A Hybrid Gradient-LMI Algorithm for Solving BMIs in Control Design Problems

Hossam Abbas, Saulat S. Chughtai, Herbert Werner

Institute of Control Systems
Hamburg University of Technology, Germany

Abstract: This paper presents an algorithm for solving optimization problems with bilinear matrix inequality constraints. The algorithm is based on a combination of gradient-based optimization and LMIs, which makes it fast and enables it to handle a large number of decision variables. It is applied to two controller synthesis problems: static output feedback controller synthesis and robust controller synthesis for linear parameter varying (LPV) systems using the idea of quadratic separation. Since the second problem has a large number of decision variables, a hybrid approach is applied, in which LMI solvers are used for the evaluation of the cost function. The algorithm is applied to two examples, and results are compared with some existing approaches.

Keywords: Gradient-based optimization, bilinear matrix inequalities, linear matrix inequalities, controller synthesis.

1. INTRODUCTION

Many controller synthesis problems can be formulated as optimization problems subject to bilinear matrix inequality (BMI) constraints. In general, these problems can be expressed in the form

\[ \min_{x} c^T x, \text{ where } x \in \mathbb{R}^n \]

s.t \[ M_0 + \sum_{i=1}^{n} x_i M_i + \sum_{1 \leq i < j \leq n} x_i x_j Q_{ij} \leq 0 \]

where, \( M_0, M_i \) and \( Q_{ij} \) are symmetric, known matrices and \( M \leq 0 \) means that \( M \) is semi-definite.

The importance of being able to solve such problems for various control applications has been recognized for many years and a number of approaches have been proposed for this purpose. Early attempts to solve these problems were based on converting them into a sequence of LMI problems. One such approach is presented in (Hassibi et al., 1997), where BMI constraints are linearized using first-order perturbations. These techniques may however be difficult to initialize if the set of feasible solutions is small. A Lagrangian-based approach is presented in (Kocvara and Stingl, 2005). However, this approach is best suited for problems where BMI constraints can be transformed into problems with nonlinear equality constraints and convex inequality constraints. An alternative is the use of stochastic optimization techniques like that proposed in (Farag and Werner, 2004). This approach is however not suitable when the number of decision variables is large.

In this paper a two-loop approach is proposed to solve the BMI problem (1). In the outer loop a linear cost function \( c^T x \) is minimized while in the inner loop a BMI constraint is applied by minimizing the spectral abscissa \( \alpha \) of the BMI matrix. The algorithm can be summarized as follows.

Step 1 Initialize the objective function with a suitable large value and fix all values of \( x_i \) that correspond to nonzero coefficients of \( c \). (In many control problems a performance index \( \gamma \) is minimized that is also a decision variable, so \( c \) consists only of zeros and a single 1, see Section 3 and 4).

Step 2 Find values for the remaining entries of \( x_i \) such that the BMI is satisfied. This step is performed using gradient-based optimization. Repeat this step \( m \) times with random initialization and select the best solution \( x \).

Step 3 If Step 2 returns a feasible solution reduce the value of the objective function and repeat, otherwise increase it.

Step 4 Iterate till no further reduction in the objective function is achieved.

In the inner loop, a semi-stochastic gradient-based optimization approach can be used if an explicit expression for \( \nabla x \alpha \) - the gradient of \( \alpha \) with respect to the decision variables \( x \) - can be found. In this work an approach presented in (Horn and Johnson, 1985) is used for this purpose. The term semi-stochastic refers to the fact that in each iteration the gradient-based search is initiated by a number of random values, and the best cost function achieved is used for the next iteration.

In this paper we illustrate the proposed algorithm by solving two BMI constrained controller synthesis problems: synthesis of a \( \mathcal{H}_\infty \) optimal static output feedback (SOF) controller, and a fixed-structure, robust controller design for a linear parameter varying (LPV) system using a technique based on quadratic separation (Chughtai and Werner, 2006). Since the latter problem has a large number of decision variables, an extension to the algorithm is also proposed to improve its speed and convergence.
The paper is organized as follows: Section 2 presents a way of constructing the gradient of the spectral abscissa of the matrix appearing in the BMI constraint. The proposed algorithm is applied to the SOF problem in section 3. Section 4 discusses an extension of the algorithm to solve robust synthesis problem for LPV systems. Section 5 presents two examples, and the results are compared with previously proposed approaches. Some concluding remarks are given in section 6.

2. PRELIMINARIES

The following result can be used to calculate the gradient of the spectral abscissa of a matrix. Consider a matrix $J(t) \in \mathbb{C}^{m \times n}$ that depends on a scalar parameter $t$. Then we have (Horn and Johnson, 1985)

**Theorem 1.** Let $J(t)$ be differentiable at $t = 0$. Assume that $\lambda$ is an algebraically simple eigenvalue of $J(0)$ and that $\lambda(t)$ is an eigenvalue of $J(t)$, for small $t$ such that $\lambda(0) = \lambda$. Let $v$ be a right eigenvector of $J(t)$ and $u$ a left eigenvector of $J(0)$ corresponding to eigenvalue $\lambda$, both normalized to 1. Then

$$\lambda'(0) = u^T J'(0)v$$

where $\lambda'$ and $J'$ denote derivatives with respect to $t$. Moreover, if $J(t) = J + tE$ for a fixed matrix perturbation, $E$, then for $t$ small

$$\lambda(J + tE) = \lambda(J) + tu^T Ev + O(t^2) \quad (2)$$

The factor $u^T Ev$ in the linear term - taken for each decision variable - can be used to compute the gradient of the spectral abscissa.

3. $\mathcal{H}_\infty$ OPTIMAL STATIC OUTPUT FEEDBACK CONTROL

This section presents the application of the algorithm outlined in Section 1 to the problem of designing $\mathcal{H}_\infty$ optimal static output feedback controllers. Consider the following plant with state space model

$$\dot{x} = Ax(t) + B_1 w(t) + B_2 u(t)$$
$$z(t) = C_1 x(t) + D_{11} w(t) + D_{12} u(t)$$
$$y(t) = C_2 x(t) + D_{21} w(t) \quad (3)$$

where $x(t) \in \mathbb{R}^n$ is the state vector, $w(t) \in \mathbb{R}^m$ denotes external inputs, $u(t) \in \mathbb{R}^{m_2}$ is the control input, $z \in \mathbb{R}^{n_1}$ is the controlled output and $y(t) \in \mathbb{R}^{p_2}$ is the measured output. $A, B_1, B_2, C_1, C_2, D_{11}, D_{12}$ and $D_{21}$ are constant matrices with appropriate dimensions, where $(A, B_2)$ is stabilizable and $(C_2, A)$ is detectable.

With a static output feedback controller

$$u(t) = Ky(t) \quad (4)$$

where $K \in \mathbb{R}^{m_2 \times p_2}$, the closed loop system is

$$A_{cl} = A + B_2 K C_2$$
$$B_{cl} = B_1 + B_2 K D_{12}$$
$$C_{cl} = C_1 + D_{12} K C_2$$
$$D_{cl} = D_{11} + D_{12} K D_{21} \quad (5)$$

The objective is to find a controller (4) such that the transfer function of the closed-loop system satisfies a $\mathcal{H}_\infty$ norm constraint

$$\| T_{uw}(s) \|_\infty < \gamma, \quad \text{for} \quad \gamma > 0 \quad (6)$$

The constraint (6) can be represented as (Cao et al., 1998),

$$\exists \gamma > 0, K :$$
$$\begin{bmatrix}
PA_{cl} + A_{cl}^T P + PB_{cl} C_{cl}^T
B_{cl}^T P
C_{cl}
D_{cl}
\end{bmatrix}
\begin{bmatrix}
-\gamma I
D_{cl}^T
-\gamma I
\end{bmatrix}
< 0, \quad (7)$$

where $A_{cl}$, $B_{cl}$, $C_{cl}$ and $D_{cl}$ are given in terms of $K$ by (16) and $P$ is a Lyapunov matrix.

The SOF $\mathcal{H}_\infty$ synthesis problem can now be expressed as follows:

$$\min_{\gamma, K} \gamma \quad \text{subject to} \quad (7) \quad \text{and} \quad (8) \quad (9)$$

Note that (7) is a BMI condition, leading to a non-convex problem. One way to solve (9) is the following.

1. For a given $\gamma$, solve

$$\min_{\gamma, K} \gamma \quad \text{subject to} \quad (7) \quad \text{and} \quad (8) \quad (10)$$

where $\gamma$ is the spectral abscissa (the largest real part of the eigenvalues of $X$), where

$$X = \begin{bmatrix}
PA_{cl} + A_{cl}^T P + PB_{cl} C_{cl}^T
B_{cl}^T P
C_{cl}
D_{cl}
\end{bmatrix}
\begin{bmatrix}
-\gamma I
D_{cl}^T
-\gamma I
\end{bmatrix}
= 0 \quad (11)$$

2. If an $\alpha < 0$ is found, reduce $\gamma$ and goto Step 1.

The minimization problem (10) can be solved using a gradient-based optimization technique. Here we use the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm which belongs to a class of Quasi-Newton methods, where a positive definite approximation of the Hessian is produced from previous iterations and the corresponding gradients.

3.1 Gradient Calculation

The main step in the above algorithm is to find the gradient of $\gamma(X)$ with respect to $P$ and $K$. This will be done by first analyzing the effect of a perturbation in $P$ on the maximum eigenvalue of $X(P)$, while keeping $K$ constant. The perturbed matrix $X$ can be factorized and expressed as follows

$$X + \delta X_P = X + Y^T \delta P I_1 + I_1^T \delta P Y + I_2^T \delta P I_2 \quad (12)$$

where $Y = [A \ B \ 0 \ 0]$, $I_1 = [I \ 0 \ 0 \ 0]$, $I_2 = [0 \ 0 \ 0 \ I]$, $I$ is the identity matrix and $\delta P$ denotes the perturbation of $P$. Then using Theorem (1) we can write

$$f(K, P + \delta P) = f(K, P) + u^T (Y^T \delta P I_1 + I_1^T \delta P Y + I_2^T \delta P I_2) \quad (13)$$

$$+ I_2^T \delta P I_2 v + O(\delta P^2)$$

Note that the inner product on the space of matrices is defined as

$$(A, B) = tr(A^T B)$$

where $tr(M)$ denotes the trace of matrix $M$. Then (13) can be expressed as,
\[ f(K,P + \delta P) = f(K,P) + tr(u^T \{ Y^T \delta P I_1 + I_1^T \delta P Y + I_2^T \delta P I_2 \} v) + O(\delta_P^2) \]
\[ = f(K,P) + \langle I_1 v u^T Y^T + Y v u^T I_1^T + I_2 v u^T I_2^T, \delta P \rangle + O(\delta_P^2) \]
\[ = f(K,P) + \langle I_1 v u^T Y^T + Y v u^T I_1^T + I_2 v u^T I_2^T, \delta P \rangle + O(\delta_P^2) \]

We therefore obtain the gradient with respect to the matrix \( P \) from (14) as
\[ \nabla_P f = \langle I_1 v u^T Y^T + Y v u^T I_1^T + I_2 v u^T I_2^T, \delta P \rangle \]

Similarly we can compute the gradient with respect to \( K \).

3.2 Algorithm

Finally an iterative algorithm to solve the SOF problem (10) can be summarized as follows.

Step 1 Choose an initial value for \( \gamma \) and an initial step size \( t < 1 \).

Step 2 Generate \( m \) initial random vectors of decision variables and solve the problem (10) \( m \) times, using (14).

Select the solution with the smallest value of \( \alpha(X) \).

Step 3 If \( \alpha(X) < 0 \), i.e. a feasible solution is found, increase the step size according to \( t = 2t \), \( t < 1 \), decrease \( \gamma \) by \( \gamma = \gamma - \gamma t \gamma \) and goto step 5).

Step 4 If \( \alpha(X) > 0 \), i.e. no feasible solution, replace the step size \( t \) by \( t/2 \) and \( \gamma \) by \( \gamma - \gamma \gamma \gamma \), goto step 2).

Step 5 If the difference between the present and the past values of \( \gamma \) less than a specified value, stop, if not goto step 2).

4. ROBUST CONTROLLER SYNTHESIS FOR LPV SYSTEMS

This section describes robust controller synthesis for LPV systems using a technique based on quadratic separators. The concept of quadratic separators can be used to reduce the conservatism of robust controller design for LPV systems when upper bounds on the rate of change of model parameters are known, see (Chughtai and Werner, 2006). This technique utilizes the fact that the existence of a quadratic separator is equivalent to the existence of a parameter-dependent Lyapunov function (Iwasaki and Shibata, 2001). The method proposed in (Chughtai and Werner, 2006) uses a hybrid evolutionary-algebraic approach for solving the non-convex problem of finding a low-order and fixed structure controller that minimizes the induced \( L_2 \) norm of the performance channel from \( w_\theta \) to \( z^p \) as shown in Fig. 1. Since the problem considered here has typically a large number of decision variables, an extension to the approach presented in previous section is proposed.

Let the closed-loop system be represented in the following LFT structure
\[ \dot{x} = Ax + B_\Delta \Delta x + B_p w_p \]
\[ z_\Delta = C_\Delta x + D_\Delta \Delta x \]
\[ z_p = C_p x + D_p w_p \]
\[ w_\Delta = \Delta z_\Delta \]

where \( w_\Delta, z_\Delta \in \mathbb{R}^l, w_p \in \mathbb{R}^d, z_p \in \mathbb{R}^v \), and \( \Delta(t) \in \mathbb{R}^{l \times 1} \), where all the system matrices depend on controller \( (K) \) parameters. Then the worst case \( L_2 \)-norm is bounded by \( \gamma \) if the condition in the following theorem holds, (Chughtai and Werner, 2006).

Theorem 2. The LPV system described by (16) is stable and has a worst-case induced \( L_2 \)-gain less than \( \gamma \) if there exist real symmetric matrices \( P > 0, S \) and \( R \) such that the following conditions hold
\[ \begin{bmatrix} A & B \\ C & D \end{bmatrix}^T \begin{bmatrix} 0 & 0 & P \\ 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} A & B \\ C & D \end{bmatrix} < 0 \]
\[ \begin{bmatrix} I \\ \Theta \\ \nabla \end{bmatrix}^T \begin{bmatrix} I \\ \Theta \\ \nabla \end{bmatrix} > 0 \]

where
\[ \Theta := \left\{ \begin{bmatrix} Y & T \end{bmatrix} : R \in S_D, S \in S_G \right\} \]

and,
\[ S_D := \{ D : D \nabla = \nabla D, D = D^T > 0 \} \]
\[ S_G := \{ G : G \nabla = \nabla G, G + G^T = 0 \} \]

where
\[ \Delta(t) = diag(q_1(t)I_{r_1}, \ldots, q_m(t)I_{r_1}) \]
\[ \hat{\Delta}(t) = diag(\hat{\delta}_1(t)I_{r_1}, \ldots, \hat{\delta}_m(t)I_{r_1}) \]
\[ \nabla := diag(\Delta, \Delta, \hat{\Delta}) \]
where the index \( r_i \) represents the multiplicity of the \( i^{th} \) uncertain parameter and all varying parameters \( q_i(t) \) and their rate of change \( \dot{q}_i(t) \) are bounded such that \( |q_i(t)| < \phi_i \) and \( |\dot{q}_i(t)| < \rho_i, \forall i = 1, \ldots, l \). Then,

\[
\begin{align*}
\Phi &:= \text{diag}(\phi_1 I_{r_1}, \ldots, \phi_l I_{r_l}) \\
\Psi &:= \text{diag}(\phi_1 I_{r_1}, \ldots, \phi_l I_{r_l}) \\
\Psi_p &:= \text{diag}(\rho_1 I_{r_1}, \ldots, \rho_l I_{r_l})
\end{align*}
\]

(22)

The theorem stated above can be used to analyze the worst-case \( L_2 \)-norm of a closed loop system. In the synthesis problem, the LMI (17) turns into a bi-linear matrix inequality in \( P, \Theta, \Gamma, \gamma \) and \( K \). The synthesis problem can be expressed as

\[
\min_{P, \Theta, K} \gamma \text{ such that (17) holds.} \tag{23}
\]

### 4.1 Algorithm

Let us define \( X \) as in (25), then an iterative algorithm to solve problem (23) can be formulated as:

**Step 1** Choose an initial value for \( \gamma \) and an initial step size \( t < 1 \).

**Step 2** Generate \( m \) initial random vectors of the decision variables representing the parameters of the controller \( K \).

**Step 3** Given \( \gamma \) and \( K \), solve \( m \) generalized eigenvalue problems by solving the LMI problems

\[
\min_{P, \Theta} \alpha(X)
\]

**Step 4** Given \( \gamma, P, \Theta \), use the BFGS algorithm to solve

\[
\min_K \alpha(X)
\]

\( m \) times

**Step 5** Select the solution with the smallest value \( \alpha(X) \).

**Step 6** If \( \alpha(X) < 0 \), i.e. feasible, replace the step size \( t \) by \( 2t \), \( t < 1 \), and \( \gamma \) by \( \gamma - \gamma \), goto step 8.

**Step 7** If \( \alpha(X) > 0 \), i.e. infeasible, replace the step size \( t \) by \( t/2 \) and \( \gamma \) by \( \gamma - \gamma \), goto step 2.

**Step 8** If the difference between the present and the past values of \( \gamma \) is less than a specified value stop, if not goto step 2.

The calculation of the gradient of \( \alpha(X) \) with respect to \( K \) can be computed using the presented idea in section 2.

**Remark:** It should be noted that the proposed algorithm is applicable for the systems which are robustly stabilizable. Under this condition the algorithm can find a controller for arbitrary value of \( \gamma \). As the algorithm proceeds the \( \gamma \) is decreased by the outer loop as long as a feasible solution is obtained in the inner loop. Hence, the \( \gamma \) will decrease monotonically and the algorithm will converge to a local minimum. Since, a global solution is not guaranteed for BMI constrained optimization problems.

### 5. EXAMPLES AND RESULTS

In this section the above algorithms are illustrated by examples, and results are compared with previously proposed techniques.

#### 5.1 Example 1: Multivariable PID \( H_\infty \) Controller Design

This example illustrates that when applied to relatively simple problems with a small number of decision variables, it yields the same performance as previously proposed techniques. The conversion of a multivariable PID \( H_\infty \) controller design problem into a SOF control problem was presented in He and Wang (2006). This method is applied here to a plant that was also used in He and Wang (2006) for illustration; the model is given in (26).

The method of Section 2 is applied, and the values of \( \gamma \) are listed together with those obtained with different methods in Table 5.1. The other methods listed are the method in (Zheng et al., 2002), the ILMI method in He and Wang (2006), PENBMI (Kocvara and Stingl, 2005) and the HEA approach (Farag and Werner, 2004). One can see that the proposed method yields the same result as the other techniques.

#### 5.2 Example 2: Robust Low-order controller using Theorem 2

This example shows that for a larger number of decision variables, the proposed method outperforms some previously published techniques. The method in Section 3 is now applied to a design example taken from (Chughtai and Werner, 2006) - control of a vertical takeoff and landing (VTOL) helicopter. The linearized longitudinal dynamics of a helicopter are given by

\[
A = \begin{bmatrix}
-0.0366 & 0.0271 & 0.0188 & -0.4555 \\
0.0482 & -1.01 & 0.0024 & -4.0208 \\
0.1002 & p_1 & -0.707 & p_2 \\
0 & 0 & 1 & 0
\end{bmatrix},
\]

\[
B = \begin{bmatrix}
p_3 & -5.5922 \\
-5.52 & 4.49 \\
0 & 0 \\
0 & 0
\end{bmatrix},
\]

\[
C^T = \begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}
\]

where parameters \( P_1 \), \( P_2 \) and \( P_3 \) are defined as:

\[
p_1 = 0.3681 + 0.05\delta_1 \\
p_2 = 1.42 + 0.01\delta_2 \\
p_3 = 3.5446 + 0.04\delta_3
\]

The parameter \( \delta_i \) depends on the flight conditions. However, for the sake of simplicity, see (Chughtai and Werner,
Table 2. Worst case $\mathcal{L}_2$-norm achieved using different approaches

<table>
<thead>
<tr>
<th>Order</th>
<th>Decision variables</th>
<th>PENBMI</th>
<th>HEA</th>
<th>Proposed approach</th>
</tr>
</thead>
<tbody>
<tr>
<td>1$^{st}$</td>
<td>6</td>
<td>Failed</td>
<td>5.9 ± 0.8</td>
<td>3.9</td>
</tr>
<tr>
<td>4$^{th}$</td>
<td>30</td>
<td>Failed</td>
<td>4.5</td>
<td></td>
</tr>
</tbody>
</table>

The proposed algorithm is applied on two examples which shows that it gives solution where standard BMI solvers like PENBMI may fail. The algorithm can further be improved by using the adaptive step size in outer loop however, this will be the topic of our future research.

REFERENCES


