Robust Robust Model Reduction

Yoram Halevi and Uri Shaked

Abstract— **The problem of linear model reduction is addressed. Given a state-space model of a linear time-invariant system, a model of prescribed order is obtained such that the H_2-norm of the difference between the transference of the two models is minimized. The reduced model is modeled as having the same order as the system but with a nonminimal observer form realization. The solution is then based on full order LMIs. The model reduction method is extended to the case where the model to be reduced suffers from parameters uncertainties that lie in a prescribed polytope. A reduced order model is obtained that achieves a prescribed upper-bound on the H_2-norm of the differences between the transference of the reduced order model and all the transferences of all the possible systems in the polytope.**

I. Introduction

Approximation of high order, complex systems by lower order, relatively simple models is one of the fundamental problems in linear system theory and has received considerable attention for many years. Since the early 1980's this problem has received renewed interest and several new state space methods such as balanced realization ([1],[2]), component cost analysis ([3]), and the Hankel norm approximations ([4]), to name a few, were suggested.

The optimal H_2 reduced order model was first derived in [5] and later in [6], by direct optimization, without imposing any structure on it. It turned out that it is given in terms of a projection into a lower order subspace and therefore the solution is sometimes referred to as the 'optimal projection'. The method has been extended to include bounds on the H_∞ error ([7]) and frequency weighting ([8]). The outcome of all of these works, were sets of nonlinearly coupled Lyapunov-like equations. Homotopic methods were used for their numerical solution, e.g. ([9], [10]) but their application is not trivial due to convergence problems and the large amount of required computation.

Other approaches to optimal H_2 order reduction were aimed at obtaining an optimization algorithm rather than a set of algebraic equations. That includes the iterative algorithm in [11] and gradient flow methods ([12]). The main problem in parametric optimization methods, such as the gradient flow, is maintaining stability. In [13] that problem was solved by restricting the reduced order model to those obtained by a symmetric projection.

Linear Matrix Inequalities(LMIs) were used for order reduction in ([14],([15],[16]). The set of conditions in those cases includes a rank condition, which is not convex. The solution involves then methods like alternating projections and semidefinite programming.

Robustness issues were addressed in several works. In [17], the reduced order model involved minimization over a class on norm bounded perturbations of the state space matrices. In [16] the uncertainty was defined by LFT with norm bounded, but otherwise arbitrary, operators . In [10] the problem of updating the reduced order model without recalculation, for given changes in the parameters,was considered.

In parallel to the introduction of the above methods for model reduction, methods for robust filtering have been developed in [18] and [19] which apply LMIs. It is the purpose of the present note to apply these methods to the problem of robust model reduction.

Notation: Throughout the paper the superscript 'T' stands for matrix transposition, $\mathcal{R}^n$ denotes the n dimensional Euclidean space, $\mathcal{R}^{n\times m}$ is the set of all $n\times m$ real matrices, and the notation $P>0$, for $P\in\mathcal{R}^{n\times n}$ means that P is symmetric and positive definite. Mathematical expectation is denoted by $\mathcal{E}$.

II. Problem formulation

We consider the following asymptotically stable linear system

$$\dot{x}=Ax+Bw,\quad y=Cx+Dw \qquad \text{(1a,b)}$$

where $x\in\mathcal{R}^n$ is the system state, $y\in\mathcal{R}^r$ is the measured output and $w\in\mathcal{R}^p$ is a standard zero mean white noise.

The system matrices are uncertain. They are supposed to belong to the following uncertainty polytope:

$$\Omega \triangleq \{(A,\ B,\ C,\ D)\,|((A,\ B,\ C,\ D)= \sum_{i=1}^{N}\alpha_i(A^{(i)},B^{(i)},C^{(i)},D^{(i)}).\ \alpha_i\geq 0,\ \sum_{i=1}^{N}\alpha_i=1\}. \qquad (2)$$

We want to obtain a robust model with some restrictions on its parameters that produces an output $\hat{y}$ which leads to a small error $\tilde{y}=y-\hat{y}$ over the entire uncertainty polytope. We seek a stationary linear time-invariant asymptotically stable model of order k with the state-space representation:

$$\dot{\hat{x}}=A_m\hat{x}+B_mw,\quad \hat{y}=C_m\hat{x}+D_mw \qquad \text{(3a,b)}$$

where $A_m,\ B_m,\ C_m$ and D_m are constant matrices of the appropriate dimensions. In the latter model some restrictions may be imposed on the matrices A_m and C_m. When $k=n$ the obtained model will be referred to as the full-order model. A reduced-order model will be obtained for $k<n$.

The above model is required to achieve a minimum upper-bound for the variance of $\tilde{y}$ over the entire uncertainty polytope. Namely, it should minimize the following performance measure:

$$J_1=\max_{\Omega}\mathcal{E}\{\tilde{y}^T\tilde{y}\} \qquad (4)$$

This work was supported by C&M Maus Chair at Tel Aviv University, Israel

Y. Halevi is with the Faculty of Mechanical Engineering, Technion-I.I.T, Haifa 32000, Israel merhy01@ tx.technion.ac.il

U. Shaked is with the School of Electrical Engineering, Tel Aviv University, Tel Aviv 69978, Israel shaked@eng.tau.ac.il

The same performance measure can be defined deterministicly by considering the integral of the impulse response $\tilde{y}$, or by using the system norm of the augmented error producing system to be defined in (5). The H_2 modeling problem we consider is thus the following:

Problem: Given a scalar $0 < \delta$, obtain an asymptotically stable model of (3) that ensures that $J_1 < \delta$.

Remark 1: In the above we required the system, and therefore the reduced order model, to be asymptotically stable. This requirement is needed in order to obtain a stable transference from w to $\tilde{y}$. In the case where the system possesses unstable modes, one can either i) separate between the contribution of the stable and the unstable modes of the system to its transference, and then find a reduced order model for the stable part, or ii) use coprime factorization and reduce each component separately. None of these methods can, however, be applied here. Generally, i) cannot be applied simultaneously to the entire polytope and ii) does not match the unstable modes exactly, thus it cannot be applied in a H_2 setting. The robustness results of this paper are therefore limited to polytopes of asymptotically stable systems.

III. The robust H_2 model

A. The full-order model

We begin our discussion by considering the system (1) which is assumed to be perfectly known. The case where its parameters are known to reside in the uncertainty polytope (2) will be treated later. It is noted that there is hardly any need to seek for a full order approximation of the system. However the results of this case are the basis for cases where some restrictions are to be imposed on the model dynamics or for uncertain systems.

Denoting $\xi = col\{x, \hat{x}\}$, (1) and (3) can be rewritten as

$$\dot{\xi} = \tilde{A}\xi + \tilde{B}w, \quad \tilde{y} = \tilde{C}\xi + \tilde{D}w \tag{5a,b}$$

where

$$\tilde{A}=\begin{bmatrix} A & 0 \\ 0 & A_m \end{bmatrix},\ \tilde{B}=\begin{bmatrix} B \\ B_m \end{bmatrix},\ \tilde{C}=\begin{bmatrix} C & -C_m \end{bmatrix},\ \tilde{D}=D-D_m. \tag{6a-d}$$

It is well known (see, e.g [21]) that using the Lyapunov function $\xi^T Q \xi$ for the system (5), $\mathcal{E}\{\tilde{y}^T\tilde{y}\} < \delta$ iff $\tilde{D} = 0$ and there exist matrices $Q \in \mathcal{R}^{2n\times 2n}$ and $Z \in \mathcal{R}^{m\times m}$ that satisfy the following inequalities:

$$\begin{bmatrix} \tilde{A}^T Q + Q\tilde{A} & Q\tilde{B} \\ * & -\delta I_p \end{bmatrix} < 0, \quad \begin{bmatrix} Z & \tilde{C} \\ \tilde{C}^T & Q \end{bmatrix} > 0, \quad trace\{Z\} < 1 \tag{7a-c}$$

The latter inequalities are linear in the variables Q and Z and can thus be used to verify whether for given $A_m,\ B_m$ C_m and D_m the error variance is less than δ. When the parameters of the model are unknown the latter inequalities are nonlinear. There exists, however, a linearizing method that reduces these inequalities to LMIs[20].

It should be noted that for the H_2-norm to be finite one should require that $D_m = D$. We can thus take $D = D_m = 0$ in the two models. Applying the latter linearization method on the inequalities (7) we obtain the following.

Theorem 1: Consider the system (1). If there exists a model (3) that achieves a cost J_1 of (4) less than a prescribed positive scalar δ, then there exist matrices $X,\ W,\ \bar{A}\ \in \mathcal{R}^{n\times n}$ $\bar{B} \in \mathcal{R}^{n\times p}$, $\bar{C} \in \mathcal{R}^{m\times n}$ and $Z \in \mathcal{R}^{m\times m}$ that satisfy the following LMIs:

$$\begin{bmatrix} A^T X + XA & \bar{A} - A^T W & XB + \bar{B} \\ * & -\bar{A} - \bar{A}^T & -WB - \bar{B} \\ * & * & -\delta I_p \end{bmatrix} < 0,$$

$$\begin{bmatrix} Z & C & \bar{C} \\ * & X & W \\ * & * & W \end{bmatrix} > 0 \quad \text{and} \quad trace\{Z\} < 1. \tag{8a-c}$$

The matrices of the model are then given by

$$A_m = -W^{-1}\bar{A}, \quad B_m = -W^{-1}\bar{B} \quad \text{and} \quad C_m = \bar{C}. \tag{9}$$

Proof: Denote

$$Q = \begin{bmatrix} X & M \\ M^T & U \end{bmatrix} \quad \text{and} \quad Q^{-1} = \begin{bmatrix} Y & N \\ N^T & V \end{bmatrix}$$

where $X,\ M,\ Y,\ N$ are $n \times n$ matrices. Multiplying (7a) by $diag\{J^T,\ I\}$ and $diag\{J,\ I\}$, from the left and the right, respectively, where:

$$J \triangleq \begin{bmatrix} I & Y \\ 0 & N^T \end{bmatrix}$$

and pre and post multiply (7b) by $diag\{I,\ \ J^T\}$ and $diag\{I,\ \ J\}$, respectively two new inequalities are obtained. Pre and post multiplying the first inequality by $diag\{\bar{T}^T,\ I\}$ and $diag\{\bar{T},\ I\}$, respectively and the second by $diag\{I,\ \bar{T}^T\}$ and $diag\{I,\ \bar{T}\}$, where $R = Y^{-1}$ and $\bar{T} = \begin{bmatrix} I & -I \\ 0 & R \end{bmatrix}$ two new inequalities are obtained. Denoting $W \triangleq X - R$, the LMIs in Th. 1 thus follow if we define:

$$\bar{A} = MA_mN^TR,\ \bar{B} = MB_m \text{ and } \bar{C} = C_mN^TR. \tag{10}$$

The relation between the latter matrices and the matrices of (3) is obtained by realizing that:

$$C_m(sI - A_m)^{-1}B_m = \bar{C}R^{-1}N^{-T}(sI - M^{-1}\bar{A}R^{-1}N^{-T})^{-1}M^{-1}\bar{B}$$
$$= \bar{C}(MN^TRs - \bar{A})^{-1}\bar{B} = -\bar{C}(Ws + \bar{A})^{-1}\bar{B}. \blacksquare$$

The above was obtained for (1) with no uncertainty in the parameters. In Section 3.2 we show how these results can be used in the case where a reduced-order model is sought. The affinity of the LMIs in (8) in A, B and C implies that the result of the theorem can be easily extended, also for full order models, to the case where the parameters of the system reside in the uncertainty polytope (2). We obtain the following.

Corollary 1: There exists a model (3) of order $k = n$ that achieves a cost J_1 of (4) less than a prescribed positive scalar δ for all the points in the polytope (2), if there exist matrices $X,\ W,\ \bar{A}\ \in \mathcal{R}^{n\times n}$, $\bar{B} \in \mathcal{R}^{n\times p}$, $\bar{C} \in \mathcal{R}^{m\times n}$ and $Z \in \mathcal{R}^{m\times m}$ that satisfy the following LMIs for $i = 1, ..., \bar{N}$.

$$\begin{bmatrix} A^{(i)T}X + XA^{(i)} & \bar{A} - A^{(i)T}W & XB^{(i)} + \bar{B} \\ * & -\bar{A} - \bar{A}^T & -WB^{(i)} - \bar{B} \\ * & * & -\delta I_p \end{bmatrix} < 0,$$
$$\begin{bmatrix} Z & C^{(i)} & \bar{C} \\ * & X & W \\ * & * & W \end{bmatrix} > 0, \quad trace\{Z\} < 1. \tag{11}$$

If a solution to the above LMIs exists the matrices of the sought model are given by (9).

The latter result applies the same Lyapunov function to all the points in Ω and it therefore entails a considerable overdesign. Recently, a parameter dependent approach has been introduced in [22] for the continuous-time case. The method of [22] also introduces a slack variable that is kept constant for all the vertices of the polytope and it leaves the decision variable Q to be chosen dependent on the vertices.

Applying the arguments used in [22] to (7a) we obtain the following.

Lemma 1: Consider the system (5) with $A_m,\ B_m$ and C_m given. The cost J_1 of (4) is less than a prescribed positive scalar δ over the polytope Ω if there exist matrices G and $Q_j \in \mathcal{R}^{2n\times 2n},\ j=1,2,...,N,$

$H \in \mathcal{R}^{(2n+p)\times(2n+p)}$ and $Z \in \mathcal{R}^{m\times m}$ that satisfy the following LMIs:

$$\begin{bmatrix} \Sigma^{(j)} & -diag\{Q_j - G^T,\, 0\} - \begin{bmatrix}\tilde{A}^{(j)T}\\ \tilde{B}^{(j)T}\end{bmatrix}\begin{bmatrix} I & 0\end{bmatrix}H \\ * & -H-H^T \end{bmatrix} < 0,$$

$$\begin{bmatrix} Z & \tilde{C}^{(j)} \\ * & Q_j \end{bmatrix} > 0,\ j=1,2,...,N,\ trace\{Z\} < 1 \quad \text{(12a-c)}$$

where $\tilde{A}^{(j)},\ \tilde{B}^{(j)}$ and $\tilde{C}^{(j)}$ are the corresponding matrices in (6) at the j-th vertex of Ω and

$$\Sigma^{(j)} = \begin{bmatrix} G^T \\ 0 \end{bmatrix}\begin{bmatrix}\tilde{A}^{(j)} & \tilde{B}^{(j)}\end{bmatrix} + \begin{bmatrix}\tilde{A}^{(j)T}\\ \tilde{B}^{(j)T}\end{bmatrix}\begin{bmatrix} G & 0\end{bmatrix} + \begin{bmatrix} 0 & 0 \\ 0 & -\delta I\end{bmatrix}.$$

Proof: If there exists Q_j that solves (7a) at the j-th vertex of the polytope, it is readily verified that the choice: $G = Q_j$ and $H = \bar{\sigma} I_{2n+p}$, $0 < \bar{\sigma} \to 0$ satisfies (12a) at this vertex point. On the other hand, if there exists a solution to (12a) at this vertex, we multiply the latter, from the left and the right, by Γ_j^T and Γ_j, where

$$\Gamma_j = \begin{bmatrix} I_{2n+p} & 0 \\ -\begin{bmatrix} I \\ 0 \end{bmatrix}\begin{bmatrix}\tilde{A}^{(j)} & \tilde{B}^{(j)}\end{bmatrix} & I_{2n+p} \end{bmatrix}.$$

The resulting inequality has then a solution only if (7a) possesses a solution Q_j.■

Applying the linearization transformation used in the proof of Theorem 1 on the inequalities of Lemma 1 we obtain the following.

Theorem 2: Consider the system (1) over the polytope Ω. There exists a model (3) that achieves a cost J_1 of (4) less than a prescribed positive scalar δ over the entire polytope if, for some positive scalar design parameters ε_1 and ε_2 there exist matrices $R,\ W,\ T,\ S,\ F_{11}^{(j)},\ F_{12}^{(j)}$ and $F_{22}^{(j)},\ j=1,2,...,N \in \mathcal{R}^{n\times n},\ S_B \in \mathcal{R}^{n\times p},\ S_C \in \mathcal{R}^{m\times n}$ and $Z \in \mathcal{R}^{m\times m}$ that satisfy the following $2N+1$ LMIs:

$$\begin{bmatrix} A^{(j)T}R+R^TA^{(j)} & * & * & * & * \\ W^TA^{(j)}+S & -S-S^T & * & * & * \\ B^{(j)T}R & B^{(j)T}W+S_B^T & -\delta I & * & * \\ \hat{F}_1^{(j)} & -F_{12}^{(j)}+W+T^T & -\epsilon_1 R^T B^{(j)} & -\epsilon_1(R+R^T) & * \\ \hat{F}_2^{(j)} & -F_{22}^{(j)}-T^T+\epsilon_2 S & -\epsilon_1 W^T B^{(j)}-\epsilon_2 S_B & -\epsilon_1 W^T-\epsilon_2 T & \epsilon_2(T+T^T) \end{bmatrix} < 0$$

$$\begin{bmatrix} Z & * & * \\ C^{(j)T}-S_c^T & F_{11}^{(j)} & * \\ S_c^T & F_{12}^{((j)T} & F_{22}^{(j)} \end{bmatrix} > 0,\ j=1,...,N,\ trace\{Z\}<1 \quad \text{(13b,c)}$$

where $\hat{F}_1^{(j)} = -F_{11}^{(j)} + R - \epsilon_1 R^T A^{(j)}$ and $\hat{F}_2^{(j)} = -F_{12}^{(j)T} - \epsilon_1 W^T A^{(j)} - \epsilon_2 S$.

If a solution to the above LMIs exists the matrices of the sought model are given by:

$$A_m = T^{-1}S, \quad B_m = T^{-1}S_B \quad \text{and} \quad C_m = S_C. \quad (14)$$

Proof: The proof follows the same lines as the one for Theorem 1. We choose $H = diag\{\epsilon G,\ \alpha I\}$ where $\epsilon = diag\{\epsilon_1 I,\ \epsilon_2 I\}$ and where ϵ_1, ϵ_2 and α are positive scalars. We partition G according to $\tilde{A}$ and have:

$$G = \begin{bmatrix} X & M \\ M_1 & U \end{bmatrix} \quad \text{and} \quad G^{-1} = \begin{bmatrix} Y & N \\ N_1 & V \end{bmatrix} \quad (15)$$

where due to the fact that in (12a) $0 < \epsilon G + G^T \epsilon$, G is nonsingular. Also we can assume, without loss of generality, that also N_1 and M_1 are nonsingular. Denoting

$$J = \begin{bmatrix} Y & I_n \\ N_1 & 0 \end{bmatrix} \quad \text{and} \quad \bar{J} = diag\{J,\ I_p,\ J,\ I_p\} \quad \text{(16a,b)}$$

we multiply (12a) by $\bar{J}^T$ and $\bar{J}$, on the left and on the right, respectively and substitute from (6) and (15). Choosing α that tends to zero we obtain:

$$\begin{bmatrix} A^{(j)}Y+Y^TA^{(j)T} & * & * & * & * \\ A^{(j)T}+X^TA^{(j)}Y+M_1^TA_mN_1 & X^TA^{(j)}+A^{(j)T}X & * & * & * \\ B^{(j)T} & B^{(j)T}X+B_m^TM_1 & -\delta I & * & * \\ -\bar{Q}_{11}+Y^T-\epsilon_1 A^{(j)}Y & -\bar{Q}_{12}-\epsilon_1 A^{(j)}+Y^TX+N_1^TM_1 & -\epsilon_1 B^{(j)} & -\epsilon_1(Y+Y^T) & * \\ -\bar{Q}_{12}^T+I-\epsilon_1 X^TA^{(j)}Y-\epsilon_2 M_1^TA_mN_1 & -\bar{Q}_{22}+X-\epsilon_1 X^TA^{(j)} & -\epsilon_1 X^TB^{(j)}-\epsilon_2 M_1^TB_m & -\epsilon_1 I-\epsilon_1 X^TY-\epsilon_2 M_1^TN_1 & -\epsilon_2(X+X^T) \end{bmatrix} < 0 \quad (17)$$

where we denote $J^TQ_jJ = \begin{bmatrix} \bar{Q}_{11} & \bar{Q}_{12} \\ \bar{Q}_{12}^T & \bar{Q}_{22} \end{bmatrix}$.

Denoting also $\tilde{J} = diag\{\begin{bmatrix} R & -R \\ 0 & I \end{bmatrix},\ I,\ \begin{bmatrix} R & -R \\ 0 & I \end{bmatrix}\}$, where $R = Y^{-1}$, we pre- and post-multiply (17) by $\tilde{J}^T$ and $\tilde{J}$, respectively and obtain (13a) where we denote

$$S = M_1^TA_mN_1R, \quad S_B = M_1^TB_m, \quad T = M_1^TN_1R$$

$$\text{and} \quad \begin{bmatrix} F_{11}^{(j)} & F_{12}^{(j)} \\ F_{12}^{(j)T} & F_{22}^{(j)} \end{bmatrix} = \begin{bmatrix} R^T & 0 \\ -R^T & I \end{bmatrix} J^TQ_jJ \begin{bmatrix} R & -R \\ 0 & I \end{bmatrix} \quad \text{(18a-d)}$$

and where $W = X - R$.

Considering next (12b), pre- and post-multiplying this inequality by $diag\{I,\ J\begin{bmatrix} R^T & 0 \\ -R^T & I \end{bmatrix}\}$ and $diag\{I,\ \begin{bmatrix} R & -R \\ 0 & I \end{bmatrix}\}$, respectively, readily leads to (13b), where

$$S_C = C_mN_1R. \quad (19)$$

If the LMIs of (8), (13a-c) possess a solution for all the vertices of Ω, the matrices of the model (3) may be derived by factorizing $NM_1 = -R^{-1}W$ and using $N_1 = M_1^{-T}TR^{-1}$.

Once M_1 and N_1 are calculated, it readily follows from (18a-c) and (19) that

$$A_m = M_1^{-T} S R^{-1} N_1^{-1},\ B_m = M_1^{-T} S_B,\ C_m = S_C R^{-1} N_1.$$

The transfer function G_m of the model is given by

$$G_m = C_m(sI - A_m)^{-1} B_m = S_c(sT - S)^{-1} S_B \quad (20)$$

and thus an alternative state space realization of the model (3) is obtained by (14).■

In Theorem 1 the matrices R and W are not necessarily symmetric. In the case where they are symmetric and $M^T = M_1$, it is easy to verify that $T = -W$.

B. The reduced-order model

The derivation in the last section was aimed at achieving a full-order model (k=n). When it is required that $k < n$ the above results no longer hold. A nonconvex optimization method has been suggested in [19] that leads to a reduced-order model solution in the case of systems without uncertainty. A much simpler solution to the reduced order model problem can be obtained, also for the uncertain case, by applying a modified version of Theorem 1.

It is obvious that if $A_m \in \mathcal{R}^{n\times n}$ and $C_m \in \mathcal{R}^{m\times n}$ in (3) had the following structure

$$A_m = \begin{bmatrix} A_{f1} & 0 \\ A_{f2} & A_{f3} \end{bmatrix} \quad \text{and} \quad C_m = \begin{bmatrix} C_{f1} & 0 \end{bmatrix} \quad (21a,b)$$

where $A_{f1} \in \mathcal{R}^{k\times k}$ and $C_{f1} \in \mathcal{R}^{m\times k}$ the state space model of (3) would be unobservable and the transfer function matrix of this model, from u to $\hat{y}$ would become $G_{reduced} = C_{f1}(sI_k - A_{f1})^{-1} B_m$. If all the eigenvalues of A_{f1} and A_{f3} reside in the left half of the complex plane then the model transference can be described by the observable k-th order triplet $\{A_{f1},\ B_m,\ C_{f1}\}$.

In order to obtain a solution to (8)-(9) of the structure in (21), all that is required is that the variable matrices $\bar{A}$ will posses the lower triangular structure of A_m in (21a), that the matrix variable $\bar{C}$ will be in the form of C_m in (21b) and that W will be block diagonal.The above arguments about the lower triangular structure lead to the following result.

Theorem 3: Consider the system (1) over the polytope Ω. There exists a model (3) of order $k < n$ which achieves a cost J_1 of (4) less than a prescribed positive scalar δ over the entire polytope if there exist matrices $X \in \mathcal{R}^{n\times n}$, $\bar{B} \in \mathcal{R}^{n\times p}$, $Z \in \mathcal{R}^{m\times m}$, $\bar{A} = \begin{bmatrix} S_1 & 0 \\ S_2 & S_3 \end{bmatrix}$, $W = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}$, $S_1,\ W_1 \in \mathcal{R}^{k\times k}$ and $S_3,\ W_2 \in \mathcal{R}^{(n-k)\times(n-k)}$ and $\bar{C} = \begin{bmatrix} C_1 & 0 \end{bmatrix}$, $C_1 \in \mathcal{R}^{m\times k}$ that satisfy (11) for all the $\bar{N}$ vertices of the uncertainty polytope.

If a solution to the latter LMIs exists, the matrices of the required reduced-order model are given by

$$A_m = -W_1^{-1} S_1,\ B_m = -\begin{bmatrix} W_1^{-1} & 0 \end{bmatrix} \bar{B} \text{ and } C_m = C_1. \quad (22)$$

The requirement imposed in Theorem 3 on W and $\bar{A}$ are conservative due to the fact that in order to achieve A_m in (9) with a lower block triangular structure it is not necessary for W and $\bar{A}$ to possess such structures. This requirement will therefore be relaxed below. The result of Theorem 3 depends on the state-space realization of the system (1) and a different minimum value of δ may be achieved if one applies the standard transformation:

$$A \to \bar{T} A \bar{T}^{-1}, \quad B \to \bar{T} B, \quad \text{and} \quad C \to C \bar{T}^{-1}$$

where $\bar{T}$ is a nonsingular matrix. An equivalent dependence on the matrix $\bar{T}$ is obtained if the original representation of (1) is used and instead of seeking the canonical structure of (21a,b), the following structure of the model matrices is sought.

$$A_m = \bar{T}^{-1} \begin{bmatrix} A_{f1} & 0 \\ A_{f2} & A_{f3} \end{bmatrix} \bar{T} \text{ and } C_m = \begin{bmatrix} C_{f1} & 0 \end{bmatrix} \bar{T} \quad (23a,b)$$

for some nonsingular matrix $\bar{T}$. In either ways, the inequalities (8a-c) become:

$$\begin{bmatrix} A^T\bar{X} + \bar{X}A & \bar{T}^T\bar{A} - A^T\bar{T}^T W & \bar{X}B + \bar{T}^T\bar{B} \\ * & -\bar{A} - \bar{A}^T & -W\bar{T}B - \bar{B} \\ * & * & -\delta I_p \end{bmatrix} < 0,$$
$$\begin{bmatrix} Z & C & \bar{C} \\ * & \bar{X} & \bar{T}^T W \\ * & * & W \end{bmatrix} > 0,\ trace\{Z\} < 1. \quad (24)$$

The latter are inequalities in the decision variables $\bar{A}$, $\bar{B}$, $\bar{C}$, $\bar{X}$, W, Z and $\bar{T}$ with the following special structure:

$$W = \begin{bmatrix} W_1 & 0 \\ 0 & W_2 \end{bmatrix}, \quad \bar{A} = \begin{bmatrix} S_1 & 0 \\ S_2 & S_3 \end{bmatrix}. \quad (25a,b)$$

These inequalities are clearly nonlinear but they become LMIs for a given $\bar{T}$. The conservative result of Theorem 3 was found for $\bar{T} = I$ and the question arises how to find the matrix $\bar{T}$ that will allow a solution of (24a-c), under (25), for the minimum value of δ. The following locally convergent iterative method is proposed which provides the minimum δ

Algorithm 1

- Start with any initial value for $\bar{T}$, say $\bar{T} = I$, $\bar{T} = A^{-1}$ or any scalar linear combination of the two.
- For the $\bar{T}$ obtained solve (24a-c), under (25), for the minimum value of δ.
- Use the matrices $\bar{A}$, $\bar{B}$ and W that were obtained in the previous step and solve (24a-c) and (25), in the decision variables $\bar{X}$, $\bar{C}$, Z and $\bar{T}$, for the minimum value of δ.
- Go to step 2 and solve the inequalities there. If the result obtained for δ is smaller than the previous value achieved in this step by less than a prescribed tolerance, stop. Otherwise continue to step 3

The latter algorithm is locally convergent since the sequence of the δ it produces is nonincreasing and is bounded from below. It is shown in Example 1 that it converges to the global minimum which is achieved in the case where the system matrices are perfectly known using the method of [10].

The above results addressed the case where the parameters of the system to be modeled by a reduced order system are perfectly known, in other words, standard model reduction. In the case where these parameters are only known to lie in the polytope Ω of (2), one can derive a corollary similar to the one obtained for Theorem 1 (Corollary 1), if a quadratic stable solution is sought, or else apply the method used to derive Theorem 3 on the results of Theorem 2. In the latter case we obtain the following.

Theorem 4: Consider the system (1) over the polytope Ω. There exists a model (3) of order $k < n$ which achieves a cost J_1 of (4) less than a prescribed positive scalar δ over the entire polytope if there exist matrices $R,\ W,\ F_{11}^{(j)},\ F_{12}^{(j)}$ and $F_{22}^{(j)}, j = 1, ..., N \in \mathcal{R}^{n\times n}$, $S_B \in \mathcal{R}^{n\times p},\ Z \in \mathcal{R}^{m\times m}$, $S = \begin{bmatrix} S_1 & 0 \\ S_2 & S_3 \end{bmatrix}$, $T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}$, $S_1,\ T_1 \in \mathcal{R}^{k\times k}$ and $S_3,\ T_2 \in \mathcal{R}^{(n-k)\times(n-k)}$ and $S_C = \begin{bmatrix} C_1 & 0 \end{bmatrix}$, $C_1 \in \mathcal{R}^{m\times k}$ that satisfy (13a-c) for all the $\bar{N}$ vertices of the uncertainty polytope.

If a solution to the latter LMIs exists, the matrices of the required reduced-order model are given by

$$A_m = T_1^{-1}S_1,\ B_m = \begin{bmatrix} T_1^{-1} & 0 \end{bmatrix} S_B \text{ and } C_m = C_1. \quad (26)$$

A corresponding realization dependent result is obtained by applying Theorem 2 and the state transformation (23a,b). The following is obtained.

Theorem 5: Consider the system (1) over the polytope Ω. There exists a model (3) of order $k < n$ which achieves a cost J_1 of (4) less than a prescribed positive scalar δ over the entire polytope if, for some positive scalar design parameters ε_1 and ε_2, there exist matrices $\bar{R},\ W,\ \bar{F}_{11}^{(j)},\ \bar{F}_{12}^{(j)}$ and $\bar{F}_{22}^{(j)}, j = 1, ..., N \in \mathcal{R}^{n\times n}$, $S_B \in \mathcal{R}^{n\times p},\ Z \in \mathcal{R}^{m\times m}$, $S = \begin{bmatrix} S_1 & 0 \\ S_2 & S_3 \end{bmatrix}$, $T = \begin{bmatrix} T_1 & 0 \\ 0 & T_2 \end{bmatrix}$, $S_1,\ T_1 \in \mathcal{R}^{k\times k}$ and $S_3,\ T_2 \in \mathcal{R}^{(n-k)\times(n-k)}$ and $S_C = \begin{bmatrix} C_1 & 0 \end{bmatrix}$, $C_1 \in \mathcal{R}^{m\times k}$ that satisfy, for a prechosen matrix $\bar{T}$, the following LMIs for all the $\bar{N}$ vertices of the uncertainty polytope:

$$\begin{bmatrix} A^{(j)T}\bar{R}+\bar{R}^TA^{(j)} & * & * & * & * \\ W^T\bar{T}A^{(j)}+S\bar{T} & -S-S^T & * & * & * \\ B^{(j)T}\bar{R} & B^{(j)T}\bar{T}^TW+S_B^T & -\delta I & * & * \\ -\bar{F}_{11}^{(j)}+\bar{R}-\epsilon_1\bar{R}^TA^{(j)} & -\bar{F}_{12}^{(j)}+\bar{T}^TW+\bar{T}^TT^T & -\epsilon_1\bar{R}^TB^{(j)} & -\epsilon_1(\bar{R}+\bar{R}^T) & * \\ -\bar{F}_{12}^{(j)T}-\epsilon_1W^T\bar{T}A^{(j)}-\epsilon_2S\bar{T} & -\bar{F}_{22}^{(j)}-T^T+\epsilon_2S & -\epsilon_1W^T\bar{T}B^{(j)}-\epsilon_2S_B & -\epsilon_1W^T\bar{T}-\epsilon_2T\bar{T} & \epsilon_2(T+T^T) \end{bmatrix} < 0$$

$$\begin{bmatrix} Z & * & * \\ C^{(j)T}-\bar{T}^TS_c^T & \bar{F}_{11}^{(j)} & * \\ S_c^T & \bar{F}_{12}^{((j)T} & \bar{F}_{22}^{(j)} \end{bmatrix} > 0,\ j=1,...,N,\ trace\{Z\}<1 \quad (27)$$

for $j = 1, 2, ..., N$.

If a solution to the latter LMIs exists, the matrices of the required reduced-order model are given by (26).

IV. EXAMPLES

4.1 Example 1: In order to examine the efficiency of Algorithm 1 we solve a very simple model reduction problem the solution to which can be compared with the result obtained in [10]. We consider the single-input-single-output plant described by the transfer function $T = 1/(s+10)^5$. A first order model is sought that best approximates the plant in the H_2-norm sense of Problem 1. Applying the following state-space model of the plant:

$$A=\begin{bmatrix} -1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 0 \\ 0 & 0 & -1 & 1 & 0 \\ 0 & 0 & 0 & -1 & 1 \\ 0 & 0 & 0 & 0 & -1 \end{bmatrix},\ B=\begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 1 \end{bmatrix},\ C=\begin{bmatrix} 1 & 0 & 0 & 0 & 0 \end{bmatrix}$$

the inequalities of (11) and (25) are first solved for $T = I_5$ using Matlab's LMI Toolbox [23]. A minimum value of $\delta = 0.1216$ is obtained in the first step of the algorithm. Starting with, say, $T = A^{-1}$ the first step provides a minimum value of $\delta = 0.0851$. Continuing with this initial value of T the algorithm converges and provides the value of $\delta = 0.0594$ (as compared to $\delta = 0.0592$ using the homotopy algorithm in citeHalevi1) with $A_m = -0.1161$, $B_m = 1.1811$ and $C_m = 0.1135$. For a convergence tolerance of 3% the algorithm arrived at a solution in just 9 iteration steps.

Example 2: Consider the system in Figure 1 where the inputs are the forces u_1 and u_2 and the outputs are the displacements y_1 and y_2. The nominal values of the parameters are $M_1 = M_2 = 1$, $K_1 = K_2 = 1$, $C_1 = 0.5$ and $C_2 = 1.5$ and the state space realization of the system is the following:

$$\dot{x}=\begin{bmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ -1 & 1 & -.5 & .5 \\ 1 & -2 & .5 & -2 \end{bmatrix}x+\begin{bmatrix} 0 & 0 \\ 0 & 0 \\ 1 & 0 \\ 0 & 1 \end{bmatrix}u,\ y=\begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \end{bmatrix}x$$

Choosing the reduced model order to be one (found using the homotopy method in [10]), the optimal H_2 model is given by:

$$\dot{x}_m=-.2950x_m+\begin{bmatrix} .8676 & .4795 \end{bmatrix}u,\ y_m=\begin{bmatrix} .8676 \\ .4795 \end{bmatrix}x_m \quad (28)$$

with a cost of $\delta_{opt} = 1.4517$, compared to the Truncated Balanced Realization method which yields in this case $\delta_{bal} = 1.8200$. Algorithm 1 which converged in 5 iteration steps lead to $\delta_1 = 1.4971$, i.e. a 3% deviation from the optimal cost and a 18% improvement on the Balanced Realization. The resulting model is described by:

$$A_m=-.3066,\ B_m=\begin{bmatrix} 2.2603 & .7881 \end{bmatrix},\ C_m=\begin{bmatrix} .3675 & .1877 \end{bmatrix}^T.$$

Next consider the robust order reduction problem where there is an uncertainty in the parameters K_1 and K_2 and each of them can assume any value in the intervals: $K_1 \in [0.9,\ 1.1]$, $K_2 \in [0.8,\ 1.2]$. Applying Corollary 1 with the structure of (21), a reduced model of order 1 is obtained which is based on a single Lyapunov function. A minimum bound of $\delta_{bound} = 2.35$ is obtained in 10 iteration steps. The resulting model is given by:

$$A_m=-.3413,\ B_m=\begin{bmatrix} 2.7294 & 1.2041 \end{bmatrix},\ C_m=\begin{bmatrix} .3241 & .1783 \end{bmatrix}^T.$$

Using Theorem 4, one obtains the following model

$$\dot{x}_m=-.2920x_m+\begin{bmatrix} 2.2094 & 1.0947 \end{bmatrix}u,\ y_m=\begin{bmatrix} .3547 \\ .2171 \end{bmatrix}x_m$$

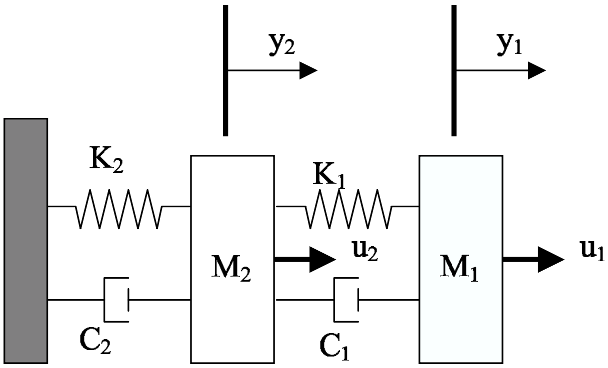

Fig. 1. The spring and mass system of Example 2

with a guaranteed cost of $\delta < 1.7673$ everywhere inside the polytope. This guaranteed cost is significantly lower than the one obtained above using Corollary 1. The actual values of δ that are achieved by applying the latter model to the four vertices of the uncertainty polytope are: 1.5932, 1.6093, 1.3654 and 1.4025.

V. Conclusion

An efficient method for robust, as well as nominal, model reduction has been introduced. It differs from previous results in both the uncertainty representation (polytopic) and the method of solution (LMI). This method is based on solving LMIs that correspond to the various vertices of the uncertainty polytope. Unlike previous results on applying LMIs to model reduction, the method does not involve rank conditions and thus can be solved by the standard algorithm. Its result depends on the state space realization of the system to be modeled. It can be significantly improved by performing an iterative search for the best realization. This search can be performed on the initial realization of the system to which a noniterative procedure is applied in order to find the reduced-order model that best approximates this specific realization, or else by applying the iterations of Algorithm 1 that iteratively finds realizations that reduce the modeling error.

The proposed method guarantees a locally minimum upper-bound on the H_2 norm of the modeling error over the entire uncertainty polytope. This bound is not necessarily tight and in many cases the maximum norm that is achieved over the entire polytope by applying the resulting reduced order model is less than the bound for which the model has been designed.

References

[1] B. C. Moore, Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction, *IEEE Trans. on Automat. Contr.*, vol. 26, pp. 17- 32, 1981.

[2] D. Enns, Model Reduction for Control Systems Design, *Ph.D. Thesis, Dept. of Aeronautics and Astronautics, Stanford Univ., CA. U.S.A*, 1984.

[3] R.E. Skelton and A. Yousuff, Component Cost Analysis of Large Scale Systems, *Int. J. of Control*, vol. 37, pp. 285-304, 1983.

[4] K. Glover, All Optimal Hankel - Norm Approximations of Linear Multivariable Systems and their L_2 error bounds, *Int. J. Control*, vol. 39, pp. 1115-1193, 1984.

[5] D. A. Wilson, Optimum Solution for Model Reduction Problem, *Proc. of IEE*, vol. 117, pp. 1161-1165, 1970.

[6] D. C. Hyland and D.S. Bernstein, The Optimal Projection Equations and the Relationships Among the Methods of Wilson, Skelton and Moore, *IEEE Trans. on Automat. Contr.*, vol. 30, pp. 1201-1211, 1985.

[7] W. M. Haddad, and D.S. Bernstein, Combined L_2/H_∞ model reduction, *Int. J. of Control*, vol. 49, pp. 1523-1535, 1989.

[8] Y. Halevi, Frequency Weighted Model Reduction via Optimal Projection Equations, *IEEE Trans. on Automat. Contr.*, vol. 37, pp. 1537-1542, 1992.

[9] D. Zigic, L.T. Watson, E.G. Collins, and D.S.Bernsein, Homotopy Methods for Solving the Optimal Projection Equations for the H_2 Reduced Order Model Problem, *Int. J. Control*, vol. 56, pp. 173-191, 1992.

[10] Y. Halevi, A. Zlochevsky and T. Gilat , Parameter-Dependent Model Order Reduction, *Int. J. of Control*, vol. 66, pp. 463-485, 1997.

[11] J.T. Spanos,M.H. Milman and D.L. Mingori, A New Algorithm for L_2 Model Reduction, *Automatica*, vol. 28(5), pp. 89-909,1992.

[12] M. Diab, W.Q. Liu and V. Sreeram, A New Approach for Frequency Weighted L_2 Model Reduction of Discrete-Time Systems, *Optimal Control Applications and Methods*, vol. 19(3), pp. 147-167, 1998.

[13] W. Y. Yan and J. Lam, An Approximate Approach to H_2 Optimal Model Reduction, *IEEE Trans. on Automat. Contr.*, vol. 44, pp. 17-32, 1999.

[14] K. M Grigoriadis, Optimal Model Reduction via Linear Matrix Inequalities: Continuous and Discrete-Time Cases, *System and Control Letters*, vol. 26, pp. 321-333, 1994.

[15] K. M Grigoriadis, L_2 and $L_2 - L_\infty$ Model Reduction via Linear Matrix Inequalities, *Int. J. of Control*, vol. 68, No. 3, pp. 485-498, 1997.

[16] C. L. Beck, J. Doyle, and K. Glover, Model Reduction of Multi-Dimensional and Uncertain Systems, *IEEE Trans. on Automat. Contr.*, Vol. 41, pp. 1466-1477, 1996.

[17] W. M. Haddad and D.S. Bernstein, Robust, Reduced-Order Modeling via the Optimal Projection Equations with Petersen-Hollot Bounds ", *IEEE Trans. on Automat. Contr.*, vol. 33, pp.591-595, 1988.

[18] H. Li and M. Fu, A Linear Matrix Inequality Approach to Robust H_∞ Filtering, *IEEE Trans. on Signal Processing*, vol. 45, pp. 2338-2350, 1997.

[19] H. D. Tuan, P. Apkarian and T. Q. Nguyen, Robust and Reduced-Order Filtering: New Characterizations and Methods," *in Proc. of the American Contr. Conf.*, Chicago, IL, 2000.

[20] C. E. de Souza and a. Trofino, An LMI Approach to the Design of Robust H_2 Filters,*in Recent Advances on Linear Matrix Inequality Methods in Control*, L. El Ghaoui and S. Niculescu (Eds.), SIAM, 1999.

[21] S. Boyd, L. El Ghaoui, E. Feron, and V. Balakrishnan, *Linear Matrix Inequality in Systems and Control Theory,* SIAM Frontier Series, April 1994.

[22] D. Peaucelle, D. Arzelier, O. Bachelier and J. Bernussou, A New Robust D-Stability Condition for Real Convex Polytopic Uncertainty, *Systems and Control Letters*, vol. 40, pp. 21-30, 2000.

[23] P. Gahinet, A. Nemirovski, A. J. Laub and M. Chilali, *LMI Control toolbox for use with MATLAB*, The Mathworks Inc., 1995.