Technical Notes and Correspondence

Stochastic Algorithms for Exact and Approximate Feasibility of Robust LMIs

Giuseppe Calafiore and Boris T. Polyak

***Abstract*—In this note, we discuss fast randomized algorithms for determining an admissible solution for robust linear matrix inequalities (LMIs) of the form $F(x, \Delta) \preceq 0$, where x is the optimization variable and Δ is the uncertainty, which belongs to a given set $\mathbf{\Delta}$. The proposed algorithms are based on uncertainty randomization: the first algorithm finds a robust solution in a finite number of iterations with probability one, if a strong feasibility condition holds. In case no robust solution exists, the second algorithm computes an approximate solution which minimizes the expected value of a suitably selected feasibility indicator function. The theory is illustrated by examples of application to uncertain linear inequalities and quadratic stability of interval matrices.**

***Index Terms*—Linear matrix inequalities (LMIs), quadratic stability, robust semidefinite programming, stochastic algorithms, uncertainty randomization.**

I. Introduction

In this note, we discuss the problem of determining feasible or approximately feasible solutions to robust linear matrix inequality (LMI) constraints, using fast iterative methods based on uncertainty randomization. Robust LMI constraints arise naturally from standard LMI problems, when uncertainty is present in the data matrices. More formally, let $\mathbb{S}$ be a prescribed subspace of $\mathbb{R}^{p,q}$, which accounts for structure in the uncertainty (e.g., block-diagonal), then we shall consider structured norm-bounded uncertainty

$$\mathbf{\Delta} \equiv \mathbf{\Delta}_\rho \doteq \{\Delta \in \mathbb{S}: \|\Delta\| \leq \rho\} \tag{1}$$

where $\|\cdot\|$ denotes any matrix norm, and ρ denotes the uncertainty radius. Alternatively, the uncertainty may be described by a set of finite cardinality N_v, i.e., $\mathbf{\Delta} \equiv \mathbf{\Delta}_{\mathrm{f}} = \{\Delta_1, \ldots, \Delta_{N_v}\}$. A robust LMI is defined as the matrix inequality

$$F(x, \Delta) \preceq 0, \qquad \forall \Delta \in \mathbf{\Delta}, \qquad x \in \mathbb{R}^m \tag{2}$$

where

$$F(x, \Delta) = F_0(\Delta) + \sum_{i=1}^{m} x_i F_i(\Delta) \tag{3}$$

and where $F_i(\Delta) = F_i^T(\Delta) \in \mathbb{R}^{n,n}$ are functions of the uncertainty Δ. Let $\mathcal{X} \subseteq \mathbb{R}^m$ be a nonempty, convex and closed set, then the robust feasibility problem is posed as

$$\text{Find } x \in \mathcal{X} \text{ such that } F(x, \Delta) \preceq 0, \qquad \forall \Delta \in \mathbf{\Delta}. \tag{4}$$

A vector $\tilde{x}$ that satisfies (2) is said a *robustly feasible* solution.

To motivate our developments, we recall that the problem of determining a robustly feasible solution to (2), with Δ restricted as in (1), is in general NP-hard; [4], [9]. When the uncertainty Δ enters $F(x, \Delta)$ in a linear fractional form (LFT), the robust feasibility problem may be dealt with using the recently emerged techniques of robust semidefinite programming, [4], [9]. Even in the special (albeit quite general) case of LFT uncertainty, robust semidefinite programming (SDP) techniques provide in general only sufficient conditions for robust feasibility. Also, these techniques transform the original uncertain problem into a convex problem of much larger size. When uncertainty is present, even problems that are originally of small size result in convex optimization problems that push to the limits the existing SDP solver codes. As an example, the problem of assessing quadratic stability for an interval matrix of dimension n may be cast as a problem of finding a common solution to $N_v = 2^{n^2}$ Lyapunov inequalities for the vertex matrices, [6]. For $n = 4$, this results in an LMI problem involving a symmetric matrix of dimension $M \times M$, with $M = 4 \times 2^{16} = 262\,144$.

Another problem related to (4), is that there may actually exist no robustly feasible solution. In this case, it may be of interest to determine a solution that minimizes some average measure of constraints violation, as detailed later. A similar approach for approximate feasibility has recently been proposed by Barmish and Scherbakov in [3].

In this note, we present simple and efficient algorithms for the computation of feasible or approximately feasible solutions for robust LMIs, based on stochastic subgradient methods, [10]. A feature of the presented theory is also the absence of assumptions on the way the uncertainty enters into $F(x, \Delta)$. This framework may therefore accommodate the cases in which the uncertainty cannot be described using the LFT formalism. The price to pay for the aforementioned enhancements is that the proposed algorithms are based on randomization in the space of the uncertainty, therefore their solution and convergence properties can only be assessed in a probabilistic sense.

The main idea is to convert LMIs into convex scalar inequalities, and compute the relative subgradients in explicit form (Section II). Then, we generate random samples of the uncertainty and work, at each iteration, with the resulting random inequality.

The first algorithm (Theorem 1) is aimed to the solution of feasible problems. This uncertainty randomization approach and its application to the solution of robust LMIs, are new to the best of the authors' knowledge. A deterministic counterpart of the first method (Theorem 1) can be traced back to Kaczmarz method for solving linear equations [12], and to Agmon–Motzkin–Shoenberg method for solving linear inequalities, [1], [13]. For nonlinear convex inequalities, it has been proposed in [16]. A version with finite convergence for linear inequalities was used by Yakubovich in the middle of the sixties for solving adaptive control problems (see [5] and references therein), and then extended to nonlinear inequalities by Fomin [11]. In the same line of research, a similar randomized approach is proposed in [17] to solve quadratic inequalities, in the context of robust linear quadratic regulator (LQR) design.

The second algorithm (Theorem 2) deals with the case when no robustly feasible solution exists. In this case, it can be employed to find a solution that minimizes an averaged measure of constraint violation. This second method is indeed an application of a general method by Nemirovskii and Yudin [14] for stochastic optimization. This method does not seem to have been exploited earlier for the solution of LMIs with uncertain data. A conference version of the present note has appeared in [8].

Manuscript received August 15, 2000; revised January 24, 2001. Recommended by Associate Editor J. Chen. The work of B. T. Polyak was supported by a NATO-CNR Guest Fellowship while visiting IRITI-CNR.

G. Calafiore is with the Dipartimento di Automatica e Informatica, Politecnico di Torino 10129 Torino, Italy (e-mail: calafiore@polito.it).

B. T. Polyak is with the Institute for Control Science, Russian Academy of Sciences, Moscow 117806, Russia (e-mail: boris@ipu.rssi.ru).

Publisher Item Identifier S 0018-9286(01)10346-6.

0018–9286/01$10.00 © 2001 IEEE

II. Preliminaries and Computation of Subgradients

We assume that the uncertainty Δ is a random matrix, with given probability distribution f_Δ over its support set $\mathbf{\Delta}$, and that it is possible to generate samples of Δ according to this probability distribution. We will not discuss here issues related to the choice of the probability distribution f_Δ. A natural choice (which also has theoretical rationale behind, see [2]) is to assume uniform distribution over $\mathbf{\Delta}$. In this case, there exist efficient algorithms for generation of uniformly distributed indices [for the case of a discrete uncertainty set (1)], for generation of random vectors uniformly distributed in ℓ_p-norm bounded balls, and for uniform generation of matrix samples in operator-norm bounded sets [for the case of matrix uncertainty (1)], see [7] and the references therein.

We introduce a scalar function $\varphi(x, \Delta)$, defined as

$$\varphi(x, \Delta) = \|F_+(x, \Delta)\| \tag{5}$$

where the notation A_+ indicates the projection of the matrix A onto the cone of positive–semidefinite matrices, as further detailed in the sequel. This function will be called a *feasibility indicator function* (FIF) in the sequel. Another (similar) class of feasibility indicators has been proposed in [3] for generic constraint sets. From the definition (5), it follows that $\varphi(x, \Delta) > 0$ if and only if $F(x, \Delta) \not\preceq 0$, and it is zero otherwise. With these premises, a vector $\tilde{x} \in \mathbb{R}^m$ is robustly feasible for (2) if and only if it satisfies $\varphi(\tilde{x}, \Delta) = 0$, for all $\Delta \in \mathbf{\Delta}$. If there exist no robustly feasible point, we say that $\tilde{x}$ is an *approximately feasible* solution for (4) if it minimizes the mean indicator function, i.e.,

$$\tilde{x} = \arg\min_{x \in \mathcal{X}} E\{\varphi(x, \Delta)\} \tag{6}$$

where E denotes expectation with respect to Δ. In this case, the function $\varphi(x, \Delta)$ measures the "distance" from feasibility, and the approximate solution $\tilde{x}$ is the one that minimizes the expectation over Δ of the constraint violation.

We introduce the notation $[\xi]_\mathcal{X}$ to denote the projection of the element ξ onto $\mathcal{X}$, i.e.,

$$\|\xi - [\xi]_\mathcal{X}\| = \min_{y \in \mathcal{X}} \|\xi - y\|.$$

Let $\mathcal{S}^n$ denote the Hilbert space of $n \times n$ real symmetric matrices, equipped with the Frobenius norm and the scalar product $\langle A, B\rangle = \mathbf{Tr}\, AB$. Let $\mathcal{C}$ be the cone of positive–semidefinite matrices, $\mathcal{C} = \{A \in \mathcal{S}^n\colon A \succeq 0\}$. The following lemma will be used in the sequel.

Lemma 1: Let A_+ be the projection of $A \in \mathcal{S}^n$ onto the cone $\mathcal{C}$, then the following conditions are equivalent:

1) $A_+ = [A]_\mathcal{C}$, i.e., A_+ is a projection of A onto $\mathcal{C}$;
2) $A = A_+ + A_-$, where $A_+ \in \mathcal{C}$, $A_- \in -\mathcal{C}$, and $\langle A_+, A_-\rangle = 0$;
3) $A_+ = RD_+R^T$, where R and $D = \mathbf{diag}(d_1, \ldots, d_n)$ are the eigenvector and eigenvalues matrices of A, respectively, that is $A = RDR^T$, and $D_+ = \mathbf{diag}(d_1^+, \ldots, d_n^+)$, being $d_i^+ = \max(d_i, 0)$.

A proof of this lemma may be found in [15]. We remark that the assertion 3) provides a direct method for explicitly computing the projection A_+.

We recall that a function $f\colon \mathbb{R}^m \to \mathcal{S}^n$ is convex (in matrix sense) if for any $x_1, x_2 \in \mathbb{R}^m$,

$$f(tx_1 + (1-t)x_2) \preceq tf(x_1) + (1-t)f(x_2), \qquad \text{for } 0 \le t \le 1.$$

The following lemmas, whose simple proofs are omitted for brevity, hold.

Lemma 2: The function $F_+(x, \Delta)$ is convex in x, in the above sense.

Lemma 3: If a function $f\colon \mathbb{R}^m \to \mathcal{C}$ is convex, then the function $g\colon \mathbb{R}^m \to \mathbb{R}$, $g(x) = \|f(x)\|$ is also convex.

As an immediate corollary, we have the following.

Lemma 4: The function $\varphi(x, \Delta) = \|F_+(x, \Delta)\|$, where $F(x, \Delta)$ is defined in (3), is convex in x.

The subgradient of this function can be found as follows.

Lemma 5: Let $\varphi(x, \Delta)$ be defined as in (5), where $F(x, \Delta)$ is defined in (3), and let

$$\nabla\varphi(x, \Delta) \doteq \frac{1}{\varphi(x, \Delta)} \begin{bmatrix} \mathbf{Tr} F_1(\Delta)F_+(x, \Delta) \\ \vdots \\ \mathbf{Tr} F_m(\Delta)F_+(x, \Delta) \end{bmatrix}$$

then

$$\partial_x\varphi(x, \Delta) = \begin{cases} \nabla\varphi(x, \Delta), & \text{if } \varphi(x, \Delta) \neq 0 \\ 0, & \text{otherwise} \end{cases}$$

is a subgradient of $\varphi(x, \Delta)$ at x.

Notice that the above subgradient does not tend to infinity when x is nearly feasible [i.e., when $\varphi(x, \Delta) \to 0$]. It can indeed be easily proved that if $F_i(x, \Delta)$ are bounded on $\mathbf{\Delta}$, then the subgradient remains bounded on any bounded set. We also emphasize that the computation of a subgradient basically requires the solution of a symmetric eigenvalue problem, therefore the iterative steps in the algorithms below can be performed efficiently. Of course, if the dimension n is very large, there may arise conditioning problems in the eigenvalues computation, that may require specific care. However, the problems of main interest in this note are those with very large number of uncertainties, but with moderate dimensions of the matrices involved.

III. Algorithm for Feasible Case

We consider first the case when a robustly feasible solution exists. In particular, we assume that a *strong feasibility* condition holds: there exist $x^* \in \mathcal{X}$, $\varepsilon > 0$ such that

$$F(x, \Delta) \preceq 0, \qquad \forall\, x \in \mathcal{X}\colon \|x - x^*\| \le \varepsilon, \quad \forall\, \Delta \in \mathbf{\Delta}. \tag{7}$$

Consider the following recursion:

$$x_{k+1} = [x_k - \lambda_k \partial_x\{\varphi(x_k, \Delta^k)\}]_\mathcal{X} \tag{8}$$

where ∂_x denotes a subgradient with respect to x, and Δ^k, $k = 0, 1, \ldots$ are i.i.d. random samples drawn from f_Δ. Define the stepsizes λ_k as

$$\lambda_k = \begin{cases} \eta \dfrac{\varphi(x_k, \Delta^k) + \varepsilon\|\partial_x\{\varphi(x_k, \Delta^k)\}\|}{\|\partial_x\{\varphi(x_k, \Delta^k)\}\|^2}, \\ \qquad \text{if } \varphi(x_k, \Delta^k) \neq 0; \\ 0, \qquad \text{otherwise} \end{cases} \tag{9}$$

where $0 < \eta < 2$ is a parameter of the algorithm. We shall also need the following technical assumption, that guarantees nonzero probability to distinguish if a vector x is a feasible solution or not.

Assumption 1: If x is not a robustly feasible solution for (2), then it must hold that

$$\Pr(F(x, \Delta) \not\preceq 0) > 0.$$

This means that the measure of the set of "bad" Δs with respect to the probability density f_Δ must be nonzero. The convergence result for the recursion (8) is stated in the following theorem.

Theorem 1: If Assumption 1 and the strong feasibility condition (7) hold, then for any $x_0 \in \mathcal{X}$, the recursion (8) finds a robustly feasible solution of (2) in a finite number of iterations with probability one.

Proof: Define

$$\overline{x} = x^* + \frac{\varepsilon}{\|\partial_x\{\varphi(x_k, \Delta^k)\}\|}\,\partial_x\{\varphi(x_k, \Delta^k)\}$$

where x^* is a robustly feasible solution. Then, due to (7), $\overline{x}$ is a feasible solution of (2) and, in particular, $\varphi(\overline{x}, \Delta^k) = 0$ for all k. Now, due to the properties of a projection

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\le \|x_k - x^* - \lambda_k\partial_x\{\varphi(x_k, \Delta^k)\}\|^2\\ &= \|x_k - x^*\|^2 + \lambda_k^2\|\partial_x\{\varphi(x_k, \Delta^k)\}\|^2\\ &\quad - 2\lambda_k(x_k - \overline{x})^T\partial_x\{\varphi(x_k, \Delta^k)\}\\ &\quad - 2\lambda_k(\overline{x} - x^*)^T\partial_x\{\varphi(x_k, \Delta^k)\}.\end{aligned}$$

We now consider the last two terms in the inequality above. Due to convexity of $\varphi(x, \Delta)$ and to the feasibility of $\overline{x}$, we obtain

$$\begin{aligned}(x_k - \overline{x})^T\partial_x\{\varphi(x_k, \Delta^k)\} &\ge \varphi(x_k, \Delta^k) - \varphi(\overline{x}, \Delta^k)\\ &\ge \varphi(x_k, \Delta^k)\end{aligned}$$

while, due to definition of $\overline{x}$

$$(\overline{x} - x^*)^T\partial_x\{\varphi(x^k, \Delta^k)\} = \varepsilon\|\partial_x\{\varphi(x_k, \Delta^k)\}\|.$$

Thus, we write

$$\begin{aligned}\|x_{k+1} - x^*\|^2 \le \|x_k - x^*\|^2 &+ \lambda_k^2\|\partial_x\{\varphi(x_k, \Delta^k)\}\|^2\\ &-2\lambda_k(\varphi(x_k, \Delta^k) + \varepsilon\|\partial_x\{\varphi(x_k, \Delta^k)\}\|).\end{aligned}$$

Now, if $\varphi(x_k, \Delta^k) > 0$, substituting in the above the value of λ_k given in (9), we get

$$\begin{aligned}\|x_{k+1} - x^*\|^2 &\le \|x_k - x^*\|^2\\ &\quad - \frac{\eta(2-\eta)(\varphi(x_k, \Delta^k) + \varepsilon\|\partial_x\{\varphi(x_k, \Delta^k)\}\|)^2}{\|\partial_x\{\varphi(x_k, \Delta^k)\}\|^2}\\ &\le \|x_k - x^*\|^2 - \varepsilon^2\eta(2-\eta).\end{aligned}$$

Therefore, if $\varphi(x_k, \Delta^k) > 0$, then we obtain

$$\|x_{k+1} - x^*\|^2 \le \|x_k - x^*\|^2 - \varepsilon^2\eta(2-\eta).$$

From this formula, we conclude that no more than $M = \|x_0 - x^*\|^2/(\varepsilon^2\eta(2-\eta))$ correction steps can be executed. On the other hand, if x_k is infeasible, then, due to Assumption 1, there is a nonzero probability to make a correction step. Thus, with probability one, the method can not terminate at an infeasible point. We therefore conclude that the algorithm must terminate after a finite number of iterations at a feasible solution. □

Remarks: In the stepsize rule (9) we have assumed that the value of ε (the radius of a ball contained in the feasibility set) is known. If it is not the case, we can replace ε in (9) by $\varepsilon_{s(k)}$, where $s(k)$ is the number of correction steps performed before the kth iteration, and $\varepsilon_s > 0$, $\varepsilon_s \to 0$, $\sum_{s=0}^{\infty}\varepsilon_s = \infty$. Then, the last inequality in the proof of the theorem writes

$$\|x_{k+1} - x^*\|^2 \le \|x_k - x^*\|^2 - \eta(2-\eta)\sum_{s=0}^{s(k)}\varepsilon_s^2.$$

From this inequality, we can again conclude that the modified algorithm terminates in a finite number of steps with probability one.

The number M of correction steps strongly depends on ε. It is confirmed by numerical simulations that the method converges very fast (often in less than 100 iterations) if the feasible set is "large enough," while for "small" feasible set convergence is much slower. In the limit case, if the feasible set shrinks to a singleton, the strong feasibility assumption is violated and the algorithm fails to converge.

Notice also that computing the projections onto X is a convex problem, which has a particularly simple solution when X is, for instance, a half-space, the positive orthant, an ellipsoid, or, in general, a closed set defined by an LMI.

In [17], the authors present a result similar in spirit to Theorem 1, but for a specific problem related to quadratic inequalities in the context of robust controller design. The present work generalizes that result to the general framework of robust LMIs.

IV. Algorithm for Approximate Feasibility

In this section, we propose an iterative randomized algorithm for determining an approximately feasible solution for (4). To this end, we specialize a result on stochastic optimization in Banach spaces of Nemirowskii and Yudin [14]. Here, we provide a new statement and proof which simplify and adapt to the robust LMI problem the original general setting of [14].

Theorem 2: Consider the robust feasibility problem (4). Let Δ be a random matrix with given probability distribution f_Δ over the support $\boldsymbol{\Delta}$. Let $\varphi(x, \Delta)$ be a feasibility indicator function, and let

$$f(x) \doteq E\{\varphi(x, \Delta)\}$$

where E denotes statistical expectation. Suppose further that $\mathcal{X}$ is a convex and closed set, $f(x)$ has a minimum point $\tilde{x}$ on $\mathcal{X}$, and $\|\partial_x\{\varphi(x, \Delta)\}\| \le \mu$, for all $x \in \mathcal{X}$, $\Delta \in \boldsymbol{\Delta}$. Given an initial vector x_0, consider the recursion

$$x_{k+1} = [x_k - \lambda_k\partial_x\{\varphi(x_k, \Delta^k)\}]_{\mathcal{X}} \tag{10}$$

where ∂_x denotes a subgradient with respect to x, and Δ^k, $k = 0, 1, \ldots$ are i.i.d. random samples drawn from f_Δ. Let

$$\overline{x}_k = \frac{m_{k-1}}{m_k}\,\overline{x}_{k-1} + \frac{\lambda_k}{m_k}\,x_k \tag{11}$$

for $\overline{x}_0 = 0$, $m_0 = 0$, $m_k = m_{k-1} + \lambda_k$, $\lambda_k > 0$. Then

$$Ef(\overline{x}_k) - f(\tilde{x}) \le C(k) \tag{12}$$

where

$$C(k) = \frac{\|x_0 - \tilde{x}\|^2 + \mu^2\sum_{i=0}^{k-1}\lambda_i^2}{2\sum_{i=0}^{k-1}\lambda_i}.$$

Moreover, if it holds that $\lambda_i \to 0$, $\sum_{i=0}^{\infty}\lambda_i = \infty$, then $\lim_{k\to\infty} Ef(\overline{x}_k) = f(\tilde{x})$.

Proof: Consider the distance from the current point x_{k+1} to the optimal point $\tilde{x}$. By the definition of projection, we have that $\|[x]_{\mathcal{X}} - x^*\| \le \|x - x^*\|$, for any x, and for any $x^* \in \mathcal{X}$, therefore

$$\begin{aligned}\|x_{k+1} - \tilde{x}\|^2 \le \|x_k - \tilde{x}\|^2 &- 2\lambda_k(x_k - \tilde{x})^T\partial\varphi(x_k, \Delta^k)\\ &+\lambda_k^2\|\partial\varphi(x_k, \Delta^k)\|^2.\end{aligned} \tag{13}$$

Now, for a convex function $g(x)$, for any x, x^* it holds that $(x - x^*)^T\partial g(x) \ge g(x) - g(x^*)$, hence $(x_k - \tilde{x})^T\partial\varphi(x_k, \Delta^k) \ge \varphi(x_k, \Delta^k) - \varphi(\tilde{x}, \Delta^k)$. On the other hand, from the boundedness condition on the subgradients we have that $\|\partial\varphi(x_k, \Delta^k)\|^2 \le \mu^2$, therefore (13) writes

$$\begin{aligned}\|x_{k+1} - \tilde{x}\|^2 &\le \|x_k - \tilde{x}\|^2\\ &\quad -2\lambda_k(\varphi(x_k, \Delta^k) - \varphi(\tilde{x}, \Delta^k)) + \lambda_k^2\mu^2.\end{aligned} \tag{14}$$

Denoting $u_k = E\|x_k - \tilde{x}\|^2$, and taking expectation of both sides of (14), we get

$$u_{k+1} \le u_k - 2\lambda_k(Ef(x_k) - f(\tilde{x})) + \lambda_k^2\mu^2$$

and $u_k \le u_0 - 2\sum_{i=0}^{k-1}\lambda_i(Ef(x_i) - f(\tilde{x})) + \mu^2\sum_{i=0}^{k-1}\lambda_i^2$, therefore

$$\sum_{i=0}^{k-1}\lambda_i(Ef(x_i) - f(\tilde{x})) \le \frac{1}{2}\left(u_0 + \mu^2\sum_{i=0}^{k-1}\lambda_i^2\right). \qquad (15)$$

It is obvious that (11) provides a recurrent version of a Cesaro mean, therefore $\overline{x}_k$ is an averaged version of the x_ks

$$\overline{x}_k = \frac{\sum_{i=0}^{k}\lambda_i x_i}{\sum_{i=0}^{k}\lambda_i}.$$

From Jensen inequality for convex functions, we then have

$$f(\overline{x}_k) = f\left(\frac{\sum_{i=0}^{k}\lambda_i x_i}{\sum_{i=0}^{k}\lambda_i}\right) \le \frac{\sum_{i=0}^{k}\lambda_i f(x_i)}{\sum_{i=0}^{k}\lambda_i}$$

therefore

$$Ef(\overline{x}_k) - f(\tilde{x}) \le \frac{\sum_{i=0}^{k}\lambda_i(Ef(x_i) - f(\tilde{x}))}{\sum_{i=0}^{k}\lambda_i}$$

and, by (15)

$$Ef(\overline{x}_k) - f(\tilde{x}) \le \frac{u_0 + \mu^2\sum_{i=0}^{k}\lambda_i^2}{2\sum_{i=0}^{k}\lambda_i}$$

which proves (12). With the further assumptions on the step sizes $\sum_{i=0}^{\infty}\lambda_i = \infty;\ \lambda_i \to 0$, it immediately follows that $Ef(\overline{x}_k) \to f(\tilde{x})$, for $k \to \infty$. □

Remark: Theorem 2 also gives an estimate of solution accuracy for a finite sample. In particular, optimizing $C(k)$ over λ_i, we get $\lambda_i = \lambda^* = (\|x_0 - \tilde{x}\|/\mu)(1/\sqrt{k})$, $i = 0, \ldots, k-1$. If the number of iterations is fixed in advance to k, then the best choice are constant step sizes $\lambda^* = \|x_0 - \tilde{x}\|/\mu\sqrt{k}$, which yields

$$Ef(\overline{x}_k) - f(\tilde{x}) \le 2\frac{\|x_0 - \tilde{x}\|\mu}{\sqrt{k}}.$$

If we do not fix *a priori* the number of iterations, a good choice for the steps is $\lambda_k = c/\sqrt{k}$, which provides asymptotically the same estimate $Ef(\overline{x}_k) - f(\tilde{x}) = O(1/\sqrt{k})$.

V. Applications

In this section, we present numerical examples of application of the proposed theory. In particular, we consider an example involving robust feasibility of uncertain linear algebraic inequalities, and problems related to quadratic stability of interval plants.

A. Uncertain Linear Inequalities

Consider a robust feasibility problem for a set of linear algebraic inequalities in the form

$$A(\Delta_A)x \le b(\Delta_b) \qquad (16)$$

where $A \in \mathbb{R}^{n,m}$, $x \in \mathbb{R}^m$, Δ_A, Δ_b are matrices that account for the structured uncertainty acting on A and b, and the $\le$ sign is to be intended element-wise. In particular, we will consider a randomly generated example with additive, independent uncertainty acting on each entry of the data A, b: $A(\Delta_A) = A_0 + \Delta_A;\ b(\Delta_b) = b_0 + \Delta_b$, where

$$A_0 = \begin{bmatrix} -12.8819 & 13.6427 & -8.1623 \\ -9.5296 & 4.8204 & 20.9407 \\ 7.7817 & -7.8707 & 0.8015 \\ -0.0633 & 7.5200 & -9.3730 \\ 5.2449 & -1.6689 & 6.3574 \end{bmatrix}$$

$$b_0 = \begin{bmatrix} 1.6820 \\ 0.5936 \\ 0.7902 \\ 0.1053 \\ -0.1586 \end{bmatrix} \qquad \Delta \doteq [\,\Delta_A \quad \Delta_b\,] \qquad \|\Delta\|_\infty \le r$$

where $\|\cdot\|_\infty$ denotes the ℓ_∞-norm (each element of the matrix is independently bounded in magnitude). It is clear that (16) can be rewritten as a robust LMI problem (see, for instance, [18]), therefore the algorithms proposed above may be applied to determine a feasible or approximately feasible solution. However, we can in this case dispense with the general matrix notation, and repeat a similar reasoning for the cone of nonnegative vectors. A feasibility indicator function is in this case simply given by $\varphi(x, \Delta) = \|(Ax - b)_+\|$, where the suffix $+$ is now to be intended as projection onto the cone of nonnegative vectors. A subgradient is easily computed following Lemma 5 as

$$\partial_x\varphi(x, \Delta) = \begin{cases} A^T(Ax - b)_+/\|(Ax - b)_+\|, & \text{if } \varphi(x, \Delta) \ne 0 \\ 0, & \text{otherwise.} \end{cases}$$

We first set $r = 0.65$, and run $N = 250$ iterations of the stochastic gradient algorithm proposed in Section IV, starting with $x_0 = 0$, and found the solution $x_N = [\,-0.2724 \quad -0.2526 \quad -0.0882\,]^T$. We then performed a Monte-Carlo test using x_N, to estimate the empirical probability of robust feasibility. Using 100 000 uniformly generated samples of the uncertainty, we obtained $\hat{p}_{feas} = 0.9989$. Then, we set $r = 0.55$, and applied the algorithm of Section IV with $\eta = 1.8$, $x_0 = 0$. This algorithm converged in less than $N = 30$ iterations to the solution $x_{N,f} = [\,-0.1697 \quad -0.1719 \quad -0.0565\,]^T$. Checking 100 000 randomly generated systems of inequalities (16), we verified all of them to be satisfied. It is interesting to note that the point x_N also satisfies all inequalities with $r = 0.55$.

B. Quadratic Stability of Interval Plant

We here consider the problem of assessing quadratic stability for a linear system described by a matrix whose elements belong to independent intervals, that is

$$A(\Delta) = A_0 + \Delta, \qquad |\Delta_{ij}| \le rS_{ij}, \qquad r > 0. \qquad (17)$$

We first use a 3 × 3 example system taken from [3], where

$$A_0 = \begin{bmatrix} -2 & -2 & 0 \\ 1 & 0 & 0 \\ 1 & 0 & -2 \end{bmatrix}$$

$$S = \begin{bmatrix} 0.651 & 0.9394 & 0.5691 \\ 0.2451 & 0.4727 & 0.1457 \\ 0.7004 & 0.4014 & 0.3141 \end{bmatrix}.$$

It is well known (see, for instance, [6]) that the system (17) is quadratically stable if and only if there exists a matrix $P \succ 0$ that simultaneously satisfies $N_v = 2^{n^2} = 512$ Lyapunov inequalities involving the vertex matrices

$$A^T(\Delta_v^k)P + PA(\Delta_v^k) \preceq 0, \qquad k = 1, \ldots, N_v$$

where Δ_v^k represents the kth vertex of the polyhedron described by $|\Delta_{ij}| \le rS_{ij}$.

We first set $r = 0.5$. In this case, we know from [3] that there exists a common solution for all the Lyapunov inequalities involving the vertices. The algorithm proposed in Section IV, drawing one uniform random vertex sample of the uncertainty at each iteration, converges in less than $N = 50$ iterations to the solution

$$P_{N,f} = \begin{bmatrix} 1.2487 & 0.8155 & 0.3177 \\ 0.8155 & 2.0443 & 0.2425 \\ 0.3177 & 0.2425 & 0.5371 \end{bmatrix}.$$

This solution may be checked to actually satisfy simultaneously all the required Lyapunov inequalities.

We then set $r = 1$. In this case it may be proved that no common solution exists for the Lyapunov inequalities. We computed an approximately feasible solution using the algorithm in Section III. After $N = 250$ iterations we obtained the solution

$$P_N = \begin{bmatrix} 1.2042 & 0.9899 & -0.2649 \\ 0.9899 & 1.7455 & -0.0967 \\ -0.2649 & -0.0967 & 0.5577 \end{bmatrix}.$$

For this solution, we computed the empirical probability of satisfaction of robust feasibility, using 100 000 uniform random samples of the uncertainty, obtaining $\hat{p}_{feas} = 0.996$. It should be remarked that the above solutions are computed in about one second on a standard workstation. Also, as already mentioned, the exact simultaneous solution of the above Lyapunov equations goes beyond the capabilities of most of the existing LMI solvers, for systems of order greater than four. As a more challenging example, we considered a system of order $n = 10$, with nominal matrix A_0 available at http://www.polito.it/~calafior/papers/data/data_stoclmi.htm. The considered system was obtained by integer truncation, and it is nominally stable. We want to determine a common Lyapunov matrix $P \succ 0$ that proves quadratic stability, for independent element-wise uncertainty on the entries of A_0, with $r = 0.5$. Using standard tools, this would require the simultaneous solution of $2^{100} \simeq 1.26\ 10^{30}$ Lyapunov inequalities.

We used the algorithm of Section III, starting from an initial point P_0 that solves the Lyapunov equation for the nominal plant, $A_0^T P_0 + P_0 A_0 = -I$. The algorithm converged in about $N = 10\,000$ iterations to the solution P_N which, for space reasons, is reported at the above web address. For this solution, we estimated the empirical probability of robust feasibility on 1 000 000 randomly selected vertices, obtaining $\hat{p}_{feas} = 1.0$.

VI. Conclusion

Two fast randomized algorithms for determining feasible or approximately feasible solutions to robust LMI problems have been discussed in this note. These techniques proved to be useful for problems which are intractable by means of standard exact LMI methods. In all cases, the solution provided by the randomized approach (which comes at very low computational cost) may serve as a good initial guess for a deterministic algorithm.

The first algorithm guarantees termination in a finite number of steps, if a robustly feasible solution exists. The second algorithm is general purpose, and can be applied in the case when we do not know in advance if a robustly feasible solution exists. Numerical tests showed that the second method has in general slower convergence than the first one. In both cases, the "goodness" of resulting solution may be checked *a posteriori* via Monte-Carlo randomization, as discussed in the examples.

It is expected that these techniques will be applied advantageously to the solution of various problems related to robust control design. Further research is needed in the direction of developing similar algorithms for the optimization of a functional under robust constraints, and to the extension of these results to general *convex* matrix inequalities.

Acknowledgment

The authors would like to thank F. Dabbene, L. El Ghaoui, P. Scherbakov, and R. Tempo for the many useful inputs and discussions.

References

[1] S. Agmon, "The relaxation method for linear inequalities," *Canad. J. Math.*, vol. 6, pp. 382–393, 1954.

[2] E. W. Bai, R. Tempo, and M. Fu, "Worst case properties of the uniform distribution and randomized algorithms for robustness analysis," *Math. Control, Sig., Syst.*, vol. 11, pp. 183–196, 1998.

[3] B. R. Barmish and P. S. Scherbakov, "On avoiding vertexization of robustness problems: The approximate feasibility concept," in *Proc. 39th IEEE Conf. Decision Control*, vol. 2, Sydney, Australia, Dec. 2000, pp. 1031–1036.

[4] A. Ben-Tal, L. El Ghaoui, and A. Nemirovskii, "Robust semidefinite programming," in *Handbook of Semidefinite Programming*, R. Saigal, L. Vandenberghe, and H. Wolkowicz, Eds. Waterloo, Canada: Kluwer, 2000.

[5] V. Bondarko and V. A. Yakubovich, "The method of recursive aim inequalities in adaptive control theory," *Int. J. Adapt. Control Signal Processing*, vol. 6, pp. 141–160, 1992.

[6] S. P. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequalities in Systems and Control Theory*. Philadelphia, PA: SIAM, 1994.

[7] G. Calafiore, F. Dabbene, and R. Tempo, "Randomized algorithms for probabilistic robustness with real and complex structured uncertainty," *IEEE Trans. Automat. Contr.*, vol. 45, pp. 2218–2235, Dec. 2000.

[8] G. Calafiore and B. T. Polyak, "Fast algorithms for exact and approximate feasibility of robust LMIs," in *Proc. 39th IEEE Conf. Decision Control*, vol. 5, Sydney, Australia, Dec. 2000, pp. 5035–5040.

[9] L. El Ghaoui, F. Oustry, and H. Lebret, "Robust solutions to uncertain semidefinite programs," *SIAM J. Optim.*, vol. 9, no. 1, pp. 33–52, 1998.

[10] Y. M. Ermoliev, *Methods of Stochastic Programming* (in Russian). Moscow, Russia: Nauka, 1976.

[11] V. N. Fomin, *Mathematical Theory of Learning Recognizing Systems* (in Russian). Leningrad, Russia: LGU Publications, 1976.

[12] S. Kaczmarz, "Angenäherte Aufslösung von Systemen Linearer Gleichungen," *Bull. Int. Acad. Polon. Sci., Lett. A*, pp. 355–357, 1937. English translation: "Approximate solution of systems of linear equations," *Int. J. Control*, vol. 57, no. 6, pp. 1269–1271, 1993.

[13] T. S. Motzkin and I. Shoenberg, "The relaxation method for linear inequalities," *Canad. J. Math.*, vol. 6, pp. 393–404, 1954.

[14] A. S. Nemirovsky and D. B. Yudin, *Information Complexity and Efficient Methods for Solution of Convex Extremal Problems* (in Russian). Moscow, Russia: Nauka, 1978. Engish translation: NY: Wiley, 1983.

[15] B. Polyak, "Gradient methods for solving equations and inequalities," *USSR Comp. Math. Math. Phys.*, vol. 4, no. 6, pp. 17–32, 1964.

[16] ——, "Minimization of nonsmooth functionals," *USSR Comp. Math. Math. Phys.*, vol. 9, no. 3, pp. 14–29, 1969.

[17] B. T. Polyak and R. Tempo, "Probabilistic robust design with linear quadratic regulators," in *Proc. 39th IEEE Conf. Decision Control*, vol. 2, Sidney, Australia, Dec. 2000, pp. 1037–1042.

[18] L. Vandenberghe and S. Boyd, "Semidefinite programming," *SIAM Rev.*, vol. 38, no. 1, pp. 49–95, Mar. 1996.