LMI relaxations in robust control (tutorial)

C.W. Scherer∗
Delft Center for Systems and Control
Delft University of Technology
Mekelweg 2, 2628 CD Delft, The Netherlands
c.w.scherer@dcsc.tudelft.nl

Abstract

This paper accompanies a survey presentation within the mini-symposium entitled “LMIs in systems and control: some recent progress and new trends”. We address the fundamental role of robust LMI problems for the analysis and synthesis of control systems. Moreover we discuss how to systematically construct families of relaxations in a unified framework if the data matrices admit linear fractional representations. We conclude by presenting recent insights into numerical tests for verifying the exactness of a given relaxation, and comment on the construction of relaxation families that are guaranteed to be asymptotically exact.

During the last fifteen years there has been a tremendous activity to identify control problems that can be translated into linear semi-definite programs which are generically formulated as follows: given a vector $c$ and symmetric matrices $F_0, F_1, \ldots, F_n$, minimize $c'x$ over all $x \in \mathbb{R}^n$ which satisfy

$$F_0 + x_1F_1 + \cdots + x_nF_n < 0.$$ 

Most classical problems such as $H_2$- and $H_\infty$-optimal control have been successfully subsumed to such formulations, but there exist a whole variety of other classes of control problems for which it is known how to achieve such a re-parametrization in an exact fashion [7, 10].

In robust control the system descriptions are assumed to be affected by either time-invariant, time-varying parametric or dynamic uncertainties. Then the data matrices $F_0(\delta), F_1(\delta), \ldots, F_n(\delta)$ are functions of some real or complex parameter $\delta$ that is only know to be contained in some set $\delta$, and the goal is to minimize $c'x$ over all $x \in \mathbb{R}^n$ such that

$$F_0(\delta) + x_1F_1(\delta) + \cdots + x_nF_n(\delta) < 0 \text{ for all } \delta \in \delta.$$ 

*Research supported by the Technology Foundation STW, applied science division of NWO and the technology programme of the Ministry of Economic Affairs.
If $F_i(\delta), i = 0, 1, \ldots, n$ depend affinely on $\delta$ and $\delta$ is the convex hull of (a moderate number of) finitely many generators, the described robust semi-definite program can be readily reduced to a standard LMI problem. The situation drastically differs in case that the dependence is non-linear in $\delta$. If the dependence is rational without pole at zero - as is often true in control - one can construct the following linear fractional representation

$$\begin{bmatrix}
F_0(\delta) \\
F_1(\delta) \\
\vdots \\
F_l(\delta)
\end{bmatrix} = \begin{bmatrix}
C_0 & D_0 \\
C_1 & D_1 \\
\vdots & \vdots \\
C_l & D_l
\end{bmatrix} \Delta(\delta)(I - A\Delta(\delta))^{-1}B + \begin{bmatrix}
D_0 \\
D_1 \\
\vdots \\
D_l
\end{bmatrix}$$

with $\Delta(\delta)$ typically being block-diagonal and depending linearly on $\delta$. If we abbreviate

$$W(x) = \begin{bmatrix}
C_0 & D_0 \\
C_1 & D_1 \\
\vdots & \vdots \\
C_l & D_l
\end{bmatrix}' \begin{bmatrix}
0 & 0 & \cdots & 0 & I \\
0 & 0 & \cdots & 0 & x_1I \\
\vdots & \vdots & \ddots & \vdots & \vdots \\
0 & I & \cdots & 0 & x_lI
\end{bmatrix} \begin{bmatrix}
C_0 & D_0 \\
C_1 & D_1 \\
\vdots & \vdots \\
C_l & D_l
\end{bmatrix}$$

and $\Delta := \{\Delta(\delta) : \delta \in \delta\}$, the robust semi-definite program amounts to minimizing $c'x$ over all $x \in \mathbb{R}^n$ with

$$\begin{bmatrix}
\Delta(I - A\Delta)^{-1}B \\
I
\end{bmatrix}' W(x) \begin{bmatrix}
\Delta(I - A\Delta)^{-1}B \\
I
\end{bmatrix} \prec 0 \quad \text{for all } \Delta \in \Delta. \quad (1)$$

Let us denote the optimal value by $\gamma_{\text{opt}}$.

For this problem it is possible to construct a relaxation for computing an upper bound as follows. We choose linear Hermitian-valued mappings $G(y)$ and $H(y)$, defined in the variable $y$ that lives in some finite-dimensional inner product space, such that

$$G(y) \preceq 0 \quad \text{implies} \quad \begin{bmatrix}
\Delta \\
I
\end{bmatrix}' H(y) \begin{bmatrix}
\Delta \\
I
\end{bmatrix} \succeq 0 \quad \text{for all } \Delta \in \Delta. \quad (2)$$

For any such pair $G, H$ consider the problem of infimizing $c'x$ over all $x, y$ with

$$G(y) \prec 0, \quad \begin{bmatrix}
I & 0 \\
A & B
\end{bmatrix}' H(y) \begin{bmatrix}
I & 0 \\
A & B
\end{bmatrix} + W(x) \prec 0.$$

Computing its optimal value denoted by $\gamma_{\text{rel}}$ amounts to solving a standard linear SDP, and it is straightforward to prove on the basis of (2) that $\gamma_{\text{opt}} \leq \gamma_{\text{rel}}$.

Before addressing properties of the relaxation in somewhat more detail, let us point out the surprising modeling power of this framework. For example it comprises robust optimization and robust linear algebra problems [11, 6]. Among the numerous practically interesting instances, consider a matrix version of the classical least-squares approximation problem for the purpose of illustration. The nominal least-squares problem can be formulated as

$$\min_X \|XM - N\|_F^2.$$
where $M$, $N$ are data matrices and $\| \cdot \|_F$ denotes the Frobenius norm. If $M$, $N$ have $l$ columns and $e_1, \ldots, e_l$ denote the standard unit vector vectors, this is equivalent to infimizing $\gamma$ over all $(\gamma, X)$ satisfying the LMI constraint

$$
\begin{pmatrix}
-gamma & (XMe_1 - Ne_1)' & \cdots & (XMe_l - Ne_l)'

XMe_1 - Ne_1 & -gammaI & 0 & \cdots \\
\vdots & & \ddots & \vdots \\
XMe_l - Ne_l & 0 & \cdots & -gammaI
\end{pmatrix} < 0.
$$

(3)

If the data matrices $M(\delta)$, $N(\delta)$ depend rationally (without pole in zero) on the uncertain parameter $\delta \in \delta$, the robust counterpart requires to minimize the worst case residual error $\sup_{\delta \in \delta} \|XM(\delta) - N(\delta)\|_F^2$.

As shown above, the robust counterpart of (3) can be easily rewritten into (1) with a linear fractional representation

$$
\begin{pmatrix}
M(\delta)e_1 \\
N(\delta)e_1 \\
\vdots \\
M(\delta)e_l \\
N(\delta)e_l
\end{pmatrix} = \Delta(\delta) \star
\begin{pmatrix}
A & B_0 \\
C_1 & D_1^M \\
D_1 & D_1^N \\
\vdots
\end{pmatrix}
$$

where $\star$ denotes the upper LFT or the star-product [38]. Obviously, the same technique applies if replacing $\| \cdot \|_F$ with other norms whose sub-level sets admit LMI representations.

If $\delta = \text{co}\{\delta^1, \ldots, \delta^g\}$, one possible relaxation [18, 27] is for example given with $y = P$ as

$$
G(y) = -\text{diag}\left(\left(\begin{array}{c} I \\ 0 \end{array}\right), \left(\begin{array}{c} \delta^k \\ I \end{array}\right)\right), \quad k = 1, \ldots, g, \quad H(y) = P
$$

since one can easily verify (2). Note that this relaxation is more flexible and more accurate than the one in [11] (since it applies to general polytopes $\delta$), at the expense of higher computational complexity.

The suggested framework nicely captures robust performance analysis such as in standard structured singular value theory and considerable generalizations thereof [23]. Recall that the transfer matrix of some performance channel $w_p \rightarrow z_p$ of the interconnection of linear time-invariant uncertain system components can be represented as

$$
\begin{pmatrix}
\frac{1}{s}I & 0 \\
0 & \delta(s)
\end{pmatrix} \star
\begin{pmatrix}
A & B_1 \\
C_1 & D_{11} \\
B_2 & D_{12} \\
C_2 & D_{21} \\
D_{22}
\end{pmatrix}
$$

with $A$ stable.

Typically the proper and stable uncertainty transfer matrix $\delta(s)$ admits some specific (often block-diagonal) structure with its size bounded in terms of its frequency response. Both structure and size constraints on $\delta(s)$ can be specified with a suitable set $\delta$ of
complex matrices by requiring $\delta(s) \in \mathbf{\delta}$ for all $s \in i\mathbb{R} \cup \{\infty\}$. Guaranteeing stability and a bound $x$ on the squared $H_\infty$-norm of $w_p \to z_p$ then amounts to (1) with

$$A = \begin{pmatrix} A & B_1 \\ C_1 & D_{11} \end{pmatrix}, \quad B = \begin{pmatrix} B_2 \\ D_{12} \end{pmatrix}, \quad W(x) = \begin{pmatrix} 0 & 0 & I \\ C_2 & D_{21} & D_{22} \end{pmatrix} \begin{pmatrix} -xI & 0 \\ 0 & I \end{pmatrix} \begin{pmatrix} 0 & 0 & I \\ C_2 & D_{21} & D_{22} \end{pmatrix}$$

and $\Delta = \{\text{diag}(\delta_0 I, \delta) : \delta_0 \in i\mathbb{R}, \delta \in \mathbf{\delta}\}$. The best possible bound is obtained by minimizing $x$, and various other performance criteria are obtained by simple substitution of the middle factor in the definition of $W(x)$. Note that the uncertainty block $\frac{1}{s}I$ results, for the algebraic test, into the complex block $\delta_0 I$ with $\delta_0 \in i\mathbb{R}$. The subsequently suggested relaxation then leads to frequency independent multipliers. It is less conservative to actually grid the imaginary axis and perform a relaxation for finitely many fixed frequencies $\delta_0 \in \{i\omega_1, \ldots, i\omega_k\}$, as usually suggested in standard $\mu$-theory. In order to avoid missing crucial frequencies [36] while still being able to control conservatism, it is preferable to cover the imaginary axis with frequency segments and to apply the relaxation to each of the segments individually. This amounts to choosing $\Delta = \{\text{diag}(\delta_0 I, \delta) : \delta_0 \in i[\omega_1, \omega_2], \delta \in \mathbf{\delta}\}$ with fixed $0 \leq \omega_1 < \omega_2 < \infty$, which can be expressed as

$$\begin{pmatrix} \delta_0 \\ 1 \end{pmatrix}' \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{pmatrix} \delta_0 \\ 1 \end{pmatrix} = 0, \quad \begin{pmatrix} \delta_0 \\ 1 \end{pmatrix}' \begin{pmatrix} -1 & i(\omega_1 + \omega_2)/2 \\ -i(\omega_1 + \omega_2)/2 & -\omega_1\omega_2 \end{pmatrix} \begin{pmatrix} \delta_0 \\ 1 \end{pmatrix} \geq 0.$$ 

For the purpose of illustration let us also assume $\delta(s) = \text{diag}(\delta_1(s), \delta_2(s) I, \delta_3)$ with a dynamic full block $\delta_1(s)$ that is positive real, a repeated dynamic block $\delta_2(s)$ satisfying $\|\delta_2\|_\infty \leq 1$ and the phase constraint $\text{arg}(\delta_2(i\mathbb{R})) \subset [-\pi/2, 0]$, and two real repeated blocks $\delta_3 = \text{diag}(uI, vI)$ with $(u, v)$ in the triangle $\text{co}\{(-1, 0), (0, 1), (0, -1)\}$. Then $\Delta$ consists of all complex matrices diag($\delta_0, \delta_1, \delta_2 I, \delta_3$) with $\delta_0$ constrained as above, the full complex matrix $\delta_1$ satisfying

$$\begin{pmatrix} \delta_1 \\ I \end{pmatrix}'\begin{pmatrix} 0 & I \\ I & 0 \end{pmatrix} \begin{pmatrix} \delta_1 \\ I \end{pmatrix} \geq 0,$$

$\delta_2 \in \mathbb{C}$ obeying the gain and phase constraints

$$\begin{pmatrix} \delta_2 \\ 1 \end{pmatrix}'\begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \delta_2 \\ 1 \end{pmatrix} \geq 0, \quad \begin{pmatrix} \delta_2 \\ 1 \end{pmatrix}'\begin{pmatrix} 0 & 1+i \\ 1-i & 0 \end{pmatrix} \begin{pmatrix} \delta_2 \\ 1 \end{pmatrix} \geq 0, \quad \begin{pmatrix} \delta_2 \\ 1 \end{pmatrix}'\begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} \begin{pmatrix} \delta_2 \\ 1 \end{pmatrix} \geq 0,$$

and $\delta_3 \in \text{co}\{\delta_3^1, \delta_3^2, \delta_3^3\}$ with $\delta_3^1 = \text{diag}(-I, 0)$, $\delta_3^2 = \text{diag}(0, I)$, $\delta_3^3 = \text{diag}(0, -I)$. Although neither of these constraints can be easily handled with standard $\mu$-tools [4], it is possible to construct a relaxation as follows [26, 15, 16, 35, 30]. Choose the variables $y = (Y_0^1, Y_0^2, Y_0^3, y_1, Y_1^2, Y_1^3, Q_3, S_3, R_3)$ and the linear mappings

$$G(y) = -\text{diag}(Y_0^2, y_1, Y_1^2, Y_1^3, -Q_3,(\delta_3^1)'Q_3(\delta_3^1) + (\delta_3^2)'S_3 + S_3\delta_3^2 + R_3, j = 1, 2, 3),$$

$$H(y) = \begin{pmatrix} \text{diag}(-Y_0^2, 0, -Y_1^2, Q_3) \text{diag}(Y_0^1 + i(\omega_1 + \omega_2)/2Y_0^2, y_1I, (1+i)Y_0^2 - iY_1^2, S_3) \\ \text{diag}(-\omega_1\omega_2Y_0^2, 0, Y_1^2, R_3) \end{pmatrix}.$$ 

Then one easily verifies that (2) holds true. For problems without uncertainties the sketched procedure allows to systematically construct variations of the KYP Lemma as recently discussed in [25, 20, 21, 12, 1, 34, 17, 3], where the latter references comprises interesting time-domain interpretations.
Let us now briefly address LPV systems that are described as

\[
\begin{pmatrix}
\dot{x}(t) \\
z_p(t)
\end{pmatrix} = \begin{pmatrix}
A(\delta(t)) & B(\delta(t)) \\
C(\delta(t)) & D(\delta(t))
\end{pmatrix}
\begin{pmatrix}
x(t) \\
w_p(t)
\end{pmatrix}
\]

with \( \delta(t) \in \delta, \ \dot{\delta}(t) \in \dot{\delta} \)

for (path-connected) sets \( \delta \subset \mathbb{R}^m \) and \( \dot{\delta} \subset \mathbb{R}^m \) that bound the parameter and its rate-of-varation respectively. Suppose that \( A(\delta_0) \) is stable for some \( \delta_0 \in \delta \). We can guarantee uniform exponential stability and an \( L_2 \)-gain from \( w_p \rightarrow z_p \) being bounded by \( \sqrt{\gamma} \) if there exists a \( C^1 \) function \( X(\delta) \) such that

\[
\begin{pmatrix}
I & 0 \\
A(\delta) & B(\delta)
\end{pmatrix}' \begin{pmatrix}
\sum_{k=1}^m \frac{\partial X}{\partial \delta_k}(\delta)\dot{\delta}_k & X(\delta) \\
X(\delta) & 0
\end{pmatrix} \begin{pmatrix}
I & 0 \\
A(\delta) & B(\delta)
\end{pmatrix} +
\begin{pmatrix}
0 & I \\
C(\delta) & D(\delta)
\end{pmatrix}' \begin{pmatrix}
-\gamma I & 0 \\
0 & I
\end{pmatrix} \begin{pmatrix}
0 & I \\
C(\delta) & D(\delta)
\end{pmatrix} < 0 \text{ for all } (\delta, \dot{\delta}) \in \delta \times \dot{\delta}.
\]

(4)

If we reduce the search for \( X(\delta) \) to a finite dimensional subspace \( \text{span}\{X_1(\delta), \ldots, X_p(\delta)\} \), we can just substitute

\[
X(\delta) = \sum_{l=1}^p x_l X_l(\delta) \quad \text{and} \quad \frac{\partial X}{\partial \delta_k}(\delta) = \sum_{l=1}^p x_l \frac{\partial X_l}{\partial \delta_k}(\delta), \quad k = 1, \ldots, m
\]

(5)

and compute the corresponding best guaranteed \( L_2 \)-gain bound by minimizing \( \gamma \) over the robust LMI constraint (4)-(5) in the variables \( (\gamma, x_1, \ldots, x_p) \). If \( A(\delta), B(\delta), C(\delta), D(\delta), X_l(\delta), l = 1, \ldots, p \) admit linear fractional representation it is routine to subsume it to our more specific formulation. At this point is it essential to stress that one has to exploit the particular structure in order to keep the size of the representation manageable. As a major advantage, this scenario covers quadratic-in-the-state Lyapunov functions which are constant [5], quadratic [37, 18], or even polynomial or rational in the parameters [9, 8]. Most importantly it nicely captures as well the design of parameter-dependent controllers for systematic gain-scheduling with rate-bounded parameter-trajectories [22, 33, 13, 2, 27].

Although all the suggested relaxations are expected to involve conservatism, it is surprisingly often true that they are actually exact, with the following precise meaning. Let us assume \( W(x) = W_0 + x_1 W_1 + \cdots + x_n W_n \) with Hermitian \( W_0, \ldots, W_n \). If the relaxation is infeasible, standard Lagrange duality allows to construct an infeasibility certificate, a pair \( (M, N) \neq 0 \) with \( (W_0, M) \geq 0 \) and

\[
M \geq 0, \ N \geq 0, \begin{pmatrix}
\langle W_1, M \rangle \\
\vdots \\
\langle W_n, M \rangle
\end{pmatrix} + c = 0, \quad H^* \begin{pmatrix}
I & 0 \\
A & B
\end{pmatrix} M \begin{pmatrix}
I & 0 \\
A & B
\end{pmatrix}' + G^*(N) = 0
\]

(6)

where \( G^*, H^* \) denote the adjoint mappings of \( G, H \) respectively. The following results shows under which conditions and how we can verify whether the original robust LMI problem is infeasible as well.

**Theorem 1** The robust LMI (1) is not feasible iff there exists some \( \Delta_0 \in \Delta \) such that

\[
\begin{pmatrix}
\Delta_0(I - A\Delta_0)^{-1}B \\
I
\end{pmatrix}' W(x) \begin{pmatrix}
\Delta_0(I - A\Delta_0)^{-1}B \\
I
\end{pmatrix} < 0 \text{ is not feasible.}
\]

(7)
\( \Delta_0 \in \Delta \) satisfies (7) iff there exist an infeasibility certificate \((M, N)\) such that

\[
\left( I - \Delta_0 \right) \begin{pmatrix} I & 0 \\ A & B \end{pmatrix} M = 0.
\]

If the relaxation is feasible, there exists dual optimal solutions, matrix pairs \((M, N)\) with \(\langle W_0, M \rangle = \gamma_{\text{rel}}\) and (6). If for any of these dual optimal solutions we can solve the equation (8) it is actually guaranteed that the relaxation is exact, and that any solution \(\Delta_0\) of (8) is a worst-case uncertainty. Moreover the converse holds true as well [29].

**Theorem 2** If the relaxation has a dual optimal solution \((M, N)\) such that there exists some \(\Delta_0 \in \Delta\) satisfying (8) then \(\gamma_{\text{opt}} = \gamma_{\text{rel}}\) and \(\Delta_0\) is a worst-case uncertainty in the sense that

\[
\inf \left\{ \langle c, x \rangle : \begin{pmatrix} \Delta_0(I - A\Delta_0)^{-1}B \\ I \end{pmatrix} W(x) \begin{pmatrix} \Delta_0(I - A\Delta_0)^{-1}B \\ I \end{pmatrix} \prec 0 \right\} = \gamma_{\text{opt}}.
\]  

(9)

Conversely if \(\gamma_{\text{opt}} = \gamma_{\text{rel}}\) and if there exists some \(\Delta_0 \in \Delta\) with (9) then there exist a dual optimal solution \((M, N)\) satisfying (8).

As in the above given examples the set \(\Delta\) often admits an LMI description. Exactness of the relaxation is now tested by just determining either an infeasibility certificate or a dual optimal solution \((M, N)\) and by computing the infimal \(\nu\) for which there exists some \(\Delta \in \Delta\) with

\[
\left\| \begin{pmatrix} I - \Delta \end{pmatrix} \begin{pmatrix} I & 0 \\ A & B \end{pmatrix} M \right\| \leq \nu.
\]

(10)

The corresponding optimal value \(\nu_{\text{rel}}\) can be clearly obtained by solving a (second) SDP. If \(\nu_{\text{rel}}\) vanishes, it is guaranteed that the relaxation is exact, and any minimizer defines a worst-case uncertainty as in the above two theorems. If \(\Delta\) is compact and the feasible set of the robust LMI problem is bounded, one can actually prove that a small but positive value of \(\nu_{\text{rel}}\) implies approximate exactness [31].

Let us finally remark that it can be shown for many of the relaxations (including the ones suggested above) that the existence of an infeasibility certificate or a dual optimal solution with \(M\) of rank one implies that \(\nu_{\text{rel}}\) vanishes [31]. On the one hand this relates the given exactness test to a variety of other rank one principles in the literature [23, 20, 14, 24, 19]. On the other hand, both theorems reveal that the rank one principle can only hold if a worst-case perturbation exists for the robust LMI problem under consideration. Moreover simple examples allow to illustrate that the presented exactness test can be often applied even if \(M\) does have a larger rank.

If the uncertainty set is polyhedral and explicitly described as the convex hull of finitely many generators, it is possible to construct a family of relaxations that can be shown (using a matrix version of a theorem of Pólya) to be asymptotically exact [28]. In a technical paper of these proceedings [32] we reveal how to construct such families for implicitly described polytopes if using matrix sum-of-squares decompositions.
References


