Extensions to an Optimization-Based Multivariable Model Reduction Method

A. Sootla, A. Rantzer

Abstract— **In this paper two extensions to an optimization based model reduction method are considered: frequency-weighted and parameter-dependent model reduction. The performed extensions are simple and natural, this fact illustrates the advantages of the method. Also, the methods may be applied using an exact model or frequency samples of a model. We construct numerical examples to illustrate both extensions.**

I. INTRODUCTION

The first extension to the optimization-based reduction framework we consider is frequency-weighted model reduction. For process models a good approximation is required only along some trajectories, i.e. around some frequencies. Controller or plant reduction can be done in a robust way with frequency-weighting (see, [1]). In the unweighted case there exist well-developed and recognized methods, like balanced truncation and optimal Hankel model reduction (see,[1]). For the frequency-weighted extensions (see, [2], [3]), however, there exists many open issues. In [4] a new state-space method was proposed that may turn out to be very useful for reduction of structured system. However, to the authors best knowledge there exist no a priori guarantee that the methods would provide a stable reduced model.

Instead of state-space form of a model one can use frequency domain information. The original problem formulation in the H_∞ space is not convex. In [5] a relaxation that makes the problem convex in the H_∞ space was proposed. One can use this method given an exact model, or a finite number of frequency samples of a model. In the second case the computational cost for large-scale systems is sufficiently lower than for state-space methods. With a sufficient number of points we can get as close as required to the solution of the optimization procedure, that uses the exact model. Error bounds for this method were obtained in [6]. In [7] a multivariable (MIMO) version of the method was proposed. We sketch main ideas of the method and the MIMO extension in section II.

In [8] a frequency-weighted extension was proposed. It consists of the method for single-input-single-output (SISO) models and an algorithm for MIMO systems that uses rank-minimization heuristics from [9]. But one could wish to obtain the algorithm with a higher degree of freedom and a relaxation gap guarantee as in the original method. We present our version of frequency-weighted extension in section III.

The second framework extension is parameter-dependent model reduction. First we have to define a notion of a simplified parameter-dependent system. We can reduce the number of parameters, the McMillan degree of the system or simplify the way how the parameters are introduced into the dynamics. We will focus on the last two problems. In this paper we only define a direction for future work. The extension is introduced in section IV to show, how convenient the optimization-based reduction framework is.

All the realization details are assembled in Appendix, in order not to complicate first sections with too much details. We talk about implementation of the algorithms and examples in Section V.

A. Sootla and A.Rantzer are with Automatic Control, LTH, Lund University, Lund, Sweden {aivar,rantzer}@control.lth.se

Notation

We will use H_∞ and H_∞^- to denote spaces of discrete-time stable and anti-stable $m \times m$ transfer function matrices, where $m \geq 1$. Operation $^\sim$ denotes an adjoint in H_∞ space: $G^\sim(z) = G^T(1/z)$, and $h^\nabla(z) = z^{-N} h^\sim(z)$, if $h(z) = \sum_{k=0}^{N} h_k z^{-k}$. G_ω is a frequency response $G(e^{\jmath\omega})$ to $\omega \in [0, \pi]$. We will use the notation $\sigma_1(G) \geq \cdots \geq \sigma_N(G)$, only when we are referring to Hankel singular values. $\overline{\sigma}(G_\omega)$ denotes the maximal singular value of the matrix G_ω. The infinity norm is computed $\|G\|_\infty = \sup_w \overline{\sigma}(G_\omega)$. The Hankel norm of a transfer function is denoted $\|\cdot\|_H$ (see for example [10]). $\deg(G)$ will denote the McMillan degree of the transfer function G.

II. PROBLEM FORMULATION AND PRELIMINARIES

First we introduce the general framework that we use for different reduction problems. It turns out, that we need to modify the method slightly to obtain frequency weighted or parameter dependent versions.

Assume, we want to solve the following optimization problem:

$$\min_{\hat{G}} \|G - \hat{G}\|_\infty, \tag{1}$$

where $\deg(G) > \deg(\hat{G})$. This problem is non-convex and the relaxation is introduced to obtain a convex one.

A. Sketch of the method.

Assume the transfer function G is scalar and we look for the reduced model in the form $\hat{G} = pq^{-1}$, where $p = \sum_{i=0}^{k} p_i z^{-i}$ and $q = \sum_{i=0}^{k} q_i z^{-i}$ are polynomials in z^{-1} of degree less or equal to k with q a Schur polynomial (No zeros outside the unit circle $\forall z : |z| \geq 1$).

We can rewrite the problem (1) as:

$$\min \gamma \quad \text{subject to} \quad |G_\omega - p_\omega q_\omega^{-1}| < \gamma \quad \forall \omega$$

ISBN 978-963-311-369-1

and after multiplying both sides with $|q|^2$ we get:

$$\min \gamma \quad \text{subject to} \quad |G_\omega q_\omega q_\omega^\sim - p_\omega q_\omega^\sim| < |q_\omega|^2\gamma \quad \forall\omega$$

This problem is not convex in variables p, q, but if we denote $a = |q|^2$ and perform the relaxation of the structure $b \leftrightarrow pq^\sim$, we can get one. a, b are parametrized as $a = \sum_{i=0}^{k} a_i(z^i + z^{-i}),\quad b = \sum_{i=0}^{k} \bar{b}_i(z^i + z^i) + \sum_{i=0}^{k-1} \bar{c}_i(z^i - z^{-i})$. So now our problem can be written as:

$$\min_{a,b} \gamma \quad \text{subject to} \quad |G_\omega a_\omega - b_\omega| < \gamma a_\omega, \quad \forall\omega$$

Using Schur complement we obtain a convenient LMI formulation of the problem:

$$\min_{a,b} \gamma \quad \text{subject to} \quad \begin{pmatrix} \gamma a_\omega & G_\omega a_\omega - b_\omega \\ * & \gamma a_\omega \end{pmatrix} > 0, \quad \forall\omega \quad (2)$$

where asterisk denotes Hermitian transpose of the upper right corner. We obtain only the denominator q from $a = qq^\sim$ (a is obtained form (2)). The numerator p of $\hat{G}$ will be obtained from:

$$\min_{p} \gamma \quad \text{subject to} \quad \|G - pq^{-1}\|_\infty < \gamma$$

In multivariable case, $\hat{G} = PQ^{-1}$, where $P = \sum_{i=0}^{k} P_i z^{-i}$ and $Q = \sum_{i=0}^{k} Q_i z^{-i}$ are $m \times m$ matrix polynomials in z^{-1}, we introduce polynomials $A = \sum_{i=0}^{k} A_i(z^i + z^{-i}),\quad B = \sum_{i=0}^{k} \bar{B}_i(z^i + z^{-i}) + \sum_{i=0}^{k-1} \bar{C}_i(z^i - z^{-i})$. We will solve the problem:

$$\min_{A,B,f} \gamma \text{ subject to} \quad (3)$$

$$\begin{pmatrix} \gamma f_\omega I & G_\omega A_\omega - B_\omega \\ * & \gamma A_\omega \end{pmatrix} > 0 \quad (4)$$

$$0 < f_\omega I \le A_\omega, \quad \forall\omega \quad (5)$$

This generalization of the SISO method is valid, since:

Lemma 2.1: If γ, A, B, f satisfy (4,5) for some G, then $\|G - BA^{-1}\|_\infty \le \gamma$.

The proof can be found in Appendix B. After solving (3-5), we calculate Q from the spectral factorization problem $A = QQ^\sim$. The solution will always exist and $\deg(\hat{G}) \le \deg(Q) = km$. The numerator P of $\hat{G}$ will be obtained from:

$$\min_{P} \gamma \quad \text{subject to} \quad \|G - PQ^{-1}\|_\infty < \gamma$$

Details can be found in Appendix A.

B. Finite dimensional problem.

The problem (3-5) has an infinite number of constraints, since the constraints (4-5) are imposed for all ω. There are two ways of obtaining the finite dimensional program. We can impose all the constraints at all the frequencies in one LMI, using the formulation of the KYP lemma described in [11]. To do that we will have to assume that f is a polynomial in z^{-1} of the same degree as A. The alternative is enforcing the constraints in a finite number of points.

Assume, we have frequency samples $G_{\omega_i} = G(e^{\jmath\omega_i})$ for $\{\omega_i\}_{i=1}^{N}$. Then we will solve the following problem:

$$\min_{A,B} \gamma \text{ subject to} \quad (6)$$

$$\begin{pmatrix} \gamma f_{\omega_i} I & G_{\omega_i} A_{\omega_i} - B_{\omega_i} \\ * & \gamma A_{\omega_i} \end{pmatrix} > 0 \quad (7)$$

$$0 < f_{\omega_i} I \le A_{\omega_i} \quad \forall\{\omega_i\}_{i=1}^{N} \quad (8)$$

It was shown in [5] that the solution of (6-8) converges to the solution of (3-5) if $N \to \infty$ and $\{\omega_i\}_{i=1}^{\infty}$ is dense in $[0, \pi]$. Numerical experiments show, that $N = O((km)^2)$ provides a good approximation of the solution to the non-sampled problem.

This technique is applicable to all the extensions of the method including the frequency weighted and the parameter dependent ones.

III. Frequency Weighted Model Reduction

A frequency weighted problem can be formulated as an optimization problem:

$$\min_{\hat{G}} \|W^o(G - \hat{G})W^i\|_\infty,$$

where G is a m-input m-output system, $\hat{G} = PQ^{-1}$ and k is the order of the polynomials P, Q. Weights W^o, W^i are considered stable with stable inverses. It is a common assumption in the literature (see, for example [1]). We follow the same pattern as in the non-weighted case and we obtain the convex program:

$$\min_{A,B} \gamma \text{ subject to} \quad (9)$$

$$\begin{pmatrix} \gamma f_\omega I & W_\omega^o(G_\omega A_\omega - B_\omega)W_\omega^i \\ * & \gamma (W_\omega^i)^\sim A_\omega W_\omega^i \end{pmatrix} > 0 \quad (10)$$

$$0 < W_\omega^i f_\omega (W_\omega^i)^\sim \le A_\omega \quad (11)$$

Here a similar to Lemma 2.1 statement can be formulated with a similar proof as well:

Lemma 3.1: If γ, A, B, f satisfy (10,11), then $\|W^o(G - BA^{-1})W^i\|_\infty \le \gamma$.

After solving (9-11), we get Q from the spectral factorization problem $A = QQ^\sim$. With a suitable parametrization of A polynomial the solution will always exist and $\deg(\hat{G}) \le \deg(Q) = km$. The numerator P of $\hat{G}$ is obtained from:

$$\min_{P} \gamma \quad \text{subject to} \quad \|W^o\left(G - PQ^{-1}\right)W^i\|_\infty < \gamma \quad (12)$$

We know that $\|W^o(G - BA^{-1})W^i\|_\infty < \gamma$, but in fact we would like to know how close is the reduced system itself to the original one.

Theorem 3.2: Consider a stable/anti-stable decomposition $BA^{-1} = PQ^{-1} + RQ_2^{-\nabla}$, where $Q_2^\sim Q_2 = A$, and R is a matrix polynomial in variable z^{-1} of degree $k-1$. Assume we have obtained $\hat{G} = PQ^{-1}, \gamma$ from (9-11), and (12), then the following statements are true:

1)
$$\|W^o(G-\hat{G})W^i\|_\infty \le \gamma\left(1+km\cdot C(W^o,W^i)\right)$$
where $C(W^o,W^i)=\frac{\sup \underline{\sigma}(W^i)\underline{\sigma}(W^o)}{\inf \underline{\sigma}(W^o)\underline{\sigma}(W^i)}$

2) Define $W^o\hat{G}_2W^i = W^oRQ_2^{-\nabla}W^i - R_2Q_2^{-\nabla}$ as the result of stable/anti-stable decomposition of $W^oRQ_2^{-\nabla}W^i$, then:
$$\|W^o(G-\hat{G}-\hat{G}_2)W^i\|_\infty \le \gamma\left(1+km\right)$$

3) $\inf_{\hat{G}} \|W^o(G-\hat{G})W^i\|_\infty \ge \gamma \ge \sigma_{km+1}(G)$, where $\sigma_{km+1}(G)$ is a $km+1$-th Hankel singular value of G.

Proof: See, Appendix C ■

In the second part of the theorem we get the error bound, which does not depended on the weights. However, the reduced system now has the McMillan degree $km+d$, where d is the McMillan degree of $\hat{G}_2$. The results should be considered as a theoretical justification for the method. The derived error bounds are rather conservative, since we use Nehari's theorem and, for example, in the first part of the theorem we use the sub-multiplicative property of the matrix norm.

IV. Reduction of Multivariable Linear Parameter Dependent Systems

We will focus on reduction of the McMillan degree of the parameter-dependent system and simplification of the parameter dependence. In our setup we do not have to decouple these problems, and we can reduce both the order, and the parameter dependence at the same time. The problem can be formulated as:
$$\min_{\hat{G}} \max_{\theta} \|G(\omega,\theta)-\hat{G}(\omega,\theta)\|_\infty$$
where $\theta \in \Theta$, the set Θ is compact in $\in \mathbb{R}^n, n \ge 1 \in \mathbb{N}$. We assume, that $\hat{G}(\omega,\theta) = P(\omega,\theta)Q^{-1}(\omega)$, where the parameter dependence introduced only in the numerator and $P(\omega,\theta)=\sum_{i=0}^k P_i(\theta)z^{-i}, \quad Q(\omega)=\sum_{i=0}^k Q_iz^{-i}$.

We follow the same pattern as earlier and we obtain the convex program:
$$\min_{A,B} \gamma \text{ subject to} \tag{13}$$
$$\begin{pmatrix} \gamma f_\omega I & G_{\omega,\theta}A_\omega - B_{\omega,\theta} \\ * & \gamma A_\omega \end{pmatrix} > 0 \tag{14}$$
$$0 < f_\omega I \le A_\omega \quad \forall \omega \in [0,\pi], \quad \theta \in \Theta \tag{15}$$
We get Q from the spectral factorization problem $A = QQ^\sim$ after solving (13-15). The numerator $P(\omega,\theta)$ of $\hat{G}(\omega,\theta)$ is calculated from:
$$\min_{P(\omega,\theta)} \gamma \quad \text{subject to} \quad \|G(\omega,\theta)-P(\omega,\theta)Q^{-1}\|_\infty < \gamma \tag{16}$$
After solving the problem on the grid on $\Theta\times[0,\pi]$, to obtain the coefficients $P_i(\theta)$ we must perform an interpolation procedure.

Theorem 4.1: Assume that we have obtained P, Q, γ from solving the program (13-15), and (16). Then the following statements are true:

1) $\|G_\theta - P_\theta Q^{-1}\|_\infty \le \gamma(1+km) \quad \forall \theta \in \Theta$
2) $\inf_{P,Q} \|G_\theta - P_\theta Q^{-1}\|_\infty \ge \gamma$.

The proof is omitted due to resemblance to one of Theorem 3.2.

To simplify the parameter dependence we need only to change the parametrization of P. For example, if we want to get a linear parameter dependence in the frequency domain, then define $P=\sum_{j=0}^k P_i^0 z^{-i} + \theta\sum_{j=0}^k P_i^1 z^{-i}$, where θ is a scalar parameter.

V. Implementation and Examples

The algorithm was implemented using LMI solver SeDuMi [12] for YALMIP [13]. In all the examples we perform bi-linear transformation in order to obtain discrete-time models. We compare the effectiveness of the method for continuous time models, using again bi-linear transformation to map the models back to continuous time. We use the uniform grid in all the examples. In examples 1 and 2 we use non-minimum phase strictly proper weights.

A. Example 1.

This example is based on the examples in [2]. We assume
$$W^o = \begin{pmatrix} W_{0.01} & 0 \\ 0 & W_{0.1} \end{pmatrix} G = \begin{pmatrix} \frac{1}{a_1(s)} & 0 \\ 0 & \frac{b_2(s)}{a_2(s)} \end{pmatrix}$$
where $a_1(s) = s^6 + 7.4641s^5 + 3.8637s^4 + 9.1416s^3 + 7.4641s^2 + 3.8637s + 1$, $a_2(s) = (s^2+0.2s+4.04)(s^2+0.2s+16.02)$, $b_2(s) = (s^2+0.2s+1.01)(s^2+0.2s+9.01)$, $W_\alpha = \frac{(s-1)^2}{s^2-2\alpha s+1}$and the input weight is identity. Enns' method is frequency weighted balanced truncation (see, [3]), LA is frequency-weighted Hankel model reduction algorithm by Latham and Anderson (see, [14]), AI and AII are the modifications of LA algorithm (see, [2]). In AI, after obtaining reduced models we perform a minimization over constant terms D of G_r. In AII, we obtain the denominator by Hankel model reduction, and then we optimize over the numerator. Lower bounds for the methods are the $l+r+1$-th Hankel singular values of the stable part of $W(s)G(s)$, where l is the order of the stable part of $W(s)$, and r is the order of the reduced function. We use $N=100$ points in the grid and $\gamma=5$ for all the experiments. We see that entry-wise for reduction of the low-scale model the proposed method acts

TABLE I
Approximation errors for entry-wise reduction.

Entry	$W_{0.1}b_2/a_2$		$W_{0.01}/a_1$
Reduction order	2	3	4
Lower bounds	2.7037	2.5261	0.0248
Enns	5.128	4.993	0.0584
LA	20.08	11.94	0.0860
AI	4.827	8.20	0.0448
AII	4.822	3.946	0.0256
Proposed Method	4.6686	3.8409	0.0253

all most as the AII algorithm, which is also optimization-based, but can not be used for the large-scale systems due to use of the Lyapunov equations.

Now we will reduce the multivariable model. For this case due to increase of decision variables we use the grid with 150 points. We reduce G to orders $4, 6$ and 8, what corresponds to reducing both entries to orders $2, 3$ and 4. In the first two cases the approximation error will be the error of reducing b_2/a_2 in the third – $1/a_1$. The approximation errors for orders $4, 6, 8$ are $4.3916, 3.8091, 0.0267$ correspondingly.

B. Example 2. Controller reduction.

This example is considered in [15]. The plant P is a NASA HiMAT aircraft model, which is a 2-input-2-output system with 6 states with 2 unstable ones. The controller K is designed in [16], is a 2-input-2-output model with 16 states. It was reduced to a 9 state model using balanced truncation without considering the weights, i.e. the plant. We will use our frequency-weighted setup to reduce the controller. The usual way to apply frequency-weighted reduction to controller simplification problem is solving (see, [1]):

$$\min_{\hat{K}} \|W^o(K - \hat{K})\|_\infty,$$

where $W^o = (I + PK)^{-1}P$. We can not apply directly our method, since there are integrator states in the controller K, so we decouple our controller first to $K = K_i + K_c$, where K_i is the integrator and K_c is the asymptotically stable part of the controller. Now we can apply our procedure to the obtained problem:

$$\min_{\hat{K}_c} \|W^o(K_c - \hat{K}_c)\|_\infty$$

We reduce K_c to 6 states, so $\hat{K} = K_i + \hat{K}_c$ has 8 states, using $N = 100$ points in the grid and tolerance of the bisection procedure is set to 10^{-3}. We also reduce the system using frequency-weighted Hankel model reduction for the minimum phase weight $\tilde{W}^o$. Controller reduction by means of balanced truncation provided unstable step responses. See, figure 1 for results.

C. Example 3. Different reduction problems for mass-spring-damper system.

The case study is mass-spring-damper system.

$$A = \begin{pmatrix} 0 & I_n \\ -M^{-1}K & -M^{-1}D \end{pmatrix} \quad B = \begin{pmatrix} 0 \\ k_m \cdot M^{-1}I_1 \end{pmatrix}$$

$$C = \begin{pmatrix} 0 & k_y \cdot M^{-1}I_1^T \end{pmatrix} \quad I_1 = \begin{pmatrix} 0 & \dots & 0 & 0 & 1 \\ 0 & \dots & 0 & 1 & 0 \end{pmatrix}$$

$$I_2 = \begin{pmatrix} 2 & -1 & & \\ -1 & \ddots & \ddots & \\ & \ddots & \ddots & -1 \\ & & -1 & 2 \end{pmatrix} \quad M = \begin{pmatrix} m_1 & & \\ & \ddots & \\ & & m_n \end{pmatrix}$$

And $K = kI_2, \quad D = dI_2$. The choice of parameters is guided by an experimental setup at LTH used for education purposes.

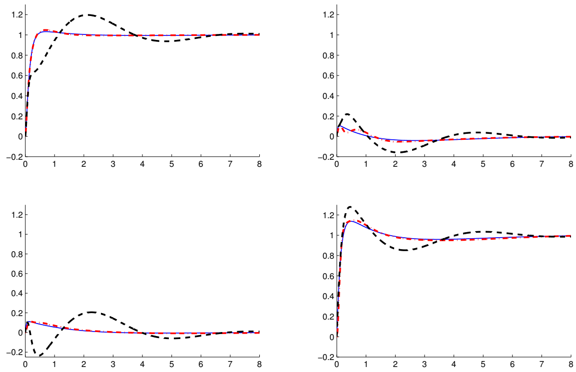

Fig. 1. Entry-wise step responses for the loop with original controller (solid), the loop with controller reduced by means of proposed method (dash-dotted), the loop with controller reduced by means of Hankel model reduction (dashed).

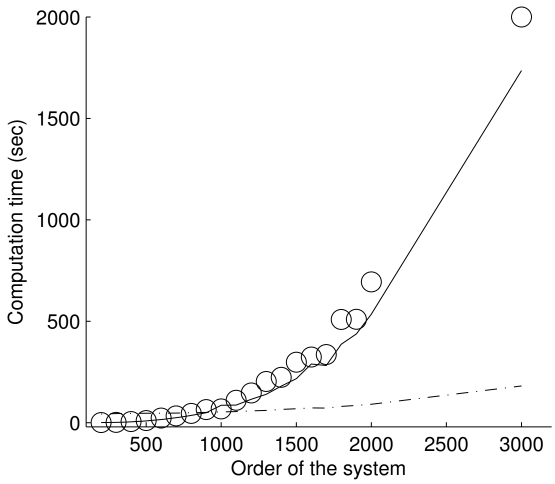

Fig. 2. Computation time (in seconds) depending on the number of states in the model. Balanced truncation (solid line), Hankel model reduction (circles) and the proposed algorithm (dashed line).

1) Weighted reduction.: The main advantage of our approach is that we use less information about the original system than Hankel model reduction and balanced truncation. There is always a computational restriction on solving high-dimensional Lyapunov equations. In this example we study the computational time of different reduction procedures. We take $k = 400, \quad d = 300, \quad m_i = 2/n \quad \forall i, \quad k_y = 3, \quad k_m = 280$, where n is the number of masses in the system. We consider input and output weights:

$$W_i = W_o = \begin{pmatrix} \frac{s+1}{s/0.001+1} & 0 \\ 0 & \frac{s/0.1+1}{s/0.001+1} \end{pmatrix}$$

Our goal is to obtain the reduced models of order 20 and compare the results for different methods. In time measuring experiments (see, figure (2)) we set tolerance of the bisection procedure to 10^{-3}, considering that $\|G\|_\infty = O(10^7)$ it seems reasonable. The starting value of $\gamma = 0.4$. The number of points in the grid is 100. With different values of γ and the tolerance, one can get different results. The computation time for state-space methods will rise as $O(n^3)$ if Lyapunov equations are solved by Hammarling method, whether for optimization-based method the cost will grow as $0(n^2)$, due to calculation of frequency samples. In the table (II) we provide approximation errors for some of the reduced systems. As the reader may notice the approximation error for balanced truncation for the system with 1200 states is much bigger than with the other methods. Since there is no stability guaranty for the reduced model (see, counter-example in [1]) it is not unusual.

TABLE II

APPROXIMATION ERRORS FOR THE ALGORITHMS.

Number of states	Balanced truncation	Hankel model reduction	Proposed method
700	$1.20 \cdot 10^{-2}$	$7.95 \cdot 10^{-2}$	$9.00 \cdot 10^{-3}$
800	$3.91 \cdot 10^{-2}$	$3.94 \cdot 10^{-2}$	$1.01 \cdot 10^{-2}$
900	$4.66 \cdot 10^{-2}$	$6.54 \cdot 10^{-2}$	$9.20 \cdot 10^{-3}$
1000	$2.50 \cdot 10^{-2}$	$1.93 \cdot 10^{-2}$	$1.30 \cdot 10^{-2}$
1200	1.48	$2.65 \cdot 10^{-1}$	$7.60 \cdot 10^{-3}$

2) Reduction of parameter depended system.: Again we use mass-spring-damper system G with $N = 6$ masses, $m_i = 0.3 \quad i = 1, \ldots, 5, 7, 8 \quad k = 1.7, \quad d = 0.9 \quad k_y = k_m = 1$ Mass m_6 is a parameter varying as: $\theta \in [0.1, 0.5]$. The parameter grid consists only of three points $\theta = \{0.1, 0.3, 0.5\}$ for simplicity. The frequency grid is uniform and consists of 60 points. The reduced parametrized model would be of order 6 and the parameter dependence is:

$$\hat{G}(\omega, \theta) = (P_0 + P_1\theta + P_2\theta^{-1})Q^{-1}.$$

Obtained error value for the parameter-depended system is $\gamma = 1.19$. A relatively big error occurs due to the fixed A polynomial for different values of parameters (see, fig 4).

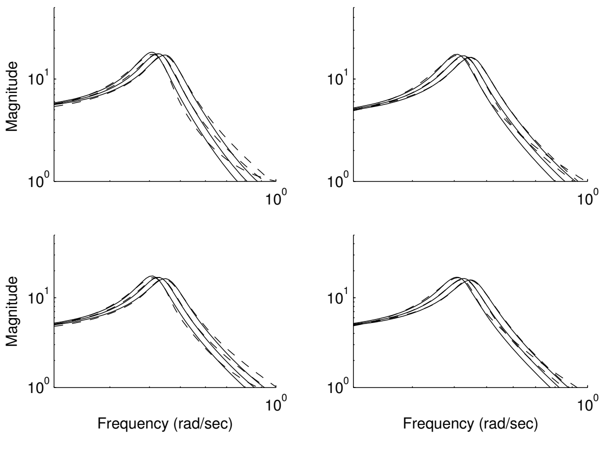

Fig. 3. Bode plots for parameter-dependent system reduction example. Dashed lines - reduced systems for different values of parameters, solid lines - original systems for different values of parameters.

VI. CONCLUSION

We presented two extensions to the optimization based model reduction framework: frequency-weighted model reduction and parameter-dependent model reduction. The

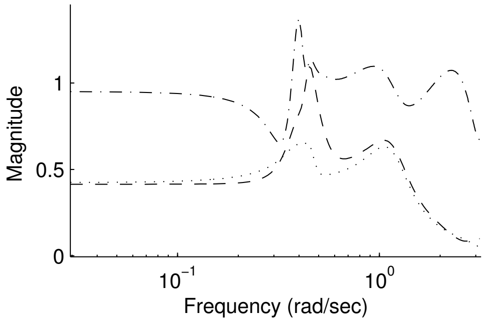

Fig. 4. $\max \sigma(G(\omega, \theta_i) - \hat{G}(\omega, \theta_i))$ for $\theta_1 = 0.1$ - dashed, $\theta_2 = 0.3$ - dotted , $\theta_3 = 0.5$ - dash-dotted.

biggest advantage is the ability to perform the optimization using only frequency samples of the model. That may be useful for large-scale systems. We showed that the frequency-weighted extension is competitive in comparison with frequency-weighted state space methods. We also showed the application of the method in problems where it is important to preserve the structure of the system. MIMO parameter model reduction is not fully developed in this paper and that should be one of the main directions of the future work.

ACKNOWLEDGMENT

This work was supported by Swedish Research Council. The authors would like to thank Johan Åkesson for valuable comments.

APPENDIX

A. Realization Details

In this section we will take a closer look on the relaxation of the structure introduced in section II-A. As we mentioned earlier transfer function BA^{-1} can have both anti-stable and stable parts, since we have relaxed the structure of B polynomial.

Assume that we made a stable/anti-stable decomposition of $BA^{-1} = PQ^{-1} + RQ_2^{\sim\nabla}$, where $R = \sum_{i=0}^{k-1} R_i z^{-i}$ and A_2 is a such that $\det(Q_2)$ is a Schur polynomial and

$$A = QQ^{\sim} = Q_2^{\sim}Q_2. \tag{17}$$

So the original problem is replaced with:

$$\|G - PQ^{-1} - RQ_2^{\sim\nabla}\|_\infty, \tag{18}$$

The relaxation needs motivation. In fact, $\min \|G - G_-\|_\infty = \|G\|_H$, where the minimization is performed over all possible anti-stable terms G_-. Moreover if we release $\Delta_- = RQ_2^{\sim\nabla}$ term we will get Hankel model reduction:

$$\min_{\Delta_-, P, Q} \|G - PQ^{-1} - \Delta_-\|_\infty = \\ = \min_{P,Q} \|G - PQ^{-1}\|_H$$

The same technique is applied to frequency-weighted problem. Next, we introduce the trigonometric parametrization of our problem, that makes optimization more numerically robust. Denote:

$$\begin{aligned}\bar{B} &= \mathbb{R}e\left\{z^k(PQ_2^{\nabla} + RQ)\right\}\\ \bar{C} &= \mathbb{I}m\left\{z^k(PQ_2^{\nabla} + RQ)\right\}\end{aligned} \tag{19}$$

The polynomials $A, \bar{B}, \bar{C}$ have certain properties to fulfil. $A(z) > 0$, if $z = e^{\jmath\omega}\ \forall\omega$, making $\det(Q)$ a Schur polynomial (the reduced system PQ^{-1} stable). The conditions $A_i = A_{-i} = A_i^T$, $\bar{B}_i = \bar{B}_{-i}, \bar{C}_i = \bar{C}_{-i}$ imply that $A, \bar{B}, \bar{C}$ are trigonometric polynomials and provide a convenient parametrization. The positivity of A_0 provides the existence of solution to spectral factorization problem, and the condition $A_0 > I$ is introduced for the normalization of calculations. $A = \sum_{i=0}^{k} A_i(z^i + z^{-i})$, $\bar{B} = \sum_{i=0}^{k} \bar{B}_i(z^i + z^{-i})$, $\bar{C} = \frac{1}{\jmath}\sum_{i=0}^{k-1} \bar{C}_i(z^i - z^{-i})$. The properties of such parametrization are following.

Lemma 1.1: There exists a one to one correspondence between Q, Q_2, P, R and $A, B = \bar{B} + \jmath\bar{C}$.

Lemma 1.2: Assume $A(z)$ is found from (13-15) and $A(z)$ is a trigonometric polynomial, positive-definite on the unit circle $(A(e^{\jmath\omega}) > 0\ \ \forall\omega \in [0, \pi])$ with $A_0 > 0$ then:

1) Spectral factorization problem (17) has a solution
2) McMillan degree of the reduced system $\hat{G} = PQ^{-1}$ is less or equal to km.

The proofs can be found in [7].

B. Proof of the Lemma 2.1

Proof: Using the Schur complement of the block A_ω of the matrix in (4), we get an inequality that is valid $\forall\omega \in [0, \pi]$:

$$\gamma_0 A_\omega I > (G_\omega A_\omega - B_\omega)'(\gamma_0 f_\omega)^{-1}(G_\omega A_\omega - B_\omega)$$

Note that inequality valid for the maximal singular value function as well, we get:

$$\gamma_0 \frac{|f_\omega|^{1/2}}{\underline{\sigma}(A_\omega)^{1/2}} > \overline{\sigma}(G_\omega - B_\omega A_\omega^{-1})$$

Using the LMI $f_\omega I \le A_\omega$ and taking the supremum of both hand sides over ω proves the statement. ■

C. Proof of the Theorem 3.2

The main idea of this proof is taken from [8].

1) The proof is based on the classical Nehari's theorem. Rewrite $\|W^o(G - BA^{-1})W^i\| \le \gamma$ as:

$$\|G - PQ^{-1} - RQ_2^{-\nabla}\|_\infty < \gamma \sup_\omega[\overline{\sigma}(W^i)\overline{\sigma}(W^o)]$$

By Nehari's theorem there exists a real matrix K such that:

$$\|RQ_2^{-\nabla} - K\|_\infty < km\gamma\left(\inf_\omega[\underline{\sigma}(W^i)\underline{\sigma}(W^o)]\right)^{-1},$$

where km is the McMillan degree of $RQ_2^{-\nabla}$. Consider the inequality for every frequency and multiply the both hand sides with $\overline{\sigma}(W^o), \overline{\sigma}(W^i)$. Using the sub-multiplicative property of singular values, we get:

$$\|W^o(RQ_2^{-\nabla} - K)W^i\|_\infty < km\gamma C(W^o, W^i),$$

Combining this inequality with $\|W^o(G - BA^{-1})W^i\| \le \gamma$ yields:

$$\begin{aligned}&\|W^o(G - (P + KQ)Q^{-1})W^i\|_\infty <\\ &< (kmC(W^i, W^o) + 1)\gamma,\end{aligned}$$

Since we obtain the numerator by minimizing over it, we bound $\|W^o(G - PQ^{-1})W^i\|_\infty$.

2) We applied Nehari's theorem without considering the weights. We can rewrite $\|W^o(G - BA^{-1})W^i\| \le \gamma$, performing the stable/anti-stable decomposition of $W^o RQ_2^{-\nabla}W^i$ yields:

$$\|W^o(G - PQ^{-1} - \hat{G}_2)W^i - R_2Q_2^{-\nabla}\|_\infty < \gamma$$

Applying the same techniques as above proves the result.

3) The result follows by applying lemma 1.1 and noting that the relaxation is related to Hankel model reduction.

REFERENCES

[1] G. Obinata and B. Anderson, *Model Reduction for Control System Design.* London,UK:Springer-Verlag, 2001.

[2] K. Zhou, "Frequency-weighted L_∞ norm and optimal Hankel norm model reduction," *Automatic Control, IEEE Transactions on*, vol. 40, no. 10, pp. 1687–1699, Oct 1995.

[3] D. F. Enns, "Model reduction with balanced realizations: An error bound and a frequency weighted generalization," *Decision and Control, 1984. The 23rd IEEE Conference on*, vol. 23, pp. 127–132, Dec. 1984.

[4] H. Sandberg, "Model reduction of linear systems using extended balanced truncation," in *Proceedings of the American Control Conference*, June 2008.

[5] A. Megretski, "H-infinity model reduction with guaranteed suboptimality bound," *American Control Conference, 2006*, pp. 448–453, June 2006.

[6] K. C. Sou, A. Megretski, and L. Daniel, "A quasi-convex optimization approach to parameterized model order reduction," in *IEEE/ACM Design Automation Conference*, June 2005.

[7] A. Sootla, A. Rantzer, and G. Kotsalis, "Multivariable optimization-based model reduction." *submitted to Automatic Control, IEEE Transaction on.*

[8] H. Sandberg and R. M. Murray, "Frequency-weighted model reduction with applications to structured models," *American Control Conference, 2007. ACC '07*, pp. 941–946, July 2007.

[9] M. Fazel, H. Hindi, and S. Boyd, "A rank minimization heuristic with application to minimum order system approximation," *American Control Conference, 2001. Proceedings of the 2001*, vol. 6, pp. 4734–4739, 2001.

[10] K. Zhou, J. C. Doyle, and K. Glover, *Robust and optimal control.* Upper Saddle River, NJ, USA: Prentice-Hall, Inc., 1996.

[11] A. Rantzer, "On the Kalman-Yakubovich-Popov lemma," *Syst. Control Lett.*, vol. 28, no. 1, pp. 7–10, 1996.

[12] J. Sturm, "Using SeDuMi 1.02, a MATLAB toolbox for optimization over symmetric cones," *Optimization Methods and Software*, vol. 11–12, pp. 625–653, 1999.

[13] J. Löfberg, "Yalmip : A toolbox for modeling and optimization in MATLAB," in *Proceedings of the CACSD Conference*, Taipei, Taiwan, 2004.

[14] G. Latham and B. Anderson, "Frequency weighted optimal hankel-norm approximation of stable transfer function," *Systems Control Lett.*, vol. 5, pp. 229–236., July 1985.

[15] M. Safonov, R. Chiang, , and H. Flashner, "H infinity control synthesis for a large space structure," in *Proc. of American Contr. Conf.*, June 1988.

[16] R. Y. Chiang and M. G. Safonov, "Robust control toolbox. the mathworks," 1992.