without loss of generality)

$$\begin{bmatrix} F_0 - \frac{1}{2}\sum_{i=1}^l F_i S_i F_i & \frac{1}{\sqrt{2}} I & \ldots & \frac{1}{\sqrt{2}} I \\ \frac{1}{\sqrt{2}} I & S_1 & * & * \\ \vdots & * & \ddots & * \\ \frac{1}{\sqrt{2}} I & * & * & S_l \end{bmatrix} \succeq 0,$$

where S_i are the $n \times n$ diagonal blocks of S, and the symbols $*$ refer to the other blocks of S. Using the elimination lemma [4], it is possible to get rid of these elements and rewrite the above as

$$\begin{bmatrix} F_0 - \frac{1}{2}\sum_{i=1} F_i S_i F_i & \frac{1}{\sqrt{2}} I \\ \frac{1}{\sqrt{2}} I & S_k \end{bmatrix} \succeq 0, \quad k = 1, \ldots, l.$$

Assuming (without loss of generality) that each S_k is positive definite, and setting

$$Q = 2F_0 - \sum_{i=1}^{l} F_i S_i F_i,$$

we get $Q \succeq S_k^{-1}$ for every k, and hence

$$2F_0 \succeq Q + \sum_{i=1}^{l} F_i Q^{-1} F_i,$$

which is precisely the result obtained from theorem 2.1 , in the case $N = 1$.

4 Special cases

In this section, we focus on several special cases when the above results yield "computationally tractable" *equivalent* forms of the robust counterpart rather than merely "tractable approximations" of it.

4.1 Linear programming with affine uncertainty

Linear programming can be treated as a very special case of Semidefinite Programming; here all considerations related to robust counterpart become especially simple. For self-contained derivation of the below results, see [2].

Consider an LP program in the form

$$\min\ c^T x \quad \text{subject to } Ax + b \geq 0 \tag{21}$$

"Simple" ellipsoidal uncertainty. Let us start with the case when the data $[A, b] \in \mathbf{R}^{m \times (n+1)}$ in (21) are affinely parameterized by perturbation vector varying in an "elliptic cylinder" – the direct sum of an ellipsoid and a linear space:

$$[A; b] \in \mathcal{U} = \left\{ [A^0; b^0] + \sum_{j=1}^{k} \xi_j [A^j; b^j] + \sum_{p=1}^{q} \zeta_p [C^p; d^p] : (\xi, \zeta) \in \mathcal{D} = \{\xi^T \xi \leq 1\} \right\}. \tag{22}$$

For this case the robust counterpart

$$\min\ c^T x \quad \text{subject to } Ax + b \geq 0 \quad \text{for all } [A; b] \in \mathcal{U} \tag{23}$$

of (21) is the program

$$\begin{array}{ll} & \min\ c^T x \\ \text{s.t.} & \\ & \begin{array}{rcl} C_i^p x + d_i^p & = & 0,\ p = 1, \ldots, q, \\ & & i = 1, \ldots, m; \\ A_i^0 x + b_i^0 & \geq & \sqrt{\sum_{j=1}^{k} (A_i^j x + b_i^j)^2}, \\ & & i = 1, 2, \ldots, m; \end{array} \end{array} \tag{24}$$

here B_i denotes i-th row (treated as a row vector) of a matrix B.

Case of $\bigcap$-ellipsoidal uncertainty. Now assume that the set $\mathcal{U}$ possible values of the data $[A;b]$ of uncertain LP problem (21) is intersection of finitely many sets of the form (22):

$$\begin{array}{rcl} [A;b]\in\mathcal{U} & = & \bigcap\limits_{s=0}^{t}\mathcal{U}_s, \\ \mathcal{U}_s & = & \{[A^{s0};b^{s0}]+\sum_{j=1}^{k_s}\xi_j[A^{sj};b^{sj}]+\sum_{p=1}^{q_s}\zeta_p[C^{sp};d^{sp}] \mid \\ & & \xi^T\xi\le 1\}, \\ & & s=0,1,\dots,t. \end{array} \tag{25}$$

Assume also that the set $\mathcal{U}$ is bounded and the ellipsoids $\mathcal{U}_s$ satisfy the following "Slater condition":

(*) *There exists $[A';b']\in\mathcal{U}$ which, for every $s\le t$, can be represented as*

$$[A;b]=[A^{s0};b^{s0}]+\sum_{j=1}^{k_s}\xi_j^s[A^{sj};b^{sj}]+\sum_{p=1}^{q_s}\zeta_p^j[C^{sp};d^{sp}]$$

with $[\xi^s]^T\xi^s<1$.

For this case the robust counterpart (23) of the uncertain LP program (21) is the program

$$\min\ c^Tx$$

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

$$(a):\qquad \mathrm{Tr}([C^{sp}]^T\lambda^{is})+[d^{sp}]^T\mu^{is} \;=\; 0,\ 1\le i\le m, 1\le s\le t, 1\le p\le q_s;$$

$$(b):\quad \sum_{s=1}^{t}\left(\mathrm{Tr}([C^{0p}]^T\lambda^{is})+[d^{0p}]^T\mu^{is}\right) \;=\; C_i^{0p}x+d_i^{0p},\ 1\le i\le m,\ 1\le p\le q_0;$$

$$(c):\qquad \sum_{s=1}^{t}\left\{\mathrm{Tr}([A^{s0}-A^{00}]^T)\lambda^{is}+[b^{s0}-b^{00}]^T\mu^{is}\right\}+A_i^{00}x+b_i^{00}$$

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

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

It is worthy of mentioning that the case of $\bigcap$-ellipsoidal uncertainty basically covers the case of affine perturbations with ellipsoidal uncertainty (4) – (3). Indeed, assume that the affine mapping (4) from the space of perturbation vectors to the space of data is an embedding. Since the set $\mathcal{D}$ given by (3) clearly is an intersection of elliptic cylinders, its image under the above embedding – i.e., the set of possible values of the perturbed data – is a $\bigcap$-ellipsoidal set. Note that this set clearly satisfies the "Slater condition" (*).

4.2 Robust quadratic programming with affine uncertainty

A (convex) quadratic problem is an optimization program of the form

$$\min\ c^T x \quad \text{subject to } -x^T[A^i]^T[A^i]x + 2[b^i]^T x + \gamma^i \geq 0,\ i = 1, \ldots, m; \tag{26}$$

such a problem can be easily reformulated as an SDP program with the data affinely depending on the data $\{A^i, b^i, \gamma^i\}_{i=1}^m$ of the original problem. Assume that the data (A^i, b^i, γ^i) of every quadratic constraint are uncertain and vary in respective ellipsoids:

$$(A^i, b^i, \gamma^i) \in \mathcal{U}_i = \left\{ (A^i, b^i, \gamma^i) = (A^{i0}, b^{i0}, \gamma^{i0}) + \sum_{j=1}^{k} \delta_j (A^{ij}, b^{ij}, \gamma^{ij}) \mid \delta^T \delta \leq 1 \right\}. \tag{27}$$

It turns out that the robust counterpart of uncertain conic quadratic program (26) – (27) is *equivalent* to the explicit semidefinite program as follows:

$$\begin{array}{ll}
 & \min\ c^T x \\
\text{s.t.} & \left[\begin{array}{c|cccc|c}
\gamma^{i0} + 2x^T b^{i0} - \lambda^i & \frac{\gamma^{i1}}{2} + x^T b^{i1} & \frac{\gamma^{i2}}{2} + x^T b^{i2} & \cdots & \frac{\gamma^{ik}}{2} + x^T b^{ik} & [A^{i0}x]^T \\
\hline
\frac{\gamma^{i1}}{2} + x^T b^{i1} & \lambda^i & & & & [A^{i1}x]^T \\
\frac{\gamma^{i2}}{2} + x^T b^{i2} & & \lambda^i & & & [A^{i2}x]^T \\
\vdots & & & \ddots & & \cdots \\
\frac{\gamma^{ik}}{2} + x^T b^{ik} & & & & \lambda^i & [A^{ik}x]^T \\
\hline
A^{i0}x & A^{i1}x & A^{i2}x & \cdots & A^{ik}x & I_{l_i}
\end{array}\right] \succeq 0, \\
 & i = 1, 2, \ldots, m
\end{array} \tag{28}$$

with variables x and additional scalar variables $\lambda^1, \ldots, \lambda^m$, I_l being the unit $l \times l$ matrix. The result may be derived from Theorem 4.1 (for independent proof, see [3]).

4.3 Robust conic quadratic programming

A conic quadratic program is an optimization program of the form

$$\min\ c^T x \quad \text{subject to } \left\| A^i x + b^i \right\|_2 \leq [d^i]^T x + \gamma^i,\ i = 1, 2, \ldots, m\}. \tag{29}$$

such a problem can be easily reformulated as a semidefinite program with the data affinely depending on the data $\{A^i, b^i, d^i, \gamma^i\}_{i=1}^m$ of the original problem. Assume that the data of (29) are uncertain and that the uncertainty is of the following specific type:

- (I) The uncertainty is "constraint-wise": the data $(A^i, b^i, d^i, \gamma^i)$ of different conic quadratic constraint independently of each other run through respective uncertainty sets $\mathcal{U}_i$;
- (II) For every i, $\mathcal{U}_i$ is the direct product of two elliptic cylinders in the spaces of (A^i, b^i)- and (d^i, γ^i)-components of the data:

$$\begin{array}{rcl}
\mathcal{U}_i & = & \mathcal{V}_i \times \mathcal{W}_i, \\
\mathcal{V}_i & = & \{[A^i; b^i] = [A^{i0}; b^{i0}] + \sum_{j=1}^{k_i} \xi_j [A^{ij}; b^{ij}] \\
 & & + \sum_{p=1}^{q_i} \zeta_p [E^{ip}; f^{ip}] \mid \xi^T \xi \leq 1\}, \\
\mathcal{W}_i & = & \{(d^i, \gamma^i) = (d^{i0}, \gamma^{i0}) + \sum_{j=1}^{k_i'} \xi_j (d^{ij}, \gamma^{ij}) \\
 & & + \sum_{p=1}^{q_i'} \zeta_p (g^{ip}, h^{ip}) \mid \xi^T \xi \leq 1\}.
\end{array} \tag{30}$$

It turns out that the robust counterpart of uncertain conic quadratic program (29) – (30) is *equivalent* to the explicit semidefinite program as follows:

$$
\begin{array}{ll}
& \min\ c^T x \\
\text{s.t.} & \\
& E^{ip}x + f^{ip} = 0,\ i = 1,\dots,m, p = 1,\dots,q_i; \\
& [g^{ip}]^T x + h^{ip} = 0,\ i = 1,\dots,m, p = 1,\dots,q_i'; \\
& \begin{bmatrix}
[d^{i0}]^T x + \gamma^{i0} - \lambda^i & [d^{i1}]^T x + \gamma^{i1} & [d^{i2}]^T x + \gamma^{i2} & \cdots & [d^{ik_i'}]^T x + \gamma^{ik_i'} \\
[d^{i1}]^T x + \gamma^{i1} & [d^{i0}]^T x + \gamma^{i0} - \lambda^i & & & \\
[d^{i2}]^T x + \gamma^{i2} & & [d^{i0}]^T x + \gamma^{i0} - \lambda^i & & \\
\vdots & & & \ddots & \cdots \\
[d^{ik_i'}]^T x + \gamma^{ik_i'} & & & \cdots & [d^{i0}]^T x + \gamma^{i0} - \lambda^i
\end{bmatrix} \succeq 0, \\
& i = 1,\dots,m; \\
& \left[\begin{array}{c|cccc|c}
\lambda^i - \mu^i & & & \cdots & & [A^{i0}x + b^{i0}]^T \\
\hline
 & \mu^i & & & & [A^{i1}x + b^{i1}]^T \\
 & & \mu^i & & & [A^{i2}x + b^{i2}]^T \\
\vdots & & & \ddots & & \cdots \\
 & & & & \mu^i & [A^{ik_i}x + b^{ik_i}]^T \\
\hline
A^{i0}x + b^{i0} & A^{i1}x + b^{i1} & A^{i2}x + b^{i2} & \cdots & A^{ik_i}x + b^{ik_i} & \lambda^i I_{l_i}
\end{array}\right] \succeq 0, \\
& i = 1,\dots,m
\end{array}
\tag{31}
$$

with variables x and additional scalar variables λ^i, μ^i, $i = 1,\dots,m$.

The result again can be derived from Theorem 4.1 (for independent derivation, see [3]).

4.4 Operator-norm bounds

In several problems, the perturbed constraint writes as $\mathbf{F}(y,\Delta) \succeq 0$, where $\mathbf{F}(y,\Delta)$ is given in the linear-fractional form (3.1.a), and Δ is a matrix bounded in norm but otherwise arbitrary. We can handle this case with the Lagrange relaxation technique described in section 3.2; the set $\mathbf{\Delta}$ has the form

$$\mathbf{\Delta} = \left\{ \Delta \in \mathbf{R}^{n_p \times n_q} \mid \|\Delta\| \le 1 \right\}, \tag{32}$$

and the corresponding linear map $\Phi_{\mathbf{\Delta}}$ is given in table 1. It turns out that the resulting SDP approximation is exact in this case (this result is proved in [15]).

Theorem 4.1 *Consider the uncertain semidefinite program with rational perturbation, described by the LFR (3.1), where the perturbation matrix is bounded but otherwise arbitrary: $\Delta \in \mathbf{\Delta}$, where $\mathbf{\Delta}$ is defined in (32).*

Then the semidefinite program

$$\max\ b^T y \text{ subject to}$$

$$\tau \ge 0,\ \begin{bmatrix} F(y) & L(y) \\ L(y)^T & 0 \end{bmatrix} \succeq \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}^T \begin{bmatrix} \tau I_{n_q} & 0 \\ 0 & -\tau I_{n_p} \end{bmatrix} \begin{bmatrix} R & D \\ 0 & I \end{bmatrix}$$

in variables y, τ, is equivalent to the robust counterpart (2).

Remark 4.1 *We note again that the above condition is strictly feasible if and only if the LFR is well-posed, meaning that* $\det(I - D\Delta) \neq 0$ *for every* $\Delta \in \mathbf{R}^{n_p \times n_q}$, $\|\Delta\| \le 1$.

5 Examples

5.1 A link with combinatorial optimization

The method we have outlined is a way to solve a non convex optimization problem. It turns out that this method is similar in spirit to the one used in SDP relaxations for combinatorial optimization.

Consider the problem

$$\max_{\delta \in \mathbf{R}^l} \delta^T W \delta \text{ subject to } \delta_i^2 = 1, \quad i = 1, \ldots, l, \tag{33}$$

where W is a given symmetric matrix (of special structure, irrelevant here). The above problem is known in the combinatorial optimization literature as "the maximum cut" (MAX-CUT) problem [7], and is proven to be NP-hard.

This problem is a robust SDP problem. First we note that, without loss of generality, we may assume $W \succ 0$, and rewrite the problem as

$$\text{minimize } x \text{ subject to } \delta^T W_c \delta \leq x \text{ for every } \delta, \ \|\delta\|_\infty \leq 1.$$

Let us now apply theorem 3.1, with

$$\mathbf{F}(y, \Delta) = \begin{bmatrix} x & \delta^T \\ \delta & W^{-1} \end{bmatrix}, \quad \Delta = \mathbf{diag}(\delta_1, \ldots, \delta_l).$$

The matrix $\mathbf{F}(y, \Delta)$ can be written in the LFR format as (3.1), with

$$F(y) = \begin{bmatrix} x & 0 \\ 0 & W^{-1} \end{bmatrix}, \quad R = \left[\begin{array}{c|c} 1 & \\ \vdots & 0 \\ 1 & \end{array}\right], \quad L = \begin{bmatrix} 0 \\ I \end{bmatrix}, \quad D = 0.$$

The perturbation set $\mathbf{\Delta}$ here is the set of diagonal $p \times p$ matrices. Theorem 3.1 shows that an upper bound on the MAX-CUT problem is given by the optimal value of the SDP

$$\min \operatorname{Tr} S \text{ subject to } W \preceq S, \ S \text{ diagonal}.$$

This upper bound is exactly the one obtained by Lovasz [10], and is dual (in the SDP sense) to the one obtained by Goemans and Williamson [7]. As shown in [7], the above relaxation is the most efficient currently available (in terms of closeness to the actual optimum, in a certain stochastic sense).

5.2 A link with Lyapunov theory in control

Semidefinite programming has many applications in control theory [4]. The basic idea is that using so-called quadratic Lyapunov functions, we may prove a number of interesting properties for uncertain dynamical systems; the search for quadratic Lyapunov functions $V(\xi) = \xi^T S \xi$ can be often written as an SDP in the matrix S.

Here we would like to show the link between classical Lyapunov theory and the Lagrange relaxations we used in section 3.2. To illustrate this link, consider the problem of checking stability (convergence to zero of every trajectory) of the dynamical system

$$\dot{q} = Aq,$$

where $q \in \mathbf{R}^n$ is the state, and A is a (constant) square matrix. Taking the Laplace transform, we obtain that stability is equivalent to matrix A having no eigenvalues with positive real part:

$$(sI - A)^H (sI - A) \succ 0 \text{ for every } s, \;\; s + s^* \geq 0.$$

The above is a robustness analysis problem (the Laplace variable s is the uncertainty). Introduce $p = sq$, we have

$$p = sq \text{ for some } s, \;\; s + s^* \geq 0$$

if and only if $pq^H + qp^H \succeq 0$. Our problem is thus to check if

$$\|p - Aq\|^2 > 0 \text{ for every } (p, q) \neq (0, 0), \;\; pq^H + qp^H \succeq 0.$$

Using Lagrange relaxation of the last matrix inequality constraint, we obtain a sufficient condition for stability: There exists a real matrix S such that

$$A^T S + SA \prec 0, \;\; S \succ 0.$$

The above condition is the Lyapunov condition for stability that is well known in control (it turns out that this condition is necessary and sufficient if A is known and constant). If S satisfies this inequality, the quadratic function $V(\xi) = \xi^T S \xi$ can be interpreted as a Lypaunov function proving stability (that is, V decreases along every trajectory). The above is easily extended to the case when the matrix A is uncertain (see [4]).

5.3 Interval computations

A basic problem in interval computations is the following. We are given a function $\mathbf{f}$ from $\mathbf{R}^l$ to $\mathbf{R}^m$, and a set confidence $\mathcal{D}$ for $\delta \in \mathbf{R}^l$, in the form of a product of intervals. We seek to estimate intervals of confidence for the components of $x = \mathbf{f}(\delta)$ when δ ranges $\mathcal{D}$. Sometimes, $\mathbf{f}$ is given in implicit form, as in the interval linear algebra problem: here, we are given matrices $\mathbf{A} \in \mathbf{R}^{n \times n}$, $\mathbf{b} \in \mathbf{R}^n$ the elements of which are only known within intervals; in other words, $[\mathbf{A}\ \mathbf{b}]$ is only known to belong to an "interval matrix set" $\mathcal{U}$. we seek to compute *intervals of confidence* for the set of solutions, if any, to the equation $\mathbf{A}x = \mathbf{b}$.

Obtaining exact estimates for intervals of confidence for the elements of solutions x, even for the "linear interval algebra" problem, is already NP-hard [17, 18].

One classical approach to this problem resorts to interval calculus, where each one of the basic operations $(+, -, x, /)$ is replaced by an "interval counterpart", and standard (eg LU) linear algebra algorithms are adapted to this new "algebra". Many refinements of this basic idea have been proposed, but the algorithms based on this idea have in general exponential complexity.

Robust semidefinite programming can be used (at least as a subproblem in a global branch and bound method) for this problem, as follows. Assume we can describe $\mathbf{f}$ explicitely as a rational function of its arguments; from lemma 3.1, we can construct (in polynomial time) a linear-fractional representation of $\mathbf{f}$, in the form

$$\mathbf{f}(\delta) = f + L\Delta \left(I - D\Delta\right)^{-1} r, \text{ where } \Delta = \operatorname{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_l I_{r_l}\right).$$

Assume first that we seek an ellipsoid of confidence for the solution, in the form $\mathcal{E} = \{x \mid (x - x_0)(x - x_0)^T \preceq P\}$, where $x_0 \in \mathbf{R}^n$ and $P \succeq 0$ (our parametrization allows for degenerate, "flat", ellipsoids, to handle cases when some components of the solution are certain). We seek to minimize the "size" of $\mathcal{E}$ subject to $\mathbf{f}(\delta) \in \mathcal{E}$ for every $\delta \in \mathcal{D}$. Measuring the size of $\mathcal{E}$ by $\operatorname{Tr} P$ (other measures are possible, as seen below), we obtain the following equivalent formulation of the problem.

$$\min_{x_0, P} \operatorname{Tr} P \text{ subject to } \begin{bmatrix} P & (\mathbf{f}(\delta) - x_0) \\ (\mathbf{f}(\delta) - x_0)^T & 1 \end{bmatrix} \succeq 0 \text{ for every } \delta \in \mathcal{D}. \tag{34}$$

The above is obviously a robust semidefinite programming problem, for which an explicit SDP counterpart (approximation) can be devised, provided $\mathcal{D}$ takes the form of a (general) ellipsoidal set. (A typical set arising in interval calculus is a product of intervals $\Pi[\underline{\delta}_i \ \overline{\delta}_i]$, where $\underline{\delta}_i$, $\overline{\delta}_i$ are given.)

The above method finds ellipsoids of confidence, but it is also possible to find intervals of confidence for the components of $\mathbf{f}(\delta)$, by modifying the objective of the above robust SDP suitably (for example, if we minimize the $(1,1)$ component of the matrix variable P instead of its trace, we will obtain an interval of confidence for the first component of $\mathbf{f}(\delta)$, when δ ranges $\mathcal{D}$).

The resulting approximations have an interesting interpretation in the context of the "linear interval algebra problem" $\mathbf{A}x = \mathbf{b}$, where $[\mathbf{A} \ \mathbf{b}]$ is an uncertain matrix, subject to "unstructured perturbations". Assume

$$[\mathbf{A} \ \mathbf{b}] \in \mathcal{U} = \{[A + \Delta A \ b + \Delta b] \ \mid \ \|[\Delta A \ \Delta b]\| \leq \rho\} ,$$

where $[A \ b] \in \mathbf{R}^{n \times (m+1)}$ and $\rho \geq 0$ are given. In this case, our results are exact, and yield a solution related to the notion of *total least squares* developed by Golub and Van Loan [8, 21]. Precisely, it can be shown that the center of the ellipsoid of confidence (corresponding to the variable x_0 in problem (34)) is of the form

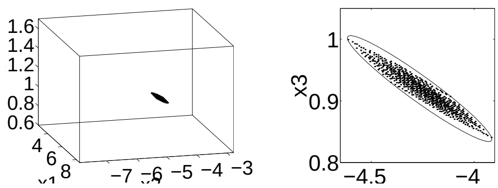
$$x_0 = (A^T A - \rho^2 I)^{-1} A^T b$$

(We assume that $\sigma_{\min}([A \ b]) \geq \rho$, otherwise the ellipsoid of confidence is unbounded. Except in degenerate cases, this guarantees the existence of the inverse in the above.) When we let $\rho = \sigma_{\min}([A \ b])$, the ellipsoid of confidence can be shown to be reduced to the singleton $\mathcal{E} = \{x_0\}$, and x_0 is the "total least squares" solution to the problem $Ax = b$.

As an example, consider the Vandermonde system

$$\begin{bmatrix} 1 & \mathbf{a}_1 & \mathbf{a}_1^2 \\ 1 & \mathbf{a}_1 & \mathbf{a}_1^2 \\ 1 & \mathbf{a}_1 & \mathbf{a}_1^2 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} = \begin{bmatrix} \mathbf{b}_1 \\ \mathbf{b}_2 \\ \mathbf{b}_3 \end{bmatrix},$$

where $\mathbf{a} \ \mathbf{b}$ are interval vectors of $\mathbf{R}^3$.



interval calculus *ellipsoid calculus*

Figure 1: Sets of confidence for an uncertain Vandermonde system.

In Figure 5.3, we show the box of confidence for the solution, computed by direct application of interval algebra; the right-hand side plots shows the ellipsoid of confidence obtained by robust semidefinite programming. We did not use elaborate algorithms to solve the problem via interval algebra, so the reader should not draw negative conclusions about it; rather, the instructive part is that the robust SDP method seems to behave well in this example.

5.4 Robust structural design

A typical problem of (static) structural design is to specify a mechanical construction capable best of all withstand a given external load. As a concrete example of this type, consider the *Truss Topology Design* (TTD) problem (for more details, see [1]).

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. Mathematically the TTD problem can be modeled by the following semidefinite program:

$$\begin{array}{lrcl} & & \min \tau & \\ \text{s.t.} & & & \\ (a) & \begin{bmatrix} \tau & f^T \\ f & \sum_{i=1}^n t_i b_i b_i^T \end{bmatrix} & \succeq & 0, \\ (b) & t & \in & P \subset \mathbf{R}^n_+, \end{array} \qquad (35)$$

with design variables $\tau \in \mathbf{R}$ and $t = (t_1, \ldots, t_n) \in \mathbf{R}^n$; t_i's are volumes of tentative bars. The data of the problem are

- vectors $b_i \in \mathbf{R}^m$; they are readily given by the geometry of the nodal set;
- vector $f \in \mathbf{R}^m$ representing the external load;
- a polytope P representing design restrictions like upper bound on the total bar volume, bounds on volumes of particular bars, etc.

In reality, the external load f should be treated as uncertain element of the data; the traditional approach to treat this uncertainty is to consider a number of "loads of interest" $f_1, \ldots, f_k$ and to optimize the worst-case, over this set of scenarios, compliance. Mathematically this approach is equivalent to replacing the LMI (35.a) (expressing the fact that τ is an upper bound on the compliance with respect to f) by k similar constraints corresponding to $f = f_1, f = f_2, \ldots, f = f_k$. A disadvantage of the "scenario approach" is that it takes care just of a restricted number of "loads of interest" and ignores "occasional" loads, even small ones; as a result, there is a risk that the resulting construction will be crushed by a small "bad" load. An example of this type is depicted on Fig. 1. 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 "nominal" optimal compliance is 1. It turns out, however, that the construction in question is highly instable: 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 3,000 times larger than the nominal one.

In order to improve design's stability, it makes sense to treat the load as uncertain element of the data varying through a "massive" uncertainty set rather than taking just a small number of "values of interest". From the mathematical viewpoint, it is convenient to deal with uncertainty set in the form of an ellipsoid centered at the origin:

$$f \in \mathcal{F} = \left\{ f = L\delta \mid \delta \in \mathcal{D} = \{\delta \in \mathbf{R}^k : \delta^T\delta \le 1\} \right\}. \tag{36}$$

Problem (35) with perturbation set given by (36) is a particular case of full matrix uncertainty (32); according to Theorem 4.1, the robust counterpart of (35) – (36) is *equivalent* to an explicit semidefinite program; this program can be finally converted to the form

$$\begin{array}{rcl} & \min \tau & \\ \text{s.t.} & & \\ \begin{bmatrix} \tau I & Q^T \\ Q & \sum_{i=1}^n t_i b_i b_i^T \end{bmatrix} & \succeq & 0, \\ t & \in & P. \end{array} \tag{37}$$

(for details, see [1]).

To illustrate the potential of the outlined approach, let us come back to the above "cantilever arm" example. In this 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. Let us pass from the nominal problem ("a singleton uncertainty set $\mathcal{F} = \{f\}$") to the problem with $\mathcal{F}$ being a "massive" ellipsoid, namely, the ellipsoid of the smallest volume containing the nominal load f and a 20-dimensional ball $B_{0.1}$ comprised of all 20-dimensional vectors ("occasional loads") of the Euclidean norm $\le 0.1\|f\|_2$.

Solving the robust counterpart (37) of the resulting uncertain SDP, we get the cantilever arm shown on Fig. 1.b).

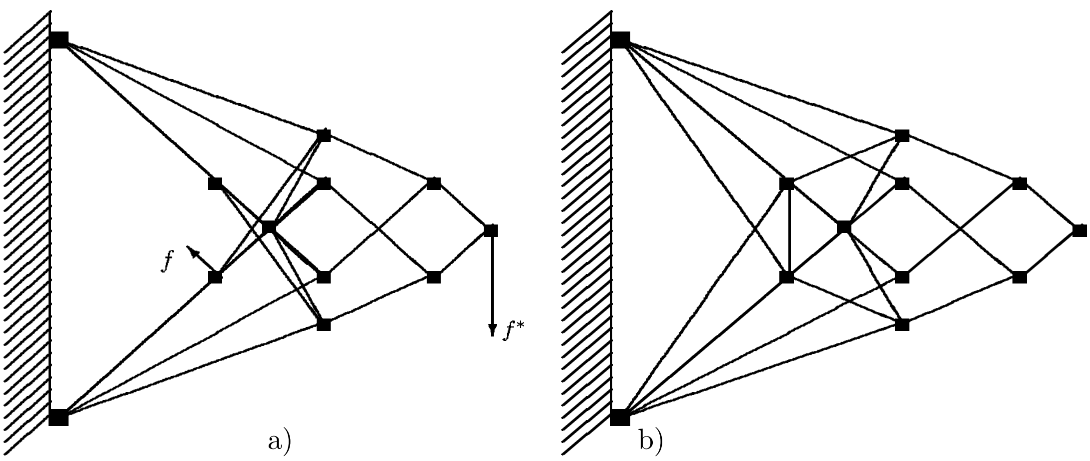


Figure 1. Cantilever arm: nominal design (left) and robust design (right)

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. $B_{0.1}$
nominal	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.

6 Concluding Remarks

We have described a general methodology to handle deterministic uncertainty in semidefinite programming, which computes robust solutions via semidefinite programming. The method handles very general (nonlinear) uncertainty structures, and uses a special Lagrange relaxation (or, in a dual form, a "rank relaxation") to obtain the approximate robust counterpart in the form of an SDP. The method is actually an extension of techniques that are well-known in several (apparently unrelated) areas, such as rank relaxations in combinatorial optimization, or Lyapunov functions in control.

In the case of affine dependence, we can estimate the quality of the resulting approximation; in some other cases, the approximation is exact.

Further work should probably concentrate on reducing the level of conservativeness as much as possible, while keeping the size of the approximate robust counterpart reasonable.

Acknowledgements

Several people—especially the Editors of the handbook!

References

[1] Ben-Tal, A. and Nemirovski, A. "Robust Truss Topology Design via Semidefinite Programming", Research Report # 4/95 (August 1995), Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion – The Israel Institute of Technology, Technion City, Haifa 32000, Israel; to appear in *SIAM Journal of Optimization*.

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

[3] Ben-Tal, A., and Nemirovski, A. "Robust Convex Programming" – Working paper (October 1995), Faculty of Industrial Engineering and Management, Technion – the Israel Institute of Technology, Technion City, Haifa 32000, Israel (to appear in *Mathematics of Operations Research*).

[4] Boyd, S., El Ghaoui, L., Feron, E., and Balakrishnan, V. *Linear Matrix Inequalities in System and Control Theory* – SIAM, Philadelphia, 1994.

[5] El-Ghaoui, L., and Lebret, H. "Robust solutions to least-squares problems with uncertain data matrices" – SIAM J. Matrix Anal. Appl., October 1997.

[6] Falk, J.E. "Exact Solutions to Inexact Linear Programs" – *Operations Research* (1976), pp. 783-787.

[7] Goemans, M. X., and Williamson, D. P. ".878-approximation for MAX CUT and MAX 2SAT" – In: *Proc.* 26^{th} *ACM Symp. Theor. Computing*, 1994, pp. 422-431.

[8] Golub, G.H., and Van Loan, C.F. *Matrix Computations.* John Hopkins University Press, 1996.

[9] Hansen, E.R. *Global optimization using interval analysis.* Marcel Dekker, NY, 1992.

[10] Lovasz, L. "On the Shannon capacity of a graph" – *IEEE Transactions on Information Theory* v. 25 (1979), pp. 355-381.

[11] Kosheler, M., and Kreinovitch, V. Interval computations web site `http://cs.utep.edu/interval-comp/main.html`, 1996.

[12] Moore, E.R. *Methods and applications of interval analysis.* SIAM, Philadelphia, 1979.

[13] Mulvey, J.M., Vanderbei, R.J. and Zenios, S.A. "Robust optimization of large-scale systems", *Operations Research* **43** (1995), 264-281.

[14] Nesterov, Yu., and Nemirovski, A. *Interior point polynomial methods in Convex Programming*, SIAM Series in Applied Mathematics, Philadelphia, 1994.

[15] Oustry, H., El Ghaoui, L., and Lebret, H. "Robust solutions to uncertain semidefinite programs", To appear in *SIAM J. of Optimization*, 1998.

[16] Rockafellar, R.T. *Convex Analysis*, Princeton University Press, 1970.

[17] Rohn, J. "Systems of linear interval equations" – *Linear Algebra and its Applications,* v. 126 (1989), pp. 39-78.

[18] Rohn, J. "Overestimations in bounding solutions of perturbed linear equations" – *Linear Algebra and its Applications,* v. 262 (1997), pp. 55-66.

[19] Singh, C. "Convex Programming with Set-Inclusive Constraints and its Applications to Generalized Linear and Fractional Programming" – *Journal of Optimization Theory and Applications*, v. 38 (1982), No. 1, pp. 33-42.

[20] Soyster, A.L. "Convex Programming with Set-Inclusive Constraints and Applications to Inexact Linear Programming" - *Operations Research* (1973), pp. 1154-1157.

[21] Van Huffel, S., and Vandewalle, J. *The total least squares problem: computational aspects and analysis.* – v. 9 of *Frontiers in applied Mathematics*, SIAM, Philadelphia, 1991.

[22] Zhou, K., Doyle, J., and Glover, K. *Robust and Optimal Control* – Prentice Hall, New Jersey, 1995.