SIAM J. OPTIM.
Vol. 9, No. 1, pp. 33–52
© 1998 Society for Industrial and Applied Mathematics

ROBUST SOLUTIONS TO UNCERTAIN SEMIDEFINITE PROGRAMS*

LAURENT EL GHAOUI†, FRANCOIS OUSTRY†, AND HERVÉ LEBRET†

Abstract. In this paper we consider semidefinite programs (SDPs) whose data depend on some unknown but bounded perturbation parameters. We seek "robust" solutions to such programs, that is, solutions which minimize the (worst-case) objective while satisfying the constraints for every possible value of parameters within the given bounds. Assuming the data matrices are rational functions of the perturbation parameters, we show how to formulate sufficient conditions for a robust solution to exist as SDPs. When the perturbation is "full," our conditions are necessary and sufficient. In this case, we provide sufficient conditions which guarantee that the robust solution is unique and continuous (Hölder-stable) with respect to the unperturbed problem's data. The approach can thus be used to regularize ill-conditioned SDPs. We illustrate our results with examples taken from linear programming, maximum norm minimization, polynomial interpolation, and integer programming.

Key words. convex optimization, semidefinite programming, uncertainty, robustness, regularization

AMS subject classifications. 93B35, 49M45, 90C31, 93B60

PII. S1052623496305717

Notation. For a matrix X, $\|X\|$ denotes the largest singular value. If X is square, $X \succeq 0$ (resp., $X \succ 0$) means X is symmetric and positive semidefinite (resp., definite). For $X \succeq 0$, $X^{1/2}$ denotes the symmetric square root of X. The notation I_p denotes the $p \times p$ identity matrix; the subscript is omitted when it can be inferred from context.

1. Introduction.

A semidefinite program (SDP) consists of minimizing a linear objective under a linear matrix inequality (LMI) constraint; precisely,

$$\mathcal{P}_0: \quad \text{minimize } c^T x \text{ subject to } F(x) = F_0 + \sum_{i=1}^{m} x_i F_i \succeq 0, \tag{1}$$

where $c \in \mathbf{R}^m - \{0\}$ and the symmetric matrices $F_i = F_i^T \in \mathbf{R}^{n\times n}, i = 0, \ldots, m$, are given. SDPs are convex optimization problems and can be solved in polynomial time with, e.g., primal-dual interior-point methods [24, 35, 26, 19, 2]. SDPs include linear programs and convex quadratically constrained quadratic programs, and arise in a wide range of engineering applications; see, e.g., [12, 1, 35, 22].

In the SDP (1), the "data" consist of the objective vector c and the matrices $F_0, \ldots, F_m$. In practice, these data are subject to uncertainty. An extensive body of work has concentrated on the sensitivity issue, in which the perturbations are assumed to be infinitesimal, and regularity of optimal values and solution(s), as functions of the data matrices, is studied. Recent references on sensitivity analysis include [30, 31, 10] for general nonlinear programs, [33] for semi-infinite programs, and [32] for semidefinite programs.

When the perturbation affecting the data of the problem is not necessarily small, a sensitivity analysis is not sufficient. For general optimization problems, a whole field

*Received by the editors June 21, 1996; accepted for publication (in revised form) September 22, 1997; published electronically October 30, 1998.
http://www.siam.org/journals/siopt/9-1/30571.html
†Ecole Nationale Supérieure de Techniques Avancées, 32, Bd. Victor, 75739 Paris, France (elghaoui@ensta.fr, oustry@ensta.fr, lebret@ensta.fr).

of study (stochastic programming) concentrates on the case where the perturbation is stochastic with known statistics. One object of this field is to study the impact of, say, a random objective on the distribution of optimal values (this problem is called the "distribution problem"). References relevant to this approach to the perturbation problem include [15, 9, 29]. We are not aware of special references for general SDPs with randomly perturbed data except for the last section of [30], some exercises in the course notes of [13], and section 2.6 in [23].

The main objective of this paper is to quantify the effect of unknown but bounded deterministic perturbation of problem data on solutions. In our framework, the perturbation is not necessarily small, and we seek a solution that is "robust," that is, remains feasible despite the allowable, not necessarily small, perturbation. Our aim is to obtain (approximate) robust solutions via SDP. Links between regularity of solutions and robustness are, of course, expected. One of our side objectives is to clarify these links to some extent. This paper extends results given in [16] for the least-squares problem.

The approach developed here can be viewed as a special case of stochastic programming in which the distribution of the perturbation is uniform.

The ideas developed in this paper draw mainly from two sources: control theory, in which we have found the tools for robustness analysis [36, 17, 12] and some recent work on sensitivity analysis of optimization problems by Shapiro [31] and Bonnans, Cominetti, and Shapiro [10].

Shortly after completion of our manuscript, we became aware of the ongoing work of Ben-Tal and Nemirovski on the same subject. In [7], they apply similar ideas to a truss topology design problem and derive very efficient algorithms for solving the corresponding robustness problem. In [8], the general problem of *tractability* of obtaining a robust solution is studied, and "tractable counterparts" of a large class of uncertain SDPs are given. The case of robust linear programming, under quite general assumptions on the perturbation bounds, is studied in detail in [6]. Our paper can be seen as a complement of [8], giving ways to cope with (not necessarily) tractable robust SDPs by means of upper bounds. (In particular, our paper handles the case of nonlinear dependence of the data on the uncertainties.) A unified treatment, and more results, will appear in [4].

The paper is divided as follows. Our problem is defined in section 2. In section 3, we show how to compute upper bounds on our problem via SDP. We give special attention to the so-called full perturbations case, for which our results are nonconservative. In section 4, we examine sensitivity of the robust solutions in the full perturbations case. We provide conditions which guarantee that the robust solution is unique and a regular function of the data matrices. We then consider several interesting examples in section 5, such as robust linear programming, robust norm minimization, and error-in-variables SDPs.

2. Problem definition.

2.1. Robust SDPs. Let $\mathbf{F}(x, \Delta)$ be a symmetric matrix-valued function of two variables $x \in \mathbf{R}^m$, $\Delta \in \mathbf{R}^{p\times q}$. In the following, we consider x to be the decision variable, and we think of Δ as a perturbation. We assume that Δ is unknown but bounded. Precisely, we assume that Δ is known to belong to a given linear subspace $\mathcal{D}$ of $\mathbf{R}^{p\times q}$, and in addition, $\|\Delta\| \leq \rho$, where $\rho \geq 0$ is given.

In section 2.2, we will be more precise about the dependence of $\mathbf{F}$ on Δ.

We define the *robust feasible set* by

$$\mathcal{X}_\rho = \left\{ x \in \mathbf{R}^m \;\middle|\; \begin{array}{c} \text{for every } \Delta \in \mathcal{D},\ \|\Delta\| \le \rho, \\ \mathbf{F}(x,\Delta) \text{ is well defined and } \mathbf{F}(x,\Delta) \succeq 0 \end{array} \right\}. \tag{2}$$

Now let $\mathbf{c}(\Delta)$ be a vector-valued rational function of the perturbation Δ, such that $\mathbf{c}(0) = c$. We consider the following min-max problem:

$$\text{minimize} \max_{\Delta\in\mathcal{D},\ \|\Delta\|\le\rho} \mathbf{c}(\Delta)^T x \text{ subject to } x \in \mathcal{X}_\rho. \tag{3}$$

From now on, we assume that the function $\mathbf{c}(\Delta)$ is independent of Δ (in other words, the objective vector c is not subject to perturbation). This is done with no loss of generality: introduce a slack variable λ and define

$$\tilde{x} = \begin{bmatrix} x \\ \lambda \end{bmatrix}, \quad \tilde{c} = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, \quad \tilde{\mathbf{F}}(\tilde{x},\Delta) = \mathbf{diag}(\mathbf{F}(x,\Delta), \lambda - \mathbf{c}(\Delta)^T x).$$

Problem (3) can be formulated as

$$\text{minimize } \tilde{c}^T\tilde{x} \text{ subject to } \tilde{x} \in \tilde{\mathcal{X}}_\rho,$$

where $\tilde{\mathcal{X}}_\rho$ is the robust feasible set corresponding to the function $\tilde{\mathbf{F}}$.

In the following, we thus consider a problem of the form

$$\mathcal{P}_\rho: \quad \text{minimize } c^T x \text{ subject to } x \in \mathcal{X}_\rho \tag{4}$$

and refer to it as a *robust semidefinite problem* (RSDP). In general, although $\mathcal{X}_\rho$ is convex, $\mathcal{P}_\rho$ is not a tractable problem—in particular, it is not an SDP. Our aim is to find a convex, inner approximation of $\mathcal{X}_\rho$ that is described by a linear matrix inequality constraint. This inner approximation is then used to find an upper bound on the optimal value of $\mathcal{P}_\rho$ by solving an SDP. In some cases, we can prove our results are nonconservative, that is, as in the so-called "full perturbation" case.

We refer to the set $\mathcal{X}_0$ (resp., problem $\mathcal{P}_0$, i.e., (1)) as the *nominal* feasible set (resp., nominal SDP). We shall assume that the nominal SDP is feasible, that is, $\mathcal{X}_0 \neq \emptyset$. Of course, the robust feasible set $\mathcal{X}_\rho$ may become empty for some $\rho > 0$; we return to this question later.

2.2. Linear-fractional representation. In this paper, we restrict our attention to functions $\mathbf{F}$ that are given by a "linear-fractional representation" (LFR):

$$\mathbf{F}(x,\Delta) = F(x) + L\Delta(I - D\Delta)^{-1}R(x) + R(x)^T(I - \Delta^T D^T)^{-1}\Delta^T L^T, \tag{5}$$

where $F(x)$ is defined in (1), $R(\cdot)$ is an affine mapping taking values in $\mathbf{R}^{q\times n}$, and $L \in \mathbf{R}^{n\times p}$ and $D \in \mathbf{R}^{q\times p}$ are given matrices. Together, the mappings $F(\cdot)$, $R(\cdot)$, the matrices L, D, the subspace $\mathcal{D}$, and the scalar ρ constitute our *perturbation model* for the nominal SDP (1).

The above class of models seems quite specialized. In fact, these models can be used in a wide variety of situations, for example, in the case where the (matrix) coefficients F_i in $\mathcal{P}_0$ are rational functions of the perturbation. The representation lemma, given below, and the examples of section 5 illustrate this point.

A constructive proof of the following result can be found in [37].

LEMMA 2.1. *For any rational matrix function* $\mathbf{M}: \mathbf{R}^k \to \mathbf{R}^{n\times c}$, *with no singularities at the origin, there exist nonnegative integers* $r_1, \ldots, r_k$, *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* $N = r_1 + \cdots + r_k$, *such that* $\mathbf{M}$ *has the following linear-fractional representation* (*LFR*): *For all* δ *where* $\mathbf{M}$ *is defined*,

$$\mathbf{M}(\delta) = M + L\Delta\left(I - D\Delta\right)^{-1} R, \text{ where } \Delta = \mathbf{diag}\left(\delta_1 I_{r_1}, \ldots, \delta_k I_{r_k}\right). \tag{6}$$

Using the LFR lemma, we may devise LFR models for SDPs, where a perturbation vector $\delta \in \mathbf{R}^k$ enters rationally in the coefficient matrices. The resulting set $\mathcal{D}$ of perturbation matrices Δ is then a set of diagonal matrices of repeated elements, as in (6). Componentwise bounds on the vector δ, such as $|\delta|_i \leq \rho$, $i = 1, \ldots, k$, translate into a norm-bound $\|\Delta\| \leq \rho$ on the corresponding matrix Δ.

2.3. A special case. We distinguish a special case for which exact (nonconservative) results can be obtained via SDP. This is when $\mathbf{F}(x, \Delta)$ is block diagonal, each block being independently perturbed—precisely, when

$$\mathbf{F}(x, \Delta) = \mathbf{diag}\big(\mathbf{F}_1(x, \Delta_1), \ldots, \mathbf{F}_L(x, \Delta_L)\big), \tag{7}$$

where each $\mathbf{F}_i(x, \Delta_i)$ assumes the form shown in section 2.2 for appropriate L_i, R_i, D_i, with $\Delta_i \in \mathbf{R}^{p_i\times p_i}$, $i = 1, \ldots, L$, and $\mathcal{D}$ consists of block-diagonal matrices of the form

$$\mathcal{D} = \left\{\Delta = \mathbf{diag}(\Delta_1, \ldots, \Delta_L), \;\middle|\; \Delta_i \in \mathbf{R}^{p_i\times q_i}\right\}.$$

We refer to this situation as the *block-full perturbation* case. When $L = 1$, we speak of the *full perturbation* case. As will be seen later, all results given for $L = 1$ can be generalized to the case $L > 1$.

3. Robust solutions for SDPs. Unless otherwise specified, we fix $\rho > 0$.

3.1. Full perturbations case. In this section, we consider the full perturbations case, that is, $\mathcal{D} = \mathbf{R}^{p\times q}$. We assume $\|D\| < \rho^{-1}$, which is a necessary and sufficient condition for $\mathbf{F}(x, \Delta)$ to be well defined for every $x \in \mathbf{R}^m$ and $\Delta \in \mathbf{R}^{p\times q}$, $\|\Delta\| \leq \rho$.

The following lemma is a simple corollary of a classic result on quadratic inequalities, referred to as the $\mathcal{S}$-procedure [12]. Its proof is detailed in [16].

LEMMA 3.1. *Let* $F = F^T$, L, R, D *be real matrices of appropriate size. We have* $\det(I - D\Delta) \neq 0$ *and*

$$F + L\Delta(I - D\Delta)^{-1}R + R^T(I - D\Delta)^{-T}\Delta^T L^T \succeq 0 \tag{8}$$

for every Δ, $\|\Delta\| \leq 1$, *if and only if* $\|D\| < 1$ *and there exists a scalar* τ *such that*

$$\begin{bmatrix} F - \tau LL^T & R^T - \tau LD^T \\ R - \tau DL^T & \tau(I - DD^T) \end{bmatrix} \succeq 0. \tag{9}$$

A direct application of the above lemma shows that, in the full perturbations case, the RSDP (4) is an SDP.

THEOREM 3.1. *When* $\mathcal{D} = \mathbf{R}^{p\times q}$, *the RSDP* (4) *and a corresponding solution* x *can be computed by solving the SDP in variables* x, τ*:*

$$\text{minimize } c^T x \text{ subject to } \begin{bmatrix} F(x) - \tau LL^T & R(x)^T - \tau LD^T \\ R(x) - \tau DL^T & \tau(\rho^{-2}I - DD^T) \end{bmatrix} \succeq 0. \tag{10}$$

Special barrier functions adapted to a conic formulation of the problem can be devised and yield an interior-point algorithm that has the same complexity as the nominal problem; see [24].

We may define the *maximum allowable perturbation level*, which is the largest number $\rho_{\max}$ such that $\mathcal{X}_\rho \neq \emptyset$ for every ρ, $0 \leq \rho \leq \rho_{\max}$ (note $\rho_{\max} > 0$ since $\mathcal{X}_0 \neq \emptyset$). Computing $\rho_{\max}$ is a (quasi-convex) generalized eigenvalue minimization problem [24, 11]:

$$\text{minimize } \lambda \text{ subject to } \begin{bmatrix} F(x) - \tau LL^T & R(x)^T - \tau LD^T \\ R(x) - \tau DL^T & \tau(\lambda I - DD^T) \end{bmatrix} \succeq 0.$$

Remark. The above exact results are readily generalized to the block-full perturbation case ($L > 1$) as defined in section 2.2.

3.2. Structured case. We now turn to the general case ($\mathcal{D}$ is now an arbitrary linear subspace). In this section, we associate with $\mathcal{D}$ the following linear subspace:

$$\mathcal{B} \triangleq \left\{ (S,T,G) \in \mathbf{R}^{p\times p} \times \mathbf{R}^{q\times q} \times \mathbf{R}^{q\times p} \;\middle|\; S\Delta = \Delta T,\ G\Delta = -\Delta^T G^T \text{ for every } \Delta \in \mathcal{D} \right\}. \tag{11}$$

As shown in [16], a general instance of problem (4) is NP-hard. Therefore, we look for upper bounds on its optimal value. The following lemma is a generalization of Lemma 3.1 that traces back to [17]. Its proof is detailed in [16].

LEMMA 3.2. *Let $F = F^T$, L, R, D be real matrices of appropriate size. Let $\mathcal{D}$ be a subspace of $\mathbf{R}^{p\times q}$, and denote by $\mathcal{B}$ the set of matrices associated with $\mathcal{D}$ as in* (11). *We have* $\det(I - D\Delta) \neq 0$ *and*

$$F + L\Delta(I - D\Delta)^{-1}R + R^T(I - D\Delta)^{-T}\Delta^T L^T \succ 0 \tag{12}$$

for every $\Delta \in \mathcal{D}$, $\|\Delta\| \leq 1$ if there exist a triple $(S,T,G) \in \mathcal{B}$ such that $S \succ 0$, $T \succ 0$, and

$$\begin{bmatrix} F - LSL^T & R^T - LSD^T + LG \\ R - DSL - GL^T & T - GD^T + DG - DSD^T \end{bmatrix} \succ 0. \tag{13}$$

Using Lemma 3.2, we obtain the following result.

THEOREM 3.2. *An upper bound on the RSDP* (4) *and a corresponding solution x can be computed by solving the SDP in variables x, S, T, G :*

$$\begin{gathered} \inf c^T x \text{ subject to } (S,T,G) \in \mathcal{B},\ S \succ 0,\ T \succ 0, \\ \begin{bmatrix} F(x) - LSL^T & R(x)^T - LSD^T + LG \\ R(x) - DSL - GL^T & \rho^{-2}T - GD^T + DG - DSD^T \end{bmatrix} \succ 0. \end{gathered}$$

Note that when the perturbation is full, the variable G is zero and S, T are of the form τI_p, τI_q, resp., for some $\tau \geq 0$. We then recover the exact results of section 3.1.

As before, we may define the maximum allowable perturbation level, which is the largest number $\rho_{\max}$ such that $\mathcal{X}_\rho \neq \emptyset$ for every ρ, $0 \leq \rho \leq \rho_{\max}$. Computing a *lower bound* on this number is a (quasi-convex) generalized eigenvalue minimization problem:

$$\begin{gathered} \inf \lambda \text{ subject to } (S,T,G) \in \mathcal{B},\ S \succ 0,\ T \succ 0, \\ \begin{bmatrix} F(x) - LSL^T & R(x)^T - LSD^T + LG \\ R(x) - DSL - GL^T & \lambda T - GD^T + DG - DSD^T \end{bmatrix} \succ 0. \end{gathered} \tag{14}$$

4. Uniqueness and regularity of robust solutions. In this section, we derive uniqueness and regularity results for the RSDP in the case of full perturbations. As before, we first take $L = 1$ (one block), that is, $\mathcal{D} = \mathbf{R}^{p\times q}$. The results of this section remain valid in the general case $L > 1$ (several blocks).

We fix ρ, $0 < \rho < \rho_{\max}$. For simplicity of notation (and without loss of generality) we take $\rho = 1$ (and thus $\rho_{\max} > 1$). For well-posedness reasons, we must assume $\|D\| < 1$. We make the further assumption that $D = 0$ (in other words, $\mathbf{F}(x, \Delta)$ is affine in Δ). In section 4.5, we show how the case $D \neq 0$ can be treated.

For full perturbations and $D = 0$, the RSDP is the SDP

$$\text{minimize } c^T x \text{ subject to } \begin{bmatrix} F(x) - \tau LL^T & R(x)^T \\ R(x) & \tau I \end{bmatrix} \succeq 0. \tag{15}$$

4.1. Hypotheses. We assume that the SDP (15) (with $D = 0$) satisfies the following hypotheses:

H1. The *Slater* condition holds, that is, the problem is strictly feasible.

H2. The problem is *inf-compact,* meaning that any unbounded sequence (x_k) of feasible points (if any) produces an unbounded sequence of objectives. An equivalent condition is that the Slater condition holds for the dual problem [28, p. 317, Thm. 30.4].

H3. (a) The nullspace of the matrix

$$\lambda R_0 + \sum_{i=1}^{m} x_i R_i$$

is independent of $(\lambda, x) \neq (0, 0)$ and not equal to the whole space.

(b) For every x,

$$\begin{bmatrix} L^T \\ R(x) \end{bmatrix} \text{ is full column-rank.}$$

Hypotheses H1 and H2 ensure, in particular, the existence of optimal points for problem (15) and its dual. Hypotheses H3(a) and (b) are difficult to check in general, but sometimes can be easily tested in practical examples, as seen in section 5. We note that H3(a) implies that $R(x) \neq 0$ for every x.

Hypothesis H1 is equivalent to Robinson's condition [27], which can be expressed in terms of

$$\mathcal{F}(x, \tau) = \begin{bmatrix} F(x) - \tau LL^T & R(x)^T \\ R(x) & \tau I \end{bmatrix}.$$

Robinson's condition is stated in [27] as the existence of $x_0 \in \mathbf{R}^m$, $\tau_0 \in \mathbf{R}$ such that

$$0 \in \text{int } \left(\mathcal{F}(x_0, \tau_0) + d\mathcal{F}(x_0, \tau_0)\mathbf{R}^{m+1} - \mathcal{S}^+_{n+q}\right),$$

where $d\mathcal{F}$ is the differential of $\mathcal{F}$, and $\mathcal{S}^+_{n+q}$ is the set of positive semidefinite matrices of order $n + q$. The equivalence between H1 and Robinson's assumption is not true, in general. Here, this equivalence stems from the fact that the problem is convex and that the cone $\mathcal{S}^+_{n+q}$ has a nonempty interior.

Remark. Hypothesis H1 holds if and only if it holds for the nominal problem (1) (recall our assumption $\rho_{\max} > 1$). Also, hypothesis H2 implies $L \neq 0$ (otherwise, we can let $\tau \to \infty$ without affecting the objective value). If H2 holds for the nominal problem and $L \neq 0$, then H2 holds for the RSDP (15).

4.2. An equivalent nonlinear program. Let $x_{\rm opt}, \tau_{\rm opt}$ be optimal for (15). Hypothesis H3(a) ensures that any τ that is feasible for (15) is nonzero (otherwise, $R(x)$ would be zero for some x). We thus have $\tau_{\rm opt} > 0$.

We introduce some notation. For $x \in \mathbf{R}^m$, $Z \in \mathbf{R}^{n\times n}$, $\tau > 0$ and $\mu \in \mathbf{R}$, define

$$d = \begin{bmatrix} c \\ 0 \end{bmatrix}, \quad y = \begin{bmatrix} x \\ \tau \end{bmatrix}, \quad Y = \mathbf{diag}(Z, \mu),$$

$$G(y) = F(x) - \tau LL^T - \frac{1}{\tau} R(x)^T R(x), \quad \mathcal{G}(y) = \mathbf{diag}(G(y), \tau - .99\tau_{\rm opt}),$$
$$\mathcal{L}(y, Y) = d^T y - \mathbf{Tr} Y G(y).$$

Using Schur complements and $\tau_{\rm opt} > 0$, we obtain that problem (15) can be rewritten as

$$\text{minimize } d^T y \text{ subject to } G(y) \succeq 0 \tag{16}$$

and that $y_{\rm opt} = [x_{\rm opt}^T \ \tau_{\rm opt}]^T$ is optimal for (16). Our aim is first to prove that the so-called quadratic growth condition [10] holds at $y_{\rm opt}$ for problem (16). Then, we will apply the results of [10] to obtain uniqueness and regularity theorems.

4.3. Checking the quadratic growth condition. Following [10], we say that the *quadratic growth condition* (QGC) holds at $y_{\rm opt}$ if there exists a scalar $\alpha > 0$ such that, for every feasible y,

$$d^T y \geq d^T y_{\rm opt} + \alpha \|y - y_{\rm opt}\|^2 + o(\|y - y_{\rm opt}\|^2).$$

Roughly speaking, this condition guarantees that $y_{\rm opt}$ is not on a facet on the boundary of the feasible set.

Define the set of dual variables associated with $y_{\rm opt}$ by

$$\mathcal{Y}(y_{\rm opt}) = \left\{ Y = \mathbf{diag}(Z, \mu) \ \middle| \ Y \succeq 0, \ \ \mathbf{Tr} Y \frac{\partial G}{\partial y_i}(y_{\rm opt}) = d_i, \ \ i = 1, \ldots, m+1 \right\}.$$

The following result is a direct consequence of a general result by Bonnans, Cominetti, and Shapiro [10]. Roughly speaking, this result states that, if an optimization problem satisfies Robinson's condition and has an optimal point, and if a certain "curvature" condition is satisfied, then the QGC holds at that point.

THEOREM 4.1. *With the notation above, if* H1 and H2 *hold, and if*

$$\exists\, Y \in \mathcal{Y}(y_{\rm opt}) \text{ such that } \nabla^2_{yy} \mathcal{L}(y_{\rm opt}, Y) > 0, \tag{17}$$

then problem (16) *satisfies the QGC.*

The following theorem is proven in appendix A.

THEOREM 4.2. *If* H1–H3 *hold, problem* (15) *satisfies the quadratic growth condition at every optimal point* $y_{\rm opt}$*. Consequently, there exists a unique solution to the SDP* (15).

Remark. Note that the QGC is satisfied independent of the objective vector. This means that the boundary of the feasible set is strictly convex (it contains no facets).

4.4. Regularity results. In problem (15), the data consist of the matrices L, and F_i, R_i, $i = 0, \ldots, m$. We seek to examine the sensitivity of the problem with respect to small variations in F_i, L_i, and R_i.

In this section, we consider matrices $L(u)$, $R_i(u)$, and $F_i(u)$, $i = 0, \ldots, m$ that are functions of class $\mathcal{C}^1$ of a (small) parameter vector u. Define

$$F(x,u) = F_0(u) + \sum_{i=1}^m x_i F_i(u), \quad R(x,u) = R_0(u) + \sum_{i=1}^m x_i R_i(u).$$

We denote by $\mathcal{P}(u)$ the corresponding problem (15), where $F(\cdot)$, $R(\cdot)$, and L are replaced by $F(\cdot,u)$, $R(\cdot,u)$, and $L(u)$. We assume that $F(\cdot,0) = F(\cdot)$, $R(\cdot,0) = R(\cdot)$, and $L(0) = L$, so that $\mathcal{P}(0)$ is (15).

We first note that, in the vicinity of $u = 0$, problem $\mathcal{P}(u)$ satisfies the hypotheses H1 and H2 if $\mathcal{P}(0)$ does. In this case, for every $\epsilon > 0$ we may define the set $\mathcal{S}_\epsilon(u)$ of ϵ-suboptimal points of $\mathcal{P}(u)$:

$$\mathcal{S}_\epsilon(u) = \left\{ x \;\middle|\; x \text{ is feasible for } \mathcal{P}(u) \text{ and } c^T x \le v(u) + \epsilon \right\},$$

where $v(u)$ is the optimal value of $\mathcal{P}(u)$.

Recall that, if $\mathcal{P}_0$ satisfies hypotheses H1 and H2, the optimal value $v(u)$ is continuous, and even directionally differentiable, at $u = 0$ [32, Thm. 5.1]. With the QGC in force, and using [31, Thm. 4.1], we can give quite complete regularity results for the robust *solutions*.

THEOREM 4.3. *If hypotheses* H1–H3 *hold for* $\mathcal{P}(0)$, *then for every* $\epsilon = O(u)$, *there exists a* $\gamma > 0$ *and a neighborhood* V *of* $u = 0$ *such that for every* $u \in V$ *and* $x \in \mathcal{S}_\epsilon(u)$, *we have*

$$\|x - x(0)\| \le \gamma \|u\|^{1/2}. \tag{18}$$

When H1–H3 hold for $\mathcal{P}(0)$, the above theorem states that every (sufficiently) suboptimal solution to $\mathcal{P}(0)$ is Hölder-stable (with coefficient 1/2). This is true, in particular, for any optimal solution of $\mathcal{P}(u)$ (that is, for $\epsilon = 0$). The fact that the theorem remains true for $\epsilon > 0$ guarantees regularity of *numerical* solutions to the RSDP. The main consequence is that even if the nominal SDP is ill conditioned (with respect to variations in the F_i's), the RSDP becomes well conditioned for every $\rho > 0$.

Now assume $\rho \neq 1$. We seek to examine the behavior of problem (10) (with $D = 0$) when the uncertainty level ρ for $0 < \rho < \rho_{\max}$ varies. This is a special case of the problem examined above, with $u = \rho$, $F(\cdot,u) = F(\cdot)$, $R(\cdot,u) = R(\cdot)$, $L(u) = \rho L$.

COROLLARY 4.1. *For every* ρ, $0 < \rho < \rho_{\max}$, *the solution to* (10) (*with* $D = 0$) *is unique and satisfies the regularity results* (*written with* $u = \rho$) *of Theorem* 4.3.

Remark. The results of this section are all valid in the block-full perturbation case ($L > 1$), as defined in section 2.2. Of course, the conditions given in H3 should be understood blockwise.

4.5. Case $D \neq 0$. When $D \neq 0$, we can get back to the case $D = 0$ as follows. Recall that we have $\|D\| < 1$ in order to ensure that $\mathbf{F}(x,\Delta)$ is defined everywhere on $\mathcal{D}$. With this assumption, we can define, for $x \in \mathbf{R}^m$ and $\tau > 0$,

$$\begin{aligned}
\tilde{L} &= L(I - D^T D)^{-1/2}, \\
\tilde{R}(x) &= (I - DD^T)^{-1/2} R(x), \\
\tilde{F}(x) &= F(x) - LD^T (I - DD^T)^{-1} R(x) - R(x)^T (I - DD^T)^{-1} L^T, \\
\tilde{G}(y) &= \tilde{F}(x) - \tau \tilde{L}\tilde{L}^T - \frac{1}{\tau} \tilde{R}(x)^T \tilde{R}(x).
\end{aligned}$$

Using Schur complements, we have, for every x and $\tau > 0$,

$$\tilde{G}(y) \succeq 0 \text{ if and only if } \begin{bmatrix} F(x) - \tau LL^T & R(x)^T - \tau LD^T \\ R(x) - \tau DL^T & \tau(I - DD^T) \end{bmatrix} \succeq 0.$$

Hypothesis H3 holds for $\tilde{L}, \tilde{R}(\cdot)$ if and only if it holds for $L, R(\cdot)$. We can then follow the steps detailed previously.

COROLLARY 4.2. *If the SDP* (10) (*with* $\rho = 1$) *satisfies* H1–H3 *and if* $\|D\| < 1$, *then the results of Theorem* 4.3 *hold.*

5. Examples.

5.1. Unstructured perturbations.

Assume

$$\mathbf{F}(x, \Delta) = F(x) + \Delta_0 + \Delta_0^T + \sum_{i=1}^{m} x_i(\Delta_i + \Delta_i^T),$$

where $\Delta = [\Delta_0 \dots \Delta_m]$. This case corresponds to the representation in section 5, with

$$L = I, \quad R(x) = \begin{bmatrix} 1 \\ x \end{bmatrix} \otimes I, \quad D = 0, \quad \mathcal{D} = \mathbf{R}^{n \times nm}. \tag{19}$$

Using Lemma 3.2, we obtain that problem (4) is equivalent to the SDP

$$\text{minimize } c^T x \text{ subject to } \begin{bmatrix} F(x) - \tau I & \begin{bmatrix} 1 & x^T \end{bmatrix} \otimes \rho I \\ \begin{bmatrix} 1 & x^T \end{bmatrix}^T \otimes \rho I & \tau I \end{bmatrix} \succeq 0. \tag{20}$$

It turns out that we may get rid of the variable τ and get back to a convex problem of the same size as that of the unperturbed problem (1). To see this, first note that every feasible variable τ in problem (20) is strictly positive. Use Schur complements to rewrite the matrix inequality in (20) as

$$F(x) \succeq \left(\tau + \rho^2 \frac{1 + \|x\|^2}{\tau}\right) I, \quad \tau > 0.$$

Minimizing (over variable τ) the scalar in the left-hand side of the above inequality shows that the RSDP (1) is equivalent to

$$\text{minimize } c^T x \text{ subject to } F(x) \succeq 2\rho\sqrt{\|x\|^2 + 1} \cdot I. \tag{21}$$

Formulation (21) is more advantageous than (20), since (21) involves a (convex) matrix inequality constraint of the same size as the original problem. As noted before, special barrier functions can be devised for this problem and yield an interior-point algorithm that has the same complexity as the original problem; see [24].

We note that, with the above choice for L, R, hypothesis H3 holds, which yields the following result.

THEOREM 5.1. *The optimal value of the RSDP* (20) *can be computed by solving the convex problem* (21). *If* (21) *satisfies hypotheses* H1 and H2, *then for every* $\rho > 0$, *the solution is unique and satisfies the regularity conditions of Theorem* 4.3.

Remark. A sufficient condition for hypotheses H1 and H2 to hold for (21) is that they hold for the nominal problem. A more restrictive sufficient condition is that the nominal feasible set $\mathcal{X}_0$ is nonempty and bounded, and $\rho < \rho_{\max}$.

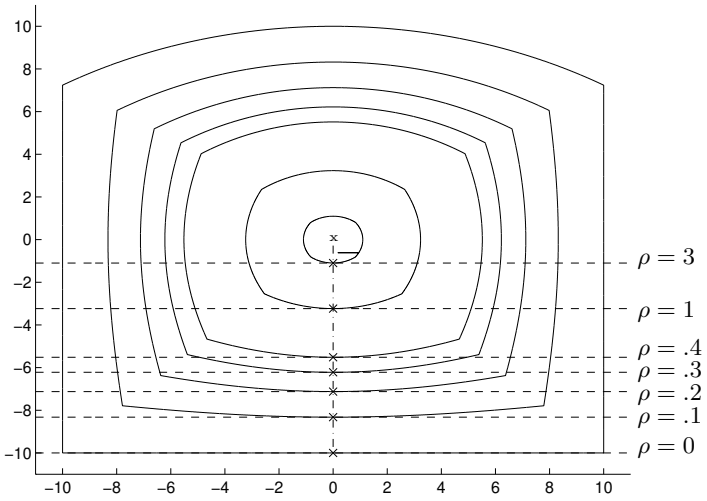

FIG. 1. *Nominal and robust solutions of an SDP, with a* 5×5 *matrix* $F(x)$. *Here* $\rho_{\max} = 5$.

5.2. Robust center of a linear matrix inequality. In this section, for $\rho > 0$, we consider the SDP (21) and corresponding feasible (convex) set $\mathcal{X}_\rho$. We assume that $\mathcal{X}_0$ is nonempty and bounded, and that $\mathcal{P}_0$ is strictly feasible. Then, for every ρ, $0 < \rho < \rho_{\max}$, $\mathcal{X}_\rho$ is nonempty and bounded, and we can define a (unique) solution $x(\rho)$ to the strictly convex problem (21).

In view of Corollary 4.1, $x(\rho)$ is a continuous function of ρ in $]0\,\rho_{\max}[$. Since $(\mathcal{X}_\rho)$ is a decreasing family of bounded sets, we may define

$$x^* = \lim_{\rho \to \rho_{\max}} x(\rho). \tag{22}$$

Note that x^* is independent on the objective vector c.

Thus, to the matrix inequality $F(x) \succeq 0$, we may associate the *robust center*, defined by (22). The robust center has the property of being the most tolerant (with respect to unstructured perturbation) among the feasible points.

An example is depicted in Fig. 1. The nominal feasible set $\mathcal{X}_0$ is described by a linear matrix inequality $F(x) \succeq 0$, where F is a 5×5 matrix. For various values of ρ, we seek to minimize x_2. The dashed lines correspond to the optimal objectives. As ρ increases, we observe that the robust feasible sets shrink. A crucial property of these robust sets is that they do not possess any straight faces, as observed in the figure. For $\rho = \rho_{\max} \simeq 5$, the robust feasible set is a singleton (in this example, $x^\star = 0$). When $\rho = 0$, the optimal solution is not unique and not continuous with respect to changes in the coefficient matrices F_i, $i = 0, 1, 2$ (although the optimal value is continuous). Since the sets $\mathcal{X}_\rho$ become strictly convex as soon as $\rho > 0$, the resulting robust solutions are continuous.

5.3. Robust linear programs. An interesting special case arises with linear programming (LP). Consider the LP

$$\text{minimize } c^T x \text{ subject to } a_i^T x \geq b_i,\ i = 1, \ldots, L.$$

Assume that the a_i's and b_i's are subject to unstructured perturbations. The perturbed value of $[a_i^T\ b_i]^T$ is $[a_i^T\ b_i]^T + \delta_i$, where $\|\delta_i\|_2 \leq \rho$, $i = 1, \ldots, L$. We seek a

robust solution to our problem, which is a special case of the block-full perturbation case referred to in section 2.2, with $\mathbf{F}$ given by (7), and

$$\mathbf{F}_i(x, \Delta_i) = a_i^T x - b_i + 2[x^T \ -1]\Delta_i, \ \ i = 1, \ldots, L,$$

where $\Delta_i = \delta_i/2$, and $\mathcal{D}$ is the set of diagonal, $L \times L$ matrices. The robust LP is

$$\text{minimize } c^T x \text{ subject to } a_i^T x - \rho\sqrt{\|x\|_2^2 + 1} \geq b_i, \ i = 1, \ldots, L. \tag{23}$$

The above program is readily written as an SDP by introducing slack variables. In fact, the robust LP is a second-order cone program (SOCP) for which efficient special-purpose interior-point methods are available [24, 20, 23].

We note that hypothesis H3 holds blockwise. This yields the following result.

THEOREM 5.2. *The optimal value of the robust LP can be computed by solving the convex problem* (23). *If the latter satisfies hypotheses* H1 and H2, *then for every* ρ, $0 < \rho < \rho_{\max}$, *the solution is unique and satisfies the regularity conditions of Theorem* 4.3.

In [6], robust linear programming is studied in detail. For a wide class of perturbation models, where the data of every linear constraint vary in an ellipsoid, explicit robust solutions are constructed using convex SOCPs. Reference [23] also provides examples of robust linear programs solved via SOCP.

5.4. Robust eigenvalue minimization. Consider the case where the nominal problem consists of minimizing the largest eigenvalue of a matrix-valued function $F(x)$:

$$\text{minimize } \lambda_{\max}(F(x)). \tag{24}$$

When $F(\cdot)$ is subject to unstructured perturbations (as defined in section 5.1), the robust version of the problem is

$$\text{minimize } \lambda + 2\rho\sqrt{\|x\|^2 + 1} \text{ subject to } \lambda I \succeq F(x),$$

or equivalently

$$\text{minimize } \lambda_{\max}(F(x)) + 2\rho\sqrt{\|x\|^2 + 1}. \tag{25}$$

Let $\rho > 0$. When written in an SDP form, the above problem satisfies the hypotheses H1–H3. From Theorem 4.3 we obtain that the solution is unique. If we consider that the data of the above problem consist of the matrices F_i, $i = 0, \ldots, m$, then we know that the corresponding solution is Hölder-stable (with coefficient 1/2). Since the problem is unconstrained, we can use a result of Shapiro [31, Thm. 3.1], by which we conclude that the solution is actually Lipschitz stable (inequality (18) holds with the exponent 1/2 replaced by 1). Finally, using the results from Attouch [3], we can show that computing the solution for $\rho \to 0$ amounts to selecting the minimum norm solution among the solutions of the nominal problem.

THEOREM 5.3. *The optimal value of the min-max problem* (24) *can be computed by solving the convex problem* (25). *For every* $\rho > 0$, *the solution is unique and is Lipschitz stable with respect to perturbations in* F_i, $i = 0, \ldots, m$. *When* $\rho \to 0$, *the solution converges to the minimum norm solution of the nominal problem* (24).

Remark. In this case, the RSDP is a regularized version of the nominal SDP, which belongs to the class of Tikhonov regularizations [34]. The regularization parameter

is 2ρ and is chosen according to some a priori information on uncertainty associated with the nominal problem's data. Taking ρ close to zero can be used as a *selection* procedure for choosing a particular (minimum norm, regular) solution among the (not necessarily unique and/or regular) solutions of the nominal problem.

Problem (25) is particularly suitable to the recent so-called $\mathcal{U}$-Newton algorithms for solving problem (24). These methods, described in [21, 25], require that the Hessian of the "smooth part" (the so-called $\mathcal{U}$-Hessian) of the objective of (24) be positive definite. For general mappings $F(\cdot)$, this property is not guaranteed. However, when looking at the robust problem (25), we see that the modified $\mathcal{U}$-Hessian is guaranteed to be positive definite for every x and $\rho > 0$. This indicates that the RSDP approach may be used to devise *robust algorithms* for solving SDPs.

5.5. Robust SOCPs. An SOCP is a problem of the form

$$\begin{array}{ll} \text{minimize} & c^T x \\ \text{subject to} & \|C_i x + d_i\| \le e_i^T x + f_i, \quad i = 1, \ldots, L, \end{array} \tag{26}$$

where $C_i \in \mathbf{R}^{n_i \times m}$, $d_i \in \mathbf{R}^{n_i}$, $e_i \in \mathbf{R}^m$, $f_i \in \mathbf{R}$, $i = 1, \ldots, L$. SOCPs can be formulated as SDPs, but special-purpose, more efficient algorithms can be devised for them; see [24, 5, 23].

Assuming that C_i, d_i, e_i, f_i are subject to linear—or even rational—uncertainty, we may formulate the corresponding RSDP as an SDP. This SDP can be written as an SOCP if the uncertainty is unstructured and affects each constraint independently.

The subject of robust SOCPs is explored in [5] in detail. Explicit SDPs that yield robust counterparts to SOCPs *nonconservatively* are given for a wide class of uncertainty structures. In some cases, albeit not all, the robust counterpart is itself an SOCP. In [16, 14], the special case of least-squares problems with uncertainty in the data is studied at length.

5.6. Robust maximum norm minimization. Several engineering problems take the form

$$\text{minimize } \|H(x)\|, \tag{27}$$

where

$$H(x) = H_0 + \sum_{i=1}^{m} x_i H_i,$$

and H_i, $i = 1, \ldots, m$ are given $p \times q$ matrices. A well-known instance of this problem is the linear least-squares problem, with $H(x) = Ax - b$. Another example is a minimal norm extension problem for a Hankel operator studied in [18], in which H_0 is a given (arbitrary) $n \times n$ Hankel matrix and H_i, $i = 1, \ldots, m$ is the $n \times n$ Hankel matrix associated with the polynomial $1/z^i$. In practice, the matrices H_i, $i = 0, \ldots, m$ are subject to perturbation, which motivates a study of the robust version of problem (27). Note that the least-squares case is extensively studied in [16].

Consider the *full perturbation case*, which occurs when each H_i is perturbed independently in a linear manner. Precisely, consider the matrix-valued function

$$\mathbf{H}(x, \Delta) = H_0 + \Delta_0 + \sum_{i=1}^{m} x_i (H_i + \Delta_i),$$

where $\Delta = [\Delta_0 \dots \Delta_m]$. For a given $\rho > 0$, we address the min-max problem

$$\min_x \max_{\|\Delta\| \le \rho} \|\mathbf{H}(x, \Delta)\|. \tag{28}$$

This problem is an RSDP for which we can get exact results using SDP. Indeed, for every $x \in \mathbf{R}^m$ and $\lambda \ge 0$, the property

$$\max_{\|\Delta\| \le \rho} \|\mathbf{H}(x, \Delta)\| \le \lambda$$

is equivalent to $\mathbf{F}(x, \lambda, \Delta) \succeq 0$ for every Δ, $\|\Delta\| \le \rho$, where

$$\mathbf{F}(x, \lambda, \Delta) = F(x, \lambda) + L\Delta R(x) + R(x)^T \Delta^T L^T,$$

where

$$F(x, \lambda) = \begin{bmatrix} \lambda I & H(x) \\ H(x)^T & \lambda I \end{bmatrix}, \quad L = \begin{bmatrix} I \\ 0 \end{bmatrix}, \quad R(x) = \begin{bmatrix} 0 & \begin{bmatrix} 1 \\ x \end{bmatrix} \otimes I \end{bmatrix}.$$

We thus write problem (28) as (4), where the perturbation set $\mathcal{D}$ is $\mathbf{R}^{p \times q}$.

Applying Theorem 3.2, we obtain that the RSDP above is equivalent to the SDP (15) (with $D = 0$). As in section 5.1, we may get rid of the variable τ and obtain the equivalent formulation

$$\text{minimize } \|H(x)\| + \rho\sqrt{\|x\|^2 + 1}. \tag{29}$$

This RSDP satisfies hypotheses H1–H3, so we conclude that the results of Theorem 4.3 hold. As in robust eigenvalue minimization, we can get improved results using [31, section 3, Thm. 3.1].

THEOREM 5.4. *The optimal value of the min-max problem* (28) *can be computed by solving the convex problem* (29). *For every* $\rho > 0$, *the solution is unique and Lipschitz stable with respect to perturbations in* H_i, $i = 0, \dots, m$. *When* $\rho \to 0$, *the solution converges to the minimum norm solution of the nominal problem* (27).

Remark. As for the RSDP arising in robust eigenvalue minimization, the robust minimum norm minimization problem is a regularized version of the nominal problem, which belongs to the class of Tikhonov regularizations.

We now consider the general case where each matrix H_i in (27) is perturbed in a structured manner. To be specific, we concentrate on the minimal norm extension problem mentioned above.

In practice, the matrix H_0 is obtained from measurement and is thus subject to error. We may assume that this matrix is constructed from an $n \times 1$ vector $h_0(\delta) = h_0 + \delta$, where δ is unknown but bounded. The perturbed matrix H_0 is of the form

$$H_0(\Delta) = H_0 + L\Delta R,$$

where L, R are given matrices (the exact form of which we do not detail), and

$$\Delta \in \mathcal{D} = \{\mathbf{diag}(\delta_1 I_1, \dots, \delta_n I_n) \mid \delta_i \in \mathbf{R}, \ \ i = 1, \dots, n\}.$$

(In the above, each δ_i corresponds to the uncertainty associated with the ith antidiagonal of H_0.) We address the min-max problem

$$\text{minimize} \max_{\Delta \in \mathcal{D},\ \|\Delta\| \le \rho} \|H(x) + L\Delta R\|, \tag{30}$$

where $\rho \ge 0$ is given.

This problem is amenable to the robustness analysis technique. Defining

$$\mathcal{S} \triangleq \left\{ \mathbf{diag}(S_1, \ldots, S_n) \;\middle|\; S_i \in \mathbf{R}^{i \times i},\ \ i = 1, \ldots, n \right\},$$

we obtain the following result.

THEOREM 5.5. *An upper bound on the objective value of the min-max problem* (30) *can be computed by solving the SDP in variables* x, S, G*:*

$$\inf \lambda \text{ subject to } \quad S = S^T,\ G = -G^T \in \mathcal{S},\ \begin{bmatrix} \lambda I - LSL^T & H(x) & LG \\ H(x)^T & \lambda I & \rho R \\ G^T L^T & \rho R^T & S \end{bmatrix} \succ 0.$$

5.7. Polynomial interpolation. This example, taken from [16], can be formulated as an RSDP with rational dependence. For given integers $n \geq 1$, k, we seek a polynomial of degree $n-1$ $p(t) = x_1 + \cdots + x_n t^{n-1}$ that interpolates given points (a_i, b_i), $i = 1, \ldots, k$, that is,

$$p(a_i) = b_i, \quad i = 1, \ldots, k.$$

If we assume that (a_i, b_i) are known exactly, we obtain a linear equation in the unknown x, with a Vandermonde structure:

$$\begin{bmatrix} 1 & a_1 & \ldots & a_1^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_k & \ldots & a_k^{n-1} \end{bmatrix}, \begin{bmatrix} x_1 \\ \vdots \\ x_n \end{bmatrix} = \begin{bmatrix} b_1 \\ \vdots \\ b_n \end{bmatrix},$$

which can be solved via standard least-squares techniques.

Now assume that the interpolation points are not known exactly. For instance, we may assume that the b_i's are known, while the a_i's are parameter dependent:

$$a_i(\delta) = a_i + \delta_i, \quad i = 1, \ldots, k,$$

where the δ_i's are unknown but bounded: $|\delta_i| \leq \rho$, $i = 1, \ldots, k$, where $\rho \geq 0$ is given. We seek a robust interpolant, that is, a solution x that minimizes

$$\max_{\|\delta\|_\infty \leq \rho} \|\mathbf{A}(\delta)x - b\|, \tag{31}$$

where

$$\mathbf{A}(\delta) = \begin{bmatrix} 1 & a_1(\delta) & \ldots & a_1(\delta)^{n-1} \\ \vdots & \vdots & & \vdots \\ 1 & a_k(\delta) & \ldots & a_k(\delta)^{n-1} \end{bmatrix}.$$

The above problem is an RSDP. Indeed, it can be shown that

$$\begin{bmatrix} \mathbf{A}(\delta) & b \end{bmatrix} = \begin{bmatrix} \mathbf{A}(0) & b \end{bmatrix} + L\Delta(I - D\Delta)^{-1}R,$$

where

$$L = \bigoplus_{i=1}^{k} \begin{bmatrix} 1 & a_i & \ldots & a_i^{n-2} \end{bmatrix}, R = \begin{bmatrix} R_1 \\ \vdots \\ R_k \end{bmatrix}, \ D = \bigoplus_{i=1}^{k} D_i, \ \ \Delta = \bigoplus_{i=1}^{k} \delta_i I_{n-1},$$

and, for each i, $i = 1, \ldots, k$,

$$R_i = \begin{bmatrix} 0 & 1 & a_i & \ldots & a_i^{n-2} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_i \\ 0 & \ldots & \ldots & 0 & 1 \end{bmatrix} \in \mathbf{R}^{(n-1)\times n},$$

$$D_i = \begin{bmatrix} 0 & 1 & a_i & \ldots & a_i^{n-3} \\ \vdots & \ddots & \ddots & \ddots & \vdots \\ \vdots & & \ddots & \ddots & a_i \\ 0 & \ldots & \ldots & 0 & 1 \end{bmatrix} \in \mathbf{R}^{(n-1)\times(n-1)}.$$

(Note that $\det(I - D\Delta) \neq 0$, since D is strictly upper triangular.) With the above notation, if we define $\mathbf{F}(x, \Delta)$ as in section 5, then problem (31) can be formulated as the RSDP (4).

With the approach described in this paper, one can compute an upper bound for the minimizing value of (31), and a corresponding suboptimal x. We do not know if the problem can be solved exactly in polynomial time, e.g., using SDP. We conjecture (as the reviewers of this paper did) that the answer is no. To motivate this claim, note that the solution to the problem of computing (31) for arbitrary *affine* functions $\mathbf{A}$ is already NP-hard [16].

5.8. Error-in-variables RSDPs. In many SDPs that arise in engineering, the variable x represents physical parameters that can be implemented with finite absolute precision only. A typical example is integer programming, where integer solutions to (linear) programs are sought. These problems (which are equivalent to integer programming) are NP-hard. We now show that we may find upper bounds on these problems using robustness analysis.

Consider, for instance, the problem of finding a solution x to the feasibility SDP

$$\text{find an integer vector } x \text{ such that } F(x) \succeq 0. \tag{32}$$

Now, consider the robust SDP

$$\begin{gathered} \text{maximize } \lambda \text{ subject to} \\ \lambda I \leq F_0 + \sum_{i=1}^{m} (x_i + \Delta x_i) F_i \text{ for every } \Delta x,\ \|\Delta x\|_\infty \leq 1/2. \end{gathered} \tag{33}$$

Assume there exists a feasible pair $(x_{\text{feas}}, \lambda)$ to the above problem, with $\lambda \geq 0$. By construction, x_{feas} satisfies $F(x_{\text{feas}}) \succeq 0$. Furthermore, any vector x chosen such that $\|x - x_{\text{feas}}\|_\infty \leq 1/2$ is guaranteed to satisfy $F(x) \succeq 0$. This is true, in particular, for x_{int}, the integer closest to x_{feas}. Thus, if we know a positive lower bound λ, and corresponding feasible point for problem (33), then we can compute an integer solution to our original problem.

Finding a lower bound for (33) and an associated feasible point can be done as follows. For i, $1 \leq i \leq m$, define $F_i = 2L_iR_i$, where $L_i, R_i^T \in \mathbf{R}^{n\times r_i}$, $r_i = \mathbf{Rank}\, F_i$. Let

$$L = \begin{bmatrix} L_1 & \ldots & L_m \end{bmatrix}, \quad R = \begin{bmatrix} R_1 \\ \vdots \\ R_m \end{bmatrix}, \text{ and } \mathcal{D} = \left\{ \Delta = \bigoplus_{i=1}^{m} \Delta x_i I_{r_i},\ \ \Delta x_i \in \mathbf{R} \right\}.$$

Problem (33) can be formulated as

$$\text{(34)} \qquad \begin{array}{ll} \text{maximize } \lambda \text{ subject to} & \lambda I \leq F(x) + L\Delta R + R^T\Delta^T L^T \\ & \text{for every } \Delta \in \mathcal{D},\ \|\Delta\| \leq 1/2. \end{array}$$

The above is a special instance of the structured problem examined in section 3.2. Define

$$\mathcal{S} \stackrel{\Delta}{=} \left\{ \mathbf{diag}(S_1, \ldots, S_m) \ \middle|\ S_i \in \mathbf{R}^{r_i \times r_i},\ \ i = 1, \ldots, n \right\}.$$

THEOREM 5.6. *A sufficient condition for an integer solution to the feasibility SDP* (32) *is that the constraints*

$$\lambda \geq 0,\ \ S = S^T \in \mathcal{S},\ \ G = -G^T \in \mathcal{S},\ \ \left[\begin{array}{cc} F(x) - \lambda I - LSL^T & (1/2)R^T + LG \\ (1/2)R - GL^T & S \end{array} \right] \succ 0$$

are feasible. If x_{feas} *is feasible for the above constraints, then any integer vector closest to* x_{feas} (*in the maximum norm sense*) *is feasible for* (32).

6. Conclusions. In this paper, we considered semidefinite programs subject to uncertainty. Assuming the latter is unknown but bounded, we have provided sufficient conditions that guarantee "robust" solutions to exist via SDPs. Under some conditions (detailed in section 4), the robust solution is unique, and not surprisingly, stable. The method can then be used to regularize possibly ill-conditioned problems. For some perturbation structures (as for unstructured perturbations), the conditions are also necessary. That is, there is no conservatism induced by the method.

The paper raises several open questions.

In our description, we have considered the problem of making the primal SDP robust, thereby obtaining upper bounds on an SDP subject to uncertainty. The dual point of view should be very interesting. One might be interested in applying the approach to the dual problem instead. Does this lead to lower bounds on the perturbed problem? Also, in some cases, the RSDP approach leads to a unique (and stable) primal solution. May we obtain a unique solution to the dual problem by making the latter robust? (This would lead to analyticity of the primal solution; see [32].)

As seen in section 5.2 the notion of robust center has, certainly, connections with the well-known analytic center; is the latter related to some robustness characterization?

It seems that the RSDP method could be useful for deriving fast and robust (stable) algorithms for solving SDPs (see section 5.4), especially in connection with maximum eigenvalue minimization.

Finally, as said in section 2.2 (Lemma 2.1), an SDP with coefficient matrices depending rationally on a perturbation vector can always be represented by an LFR model. Now, this LFR model is not unique. However, the results given here (for example, Theorem 3.2) hinge on a particular linear-fractional representation for a perturbed SDP. Hence we have the question: are our results independent of the chosen representation? We partially answer this difficult question in Appendix B.

Appendix A. Proof of Theorem 4.2. We take the notation of section 4. Let $Y = \mathbf{diag}(Z, \mu)$ be dual variables associated with $(x_{\text{opt}}, \tau_{\text{opt}})$ that are optimal (their existence is guaranteed by H1 and H2). Then, $Y \in \mathcal{Y}(y_{\text{opt}})$. Let us show that condition (17) holds for this choice of Y.

Since the problem satisfies H1 and H2, the complementarity conditions hold; therefore, the (optimal) dual variable μ associated with the constraint $\tau = \tau_{\rm opt}$ is zero. Consequently the variable Z is nonzero (recall $c \neq 0$). Using

$$\mathbf{Tr} Y \frac{\partial G}{\partial y_{m+1}}(y_{\rm opt}) = d_{m+1} = 0,$$

we obtain

$$\tau_{\rm opt}^2 \mathbf{Tr} L L^T Z = \mathbf{Tr} R(x_{\rm opt})^T R(x_{\rm opt}) Z.$$

From $\tau_{\rm opt} \neq 0$ (implied by H3(a)), and using hypothesis H3(b), we can show that

$$\mathbf{Tr} L L^T Z = 0 \text{ and } \mathbf{Tr} R(x_{\rm opt})^T R(x_{\rm opt}) Z = 0$$

are impossible for $Z \succeq 0$, $Z \neq 0$. This yields $\mathbf{Tr} R(x_{\rm opt})^T R(x_{\rm opt}) Z > 0$.

Now let $\xi \in \mathbf{R}^m$ and $\lambda \in \mathbf{R}$, and define

$$\Phi(\xi, \lambda) = d^2 \mathcal{G}(x_{\rm opt}, \tau_{\rm opt})[(\xi, \lambda), (\xi, \lambda)],$$
$$\phi(\xi, \lambda) = \begin{bmatrix} \xi \\ \lambda \end{bmatrix}^T \nabla^2_{yy} \mathcal{L}(y_{\rm opt}, Y) \begin{bmatrix} \xi \\ \lambda \end{bmatrix} = -\mathbf{Tr} Z \Phi(\xi, \lambda).$$

We have, for every i, j, $1 \leq i, j \leq m$,

$$\frac{\partial G}{\partial x_i} = F_i - \frac{1}{\tau}(R(x)^T R_i + R_i^T R(x)), \quad \frac{\partial G}{\partial \tau} = -L L^T + \frac{1}{\tau^2} R(x)^T R(x),$$
$$\frac{\partial^2 G}{\partial x_i \partial x_j} = -\frac{1}{\tau}(R_j^T R_i + R_i^T R_j), \quad \frac{\partial^2 G}{\partial x_i \partial \tau} = \frac{1}{\tau^2}(R(x)^T R_i + R_i^T R(x)),$$
$$\frac{\partial^2 G}{\partial \tau^2} = -\frac{2}{\tau^3} R(x)^T R(x).$$

By summation, we have

$$\begin{aligned} -\Phi(\xi, \lambda) &= \frac{2}{\tau_{\rm opt}}(R(\xi) - R(0))^T (R(\xi) - R(0)) + 2\frac{\lambda^2}{\tau_{\rm opt}^3} R(x_{\rm opt})^T R(x_{\rm opt}) \\ &\quad - \frac{\lambda}{\tau_{\rm opt}^2}\left(R(x_{\rm opt})^T (R(\xi) - R(0)) + (R(\xi) - R(0))^T R(x_{\rm opt})\right) \\ &= \frac{\lambda^2}{\tau_{\rm opt}^3} R(x_{\rm opt})^T R(x_{\rm opt}) + \frac{1}{\tau_{\rm opt}} \mathcal{R}^T \mathcal{R} + \frac{1}{\tau_{\rm opt}}(R(\xi) - R(0))^T (R(\xi) - R(0)), \end{aligned}$$

where $\mathcal{R} = R(\xi) - R(0) - \dfrac{\lambda}{\tau_{\rm opt}} R(x_{\rm opt})$. We obtain finally,

$$\begin{aligned} \phi(\xi, \lambda) &= \frac{1}{\tau_{\rm opt}} \mathbf{Tr} Z \mathcal{R}^T \mathcal{R} + \frac{\lambda^2}{\tau_{\rm opt}^3} \mathbf{Tr} Z R(x_{\rm opt})^T R(x_{\rm opt}) \\ &\quad + \frac{1}{\tau_{\rm opt}} \mathbf{Tr} Z (R(\xi) - R(0))^T (R(\xi) - R(0)). \end{aligned}$$

If $\phi(\xi, \lambda) = 0$, then $\lambda = 0$ (from $\mathbf{Tr} R(x_{\rm opt})^T R(x_{\rm opt}) Z > 0$), and thus $\mathbf{Tr} Z \mathcal{R}^T \mathcal{R} = 0$ with $\mathcal{R} = R(\xi) - R(0)$. Since $Z \succeq 0$, this means that every column of $Z^{1/2}$ belongs to the nullspace of $R(\xi) - R(0)$. Now assume $\xi \neq 0$. By hypothesis H3(a), we obtain that every column of $Z^{1/2}$ also belongs to the nullspace of $R(x_{\rm opt})$, which contradicts

$\mathbf{Tr} R(x_{\rm opt})^T R(x_{\rm opt}) Z > 0$. We conclude that $\nabla^2_{yy}\mathcal{L}$ is positive definite at $(y_{\rm opt}, Y)$. Thus, problem (15) satisfies the QCG.

Appendix B. Invariance with respect to the LFR model. In this section, we show that the sufficient conditions obtained in this paper are, in some sense, independent of the LFR model used to describe the perturbation structure.

Consider a function $\mathbf{F}$ taking values in the set of symmetric matrices having an LFR such as that in section 5. This function can be written in a more symmetric form,

$$\mathbf{F}(\Delta) = F + \tilde{L}\tilde{\Delta}(I - D\tilde{\Delta})^{-1}\tilde{L}^T, \tag{35}$$

where we have dropped the dependence on x for convenience, and

$$\tilde{L} = \left[\begin{array}{cc} L & R^T \end{array}\right], \quad \tilde{D} = \left[\begin{array}{cc} 0 & D^T \\ D & 0 \end{array}\right], \quad \tilde{\Delta} = \left[\begin{array}{cc} 0 & \Delta \\ \Delta^T & 0 \end{array}\right].$$

It is easy to check that, if an invertible matrix Z satisfies the relation $Z\tilde{\Delta}Z^T = \tilde{\Delta}$ for every $\Delta \in \mathcal{D}$, then

$$\mathbf{F}(\Delta) = F + (\tilde{L}Z)\tilde{\Delta}(I - (Z^T\tilde{D}Z)\tilde{\Delta})^{-1}(\tilde{L}Z)^T.$$

In other words, the "scaled" triple $(F, (\tilde{L}Z), (Z^T\tilde{D}Z))$ can be used to represent $\mathbf{F}$ instead of $F, \tilde{L}, \tilde{D}$ in (35). By spanning valid scaling matrices Z, we span a subset of all LFR models that describe $\mathbf{F}$.

A valid scaling matrix Z can be constructed as follows. Let $(S, T, G) \in \mathcal{B}$, and define

$$Z = \left[\begin{array}{cc} T^{-1/2} & 0 \\ 0 & S^{1/2} \end{array}\right] \left[\begin{array}{cc} I & G \\ 0 & I \end{array}\right].$$

It turns out that such a Z satisfies the relation $Z\tilde{\Delta}Z^T = \tilde{\Delta}$ for every $\Delta \in \mathcal{D}$.

Using the above facts, we can show that if condition (13) is true for the original LFR model F, L, R, D with appropriate S, T, G, then it is also true for the scaled LFR obtained using any scaling matrix Z such as that above, for appropriate matrices $\tilde{S}$, $\tilde{G}$, $\tilde{T}$. That is, the condition is independent of the scaling Z.

In this sense, the conditions we obtained are independent of the LFR used to represent the perturbation structure.

Acknowledgments. This paper has benefitted from many stimulating discussions with several colleagues, including Aharon Ben-Tal, Stephen Boyd, Arkadii Nemirovski, Michael Overton, and particularly, Lieven Vandenberghe (who pointed out a mistake just before the final version was sent). Last but not least, the authors would like to thank the editor and reviewers for their very helpful comments and revisions.

REFERENCES

[1] F. Alizadeh, *Interior point methods in semidefinite programming with applications to combinatorial optimization*, SIAM J. Optim., 5 (1995), pp. 13–51.

[2] F. Alizadeh, J.-P. A. Haeberly, and M. L. Overton, *Primal-dual interior-point methods for semidefinite programming: Convergence rates, stability and numerical results*, SIAM J. Optim., 8 (1998), pp. 746–768.

[3] H. Attouch, *Viscosity Solutions of Optimization Problems*, Tech. Report 07, Dépt. des Sciences Mathématiques, Université Montpellier 2, France, 1994.

[4] A. Ben-Tal, L. E. Ghaoui, and A. Nemirovski, *Robust semidefinite programming*, in Semidefinite Programming and Applications, to appear.

[5] A. Ben-Tal and A. Nemirovski, *Robust Convex Programming*, Tech. Report 1/95, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion, Israel Institute of Technology, Technion City, Haifa 32000, Israel, 1995; Math. Oper. Res., to appear.

[6] A. Ben-Tal and A. Nemirovski, *Robust Solutions to Uncertain Linear Programs*, Tech. Report 6/95, Optimization Laboratory, Faculty of Industrial Engineering and Management, Technion, Israel Institute of Technology, Technion City, Haifa 32000, Israel, 1995; Oper. Res. Lett., to appear.

[7] A. Ben-Tal and A. Nemirovski, *Robust truss topology design via semidefinite programming*, SIAM J. Optim., 7 (1997), pp. 991–1016.

[8] A. Ben-Tal and A. Nemirovski, *Robust convex programming*, IMA J. Numer. Anal., 1998, to appear.

[9] B. Bereanu, *Some Numerical Methods in Stochastic Programming Under Risk and Uncertainty*, Academic Press, New York, 1980, Ch. 11, pp. 169–205.

[10] J. F. Bonnans, R. Cominetti, and A. Shapiro, *Sensitivity Analysis of Optimization Problems under Second Order Regular Constraints*, Tech. Report 2989, INRIA, 1996; Math. Oper. Res., to appear.

[11] S. Boyd and L. El Ghaoui, *Method of centers for minimizing generalized eigenvalues*, Linear Algebra and Appl., special issue on Numerical Linear Algebra Methods in Control, Signals and Systems, 188 (1993), pp. 63–111.

[12] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in System and Control Theory*, Studies in Applied Mathematics series, SIAM, Philadelphia, PA, 1994.

[13] S. P. Boyd and L. Vandenberghe, *Introduction to convex optimization with engineering applications*, lecture notes for ee392x, Stanford University, Stanford, CA, 1996. Available via anonymous ftp at `isl.stanford.edu/pub/boyd`.

[14] S. Chandrasekaran, G. H. Golub, M. Gu, and A. H. Sayed, *An efficient algorithm for a bounded errors-in-variables model*, SIAM J. Matrix Anal. Appl., to appear.

[15] M. Dempster, *Stochastic Programming*, Academic Press, New York, 1980.

[16] L. El Ghaoui and H. Lebret, *Robust solutions to least-squares problems with uncertain data*, SIAM J. Matrix Anal. Appl., 18 (1997), pp. 1035–1064.

[17] M. K. H. Fan, A. L. Tits, and J. C. Doyle, *Robustness in the presence of mixed parametric uncertainty and unmodeled dynamics*, IEEE Trans. Automat. Control, 36 (1991), pp. 25–38.

[18] J. W. Helton and H. J. Woerdeman, *Symmetric Hankel operators: Minimal norm extensions and eigenstructures*, Linear Algebra Appl., 185 (1993), pp. 1–19.

[19] M. Kojima, M. Shida, and S. Shindoh, *Global and Local Convergence of Predictor-Corrector-Interior-Point Algorithm for Semidefinite Programming*, Tech. Report B-308, Dept. of Information Sciences, Tokyo Institute of Technology, 2-12-1 Oh-Okayama, Meguru-ku, Tokyo 152, Japan, 1995.

[20] H. Lebret, *Synthèse de diagrammes de réseaux d'antennes par optimisation convexe*, Ph.D. thesis, Université de Rennes I, Nov. 1994.

[21] C. Lemaréchal, F. Oustry, and C. Sagastizábal, *The $\mathcal{U}$-Lagrangian of a convex function*, Trans. Amer. Math. Soc., to appear.

[22] A. S. Lewis and M. L. Overton, *Eigenvalue optimization*, Acta Numerica, 5 (1996), pp. 149–190.

[23] M. S. Lobo, L. Vandenberghe, S. Boyd, and H. Lebret, *Second-order cone programming: Interior-point methods and engineering applications*, Linear Algebra Appl., submitted.

[24] Y. Nesterov and A. Nemirovskii, *Interior Point Polynomial Methods in Convex Programming: Theory and Applications*, Studies in Applied Mathematics series, SIAM, Philadelphia, PA, 1994.

[25] F. Oustry, *The $\mathcal{U}$-Lagrangian of the maximum eigenvalue function*, SIAM J. Optim., to appear.

[26] F. A. Potra and R. Sheng, *On homogeneous interior-point algorithms for semidefinite programming*, Optim. Methods Softw., 9 (1998), pp. 161–184.

[27] S. Robinson, *Stability theorems for systems of inequalities, part* II: *Differentiable nonlinear systems*, SIAM J. Numer. Anal., 13 (1976), pp. 497–513.

[28] R. T. Rockafellar, *Convex Analysis*, 2nd ed., Princeton Univ. Press, Princeton, NJ, 1970.

[29] R. Rubinstein and A. Shapiro, *Discrete Event Systems*, John Wiley, New York, 1993.
[30] A. Shapiro, *Perturbation theory of nonlinear programs when the set of optimal solutions is not a singleton*, Appl. Math. Optim., 18 (1988), pp. 215–229.
[31] A. Shapiro, *Perturbation analysis of optimization problems in Banach spaces*, Numer. Funct. Anal. Optim., 13 (1992), pp. 97–116.
[32] A. Shapiro, *First and second order analysis of nonlinear semidefinite programs. Semidefinite programming*, Math. Programming Ser. B, 77 (1997), pp. 301–320.
[33] A. Shapiro and M. K. H. Fan, *On eigenvalue optimization*, SIAM J. Optim., 5 (1995), pp. 552–568.
[34] A. Tikhonov and V. Arsenin, *Solutions of Ill-Posed Problems*, John Wiley, New York, 1977.
[35] L. Vandenberghe and S. Boyd, *Semidefinite programming*, SIAM Rev., 38 (1996), pp. 49–95.
[36] V. A. Yakubovich, *The solution of certain matrix inequalities in automatic control theory*, Soviet Math. Dokl., 3 (1962), pp. 620–623.
[37] K. Zhou, J. Doyle, and K. Glover, *Robust and Optimal Control*, Prentice–Hall, Upper Saddle River, NJ, 1995.