$\mathcal{H}_2$ Model Reduction Using LMIs

Masaaki Kanno

CUED/F-INFENG/TR.519
March 2005

$\mathcal{H}_2$ Model Reduction Using LMIs

Masaaki Kanno
University of Cambridge
Department of Engineering
Trumpington Street
Cambridge CB2 1PZ
United Kingdom
E-mail: mk303@eng.cam.ac.uk

March 16, 2005

Abstract

The problem of $\mathcal{H}_2$ reduced order approximation is considered in this report. The result covers both continuous and discrete time MIMO systems. Necessary and sufficient conditions for the existence of an approximant within a specified error are given in terms of a set of LMIs and a matrix rank constraint. A heuristic algorithm which uses the alternating projection method is proposed and a method of finding a starting point is suggested. Three numerical examples are employed to show the effectiveness of the choice of starting points and the capability of the algorithm to find at least as good approximants as other methods.

1 Introduction

There has been a significant interest in the model reduction problem, namely, the problem of approximating a high order system by a lower order, thus simpler, system. One of the most commonly used methods is the balanced truncation method [21]. The procedure is relatively simple and also the method is extensively studied [11]. Another popular model reduction method is the Hankel norm approximation method [8], which also has a constructive way of finding reduced order models, or approximants. It is common practice to employ an error between the original high order system and the obtained reduced order model in some sense as an index of how good the approximant is. For both methods upper bounds of the error in the $\mathcal{H}_\infty$ sense (and also a lower bound for the Hankel norm approximation method) are explicitly expressed in terms of the Hankel singular values of the original system. These methods do not in general produce optimal approximants in the $\mathcal{H}_\infty$ sense and several methods for $\mathcal{H}_\infty$ optimal model reduction are developed, e.g., [5, 9].

In this report the model reduction problem in terms of the $\mathcal{H}_2$-norm is considered. This norm has an attractive aspect in that minimizing the $\mathcal{H}_2$ error means minimizing the $\mathcal{H}_2$ error in the impulse response as well as minimizing the $\mathcal{H}_2$ error in the frequency response [25] and the $\mathcal{H}_2$ model reduction problem has also attracted considerable attention over four decades.

A number of approaches, e.g., [2, 7, 16, 25], just to name a few, use first order necessary conditions for optimality in one way or another and develop optimization algorithms to find solutions to resulting nonlinear equations. Most of the methods in this direction are only applicable to the single input single output (SISO) case. Furthermore it is argued [14, 27] that whether the global optimum is always achievable is unclear in the continuous time case (while it is shown to exist in the discrete time case [1]) and that, in the case of nonexistence of the optimum, these approaches can only find local optima which may be far from the true (global) optimum.

Even if the existence of the global optimum is guaranteed, optimization methods based on search algorithms can have difficulties [12]: There may be one or more local optima and it is difficult to guarantee that the obtained solution is close to the global optimum. Moreover there is usually no guarantee that the chosen stopping criterion for such a search algorithm is appropriate. To overcome these problems, several algorithms based on algebraic methods have been proposed that directly solve a set of nonlinear equations [12, 19, 22]. These approaches seem to have potential (in cases where the optima are achievable), but computation cost required for such approaches is still high and structural properties of the problem seem to require further exploitation for algorithmic development, which prevent them from becoming useful alternatives in practice at this moment.

A different type of approaches has emerged recently. In [14, 27], it is proposed to solve slightly modified problems for the continuous time case, where the global optimum is proven to exist and the use of a search algorithm makes sense. Those methods can deal with the multi input multi output (MIMO) case and thus favourable compared to many other methods in this respect. A problem of those methods may be the difficulty of measuring the conservativeness of the obtained result due to the modification of the problem.

In this report an $\mathcal{H}_2$ model reduction method for the MIMO case based on linear matrix inequality (LMI) techniques is developed. Unlike other methods this approach allows both continuous time and discrete time cases to be treated in a unified manner, as in [9] for $\mathcal{H}_\infty$ model reduction. Necessary and sufficient conditions for the existence of suboptimal approximants are expressed in bilinear matrix inequality (BMI) form, which will then be converted to a set of LMIs and a (non-convex) matrix rank constraint. An algorithm using the alternating projection method is proposed to solve this problem. Due to the non-convex property of the problem, the suggested method does not guarantee global convergence. However numerical examples show that, from starting points computed by a method which is also proposed in this report, this method can yield approximants at least as good as those computed by other methods. It is emphasized that this method deals with the original problem rather than a modified one and thus is not affected by the potential conservativeness resulting from modification of the problem. Also the algorithm essentially solves suboptimal problems and hence avoids the issue of existence/nonexistence of the optimal solution. Moreover a search is carried out for the feasible $\mathcal{H}_2$ error by executing feasible tests and therefore can be terminated when a desired difference between the achieved error and the (local) optimum is reached.

The structure of this report is as follows. In Section 2, the $\mathcal{H}_2$ model reduction problem is formally stated and necessary and sufficient conditions for the existence of a reduced order model within a specified error are derived for both continuous and discrete time cases. Section 3 reviews the alternating projection method which is used for various controller synthesis problems with fixed controller order, and also proposes an algorithm for the $\mathcal{H}_2$ model reduction problem which makes use of this method. Since the above algorithm is heuristic due to the non-convexity of the necessary and sufficient conditions derived in Section 2, the choice of starting points has a significant effect on the practicality of the proposed algorithm. This is the topic of Section 4. In Section 5, three numerical examples are employed to demonstrate the choice of starting points and the algorithm. In particular the third example shows that the obtained results are at least as good as those computed by other methods. Some concluding remarks are made in Section 6.

Notation: Given a matrix $A \in \mathbb{C}^{n\times m}$, A^* denotes the complex conjugate transpose. The trace of a square matrix is denoted by $\mathrm{tr}\,\{\cdot\}$. The set of all symmetric matrices in $\mathbb{C}^{n\times n}$ is denoted by $\mathcal{S}_n$. The notation $>$ (resp., $\geq$) is used to denote the positive definiteness (resp., semi-definiteness) of a square symmetric matrix, and $A > B$ (resp., $A \geq B$), $A, B \in \mathcal{S}_n$ means $A - B > 0$ (resp., $A - B \geq 0$). Also, $A < 0$ (resp., $A \leq 0$) means $-A > 0$ (resp., $-A \geq 0$). For brevity, $()^*$ is used to denote the complex conjugate transpose of the preceding matrix, e.g., for a square matrix A, $A + ()^* = A + A^*$. For a matrix $A \in \mathbb{R}^{n\times m}$ with rank r, $A^\perp \in \mathbb{R}^{(n-r)\times n}$ is a matrix such that $A^\perp A = 0$ and $A^\perp (A^\perp)^* > 0$. The $\mathcal{H}_2$-norm of a stable system is denoted by $\|\cdot\|_2$.

2 $\mathcal{H}_2$ Model Reduction

The $\mathcal{H}_2$ optimal model reduction problem is stated as follows: Given a stable system G of McMillan degree n with q inputs and p outputs, find a stable system $\hat{G}$ of McMillan degree $r\,(< n)$ with the same numbers of inputs and outputs that minimizes the $\mathcal{H}_2$-norm of the error system $E = G - \hat{G}$, i.e., minimizes the error $\|E\|_2 = \left\|G - \hat{G}\right\|_2$. Under the same set-up, the $\mathcal{H}_2$ suboptimal model reduction problem is stated as: Given $\gamma\,(> 0)$, find, if it exists, $\hat{G}$ that achieves the $\mathcal{H}_2$ error less than γ, i.e., achieves $\left\|G - \hat{G}\right\|_2 < \gamma$. Without loss of generality, it can be assumed that both G and $\hat{G}$ are strictly proper.

2.1 Continuous Systems

In this subsection the continuous time case is considered and necessary and sufficient conditions for the existence of a reduced order model achieving a specified error are derived. Let state space realizations of $G(s)$ and $\hat{G}(s)$ be

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right], \tag{1}$$

$$\hat{G}(s) = \left[\begin{array}{c|c} \hat{A} & \hat{B} \\ \hline \hat{C} & 0 \end{array}\right] \tag{2}$$

where $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times q}$, $C \in \mathbb{R}^{p\times n}$, $\hat{A} \in \mathbb{R}^{r\times r}$, $\hat{B} \in \mathbb{R}^{r\times q}$, $\hat{C} \in \mathbb{R}^{p\times r}$. A state space realization of the error system is

$$E(s) = G(s) - \hat{G}(s) = \left[\begin{array}{cc|c} A & 0 & B \\ 0 & \hat{A} & \hat{B} \\ \hline C & -\hat{C} & 0 \end{array}\right] =: \left[\begin{array}{c|c} A_E & B_E \\ \hline C_E & 0 \end{array}\right]. \tag{3}$$

Then the $\mathcal{H}_2$ optimal model reduction problem can be expressed as:

$$\begin{aligned} &\text{minimize } \gamma\,(> 0) \\ &\text{subject to } A_E P + P A_E^* + B_E B_E^* < 0\,, && (4)\\ &\qquad P > 0\,, && (5)\\ &\qquad \mathrm{tr}\left\{C_E P C_E^*\right\} < \gamma^2 && (6) \end{aligned}$$

where the positive definiteness of $P \in \mathcal{S}_{n+r}$ is to guarantee (in fact, is equivalent to) the stability of $\hat{G}(s)$ under the assumption that (A_E, B_E) is controllable. Partition P conformally with A_E and write

$$P = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix}$$

where $P_{11} \in \mathcal{S}_n$, $P_{12} \in \mathbb{R}^{n\times r}$, $P_{22} \in \mathcal{S}_r$. Then, from the Schur complement formula [3, pp. 7-8], inequality (4) is equivalent to

$$\begin{bmatrix} A_E P + P A_E^* & B_E \\ B_E^* & -I \end{bmatrix} = \begin{bmatrix} AP_{11} + P_{11}A^* & AP_{12} + P_{12}\hat{A}^* & B \\ P_{12}^* A^* + \hat{A}P_{12}^* & \hat{A}P_{22} + P_{22}\hat{A}^* & \hat{B} \\ B^* & \hat{B}^* & -I \end{bmatrix} < 0 . \tag{7}$$

Using a slack variable $W \in \mathcal{S}_n$, inequalities (5) and (6) can be expressed as

$$\mathrm{tr}\,\{W\} < \gamma^2 , \tag{8}$$

$$\begin{bmatrix} W & C_E P \\ P C_E^* & P \end{bmatrix} = \begin{bmatrix} W & CP_{11} - \hat{C}P_{12}^* & CP_{12} - \hat{C}P_{22} \\ P_{11}C^* - P_{12}\hat{C}^* & P_{11} & P_{12} \\ P_{12}^* C^* - P_{22}\hat{C}^* & P_{12}^* & P_{22} \end{bmatrix} > 0 . \tag{9}$$

It is observed that neither (7) nor (9) is an LMI in P_{11}, P_{12}, P_{22}, $\hat{A}$, $\hat{B}$, $\hat{C}$ since there are bilinear terms such as $\hat{A}P_{12}$.

Now those conditions are expressed with respect to two decision variables (symmetric matrices) by eliminating $\hat{A}$, $\hat{B}$, $\hat{C}$. The $\mathcal{H}_2$ model reduction problem is a special case of the $\mathcal{H}_2$ optimal controller synthesis problem and therefore the following result may readily be obtained from [24]. The result is included for completeness and also due to the straightforwardness of the proof. Further the result is expressed in a form suited for the method developed later.

Theorem 1 *Consider a stable continuous time system* $G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *of McMillan degree* n. *There exists a stable continuous time system* $\hat{G}(s)$ *of McMillan degree at most* r *that satisfies* $\left\|G(s) - \hat{G}(s)\right\|_2 < \gamma$ *if and only if there exist* $X, Z \in \mathcal{S}_n$ *satisfying*

$$AX + XA^* + BB^* < 0 , \tag{10}$$

$$A(X - Z) + (X - Z)A^* < 0 , \tag{11}$$

$$\mathrm{tr}\left\{C(X - Z)C^*\right\} < \gamma^2 , \tag{12}$$

$$Z \geq 0 , \tag{13}$$

$$\mathrm{rank}\, Z \leq r . \tag{14}$$

Proof

The elimination lemma [3, pp. 32-33] is used to eliminate $\hat{A}$, $\hat{B}$, $\hat{C}$ from inequalities (7) and (9). Inequality (7) can be written as

$$\begin{aligned}
&\begin{bmatrix} AP_{11}+P_{11}A^* & AP_{12} & B \\ P_{12}^*A^* & 0 & 0 \\ B^* & 0 & -I \end{bmatrix} + \begin{bmatrix} 0 & P_{12}\hat{A}^* & 0 \\ \hat{A}P_{12}^* & \hat{A}P_{22}+P_{22}\hat{A}^* & \hat{B} \\ 0 & \hat{B}^* & 0 \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} AP_{11}+P_{11}A^* & AP_{12} & B \\ P_{12}^*A^* & 0 & 0 \\ B^* & 0 & -I \end{bmatrix}}_{\mathcal{L}_1^c(P)} + \begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix} \begin{bmatrix} \hat{A} & \hat{B} \end{bmatrix} \begin{bmatrix} P_{12}^* & P_{22} & 0 \\ 0 & 0 & I \end{bmatrix} + ()^* < 0 \,. \qquad (15)
\end{aligned}$$

By noting that

$$\begin{bmatrix} 0 \\ I \\ 0 \end{bmatrix}^{\perp} = \begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix}, \qquad \begin{bmatrix} P_{12} & 0 \\ P_{22} & 0 \\ 0 & I \end{bmatrix}^{\perp} = \begin{bmatrix} I & -P_{12}P_{22}^{-1} & 0 \end{bmatrix},$$

$\begin{bmatrix} \hat{A} & \hat{B} \end{bmatrix}$ can be eliminated from (15). (Notice that (9) guarantees that P_{22} is nonsingular.) There exists $\begin{bmatrix} \hat{A} & \hat{B} \end{bmatrix}$ that satisfies (15) if and only if there exist P_{11}, P_{12}, P_{22} such that

$$\begin{bmatrix} I & 0 & 0 \\ 0 & 0 & I \end{bmatrix} \mathcal{L}_1^c(P) \begin{bmatrix} I & 0 \\ 0 & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} AP_{11}+P_{11}A^* & B \\ B^* & -I \end{bmatrix} < 0 \,, \qquad (16)$$

$$\begin{aligned}
&\begin{bmatrix} I & -P_{12}P_{22}^{-1} & 0 \end{bmatrix} \mathcal{L}_1^c(P) \begin{bmatrix} I \\ -P_{22}^{-1}P_{12}^* \\ 0 \end{bmatrix} \\
&\quad = A(P_{11}-P_{12}P_{22}^{-1}P_{12}^*) + (P_{11}-P_{12}P_{22}^{-1}P_{12}^*)A^* < 0 \,. \qquad (17)
\end{aligned}$$

Inequality (16) is equivalent to

$$AP_{11}+P_{11}A^* + BB^* < 0 \,. \qquad (18)$$

Similarly, inequality (9) can be written as

$$\begin{aligned}
&\begin{bmatrix} W & CP_{11} & CP_{12} \\ P_{11}C^* & P_{11} & P_{12} \\ P_{12}^*C^* & P_{12}^* & P_{22} \end{bmatrix} - \begin{bmatrix} 0 & \hat{C}P_{12}^* & \hat{C}P_{22} \\ P_{12}\hat{C}^* & 0 & 0 \\ P_{22}\hat{C}^* & 0 & 0 \end{bmatrix} \\
&= \underbrace{\begin{bmatrix} W & CP_{11} & CP_{12} \\ P_{11}C^* & P_{11} & P_{12} \\ P_{12}^*C^* & P_{12}^* & P_{22} \end{bmatrix}}_{\mathcal{L}_2^c(P)} - \begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix} \hat{C} \begin{bmatrix} 0 & P_{12}^* & P_{22} \end{bmatrix} + ()^* > 0 \,. \qquad (19)
\end{aligned}$$

Since

$$
\begin{bmatrix} I \\ 0 \\ 0 \end{bmatrix}^{\perp} = \begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \end{bmatrix}, \qquad \begin{bmatrix} 0 \\ P_{12} \\ P_{22} \end{bmatrix}^{\perp} = \begin{bmatrix} I & 0 & 0 \\ 0 & I & -P_{12}P_{22}^{-1} \end{bmatrix},
$$

there exists $\hat{C}$ that satisfies (19) if and only if there exist P_{11}, P_{12}, P_{22} such that

$$
\begin{bmatrix} 0 & I & 0 \\ 0 & 0 & I \end{bmatrix} \mathcal{L}_2^c(P) \begin{bmatrix} 0 & 0 \\ I & 0 \\ 0 & I \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix} > 0 \; , \tag{20}
$$

$$
\begin{aligned}
&\begin{bmatrix} I & 0 & 0 \\ 0 & I & -P_{12}P_{22}^{-1} \end{bmatrix} \mathcal{L}_2^c(P) \begin{bmatrix} I & 0 \\ 0 & I \\ 0 & -P_{12}P_{22}^{-1} \end{bmatrix} \\
&= \begin{bmatrix} W & C(P_{11} - P_{12}P_{22}^{-1}P_{12}^*) \\ (P_{11} - P_{12}P_{22}^{-1}P_{12}^*)C^* & P_{11} - P_{12}P_{22}^{-1}P_{12}^* \end{bmatrix} > 0 \; .
\end{aligned} \tag{21}
$$

Inequality (20) is equivalent to

$$
P_{11} - P_{12}P_{22}^{-1}P_{12}^* > 0 \; , \tag{22}
$$

$$
P_{22} > 0 \; . \tag{23}
$$

Furthermore, inequality (21) is equivalent to

$$
P_{11} - P_{12}P_{22}^{-1}P_{12}^* > 0 \; , \tag{24}
$$

$$
W - C(P_{11} - P_{12}P_{22}^{-1}P_{12}^*)C^* > 0 \; . \tag{25}
$$

Now, inequalities (8) and (25) imply

$$
\operatorname{tr}\left\{C(P_{11} - P_{12}P_{22}^{-1}P_{12}^*)C^*\right\} < \gamma^2 \; , \tag{26}
$$

while this implies the existence of W that satisfies (8) and (25). Note that inequality (17) requires inequalities (22) and (24) to hold since A is stable. Therefore, inequalities needed are (18), (17), (23) and (26). By writing $X = P_{11}$, $Z = P_{12}P_{22}^{-1}P_{12}^*$, inequalities (18), (17), (26), respectively, can be written as (10), (11), (12), respectively. Also the form of Z and (23) implies (13) and (14). Conversely, any Z satisfying (13) and (14) can be decomposed in the required form by, e.g., eigenvalue-eigenvector decomposition. This concludes the proof.

□

While inequalities (10)-(13) are convex constraints, the rank constraint (14) is not. An optimization problem/a feasibility problem under those constraints is a non-convex problem. Thus interior-point algorithms used for (convex) LMI feasibility/optimization problems cannot be employed and this makes the $\mathcal{H}_2$ model reduction problem a difficult task. This is not surprising since a number of reduced order controller synthesis problems, which, without controller order constraints, would be formulated as convex feasibility/optimization problems, yield non-convex feasibility/optimization problems, and the problem considered here is a special case of the $\mathcal{H}_2$ optimal controller synthesis problem with fixed controller order.

If X and Z that satisfy (10)-(14) are found, then a reduced order model that achieves the error less than γ can be obtained by firstly computing P_{12}, P_{22} from a decomposition of Z and then solving an LMI feasibility problem (7), (8), (9) for $\hat{A}$, $\hat{B}$, $\hat{C}$.

2.2 Discrete Systems

Now consider the discrete time case. Suppose that state space realizations of the original system $G(z)$, the reduced order approximant $\hat{G}(z)$ and the error system $E(z)$ are given as in the right hand sides of (1), (2) and (3), respectively. The model reduction problem can be expressed identically to the continuous time case except for (4), which is to be replaced with

$$A_E P A_E^* - P + B_E B_E^* < 0 \,. \tag{27}$$

Similar to the continuous time case, necessary and sufficient conditions with respect to two symmetric matrices are derived.

Theorem 2 *Consider a stable discrete time system* $G(z) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$ *of McMillan degree* n. *There exists a stable discrete time system* $\hat{G}(z)$ *of McMillan degree at most* r *that satisfies* $\left\|G(z) - \hat{G}(z)\right\|_2 < \gamma$ *if and only if there exist* $X, Z \in \mathcal{S}_n$ *satisfying*

$$AXA^* - X + BB^* < 0 \,, \tag{28}$$

$$A(X-Z)A^* - (X-Z) < 0 \,, \tag{29}$$

$$\operatorname{tr}\left\{C(X-Z)C^*\right\} < \gamma^2 \,, \tag{30}$$

$$Z \geq 0 \,, \tag{31}$$

$$\operatorname{rank} Z \leq r \,. \tag{32}$$

Proof

Inequalities (5) and (27) are equivalent to

$$\begin{bmatrix} P & -A_E P & -B_E \\ -PA_E^* & P & 0 \\ -B_E^* & 0 & I \end{bmatrix} = \begin{bmatrix} P_{11} & P_{12} & -AP_{11} & -AP_{12} & -B \\ P_{12}^* & P_{22} & -\hat{A}P_{12}^* & -\hat{A}P_{22} & -\hat{B} \\ -P_{11}A^* & -P_{12}\hat{A}^* & P_{11} & P_{12} & 0 \\ -P_{12}^*A^* & -P_{22}\hat{A}^* & P_{12}^* & P_{22} & 0 \\ -B^* & -\hat{B}^* & 0 & 0 & I \end{bmatrix} > 0 \,. \tag{33}$$

This can be written as

$$
\begin{bmatrix} P_{11} & P_{12} & -AP_{11} & -AP_{12} & -B \\ P_{12}^* & P_{22} & 0 & 0 & 0 \\ -P_{11}A^* & 0 & P_{11} & P_{12} & 0 \\ -P_{12}^*A^* & 0 & P_{12}^* & P_{22} & 0 \\ -B^* & 0 & 0 & 0 & I \end{bmatrix} - \begin{bmatrix} 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & \hat{A}P_{12}^* & \hat{A}P_{22} & \hat{B} \\ 0 & P_{12}\hat{A}^* & 0 & 0 & 0 \\ 0 & P_{22}\hat{A}^* & 0 & 0 & 0 \\ 0 & \hat{B}^* & 0 & 0 & 0 \end{bmatrix}
$$

$$
= \underbrace{\begin{bmatrix} P_{11} & P_{12} & -AP_{11} & -AP_{12} & -B \\ P_{12}^* & P_{22} & 0 & 0 & 0 \\ -P_{11}A^* & 0 & P_{11} & P_{12} & 0 \\ -P_{12}^*A^* & 0 & P_{12}^* & P_{22} & 0 \\ -B^* & 0 & 0 & 0 & I \end{bmatrix}}_{\mathcal{L}_1^d(P)} - \begin{bmatrix} 0 \\ I \\ 0 \\ 0 \\ 0 \end{bmatrix} \begin{bmatrix} \hat{A} & \hat{B} \end{bmatrix} \begin{bmatrix} 0 & 0 & P_{12}^* & P_{22} & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix} + ()^*
$$

$$
> 0\ . \qquad (34)
$$

Using

$$
\begin{bmatrix} 0 \\ I \\ 0 \\ 0 \\ 0 \end{bmatrix}^{\perp} = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix}, \qquad \begin{bmatrix} 0 & 0 \\ 0 & 0 \\ P_{12} & 0 \\ P_{22} & 0 \\ 0 & I \end{bmatrix}^{\perp} = \begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & -P_{12}P_{22}^{-1} & 0 \end{bmatrix},
$$

it is seen that there exists $\begin{bmatrix} \hat{A} & \hat{B} \end{bmatrix}$ that satisfies (34) if and only if there exist P_{11}, P_{12}, P_{22} such that

$$
\begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & 0 & I & 0 & 0 \\ 0 & 0 & 0 & I & 0 \\ 0 & 0 & 0 & 0 & I \end{bmatrix} \mathcal{L}_1^d(P) \begin{bmatrix} I & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 \\ 0 & 0 & I & 0 \\ 0 & 0 & 0 & I \end{bmatrix} = \begin{bmatrix} P_{11} & -AP_{11} & -AP_{12} & -B \\ -P_{11}A^* & P_{11} & P_{12} & 0 \\ -P_{12}^*A^* & P_{12}^* & P_{22} & 0 \\ -B^* & 0 & 0 & I \end{bmatrix} > 0\ , \qquad (35)
$$

$$
\begin{bmatrix} I & 0 & 0 & 0 & 0 \\ 0 & I & 0 & 0 & 0 \\ 0 & 0 & I & -P_{12}P_{22}^{-1} & 0 \end{bmatrix} \mathcal{L}_1^d(P) \begin{bmatrix} I & 0 & 0 \\ 0 & I & 0 \\ 0 & 0 & I \\ 0 & 0 & -P_{22}^{-1}P_{12}^* \\ 0 & 0 & 0 \end{bmatrix}
$$

$$
= \begin{bmatrix} P_{11} & P_{12} & -A(P_{11} - P_{12}P_{22}^{-1}P_{12}^*) \\ P_{12}^* & P_{22} & 0 \\ -(P_{11} - P_{12}P_{22}^{-1}P_{12}^*)A^* & 0 & P_{11} - P_{12}P_{22}^{-1}P_{12}^* \end{bmatrix} > 0\ . \qquad (36)
$$

Moreover, inequality (35) is equivalent to

$$P_{11} - \begin{bmatrix} AP_{11} & AP_{12} & B \end{bmatrix} \begin{bmatrix} P_{11} & P_{12} & 0 \\ P_{12}^* & P_{22} & 0 \\ 0 & 0 & I \end{bmatrix}^{-1} \begin{bmatrix} P_{11}A^* \\ P_{12}^*A^* \\ B^* \end{bmatrix} > 0 \,, \tag{37}$$

$$\begin{bmatrix} P_{11} & P_{12} \\ P_{12}^* & P_{22} \end{bmatrix} > 0 \,. \tag{38}$$

A straightforward calculation shows that (37) can be written as

$$AP_{11}A^* - P_{11} + BB^* < 0 \,. \tag{39}$$

Also inequality (38) is equivalent to

$$P_{11} - P_{12}P_{22}^{-1}P_{12}^* > 0 \,, \tag{40}$$

$$P_{22} > 0 \,. \tag{41}$$

Inequality (36) is equivalent to

$$P_{11} - \begin{bmatrix} P_{12} & -A(P_{11} - P_{12}P_{22}^{-1}P_{12}^*) \end{bmatrix} \begin{bmatrix} P_{22} & 0 \\ 0 & P_{11} - P_{12}P_{22}^{-1}P_{12}^* \end{bmatrix}^{-1} \begin{bmatrix} P_{12}^* \\ -(P_{11} - P_{12}P_{22}^{-1}P_{12}^*)A^* \end{bmatrix} > 0 \tag{42}$$

and inequalities (40)-(41). Inequality (42) can further be modified to

$$A(P_{11} - P_{12}P_{22}^{-1}P_{12}^*)A^* - (P_{11} - P_{12}P_{22}^{-1}P_{12}^*) < 0 \,. \tag{43}$$

Note that the stability of A along with (43) implies (40). Hence, inequalities (39), (43), (41) and (6) are only needed. Similar to the continuous time case, it can be shown that (6) can be replaced with (26). Finally, by writing $X = P_{11}$, $Z = P_{12}P_{22}^{-1}P_{12}^*$, it is deduced that the stated conditions are necessary and sufficient conditions for the existence of a reduced order approximant.

□

As is the case with continuous time systems, inequalities (28)-(31) are LMIs and thus convex, but the rank constraint (32) is not. It makes the problem a non-convex one.

3 $\mathcal{H}_2$ Model Reduction Algorithm Using the Alternating Projection Method

Not only the $\mathcal{H}_2$ model reduction problem but also more general controller synthesis problems with fixed controller order are formulated as LMIs with rank constraints. To tackle those control problems with controller order constraints, various heuristic algorithms have been proposed, e.g., [6, 10, 17, 18], though there is no guarantee for convergence of those algorithms to a solution.

In this report the alternating projection method [10] is employed for finding reduced order models. Consider a pair of sets $\mathcal{C}_1$ and $\mathcal{C}_2$ in the space $\mathcal{S}_n \times \mathcal{S}_n$ and assume that the intersection of these sets is non-empty. The feasibility problem of finding an element in the intersection $\mathcal{C}_1 \cap \mathcal{C}_2$ is considered. Denote by $P_{\mathcal{C}_i}$ the orthogonal projection operator onto the set $\mathcal{C}_i$. Suppose that $\mathcal{C}_1$ and $\mathcal{C}_2$ are closed and convex. Then, starting from any element (X_0, Z_0) in the space, the sequence of alternating projections

$$\begin{aligned}
(X_1, Z_1) &= P_{\mathcal{C}_1}(X_0, Z_0) \\
(X_2, Z_2) &= P_{\mathcal{C}_2}(X_1, Z_1) \\
&\vdots \\
(X_{2m-1}, Z_{2m-1}) &= P_{\mathcal{C}_1}(X_{2m-2}, Z_{2m-2}) \\
(X_{2m}, Z_{2m}) &= P_{\mathcal{C}_2}(X_{2m-1}, Z_{2m-1}) \\
&\vdots
\end{aligned}$$

always converges to an element in the intersection $\mathcal{C}_1 \cap \mathcal{C}_2$. In case the intersection is empty, the sequence does not converge.

When either $\mathcal{C}_1$ or $\mathcal{C}_2$ is non-convex, (global) convergence is not guaranteed. However local convergence is guaranteed, i.e., if a starting point is in a neighbourhood of a feasible solution, the alternating projection method can yield a sequence converging to an element in the intersection.

In the case of the $\mathcal{H}_2$ model reduction problem, $\mathcal{C}_1$ can be taken as

$$\mathcal{C}_1 = \{(X, Z) | X \in \mathcal{S}_n, Z \in \mathcal{S}_n, (10), (11), (12), (13)\}$$

in the continuous time case, and

$$\mathcal{C}_1 = \{(X, Z) | X \in \mathcal{S}_n, Z \in \mathcal{S}_n, (28), (29), (30), (31)\}$$

in the discrete time case. Also,

$$\mathcal{C}_2 = \{(X, Z) | X \in \mathcal{S}_n, Z \in \mathcal{S}_n, \text{rank } Z \le r\}$$

for either case. Note that $\mathcal{C}_1$ is convex while $\mathcal{C}_2$ is not.

By equipping the space $\mathcal{S}_n \times \mathcal{S}_n$ with the inner product

$$\langle (X_1, Z_1), (X_2, Z_2) \rangle = \text{tr}\{X_1 X_2\} + \text{tr}\{Z_1 Z_2\} \; ,$$

the orthogonal projection of $(X_0, Z_0) \in \mathcal{S}_n \times \mathcal{S}_n$ onto $\mathcal{C}_1$ can be found by solving the following (convex) optimization problem [10]:

$$\begin{aligned} \text{minimize} \;\; & \operatorname{tr}\{S + T\} \\ \text{subject to} \;\; & \begin{bmatrix} S & (X - X_0) \\ (X - X_0) & I \end{bmatrix} \geq 0 \;, \\ & \begin{bmatrix} T & (Z - Z_0) \\ (Z - Z_0) & I \end{bmatrix} \geq 0 \;, \\ & (X, Z) \in \mathcal{C}_1 \;, \quad S, T \in \mathcal{S}_n \;. \end{aligned}$$

This minimization problem can be solved by using standard numerical algorithms. The pair (X, Z) that solves it is the projection to be sought.

Now consider the projection $P_{\mathcal{C}_2}(X_0, Z_0)$. Since $\mathcal{C}_2$ is not convex, there may be more than one matrix pair that minimize the distance from (X_0, Z_0). Let $Z_0 = U\Sigma V^*$ be a singular value decomposition of Z_0. Then a projection of (X_0, Z_0) onto $\mathcal{C}_2$ is given by

$$P_{\mathcal{C}_2}(X_0, Z_0) = (X_0, U\Sigma_r V^*)$$

where Σ_r is a diagonal matrix obtained from Σ by replacing the $(n - r)$ smallest diagonal elements of Σ by zero [13, Section 7.4].

Now the following algorithm is suggested for $\mathcal{H}_2$ model reduction.

1. Find $X, Z \in \mathcal{S}_n$ and an $\mathcal{H}_2$-norm bound γ that satisfy (10)-(14) in the continuous time case (resp., (28)-(32) in the discrete time case).

2. Reduce γ. Find $X, Z \in \mathcal{S}_n$ that satisfy (10)-(14) (resp., (28)-(32)) using the alternating projection method, taking (X, Z) from the previous step as a starting point.

3. If successful, go back to Step 2. Otherwise, compute an approximant from the best (X, Z) available by solving a feasibility problem (7) (resp., (33)), (8) and (9).

It is also possible to use a bisection method with respect to γ.

It is repeated that the presented algorithm is heuristic since $\mathcal{C}_2$ is non-convex and thus the alternating projection method becomes heuristic. Hence this method may not provide a suboptimal approximant whose achieved error is as close to the optimal error as desired. It is of great significance to find a nice starting point in Step 1. This is because it can determine whether an approximant which is close to the global optimum will be obtained. Also, Step 2 is not in general an inexpensive task and it is desired to have initial γ close to the optimal γ, which may be achieved by having (X, Z) close to the optimum (or a suboptimum achieving practically the optimal error in the case where the optimum cannot be achieved). The choice of the starting point is the topic of the next section.

4 Choice of Starting Points

4.1 Choosing from the Balanced Realization

From now on, it is supposed that the given state space realization is balanced. This is a sensible assumption from the practical point of view since the use of balanced realizations in general improves the reliability of numerical computation and thus is common practice. In such a case the Gramian of the form

$$\Sigma = \begin{bmatrix} \sigma_1 & 0 & \cdots & 0 \\ 0 & \sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \sigma_n \end{bmatrix}$$

satisfies

$$A\Sigma + \Sigma A^* + BB^* = 0\ , \qquad A^*\Sigma + \Sigma A + C^*C = 0$$

in the continuous time case, or

$$A\Sigma A^* - \Sigma + BB^* = 0\ , \qquad A^*\Sigma A - \Sigma + C^*C = 0$$

in the discrete time case. The diagonal elements of Σ are called Hankel singular values and ordered as

$$\sigma_1 \geq \sigma_2 \geq \cdots \sigma_n \geq 0\ .$$

Model reduction by balanced truncation [21] is carried out by discarding modes corresponding to small Hankel singular values.

To make the left hand side of (12) (or (30)) small, a straightforward idea is to make $X - Z$ "small". From (10) (or (28)), it is observed that $X > \Sigma$ [26]. Inequality (11) (or (29)) implies that $X > Z$, and, if X is taken to be very close to Σ, the i-th diagonal element of Z is smaller than σ_i. Write $C = \begin{bmatrix} c_1 & c_2 & \cdots & c_n \end{bmatrix}$ where c_i is a column vector and observe that

$$\mathrm{tr}\left\{C\Sigma C^*\right\} = \mathrm{tr}\left\{C^*C\Sigma\right\} = \sum_i (c_i^* c_i)\sigma_i\ . \tag{44}$$

If Z is assumed to be a diagonal matrix whose r (diagonal) elements are only nonzero, and the only constraint is $\Sigma > Z$, then the optimal choice of Z is obtained by identifying the $n - r$ smallest terms of $(c_i^* c_i)\sigma_i$, $i = 1, \ldots, n$, and letting the corresponding diagonal elements of Σ be zero.

Unfortunately such Z does not in general satisfy inequality (11) (or (29)). An option may be to restrict the structure of Z and minimize γ under LMIs (10)-(13) (or (28)-(31)). Another method is to use the balanced realization. This has an effect similar to having structured Z. As is mentioned above, a balanced realization is often obtained beforehand, so it can be used without extra cost. However, instead of the ordinary balanced truncation, modes whose contributions to (44) are small are truncated. Indeed this idea can be used to find upper bounds of $\mathcal{H}_2$ error [20, 27].

Once a reduced order model and the $\mathcal{H}_2$-norm of the error system, which will serve as initial γ, are obtained in this way, the controllability Gramian of the error system is computed, from which initial X and Z are obtained. Note that, if the modes are chosen such that the retained modes and the truncated modes do not share the same Hankel singular values, then the reduced order system is stable [23], [28, Theorem 21.29] and thus inequalities (10) and (11) (or (28) and (29)) are automatically satisfied.

In fact it is observed that a good starting point is not necessarily obtained from the modes of the r largest contributions. In the worst case, $\binom{n}{r}$ combinations of modes are to be examined. Nevertheless, in practice, combinations containing modes with small contributions may be ignored and the number of combinations to be tested can be greatly reduced.

4.2 Improving Starting Points

It is observed that, once an initial approximant is obtained from the balanced realization as in the previous subsection, a better approximant may be obtained by solving a feasibility problem with LMIs. The set of LMIs (7) (or (33)), (8) and (9) is to be solved for $\hat{A}$, $\hat{B}$, $\hat{C}$ where P is replaced with the controllability Gramian of the error system and γ in the right hand side of (8) is replaced with the $\mathcal{H}_2$-norm of the error system. Notice that, once P is fixed, the set of inequalities are LMIs with respect to $\hat{A}$, $\hat{B}$, $\hat{C}$.

The above method usually gives a reduced order model that achieves a smaller error. The effect is sometimes trivial. (If the model obtained from the balanced realization is close to the (local) minimum, there is little room for improvement.) However the required computation cost is relatively small compared to the alternating projection method, so it is worth carrying out.

5 Numerical Examples

In this section three numerical examples are presented. The first two examples demonstrate the effectiveness of the choice of starting points in Section 4, and the third example shows that the algorithm proposed in Section 3 works at least as well as other methods. The error achieved by an approximant is shown in the relative error, i.e.,

$$\mathcal{J}(G, \hat{G}) = \frac{\left\| G - \hat{G} \right\|_2}{\|G\|_2} .$$

Example 1. This second order continuous time SISO example is taken from [25, Example 2]:

$$G(s) = \frac{10000s + 5000}{s^2 + 5000s + 25} \,.$$

A first order approximant is sought. Its Hankel singular values σ_i, $i = 1, 2$, are

$$99.000,\; 0.99990,$$

but, $(c_i^* c_i)\sigma_i$, $i = 1, 2$, are

$$102.05,\; 9997.9.$$

It is seen that the contribution of the mode corresponding to the second (i.e., the smaller) Hankel singular value to (44) is the larger. Indeed,

$$\mathcal{J}(G(s), \hat{G}_1(s)) = 0.99494,$$
$$\mathcal{J}(G(s), \hat{G}_2(s)) = 0.0985088$$

where $\hat{G}_i(s)$ is the approximant obtained by retaining the i-th mode.

An improved starting point is computed from $\hat{G}_2(s)$ using the method in Subsection 4.2. However the improvement is practically zero and the error the new approximant $\hat{G}_2'(s)$ achieves is

$$\mathcal{J}(G(s), \hat{G}_2'(s)) = 0.0985086.$$

The algorithm described in Section 3 hardly improves this, but this is natural since the optimal relative error is 0.0985 [25].

Example 2. This discrete time system is taken from [22]:

$$G(z) = \left[\begin{array}{cccc|cc} 0 & 1 & 0 & 0 & \frac{1}{2} & \frac{-3}{4} \\ 0 & 0 & 1 & 0 & \frac{383}{2080} & \frac{279}{1040} \\ 0 & 0 & 0 & 1 & \frac{1839}{8320} & \frac{-1317}{4160} \\ 0 & \frac{-1}{8} & \frac{1}{2} & \frac{1}{4} & \frac{1419}{33280} & \frac{99}{1280} \\ \hline 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 \end{array}\right].$$

This is a 4th order MIMO system, with 2 inputs and 2 outputs. A 2nd order approximant is to be found. Its Hankel singular values σ_i, $i = 1, \ldots, 4$, are

$$1.6752,\; 0.90114,\; 0.16511,\; 0.083916,$$

and, $(c_i^* c_i)\sigma_i$, $i = 1, \ldots, 4$, are

$$2.0191,\; 0.18054,\; 0.0045092,\; 0.00082210.$$

In this example the order of the contributions is identical to that of the Hankel singular values. Indeed the approximant $\hat{G}_{12}(z)$, constructed by retaining the 1st and 2nd modes and discarding the 3rd and 4th modes, achieves the smallest error among $\binom{4}{2} = 6$ possible approximants of order 2:

$$\mathcal{J}(G(z), \hat{G}_{12}(z)) = 0.089277.$$

The method in Subsection 4.2 is carried out to find an improved starting point, but again the improvement is not significant:

$$\mathcal{J}(G(z), \hat{G}'_{12}(z)) = 0.089119.$$

Also the algorithm described in Section 3 hardly improves this since $\hat{G}'_{12}(z)$ is nearly optimal; The optimal approximant given in [22] achieves

$$\mathcal{J}(G(z), \hat{G}_{\mathrm{opt}}(z)) = 0.089039.$$

Notice that

$$\frac{\mathcal{J}(G(z), \hat{G}'_{12}(z))}{\mathcal{J}(G(z), \hat{G}_{\mathrm{opt}}(z))} = 1.00089675.$$

Example 3. A more realistic example, the System AUTM in [15], is examined. The system to be approximated is a 12th order continuous time MIMO system, with 2 inputs and 2 outputs. The system data is given in Appendix A. The Hankel singular values σ_i, $i = 1, 2, \ldots, 12$, are

$$\begin{gathered} 7.1833,\ 1.4904,\ 0.92791,\ 0.58756,\ 0.46331,\ 0.23683,\ 0.16132, \\ 0.093582,\ 0.56596 \times 10^{-3},\ 0.20608 \times 10^{-4},\ 0.14124 \times 10^{-5},\ 0.34341 \times 10^{-7}, \end{gathered}$$

and, $(c_i^* c_i)\sigma_i$, $i = 1, 2, \ldots, 12$, are

$$\begin{gathered} 17.001,\ 3.6844,\ 0.23558,\ 0.076721,\ 2.5247,\ 0.30031,\ 0.065789, \\ 0.20424,\ 4.0153 \times 10^{-6},\ 1.1683 \times 10^{-8},\ 9.8808 \times 10^{-13},\ 2.1223 \times 10^{-15}. \end{gathered}$$

It is observed that modes corresponding to large Hankel singular values in general have large contributions in (44), but the orders are slightly different.

Reduced order models of McMillan degrees 4, 5 and 6 are sought. First a 4th order approximant is found. In this case the best initial approximant is obtained by discarding all the modes except for those corresponding to the 1st, 2nd, 3rd and 5th Hankel singular values, or the 1st, 2nd, 3rd and 5th largest $(c_i^* c_i)\sigma_i$:

$$\mathcal{J}(G(z), \hat{G}_{1235}(s)) = 0.15231.$$

It is pointed out that the approximant constructed from the modes corresponding to the 4 largest $(c_i^* c_i)\sigma_i$ yields a much worse error:

$$\mathcal{J}(G(z), \hat{G}_{1257}(s)) = 0.36873.$$

By means of the method in Subsection 4.2, a slight improvement of the starting point is made:

$$\mathcal{J}(G(z), \hat{G}'_{1235}(s)) = 0.15171.$$

Finally the $\mathcal{H}_2$ model reduction algorithm in Section 3 is invoked and an approximant $\hat{G}^4_{\mathrm{APM}}(s)$ that achieves the following error is obtained:

$$\mathcal{J}(G(z), \hat{G}^4_{\mathrm{APM}}(s)) = 0.13494.$$

The data of the approximant is provided in Appendix B. It is pointed out that this is a slight improvement over the approximant $\hat{G}^4_{\mathrm{YL}}(s)$ reported in [27]:

$$\mathcal{J}(G(z), \hat{G}^4_{\mathrm{YL}}(s)) = 0.1354.$$

The same procedure is executed for $r = 5$. By retaining the modes corresponding to the 1st, 2nd, 3rd, 5th and 8th Hankel singular values, or the 1st, 2nd, 3rd, 5th and 6th largest $(c_i^* c_i)\sigma_i$, the best initial approximant is obtained:

$$\mathcal{J}(G(z), \hat{G}_{12358}(s)) = 0.10831.$$

This is improved a little by the method described in Subsection 4.2:

$$\mathcal{J}(G(z), \hat{G}'_{12358}(s)) = 0.10634.$$

Finally the model reduction algorithm finds an approximant achieving

$$\mathcal{J}(G(z), \hat{G}^5_{\mathrm{APM}}(s)) = 0.078078.$$

Again a slight improvement over the result in [27] is observed:

$$\mathcal{J}(G(z), \hat{G}^5_{\mathrm{YL}}(s)) = 0.0795.$$

For $r = 6$, the best initial approximant is constructed by keeping the modes corresponding to the 1st, 2nd, 3rd, 4th, 5th and 8th Hankel singular values, or the 1st, 2nd, 3rd, 5th, 6th and 7th largest $(c_i^* c_i)\sigma_i$. It achieves the error

$$\mathcal{J}(G(z), \hat{G}_{123458}(s)) = 0.078882.$$

The method in Subsection 4.2 yields a slightly better approximation:

$$\mathcal{J}(G(z), \hat{G}'_{123458}(s)) = 0.076574.$$

Finally the model reduction algorithm finds an approximant whose error is

$$\mathcal{J}(G(z), \hat{G}^{6}_{\mathrm{APM}}(s)) = 0.052709.$$

This reduced order model is better than the one reported in [27]:

$$\mathcal{J}(G(z), \hat{G}^{6}_{\mathrm{YL}}(s)) = 0.0541.$$

6 Conclusion

This report has considered the $\mathcal{H}_2$ model reduction problem. Necessary and sufficient conditions for the existence of an $\mathcal{H}_2$ suboptimal reduced order model are derived for both continuous and discrete time cases by means of LMI techniques. The resulting constraints are non-convex and do not allow globally convergent algorithms to be developed. A heuristic algorithm is proposed which utilizes the alternating projection method. Along with the suggested method for choosing starting points, this algorithm can find suboptimal approximants which are as good as those computed by previously proposed methods, which is demonstrated by numerical examples.

This algorithm is believed to have several advantages. It relies on off-the-shelf routines and requires rather simple programming. It covers both continuous and discrete time systems; The difference in the programs is trivial. The conditions share the same structure—several LMIs and a matrix rank constraint—as various controller synthesis problems with fixed controller order. Observe in particular some similarity of the conditions in [5] for $\mathcal{H}_\infty$ model reduction and those in [4] for model reduction in the ν-gap metric. Research on the solution of such problems is one of areas where intensive studies are carried out and the potential for the development of efficient, numerically reliable algorithms to tackle those problems including this approach to $\mathcal{H}_2$ model reduction can be large.

A The System AUTM

In Example 3 of this report, the system AUTM, a 2-input, 12-state, 2-output model of an automotive gas turbine, is used, which is studied in [15]. The following is a state-space representation of this model, taken from [15]:

$$G(s) = \left[\begin{array}{c|c} A & B \\ \hline C & 0 \end{array}\right]$$

where

$A =$

$$\begin{bmatrix} 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -0.202 & -1.15 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -2.360 & -13.6 & -12.8 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -1.62 & -9.4 & -9.15 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -188.0 & -111.6 & -116.4 & -20.8 \end{bmatrix},$$

$$B = \begin{bmatrix} 0 & 1.0439 & 0 & 0 & -1.794 & 0 & 0 & 1.0439 & 0 & 0 & 0 & -1.794 \\ 0 & 4.1486 & 0 & 0 & 2.6775 & 0 & 0 & 4.1486 & 0 & 0 & 0 & 2.6775 \end{bmatrix}^T,$$

$$C = \begin{bmatrix} 0.2640 & 0.8060 & -1.420 & -15.00 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 4.9000 & 2.1200 & 1.9500 & 9.3500 & 25.800 & 7.1400 & 0 \end{bmatrix}.$$

B Reduced Order Models Obtained in Example 3

The obtained approximants of McMillan degrees 4, 5 and 6 for the system in Example 3 are as follows:

$$G_r(s) = \left[\begin{array}{c|c} A_r & B_r \\ \hline C_r & 0 \end{array}\right], \; r = 4, 5, 6$$

where

$$A_4 = \begin{bmatrix} -0.1632 & -0.003010 & 0.01469 & 0.02672 \\ -0.02222 & -0.8835 & 0.1378 & 0.2156 \\ -0.3549 & -1.193 & -0.1657 & -1.197 \\ 9.138 & -1.434 & 2.953 & -8.384 \end{bmatrix},$$

$$B_4 = \begin{bmatrix} 0.06693 & 0.2025 \\ 1.060 & -0.3248 \\ 0.6186 & -0.04439 \\ -1.417 & -6.133 \end{bmatrix},$$

$$C_4 = \begin{bmatrix} 3.885 & 2.160 & -0.0002362 & -0.4002 \\ 10.34 & -0.7046 & -0.4843 & -1.009 \end{bmatrix},$$

$$A_5 = \begin{bmatrix} -0.1645 & -0.005458 & 0.01376 & -0.02903 & 0.001243 \\ -0.004979 & -0.7658 & 0.1608 & -0.2112 & -0.1891 \\ -0.3550 & -1.156 & -0.1595 & 1.226 & -0.04172 \\ -9.036 & 1.864 & -2.882 & -8.297 & -0.1604 \\ 1.028 & 22.35 & 7.137 & -2.497 & -10.38 \end{bmatrix},$$

$$B_5 = \begin{bmatrix} 0.06813 & 0.2029 \\ 0.9769 & -0.3062 \\ 0.6061 & -0.03843 \\ 1.136 & 6.160 \\ -12.88 & 3.566 \end{bmatrix},$$

$$C_5 = \begin{bmatrix} 3.874 & 2.192 & 0.008667 & 0.4068 & 0.1555 \\ 10.35 & -0.6900 & -0.4799 & 1.010 & -0.07624 \end{bmatrix},$$

$$A_6 = \begin{bmatrix} -0.1651 & -0.004383 & 0.01380 & 0.002821 & -0.03048 & -0.0002774 \\ 0.02033 & -0.8239 & 0.1817 & 0.1418 & -0.09048 & 0.1308 \\ -0.4088 & -1.139 & -0.1363 & -0.1050 & 1.040 & 0.04388 \\ 0.7130 & -0.6050 & 0.1694 & -0.07483 & -0.3059 & -0.07212 \\ -9.010 & 1.409 & -2.841 & 2.210 & -8.148 & 0.01207 \\ 1.602 & -25.53 & -7.066 & -1.225 & 7.046 & -12.31 \end{bmatrix},$$

$$B_6 = \begin{bmatrix} 0.06731 & 0.2036 \\ 0.9984 & -0.3296 \\ 0.5602 & 0.01618 \\ -0.008500 & -0.5092 \\ 1.281 & 6.097 \\ 14.05 & -5.727 \end{bmatrix},$$

$$C_6 = \begin{bmatrix} 3.855 & 2.238 & -0.05627 & -0.1745 & 0.3806 & -0.1526 \\ 10.36 & -0.6923 & -0.4478 & -0.05021 & 1.014 & 0.06293 \end{bmatrix}.$$

References

[1] L. Baratchart. Existence and generic properties of L^2 approximants for linear systems. *IMA Journal of Mathematical Control and Information*, 3:89–101, 1986.

[2] L. Baratchart, M. Cardelli, and M. Olivi. Identification and rational L^2 approximation: A gradient algorithm. *Automatica*, 27(2):413–417, March 1991.

[3] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan. *Linear Matrix Inequalities in System and Control Theory*, volume 15 of *SIAM Studies in Applied Mathematics*. Society for Industrial and Applied Mathematics, Philadelphia, PA, 1994.

[4] M. Cantoni. On model reduction in the ν-gap metric. In *Proceedings of the 40th IEEE Conference on Decision and Control*, pages 3665–3670, Orlando, Florida, December 2001.

[5] Y. Ebihara and T. Hagiwara. On $\mathcal{H}_\infty$ model reduction using LMIs. *IEEE Transactions on Automatic Control*, 49(7):1187–1191, July 2004.

[6] L. El Ghaoui, F. Oustry, and M. AitRami. A cone complementarity linearization algorithm for static output-feedback and related problems. *IEEE Transactions on Automatic Control*, 42(8):1171–1176, August 1997.

[7] A. Ferrante, W. Krajewski, A. Lepschy, and U. Viaro. Convergent algorithm for L_2 model reduction. *Automatica*, 35(1):75–79, January 1999.

[8] K. Glover. All optimal Hankel-norm approximations of linear multivariable systems and their L^∞-error bounds. *International Journal of Control*, 39(6):1115–1193, June 1984.

[9] K. M. Grigoriadis. Optimal H_∞ model reduction via linear matrix inequalities: Continuous- and discrete-time cases. *Systems and Control Letters*, 26(5):321–333, December 1995.

[10] K. M. Grigoriadis and E. B. Beran. Alternating projection algorithms for linear matrix inequalities problems with rank constraints. In L. El Ghaoui and S.-I. Niculescu, editors, *Advances in Linear Matrix Inequality Methods in Control*, chapter 13, pages 251–267. Society for Industrial and Applied Mathematics, 2000.

[11] S. Gugercin and A. C. Antoulas. A survey of model reduction by balanced truncation and some new results. *International Journal of Control*, 77(8):748–766, May 2004.

[12] B. Hanzon and J. M. Maciejowski. Constructive algebra methods for the L_2-problem for stable linear systems. *Automatica*, 32(12):1645–1657, December 1996.

[13] R. A. Horn and C. R. Johnson. *Matrix Analysis*. Cambridge University Press, Cambridge, 1985.

[14] X.-X. Huang, W.-Y. Yan, and K. L. Teo. H_2 near-optimal model reduction. *IEEE Transactions on Automatic Control*, 46(8):1279–1284, August 2001.

[15] Y. S. Hung and A. G. J. MacFarlane. *Multivariable Feedback: A Quasi-Classical Approach*, volume 40 of *Lecture Notes in Control and Information Sciences*. Springer-Verlag, Berlin, 1982.

[16] L. Meier III and D. G. Luenberger. Approximation of linear constant systems. *IEEE Transactions on Automatic Control*, 12(5):585–588, October 1967.

[17] T. Iwasaki. The dual iteration for fixed-order control. *IEEE Transactions on Automatic Control*, 44(4):783–788, April 1999.

[18] T. Iwasaki and R. E. Skelton. The XY-centering algorithm for the dual LMI problem: A new approach to fixed-order control design. *International Journal of Control*, 62(6):1257–1272, December 1995.

[19] D. Jibetean. *Algebraic Optimization with Applications to System Theory*. PhD thesis, Vrije Universiteit, Amsterdam, 2003.

[20] P. T. Kabamba. Balanced gains and their significance for L^2 model reduction. *IEEE Transactions on Automatic Control*, 30(7):690–693, July 1985.

[21] B. C. Moore. Principal component analysis in linear systems: Controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, 26(1):17–32, February 1981.

[22] R. L. M. Peeters, B. Hanzon, and D. Jibetean. Optimal $\mathcal{H}_2$ model reduction in state-space: A case study. In *Proceedings of the European Control Conference ECC 2003*, Cambridge, U.K., September 2003.

[23] L. Pernebo and L. M. Silverman. Model reduction via balanced state space representations. *IEEE Transactions on Automatic Control*, 27(2):382–387, April 1982.

[24] C. Scherer, P. Gahinet, and M. Chilali. Multiobjective output-feedback control via LMI optimization. *IEEE Transactions on Automatic Control*, 42(7):896–911, July 1997.

[25] J. T. Spanos, M. H. Milman, and D. L. Mingori. A new algorithm for L_2 optimal model reduction. *Automatica*, 28(5):897–909, September 1992.

[26] H. K. Wimmer. Monotonicity of maximal solutions of algebraic Riccati equations. *Systems and Control Letters*, 5(5):317–319, April 1985.

[27] W. Yan and J. Lam. An approximate approach to H^2 optimal model reduction. *IEEE Transactions on Automatic Control*, 44(7):1341–1358, July 1999.

[28] K. Zhou, J. C. Doyle, and K. Glover. *Robust and Optimal Control*. Prentice-Hall, Upper Saddle River, NJ, 1996.