Optimal time-domain moment matching with partial placement of poles and zeros

Ionescu, Tudor; Iftime, Orest; Necoara, I.

IMPORTANT NOTE: You are advised to consult the publisher's version (publisher's PDF) if you wish to cite from it. Please check the document version below.

Document Version
Publisher's PDF, also known as Version of record

[Link to publication in University of Groningen/UMCG research database](#)

Optimal time-domain moment matching
with partial placement of poles and zeros

T. C. Ionescu, O. V. Iftime and I. Necoara

Abstract— In this paper we consider a minimal, linear, time-invariant (LTI) system of order n, large. Our goal is to compute an approximation of order $\nu < n$ that *simultaneously* matches ν moments, has ℓ poles and k zeros fixed, with $\ell + k < \nu$, and achieves minimal H_2 norm of the approximation error. For this, in the family of ν order parametrized models that match ν moments we impose $\ell + k$ linear constraints yielding a subfamily of models with ℓ poles and k zeros imposed. Then, in the subfamily of ν order models matching ν moments, with ℓ poles and k zeros imposed we propose an optimization problem that provides the model yielding the minimal H_2-norm of the approximation error. We analyze the first-order optimality conditions of this optimization problem and compute explicitly the gradient of the objective function in terms of the controllability and the observability Gramians of the error system. We then propose a gradient method that finds the (optimal) *stable* model, with fixed ℓ poles and k zeros.

I. INTRODUCTION

In moment matching-based model reduction techniques, the approximation is yielded by constructing a lower degree rational function that approximates the original transfer function of high dimension, see e.g. [1]. The low degree rational function matches a number of terms of the series expansion of the original transfer function at various complex points. It has also been shown that properties of the original transfer function are preserved/inherited if the matching points are chosen in a certain way. For instance, in Antoulas [2] and Sorensen [3] the problem of moment matching with preservation of physical properties such as passivity/positive realness has been tackled. Also, in [4], [5] the problem of the moment matching-based approximation yielding the lowest H_2 norm of the approximation error has been solved resulting in a numerically efficient algorithm called Iterative Rational Krylov Approximation (IRKA).

Recently, for linear, time-invariant (LTI) systems, a system theoretic, Sylvester equation (time-domain) approach to moment matching has been taken in [6], [7]. In short, the notion of moment of an LTI minimal system has been related to the unique solution of a Sylvester equation, see also, e.g., [8], [9], for earlier results. Used for model reduction, the Sylvester equation approach yields simple and direct characterizations of all parameterized, reduced order models that match a prescribed set of moments of a given system at a set of *finite* interpolation points. The free parameters can be selected to identify the models that meet additional requirements or constraints. Hence, the resulting families of models of order ν that match ν moments of the given n-th order system contain subfamilies of ν order models that satisfy desired systemic properties. For instance, in [6] the ν free parameters a selected such stability and relative degree are preserved, in [10] the ν parameters are chosen to preserve the port-Hamiltonian physical structure in [7] the ν parameters are selected to find the lowest order minimal model and in [11] the ν parameters are computed to find the model that matches 2ν moments as well as the model that matches ν moments of te given system and ν moments of its first order derivative. Recently, in [12], [13] the ν parameters are computed, using gradient-based optimization algorithms, to find the model that achieves the lowest H_2 norm of the approximation error. Note that all the above mentioned models are computed irrespective of the choice of interpolation points, but using *all* the ν degrees of freedom available.

The resulting ν order models matching ν moments may satisfy additional properties using all the ν degrees of freedom. In control there exist algorithm to place a number of $\mu < \nu$ poles with a ν order controller, see [14]. To the best of our knowledge there are no results providing ν order models that match ν moments of the given system and satisfy more additional physical properties at once. In detail, one may seek a ν order approximation which is stable and, e.g., decreases the H_2 norm of the error simultaneously. Furthermore, preservation of ν poles (modes) of the original system may not be enough to capture the input-output behaviour of the given system. The zeros of the approximations may differ significantly from those of the original system. In particular, assuming the original system to be of minimum phase, right half-plane zeros, which are not present in the original system, may appear in the approximation altering the desired input-output behaviour. First steps have been made in simulation in [15] where a number of ν poles and zeros have been preserved simultaneously. Furthermore, one may impose on the approximation ℓ fixed poles and k fixed zeros (possibly of the given system) and use the rest of the degrees of freedom to solve an optimal H_2 error norm problem.

T.C. Ionescu is with the Dept. of Automatic Control & Systems Eng., Politehnica Univ. of Bucharest, 060042 Bucharest, Romania, and with the "G. Mihoc-C. Iacob" Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy, 050711 Bucharest, Romania. E-mail: tudor.ionescu@acse.pub.ro.

O. V. Iftime is with the Department of Economics, Econometrics and Finance, Faculty of Economics and Business, University of Groningen, Nettelbosje 2, 9747 AE, Groningen, The Netherlands. E-mail: o.v.iftime@rug.nl

I. Necoara is with the Department of Automatic Control and Systems Engineering, University Politehnica Bucharest, 060042 Bucharest, Romania. E-mail: ion.necoara@acse.pub.ro.

We consider a single input-single output, linear, time-invariant (LTI) stable, minimal n-th order system and the family of ν order models that match ν moments of the given system. First we provide the explicit linear constraints such that $\ell \leq \nu$ poles and $k \leq \nu$ are imposed. We then formulate an optimisation problem to find the (unique) model of order ν that matches ν moments of the given system, has ℓ poles and k zeros fixed, with $\ell + k < \nu$, and achieves the optimal H_2 error norm, *simultaneously*. The problem is solved and the solution is found in the subfamily of models employing a partial minimization approach and then using a gradient method-based optimization algorithm.

In detail, for a fixed set of interpolation points and a set of $\ell + k < \nu$ linear constraints (through fixing poles and zeros) the proposed procedure is seeking the approximation yielding the minimal H_2-norm of the approximation error. We analyze the (necessary) first-order optimality conditions and compute explicitly the gradient of the objective function in terms of the controllability and the observability Gramians of a minimal realization of the error system. Based on the gradient expression for the objective function we propose a gradient method to find the free parameters that yield the minimum approximation error, and the reduce model *stable*, with ℓ fixed poles and k fixed zeros.

Content: The paper is organized as follows. In Section II we briefly recall the Sylvester equation moment matching and the computation of the family of ν order models that match ν moments of the given system. In Section III, we give the set of linear constraints such that the ν order models matching ν moments have ℓ poles fixed (Section III-A) and k zeros fixed (Section III-B). In Section IV, we formulate the optimal H_2-norm model reduction problem, recast it as an optimization problem with a Gramian-based written cost function, and derive the corresponding first-order optimality conditions. We further compute an (optimal) solution to the optimization problem using a gradient method. In Section V we illustrate the theory on a benchmark test case. The paper ends with Conclusions.

II. PRELIMINARIES

In this section we briefly review the computation of the family of ν order models matching ν moments of a stable LTI system. For more details see e.g., [6], [7].

A. Linear systems

Consider a single input-single output (SISO) linear time-invariant (LTI) minimal system:

$$\Sigma : \quad \dot{x} = Ax + Bu, \quad y = Cx, \tag{1}$$

with the state $x \in \mathbb{R}^n$, the input $u \in \mathbb{R}$ and the output $y \in \mathbb{R}$. The transfer function of (1) is:

$$K(s) = C(sI - A)^{-1}B, \quad K : \mathbb{C} \to \mathbb{C}. \tag{2}$$

Throughout the rest of the paper we assume that the system (1) is stable, that is $\sigma(A) \subset \mathbb{C}^-$. For the sake of clarity we consider the SISO case. However the results can be extended to the multiple input-multiple output case.

B. Sylvester equation-based time-domain moment matching

Assume that (1) is a minimal realization of the transfer function $K(s)$. The moments of (2) are defined as follows.

Definition 1. [1], [6] The k-moment of system (1), with the transfer function $K(s)$ described by (2), at $s_1 \in \mathbb{R}$ is

$$\eta_k(s_1) = [(-1)^k/k!] \cdot \left[d^k K(s)/ds^k \right]_{s=s_1} \in \mathbb{R}.$$

Pick suitable points $s_1, \ldots, s_\nu \in \mathbb{R}$ and let the matrices $S \in \mathbb{R}^{\nu \times \nu}$, with the spectrum $\sigma(S) = \{s_1, \ldots, s_\nu\}$, and $L = [1 \ 1 \ \ldots \ 1] \in \mathbb{R}^{1 \times \nu}$, such that the pair (L, S) is observable. Let $\Pi \in \mathbb{R}^{n \times \nu}$ be the solution of the Sylvester equation:

$$A\Pi + BL = \Pi S. \tag{3}$$

Furthermore, since the system is minimal, assuming that $\sigma(A) \cap \sigma(S) = \emptyset$, then Π is the unique solution of the equation (3) and $\text{rank } \Pi = \nu$, see e.g. [16]. Then, the moments of (1) are characterised as follows:

Proposition 1. *[6] The moments of system (1) at the interpolation points $\{s_1, s_2, ..., s_\ell\} = \sigma(S)$ are in one-to-one relation[1] with the elements of the matrix $C\Pi$.*

The following proposition gives necessary and sufficient conditions for a low-order system to achieve moment matching:

Proposition 2. *[6] Consider the LTI system:*

$$\dot{\xi} = F\xi + Gu, \quad \psi = H\xi, \tag{4}$$

with $F \in \mathbb{R}^{\nu \times \nu}$, $G \in \mathbb{R}^\nu$ and $H \in \mathbb{R}^{p \times \nu}$, and the corresponding transfer function:

$$K_G(s) = H(sI - F)^{-1}G. \tag{5}$$

Fix $S \in \mathbb{R}^{\nu \times \nu}$ and $L \in \mathbb{R}^{1 \times \nu}$, such that the pair (L, S) is observable and $\sigma(S) \cap \sigma(A) = \emptyset$. Furthermore, assume that $\sigma(F) \cap \sigma(S) = \emptyset$. The reduced system (4) matches the moments of (1) at $\sigma(S)$ if and only if:

$$HP = C\Pi, \tag{6}$$

where the invertible matrix $P \in \mathbb{R}^{\nu \times \nu}$ is the unique solution of the Sylvester equation $FP + GL = PS$.

We are now ready to present a family of ν reduced order models parameterized in G that match ν moments of the given system (1). The reduced system:

$$\Sigma_G : \dot{\xi} = (S - GL)\xi + Gu, \quad \psi = C\Pi\xi, \tag{7}$$

with the transfer function

$$K_G(s) = C\Pi(sI - S + GL)^{-1}G, \tag{8}$$

describes a family of ν order models that achieve moment matching at $\sigma(S)$ *fixed* satisfying the following properties and constraints:

1) Σ_G is parameterized in $G \in \mathbb{R}^\nu$,
2) $\sigma(S - GL) \cap \sigma(S) = \emptyset$.

---

[1]By one-to-one relation between a set of moments and the elements of a matrix, we mean that the moments are uniquely determined by the elements of the matrix.

C. Time-domain moment matching for MIMO systems

The results can be directly extended to the MIMO case, see, e.g., [7] for more details. Consider a MIMO system (1), with input $u(t) \in \mathbb{R}^m$, output $y(t) \in \mathbb{R}^p$ and the transfer function $K(s) \in \mathbb{C}^{p \times m}$. Let $S \in \mathbb{C}^{\nu \times \nu}$ and $L = [l_1 \ l_2 \ ... \ l_\nu] \in \mathbb{C}^{m \times \nu}$, $l_i \in \mathbb{C}^m$, $i = 1, ..., \nu$, be such that the pair (L, S) is observable. Let $\Pi \in \mathbb{C}^{n \times \nu}$ be the unique solution of the Sylvester equation (3). Then the moments $\eta(s_i) = K(s_i)l_i$, $\eta(s_i) \in \mathbb{C}^p$, $i = 1, ..., \nu$ of at $\{s_1, ..., s_\nu\} = \sigma(S)$ are in one-to-one relation with $C\Pi$. The model reduction problem for MIMO systems boils down to finding a ν-th order model described by the equations (4), with the transfer function K_G as in (5), $G \in \mathbb{R}^{\nu \times m}$ which satisfies the right tangential interpolation conditions [17]

$$K(s_i)l_i = \widehat{K}(s_i)l_i, \ i = 1, ..., \nu. \tag{9}$$

It immediately follows that the solution to this problem is provided by a direct application of Proposition 2, i.e., a class of reduced order MIMO models that achieve moment matching in the sense of satisfying the tangential interpolation conditions (9) is given by $\Sigma_G = (S - GL, G, C\Pi)$ as in (7). Hence, without loss of generality, throughout the rest of the paper we discuss the SISO case, i.e., $m = p = 1$, the results being easily extended to tangential interpolation for MIMO systems.

III. PARTIAL POLE AND ZERO PLACEMENT AS CONSTRAINTS

In this section we derive linear relations parametrized in G yielding the subfamily of ν order models that preserve ℓ poles and k zeros of the given system.

A. Pole placement constraints

In this section, we place ℓ poles of the reduced order, for example in some of the poles of the original system, by properly selecting G. Consider an LTI system (1) and the class of reduced ν order models Σ_G from (7) that match ν moments of (1) at $\sigma(S)$. Let $\lambda_i \in \mathbb{C}$, $i = 1, ..., \ell$, $\ell \leq \nu$ be such that $\lambda_i \notin \sigma(S)$. Then λ_i are poles of Σ_G if $\det(\lambda_i I - S + GL) = 0$, $i = 1 : \ell$. To this end, let $Q \in \mathbb{R}^{\ell \times \ell}$ be a matrix such that $\sigma(Q) = \{\lambda_1, ..., \lambda_\ell\}$. Furthermore, consider $\widetilde{C} \in \mathbb{R}^{1 \times n}$ such that $\widetilde{C}\Pi = 0$, where Π solves (3), and let $\Upsilon \in \mathbb{C}^{\ell \times n}$ be the unique solution of the Sylvester equation

$$Q\Upsilon = \Upsilon A + R\widetilde{C}, \tag{10}$$

with $R \in \mathbb{R}^\ell$ any matrix such that the pair (Q, R) is controllable. The next result gives G such that $\{\lambda_1, ..., \lambda_\ell\} = \sigma(Q) \subseteq \sigma(S - GL)$.

Proposition 3. *Let Σ_G from (7) be a ν order model that matches the moments of (1) at $\sigma(S)$, with the transfer function $K_G(s)$ as in (8). Furthermore, let $Q \in \mathbb{R}^{\ell \times \ell}$ be a matrix such that $\sigma(Q) = \{\lambda_1, ..., \lambda_\ell\}$ and $R \in \mathbb{R}^\ell$ any matrix such that the pair (Q, R) is controllable. If G is a solution of the matrix equation*

$$\Upsilon\Pi G = \Upsilon B, \tag{11}$$

with $\Upsilon \in \mathbb{C}^{\ell \times n}$, the unique solution of (10), then $K_G(s)$ has ℓ poles at $\{\lambda_1, ..., \lambda_\ell\} = \sigma(Q)$.

Remark 1. Note that if $\ell = \nu$, then all ν poles of $K_G(s)$ are placed at $\{\lambda_1, ..., \lambda_\nu\} = \sigma(Q)$, by

$$G = (\Upsilon\Pi)^{-1}\Upsilon B. \tag{12}$$

Choosing $Q = \text{diag}\{\lambda_1, ..., \lambda_\nu\}$ and $S = \text{diag}\{s_1, ..., s_\nu\}$, then Υ and Π are the left and the right Krylov projections, respectively, and the result in [18, Lemma 2.1] is a particular case of Proposition 3.

Explicit algebraic constraints for poles:: Let $S = \text{diag}\{s_1, ..., s_\nu\}$ and $L = [1 \ ... \ 1] \in \mathbb{R}^{1 \times \nu}$. Then $\{\lambda_1, ..., \lambda_\ell\}$ are poles of $K_G(s)$ if and only if G satisfies the equation

$$1 + LD_k^{-1}G = 0, \quad \forall k = 1 : \ell. \tag{13}$$

Stable approximations: Consider the families of approximations Σ_G described by the equations (7) and the problem of finding G such that the reduced order system is asymptotically stable. The goal is achieved by selecting G such that $\sigma(S - GL) \cap \sigma(A) = \emptyset$ and $\sigma(S - GL) = \{\lambda_1, ..., \lambda_\nu\} \subset \mathbb{C}^-$. Note that, by the observability of the pair (L, S), there exists a unique matrix G such that this condition holds, see [6] for more details.

B. Zero placement constraints

Consider a system (1) and the family of ν order models Σ_G that approximate (1) by matching ν moments. Let $z_1, ..., z_k \in \mathbb{C}$, $k \leq \nu$. By, e.g., [6], [19], [15], there exists a subfamily of models Σ_G, with the property that the set of zeros of each model contains $z_1, ..., z_k$. Equivalently, there exists G such that

$$\det \begin{bmatrix} z_i I - S & G \\ C\Pi & 0 \end{bmatrix} = 0, \quad i = 1 : k. \tag{14}$$

Now, let $G = [g_1 \ g_2 \ ... \ g_\nu]^T \in \mathbb{R}^\nu$. Directly following arguments from the proof of [6, Theorem 3], condition (14) is equivalent to a system of k equations with ν unknowns $g_1, ..., g_\nu$, given by

$$(-1)^\nu [-g_1\zeta_1(z_1) + g_2\zeta_2(z_1) + \cdots + (-1)^\nu g_\nu\zeta_\nu(z_1)] = 0,$$
$$(-1)^\nu [-g_1\zeta_1(z_2) + g_2\zeta_2(z_2) + \cdots + (-1)^\nu g_\nu\zeta_\nu(z_2)] = 0,$$
$$\vdots \tag{15}$$
$$(-1)^\nu [-g_1\zeta_1(z_k) + g_2\zeta_2(z_k) + \cdots + (-1)^\nu g_\nu\zeta_\nu(z_k)] = 0,$$

with $\zeta_j(s)$ polynomials of degree $\nu - 1$, $j = 1 : \nu$.

Explicit algebraic constraints for zeros:: Let $S = \text{diag}\{s_1, ..., s_\nu\}$, $L = [1 \ ... \ 1] \in \mathbb{R}^{1 \times \nu}$ and explicitly write $C\Pi = [\eta_1 \ \eta_2 \ \eta_3 \ ... \ \eta_\nu]$. Then $\{z_1, ..., z_k\}$ are poles

of $K_G(s)$ if and only if (14) is satisfied, i.e.,

$$\begin{vmatrix} \gamma_{j1} & 0 & 0 & \dots & 0 & g_1 \\ 0 & \gamma_{j2} & 0 & \dots & 0 & g_2 \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \dots & \gamma_{j\nu} & g_\nu \\ \eta_1 & \eta_2 & \eta_3 & \dots & \eta_\nu & 0 \end{vmatrix} = 0, \qquad (16)$$

$$\gamma_{ji} = z_j - s_i, \ i = 1 : \nu, \ j = 1 : k.$$

First, note that $\gamma_{ji} \neq 0$ for all i, j. Successively decomposing the determinant in (16) by the last column and computing the resulting minors performing row decomposition yields

$$\sum_{i=1}^{\nu} \eta_i g_i \prod_{l=1:\nu, l\neq i} \gamma_{jl} = 0, \quad j = 1 : k. \qquad (17)$$

Then, dividing (17) by $\prod_{l=1:\nu} \gamma_{jl} \neq 0$ leads to the following linear equations in G:

$$\sum_{i=1}^{\nu} \frac{\eta_i}{\gamma_{ji}} g_i = 0, \quad j = 1 : k. \qquad (18)$$

IV. Optimal H_2 model reduction with poles and zeros constraints–an optimization approach

Problem 1. Consider a SISO LTI system (1) of order $n \in \mathbb{N}$, with the transfer function $K(s)$ as in (2). Let $\nu < n$, $\nu \in \mathbb{N}$, $S \in \mathbb{R}^{\nu \times \nu}$ with the spectrum $\sigma(S) = \{s_1, \dots, s_\nu\}$ and $L \in \mathbb{R}^{1 \times \nu}$ be such that the pair (L, S) is observable. Let Π be the (unique) solution of the Sylvester equation (3). Furthermore, consider the family of parametrized ν order models Σ_G from (7), with the transfer function $K_G(s)$ as in (8) that match ν moments $C\Pi$ of (1) at $\sigma(S)$, for all $G \in \mathbb{R}^\nu$, such that

$$\sigma(S) \cap \sigma(S - GL) = \emptyset. \qquad (19)$$

Let $\ell, k \in \mathbb{N}$ such that $\ell + k < \nu$. Find $G = [g_1 \ g_2 \ \dots \ g_\nu]^T \in \mathbb{R}^\nu$ such that the following constrains are satisfied simultaneously:

i) Σ_G has ℓ prescribed poles, i.e., $\{\lambda_1, \dots, \lambda_\ell\} \subseteq \sigma(S - GL)$ or, equivalently, equation (11) holds;
ii) Σ_G has k prescribed zeros, i.e., $\{z_1, \dots, z_k\}$ such that $\det \begin{bmatrix} z_i I - S & G \\ C\Pi & 0 \end{bmatrix} = 0$, $i = 1 : k$ or, equivalently, equation (15) holds;
iii) $\|K - K_G\|_2$ is minimal;
iv) condition (19) holds.

We now discuss each of the conditions i)–iv) separately. Conditions i) and ii) are solved employing relations (11) and (15), respectively. Regarding condition iii), Problem 1 can be recast in terms of the computation of the H_2-norm using the Gramians of the realization of the *error system*

$$\mathbf{K} = K - K_G,$$

with K_G from (8) parameterized in G. Let $(\mathbf{A}, \mathbf{B}, \mathbf{C})$ be a state-space realization of the error transfer function $\mathbf{K}$

$$\mathbf{K}(s) = \mathbf{C}(sI - \mathbf{A})^{-1}\mathbf{B},$$

where

$$\mathbf{A} = \begin{bmatrix} A & 0 \\ 0 & S - GL \end{bmatrix}, \ \mathbf{B} = \begin{bmatrix} B \\ G \end{bmatrix}, \ \mathbf{C} = C \begin{bmatrix} I & -\Pi \end{bmatrix}. \quad (20)$$

Denote the controllability and the observability Gramians of (20) by $\mathbf{W}$ and $\mathbf{M}$, respectively. Then, they are solutions of the Lyapunov equations

$$\mathbf{AW} + \mathbf{WA}^T + \mathbf{BB}^T = 0, \qquad (21a)$$
$$\mathbf{A}^T\mathbf{M} + \mathbf{MA} + \mathbf{C}^T\mathbf{C} = 0, \qquad (21b)$$

Let us also partition the controllability and observability Gramians $\mathbf{W}, \mathbf{M}$ as in matrix $\mathbf{A}$

$$\mathbf{W} = \begin{bmatrix} W_{11} & W_{12} \\ W_{12}^T & W_{22} \end{bmatrix}, \ \mathbf{M} = \begin{bmatrix} M_{11} & M_{12} \\ M_{12}^T & M_{22} \end{bmatrix}. \qquad (22)$$

By [4], condition iii) can be written explicitly as

$$\min_{G, \mathbf{M}} \text{Trace}(\mathbf{B}^T\mathbf{MB}).$$

Constraint iv) can be easily satisfied by properly choosing the matrix pair (L, S) with $\sigma(S) \subset \mathbb{C}^+ \cup \mathbb{C}^0$ and by finding G such that the resulting approximation Σ_G is stable. For instance, we can pick $S = \text{diag}(s_1, s_2, \dots, s_\nu)$ and $L = [1 \ 1 \ \dots \ 1]$, with $s_i \in \mathbb{C}^+ \cup \mathbb{C}^0$ such that condition iv) holds.

Equivalent formulation of Problem 1. It can be written as the nonconvex optimization problem:

$$\min_{G, \mathbf{M}} \text{Trace}(\mathbf{B}^T\mathbf{MB}) \qquad (23)$$
$$\text{s.t.} : \ \sigma(S - GL) \subset \mathbb{C}^-, \ \mathbf{A}^T\mathbf{M} + \mathbf{MA} + \mathbf{C}^T\mathbf{C} = 0,$$
$$G \text{ satisfies the linear systems (11) and (15).} \quad (24)$$

For simplicity, to emphasize the dependency on G, let

$$\mathcal{A}(G) := \begin{bmatrix} A & 0 \\ 0 & S - GL \end{bmatrix}, \ \mathcal{B}(G) := \begin{bmatrix} B \\ G \end{bmatrix} \begin{bmatrix} B \\ G \end{bmatrix}^T.$$

Equivalently, the problem (23) can be rewritten as

$$\min_{G} \text{Trace}(\mathbf{M}(G)\mathcal{B}(G)) \qquad (25)$$
$$\text{s.t.} : \ \sigma(S - GL) \subset \mathbb{C}^-, \qquad (26)$$
$$\mathcal{A}(G)^T\mathbf{M}(G) + \mathbf{M}(G)\mathcal{A}(G) + \mathbf{C}^T\mathbf{C} = 0,$$
$$G \text{ satisfies the linear systems (11) and (15).} \quad (27)$$

A. Gradient method for solving Problem 1

Let $\mathcal{R} = \{G : \sigma(S - GL) \subset \mathbb{C}^-, \ (11) \text{ and } (15) \text{ hold}\}$. Consider the following partial minimization for (23):

$$(23) = \min_{G \in \mathcal{R}} \left(\min_{\mathbf{M}:\mathbf{A}^T\mathbf{M} + \mathbf{MA} + \mathbf{C}^T\mathbf{C} = 0} \text{Trace}(\mathbf{B}^T\mathbf{MB}) \right).$$

If $S - GL$ and A are stable, then it follows from basic results for Lyapunov equations that there exists unique $\mathbf{M} = \mathbf{M}(G) \succeq 0$ solution of:

$$\mathbf{A}^T\mathbf{M} + \mathbf{MA} + \mathbf{C}^T\mathbf{C} = 0.$$

Hence, for any G stabilizable, the partial minimization in $\mathbf{M}$ leads to an optimal value:

$$f(G) = \text{Trace}\left(\begin{bmatrix} B \\ G \end{bmatrix}^T \mathbf{M}(G) \begin{bmatrix} B \\ G \end{bmatrix}\right),$$

where $\mathbf{M}(G)$ is unique solution of Lyapunov equation:

$$\begin{bmatrix} A & 0 \\ 0 & S-GL \end{bmatrix}^T \mathbf{M} + \mathbf{M}\begin{bmatrix} A & 0 \\ 0 & S-GL \end{bmatrix} + \mathbf{C}^T\mathbf{C} = 0, \quad (28)$$

with $C_V = C\Pi = CVT$. Explicitly, in terms of G, we have:

$$\min_G f(G) \quad \left(= \text{Trace}\left(\begin{bmatrix} B \\ G \end{bmatrix}^T \mathbf{M}(G) \begin{bmatrix} B \\ G \end{bmatrix}\right)\right) \quad (29)$$

$$\text{s.t.:} \quad \sigma(S-GL) \subset \mathbb{C}^-, \text{ (11), (15) and (30) hold,}$$

where $\mathbf{M}(G)$ is the solution of the Lyapunov equation:

$$\begin{bmatrix} A & 0 \\ 0 & S-GL \end{bmatrix}^T M(G) + M(G)\begin{bmatrix} A & 0 \\ 0 & S-GL \end{bmatrix} + \mathbf{C}^T\mathbf{C} = 0. \quad (30)$$

For solving the equivalent non-convex optimization problem (29) we can apply any first- or second-order optimization method. For this type of optimization scheme we need to compute the gradient and even the Hessian of the objective function f. Below we provide the expression of the gradient of the objective function of (29):

$$f(G) = \text{Trace}\left(\mathbf{M}(G)\mathcal{B}(G)\right).$$

Theorem 1. *The objective function f of (29) is differentiable on the set of stable matrices $\mathcal{D}_{(SL)}$ and the gradient of f at any $G \in \mathcal{D}_{(SL)}$ is given by*

$$\nabla f(G) = 2(-M_{12}^T(G)W_{12}(G)L^T - M_{22}(G)W_{22}(G)L^T$$
$$+ M_{12}^T(G)B + M_{22}(G)G), \quad (31)$$

where $\mathbf{M}(G)$ solves the Lyapunov equation (30) and $\mathbf{W}(G)$ solves the Lyapunov equation

$$\begin{bmatrix} A & 0 \\ 0 & S-GL \end{bmatrix}\mathbf{W}(G) + \mathbf{W}(G)\begin{bmatrix} A & 0 \\ 0 & S-GL \end{bmatrix}^T + \mathcal{B}(G) = 0. \quad (32)$$

The proof of this theorem follows arguments similar to [12], [13]. The result of Theorem 1 yields the necessary first-order optimality condition for the model reduction Problem 1 expressed in terms of the optimization problem (29).

Lemma 1. *If $G \in \mathcal{R}$ solves the optimization problem (29) corresponding to the model reduction Problem 1, then*

$$M_{12}^T(G)W_{12}(G)L^T + M_{22}(G)W_{22}(G)L^T$$
$$= M_{12}^T(G)B + M_{22}(G)G,$$

where $\mathbf{M}(G)$ solves the Lyapunov equation (30) and $\mathbf{W}(G)$ solves the Lyapunov equation (32).

We can replace the set $\mathcal{R}$ with any sublevel set:

$$\mathcal{N}_{(SL)}^{G_0} = \{G \in \mathcal{R} : f(G) \leq f(G_0)\},$$

where $G_0 \in \mathcal{R}$ is any initial stable reduced order system matrix. Using similar arguments as in [20] we can show that $\mathcal{N}_{(SL)}^{G_0}$ is a compact set. Then, the theorem of Weierstrass implies that for any given matrix $G_0 \in \mathcal{R}$, the model

reduction Problem 1 given by optimization formulation (23) or equivalently (29) has a global minimum in the sublevel set $\mathcal{N}_{(SL)}^{G_0}$. We can also show that the gradient $\nabla f(G)$ is Lipschitz continuous on the compact sublevel set $\mathcal{N}_{(SL)}^{G_0}$. Let us briefly sketch the proof of this statement. First we observe that $\mathbf{M}(G)$ and $\mathbf{W}(G)$ are continuous functions. Moreover, there exists finite $\ell_M > 0$ such that:

$$\|\mathbf{M}(G) - \mathbf{M}(\bar{G})\| \leq \ell_M\|G - \bar{G}\| \quad \forall G, \bar{G} \in \mathcal{N}_{(SL)}^{G_0}.$$

Then, using the expression of $\nabla f(G)$, compactness of $\mathcal{N}_{(SL)}^{G_0}$, continuity of $\mathbf{M}(G)$ and $\mathbf{W}(G)$, and previous relation we conclude that there exists finite $\ell_f > 0$ such that:

$$\|\nabla f(G) - \nabla f(\bar{G})\| \leq \ell_f\|G - \bar{G}\| \quad \forall G, \bar{G} \in \mathcal{N}_{(SL)}^{G_0}.$$

This property of the gradient is useful when analyzing the convergence behavior of gradient-based algorithms. We propose to apply the gradient method for solving the smooth optimization problem (29), having the following iteration:

$$G_{k+1} = \Pi_{\widehat{\mathcal{R}}}(G_k - \alpha_k\nabla f(G_k)), \quad \alpha_k > 0, \quad (33)$$

where the linear subspace $\widehat{\mathcal{R}} = \{G : (11) \text{ and } (15) \text{ hold}\}$, $\Pi_{\widehat{\mathcal{R}}}(\cdot)$ denotes the ortogonal projection onto the linear subspace $\widehat{\mathcal{R}}$ and the stepsize α_k can be chosen using a line search procedure or constant in the interval $(0, 2/\ell_f)$. It is easy to see that if G_k is stabilizing, it follows that also the new iterate G_{k+1} has the same property provided that α_k is sufficiently small. Moreover, since $\nabla f(G)$ is Lipschitz, from standard optimization theory it follows that we have sublinear convergence rate of order $\mathcal{O}(1/k)$, see [21] for more details.

V. Illustrative numerical example

Consider the cart system controlled by a double-pendulum controller, with 6 states, with the matrices $A \in \mathbb{R}^{6\times6}, B \in \mathbb{R}^{6\times1}$ and $C \in \mathbb{R}^{1\times6}$, see [11] for the explicit matrices. The poles and the zeros of the system are $\{-1.6 + 6.63j, -1.6 - 6.63j, -0.74 + 3.48j, -0.74 - 3.48j, -0.16 + 0.55j, -0.16 - 0.55j\}$ and $\{-1.28 + 5.63j, -1.28 - 5.63j, -0.22 + 2.39j, -0.22 - 2.39j\}$, respectively, i.e., the system is stable and of minimum phase. Let

$$S = \text{diag}(0, 1/4, 1/2), \text{ and } L = \begin{bmatrix} 1 & 1 & 1 \end{bmatrix}.$$

Note that the pair (L, S) is observable. A family of third order models is described by Σ_G as in (7) with:

$$F = \begin{bmatrix} -g_1 & -g_1 & -g_1 \\ -g_2 & 0.5-g_2 & -g_2 \\ -g_3 & -g_3 & 0.25-g_3 \end{bmatrix}, \quad G = \begin{bmatrix} g_1 \\ g_2 \\ g_3 \end{bmatrix} \quad (34)$$

and $H = C\Pi = [1 \ 0.69 \ 0.45]$.

Using the gradient method proposed in Section IV to solve Problem 1, with constraints given by fixing a pole at -1.6 and a zero at -1.28, yields a third order model Σ_G with:

$$F = \begin{bmatrix} -3.95 & -3.95 & -3.95 \\ 13.64 & 13.89 & 13.64 \\ -12.39 & -12.39 & -11.89 \end{bmatrix}, \quad G = \begin{bmatrix} -1.6 \\ -0.18 \\ -0.18 \end{bmatrix}, \quad (35)$$

$$H = [1 \ 0.69 \ 0.45].$$

1773

The H_2 norm of the approximation error achieved by (35) is $8.7 \cdot 10^{-3}$. The poles of (35) are $\{-1.6, -0.18 + 0.53j, -0.18 - 0.53j\}$ and the zeros are $\{-1.28, -45.92\}$, i.e., a stable and minimum phase third order approximation. We compare approximation (35) with a third order model Σ_G with G given by (12), such that three poles are fixed at $\{-1.6, -0.16 - 0.55j, -0.16 + 0.55j\}$. Note that two of the poles are chosen from the poles of the given system. The resulting approximation is a model Σ_G as in (34) with:

$$F = \begin{bmatrix} -39.84 & -39.84 & -39.84 \\ 116.11 & 116.36 & 116.11 \\ -92.65 & -92.65 & -92.15 \end{bmatrix}, \; G = \begin{bmatrix} 39.84 \\ -116.11 \\ 92.65 \end{bmatrix}, \quad (36)$$
$$H = [1 \; 0.69 \; 0.45].$$

The H_2 norm of the approximation error achieved by (36) is $1.02 \cdot 10^{-2}$. The poles of the system (36) are indeed $\{-1.6, -0.16 - 0.55j, -0.16 + 0.55j\}$, as prescribed. The zeros are $\{-0.17 + 3.14j, -0.17 - 3.14j\}$. Hence (36) is a stable and minimum phase third order approximation.

We compare approximation (35) with a third order model Σ_G with G the solution of a square third order linear system yielded by (13) and (18), such that two poles and one zero are fixed, at $\{-1.6, -0.16\}$ and -1.28, respectively. The resulting approximation is a model Σ_G as in (34) with:

$$F = \begin{bmatrix} -0.03 & -0.03 & -0.03 \\ 3.21 & 3.4563 & 3.2063 \\ -5.7 & -5.7 & -5.2 \end{bmatrix}, \; G = \begin{bmatrix} 0.03 \\ -3.21 \\ 5.7 \end{bmatrix}, \quad (37)$$
$$H = [1 \; 0.69 \; 0.45].$$

The H_2 norm of the approximation error achieved by (37) is $2.8 \cdot 10^{-1}$. The poles of the system (37) are $\{-1.6, -0.02, -0.16\}$ and the zeros are $\{-1.28, -0.01\}$. Hence (37) is a stable and minimum phase third order approximation.

Table V shows that the (35) yielded by solving Problem 1 yields the best approximation error. Models (36) and (37) preserve desired poles and zeros, but with higher error. Furthermore, without optimization one must carefully select the desired constraints to achieve good results.

Third order model	H_2-norm of the error
Model (35), with a pole at -1.6 and a zero at -1.28, solution of Problem 1	$8.7 \cdot 10^{-3}$
Model (36), with poles at $\{-1.6, -0.16 - 0.55j, -0.16 + 0.55j\}$	$1.02 \cdot 10^{-2}$
Model (37) with two poles at $\{-1.6, -0.16\}$ and a zero at -1.28	$2.8 \cdot 10^{-1}$

TABLE I

H_2-NORMS OF THE APPROXIMATION ERRORS FOR DIFFERENT SCENARIOS.

VI. CONCLUSIONS

In this paper we have formulated an optimization problem with respect to H_2-norm minimal error approximation in a family of reduced order models that match a set of fixed ν moments. For this optimization problem we have derived first-order optimality conditions and a solution has been developed in terms of the gradient method. Using the cart controlled by a double pendulum benchmark example, we have shown the efficiency of our results.

REFERENCES

[1] A. C. Antoulas, *Approximation of large-scale dynamical systems*. Philadelphia: SIAM, 2005.
[2] ——, "A new result on passivity preserving model reduction," *Systems & Control Letters*, vol. 54, pp. 361–374, 2005.
[3] D. C. Sorensen, "Passivity preserving model reduction via interpolation of spectral zeros," *Systems & Control Letters*, vol. 54, pp. 347–360, 2005.
[4] S. Gugercin, A. C. Antoulas, and C. A. Beattie, "H_2 model reduction for large-scale dynamical systems," *SIAM J. Matrix Analysis & App.*, vol. 30, no. 2, pp. 609–638, 2008.
[5] B. Anic, C. A. Beattie, S. Gugercin, and A. C. Antoulas, "Interpolatory weighted-H_2 model reduction," *Automatica*, vol. 49, pp. 1275–1280, 2013.
[6] A. Astolfi, "Model reduction by moment matching for linear and nonlinear systems," *IEEE Trans. Autom. Contr.*, vol. 50, no. 10, pp. 2321–2336, 2010.
[7] T. C. Ionescu, A. Astolfi, and P. Colaneri, "Families of moment matching based, low order approximations for linear systems," *Systems & Control Letters*, vol. 64, pp. 47–56, 2014.
[8] K. Gallivan, A. Vandendorpe, and P. van Dooren, "Model reduction and the solution of Sylvester equations," in *Proc. 17th Int. Symposium on MTNS*, 2006.
[9] K. Gallivan, A. Vandendorpe, and P. V. Dooren, "Sylvester equations and projection based model reduction," *J. Comp. Appl. Math.*, vol. 162, pp. 213–229, 2004.
[10] T. C. Ionescu and A. Astolfi, "Families of moment matching based, structure preserving approximations for linear port Hamiltonian systems," *Automatica*, vol. 49, pp. 2424–2434, 2013.
[11] T. C. Ionescu, "Two-sided time-domain moment matching for linear systems," *IEEE Trans. Autom. Contr.*, vol. 61, no. 9, pp. 2632–2637, 2016.
[12] I. Necoara and T. Ionescu, "Optimal H_2 moment matching-based model reduction for linear systems by (non)convex optimization," 2018, preprint arXiv:1811.07409.
[13] ——, "H_2 model reduction of linear network systems by moment matching and optimization," *IEEE Trans. Autom. Contr.*, vol. 65, no. 12, 2020.
[14] S. Datta, D. Chakraborty, and B. Chaudhuri, "Partial pole placement with controller optimization," *IEEE Transactions on Automatic Control*, vol. 57, no. 4, pp. 1051–1056, 2012.
[15] O. V. Iftime and T. C. Ionescu, "On an approximation with prescribed zeros of SISO abstract boundary control systems," in *European Control Conf.*, 2013, pp. 2104–2109.
[16] E. de Souza and S. P. Bhattacharyya, "Controllability, observability and the solution of $AX - XB = C$," *Linear Algebra & Its App.*, vol. 39, pp. 167–188, 1981.
[17] K. Gallivan, A. Vandendorpe, and P. V. Dooren, "Model reduction of MIMO systems via tangential interpolation," *SIAM J. Matrix Anal. Appl.*, vol. 26, no. 2, pp. 328–349, 2004.
[18] A. C. Antoulas, "On pole placement in model reduction," *Automatisierungstechnik*, vol. 55, pp. 443–448, 2007.
[19] T. C. Ionescu and O. V. Iftime, "Moment matching with prescribed poles and zeros for infinite-dimensional systems," in *Proc. American Control Conf.*, 2012, pp. 1412–1417.
[20] H. Toivonen, "A globally convergent algorithm for the optimal constant output feedback problem," *Int. J. of Control*, vol. 41, pp. 1589–1599, 1985.
[21] Y. Nesterov, *Introductory Lectures on Convex Optimization: A Basic Course*. Kluwer, 2004.