Preprints of the 18th IFAC World Congress
Milano (Italy) August 28 - September 2, 2011

Gramian-based error bound in model reduction by Krylov subspace methods

Thomas Wolf *, **Heiko Panzer** *, **Boris Lohmann** *

* *Institute of Automatic Control, Technische Universität München*
Boltzmannstr. 15, D-85748 Garching, Germany
e-mail: {thomas.wolf, panzer, lohmann}@tum.de

Abstract: In this paper, a factorization of the error system resulting from model order reduction by Krylov subspace methods is presented. It is shown that the first factor is a system equal to the original one except for the input vector, while the second factor is of reduced order. This enables a new analysis of the error system: Provided that the Observability Gramian of the original system has once been calculated, an $\mathcal{H}_2$ error bound can be computed with negligible numerical effort for *any* reduced model resulting from Krylov subspace methods. The new results can therefore be ideally applied to SVD-Krylov methods where the Gramian is available anyway. In addition, it is proven that the bound even matches the exact error value for $\mathcal{H}_2$ optimal reduced order models. A numerical example is used to demonstrate the potential of the new error bound.

Keywords: Model reduction, Krylov subspace methods, Error estimation

1. INTRODUCTION

Model order reduction (MOR) aims at the approximation of a dynamical system by another model of reduced order. When the task is to investigate the dynamics, to design a controller or to optimize certain parameters of complex systems, this procedure is often inevitable, because accurate modeling often leads to large-scale models with a high number of ordinary differential equations. For the purpose of reducing linear time-invariant (LTI) dynamical systems in state-space realization, several MOR approaches have been shown to be well-suited, among them SVD-based methods, modal truncation and Krylov subspace methods, each of which features specific advantages and disadvantages.

Truncated Balanced Realization (TBR) [Moore (1981)] for instance is based on a state transformation that balances the state variables w.r.t. the energy flow from input to output. It preserves stability in the reduced system and provides an a priori $\mathcal{H}_\infty$ error bound on the approximation accuracy. However, two Lyapunov equations have to be solved for the calculation of the so-called *Controllability* and *Observability Gramian*, respectively, which is comparatively expensive and requires considerable numerical effort with increasing system order.

Krylov subspace methods [Villemagne and Skelton (1987)], on the other hand, have proven to be well applicable also for the reduction of very large-scale systems, but suffer – among other disadvantages, like a possible loss of stability – from the lack of a general error bound. Existing approaches towards this aim, presented e. g. in [Bai et al. (1999); Grimme (1997)], are only valid under special conditions or offer an approximation of the error. In [Bechtold et al. (2005)], heuristic error indicators are presented that are based on the observation that some system quantities of the original system are approximated by the reduced system; however, no theoretical proof is given. An error-bound for the reduction of a special class of Finite Element models is described in [Konkel et al. (2008)]; the method is though restricted to a single expansion point and orthogonal projection.

In this paper, the error of the approximation in Krylov based MOR is investigated. It will be shown that the error system can be factorized into a product of two systems, having the order of the original system and of the reduced system, respectively. This significantly facilitates further investigation of the error. Assuming that the Observability Gramian of the original system is known, this leads to a bound on the $\mathcal{H}_2$ error, computed by simple matrix-vector multiplication. Additionally, it is shown that for the case of $\mathcal{H}_2$ optimal reduction, the bound delivers the exact value of the error.

The outline of the paper is as follows: In section 2, some preliminaries and the problem formulation are presented, whereas the factorization of the error system as the main contribution is shown in section 3. The results are employed to derive the upper bound on the reduction error in section 4, while its suitability is demonstrated at a technical example in section 5.

2. PRELIMINARIES AND PROBLEM STATEMENT

In this section some preliminaries are given, together with a concise problem formulation.

2.1 Projection-based order reduction

Consider the single-input single-output (SISO) LTI dynamical system $G(s)$ in descriptor form

$$\begin{aligned} \mathbf{E}\,\dot{\mathbf{x}}(t) &= \mathbf{A}\,\mathbf{x}(t) + \mathbf{b}\,u(t), \\ y(t) &= \mathbf{c}\,\mathbf{x}(t), \end{aligned} \tag{1}$$

of order $n \in \mathbb{N}$, where $\mathbf{E},\ \mathbf{A} \in \mathbb{R}^{n\times n}$, $\mathbf{b} \in \mathbb{R}^n$ and $\mathbf{c} \in \mathbb{R}^{1\times n}$ are matrices with constant coefficients; $u(t)$, $y(t) \in \mathbb{R}$ and $\mathbf{x}(t) \in \mathbb{R}^n$ are, respectively, the input, output and state vector of the system. Assume the matrices $\mathbf{E}$ and $\mathbf{A}$ to be of full rank, $\det(\mathbf{E}) \neq 0$, $\det(\mathbf{A}) \neq 0$.

In projection-based model order reduction, the original state variables $\mathbf{x}(t)$ are approximated by a vector $\mathbf{x}_r(t) \in \mathbb{R}^q$ of reduced dimension $q \ll n$:

$$\mathbf{x}(t) \approx \mathbf{V}\mathbf{x}_r(t), \tag{2}$$

where $\mathbf{V} \in \mathbb{R}^{n\times q}$ is a suitable matrix of full column rank. The reduced model $G_r(s)$ to approximate the input-output behavior of system (1) is generally obtained by:

$$\begin{aligned} \overbrace{\mathbf{I}_q}^{\mathbf{E}_r}\,\dot{\mathbf{x}}_r(t) &= \overbrace{\mathbf{W}^T\mathbf{A}\mathbf{V}}^{\mathbf{A}_r}\,\mathbf{x}_r(t) + \overbrace{\mathbf{W}^T\mathbf{b}}^{\mathbf{b}_r}\,u(t), \\ y_r(t) &= \underbrace{\mathbf{c}\mathbf{V}}_{\mathbf{c}_r}\,\mathbf{x}_r(t), \end{aligned} \tag{3}$$

where $\mathbf{W} \in \mathbb{R}^{n\times q}$ is a matrix of full column rank such that $\mathbf{W}^T\mathbf{E}\mathbf{V} = \mathbf{I}_q \in \mathbb{R}^{q\times q}$ is identity. Note that this is not restrictive, since the input-output behavior of the reduced system (3) is uniquely determined by the subspace spanned by the columns of $\mathbf{V}$ and $\mathbf{W}$ – denoted as "$\mathrm{sp}(\mathbf{V})$" and "$\mathrm{sp}(\mathbf{W})$", respectively. Thus, it is always possible to find appropriate bases for $\mathrm{sp}(\mathbf{V})$ and $\mathrm{sp}(\mathbf{W})$, such that $\mathbf{W}^T\mathbf{E}\mathbf{V} = \mathbf{I}_q$ holds true. This way of model reduction involves a Petrov-Galerkin type projection [Grimme (1997)], where the projector $\boldsymbol{\mathcal{P}}$ is defined as

$$\boldsymbol{\mathcal{P}} = \mathbf{E}\mathbf{V}\mathbf{W}^T \tag{4}$$

and maps all vectors onto $\mathrm{sp}(\mathbf{E}\mathbf{V})$, orthogonally to $\mathrm{sp}(\mathbf{W})$. The projector thus allows to decompose an arbitrary vector $\mathbf{r} \in \mathbb{R}^n$ into one component $\boldsymbol{\mathcal{P}}\mathbf{r} \in \mathrm{sp}(\mathbf{E}\mathbf{V})$ and a remainder $\mathbf{r}_\perp := (\mathbf{I} - \boldsymbol{\mathcal{P}})\,\mathbf{r}$ closing the vector chain: $\mathbf{r} = \boldsymbol{\mathcal{P}}\mathbf{r} + \mathbf{r}_\perp$. The notation $\mathbf{r}_\perp$ is motivated by the fact that $\mathbf{W}^T\mathbf{r}_\perp = \mathbf{0}$ and thus $\mathbf{r}_\perp$ is orthogonal to $\mathrm{sp}(\mathbf{W})$.

2.2 System norms

Several norms are available to quantify dynamical systems, such as the $\mathcal{H}_2$ norm, defined as:

$$\|G(s)\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G(j\omega)|^2\,\mathrm{d}\omega. \tag{5}$$

In [Doyle et al. (1992)] a helpful way for calculating the $\mathcal{H}_2$ norm of a system $G(s)$ is presented

$$\|G(s)\|_2^2 = \mathbf{b}^T\mathbf{Q}\mathbf{b}, \tag{6}$$

where $\mathbf{Q} \in \mathbb{R}^{n\times n}$ is the *Observability Gramian* solving the generalized *Lyapunov*-equation:

$$\mathbf{A}^T\mathbf{Q}\mathbf{E} + \mathbf{E}^T\mathbf{Q}\mathbf{A} + \mathbf{c}^T\mathbf{c} = \mathbf{0}. \tag{7}$$

2.3 Krylov subspace methods

One possible approach for the calculation of $\mathbf{V}$ and $\mathbf{W}$ are the Krylov subspace methods, also known as *Moment Matching*. For the review of this method, some definitions are introduced. The moments $m_i^{s_0}$ of the dynamical system (1) are defined as the negative Taylor coefficients of the transfer function $G(s)$ around an expansion point s_0:

$$G(s) = \mathbf{c}\,(s\mathbf{E} - \mathbf{A})^{-1}\,\mathbf{b} = -\sum_{i=0}^{\infty} m_i^{s_0}\,(s - s_0)^i\,. \tag{8}$$

Associated with an expansion point s_0, two types of Krylov vectors are defined:

$$\mathbf{v}_K^j := \left((\mathbf{A} - s_0\mathbf{E})^{-1}\,\mathbf{E}\right)^{j-1}(\mathbf{A} - s_0\mathbf{E})^{-1}\,\mathbf{b}, \tag{9}$$

$$\mathbf{w}_K^j := \left((\mathbf{A} - s_0\mathbf{E})^{-T}\,\mathbf{E}^T\right)^{j-1}(\mathbf{A} - s_0\mathbf{E})^{-T}\,\mathbf{c}^T, \tag{10}$$

generating the input and output Krylov sequences, $\mathcal{K}_I^q$ and $\mathcal{K}_O^q$, respectively, of order q:

$$\mathcal{K}_I^q(s_0) := \mathrm{sp}\left\{\mathbf{v}_K^1, \mathbf{v}_K^2, \ldots, \mathbf{v}_K^q\right\} \tag{11}$$

$$\mathcal{K}_O^q(s_0) := \mathrm{sp}\left\{\mathbf{w}_K^1, \mathbf{w}_K^2, \ldots, \mathbf{w}_K^q\right\}. \tag{12}$$

Note that these Krylov sequences are usually numerically ill-conditioned. Typically, the *Arnoldi* algorithm employing a *Gram-Schmidt* orthogonalization is employed in order to calculate a basis for the subspace spanned by the Krylov sequences $\mathcal{K}_I^q$ and $\mathcal{K}_O^q$. The main result of moment matching is stated in the following theorem.

Theorem 1. [Villemagne and Skelton (1987)] Constructing the projection matrices $\mathbf{V}$ and $\mathbf{W}$ as Krylov subspaces

$$\mathrm{sp}\left\{\mathcal{K}_I^{q_1}(s_1), \ldots, \mathcal{K}_I^{q_L}(s_L)\right\} \subseteq \mathrm{sp}(\mathbf{V}), \tag{13}$$

$$\mathrm{sp}\left\{\mathcal{K}_O^{q_{L+1}}(s_{L+1}), \ldots, \mathcal{K}_O^{q_{2L}}(s_{2L})\right\} \subseteq \mathrm{sp}(\mathbf{W}), \tag{14}$$

the first $q_i,\ i = 1 \ldots 2L$ moments around the expansion points s_i of the original (1) and the reduced model (3) match. Altogether, $2q = \sum_{i=1}^{2L} q_i$ moments match, where $\mathbf{V}$ and $\mathbf{W}$ are of same rank $q = \sum_{i=1}^{L} q_i = \sum_{i=L+1}^{2L} q_i$.

2.4 $\mathcal{H}_2$ optimal reduction and SVD-Krylov methods

An important task in Krylov based MOR is the choice of expansion points. The ICOP algorithm [Eid (2009)] for instance is suitable for the computation of a single expansion point that minimizes a Laguerre-based optimality criterion. In [Gugercin et al. (2008)] the IRKA algorithm is introduced, giving a local solution to the $\mathcal{H}_2$ minimization problem. First-order necessary conditions for $\mathcal{H}_2$ optimality are given, namely that $G_r(s)$ interpolates $G(s)$ and its first derivative at the mirror images of the reduced system poles. The expansion points used to construct the projection matrices $\mathbf{V}$ and $\mathbf{W}$ are iteratively chosen in order to fulfill the required interpolation condition.

The ISRK algorithm [Gugercin (2008)] characterizes an *SVD-Krylov* method, where the matrix $\mathbf{V}$ is typically calculated as the basis of a Krylov subspace (13) and $\mathbf{W}$ is chosen such that $\mathrm{sp}(\mathbf{W}) = \mathrm{sp}(\mathbf{Q}\mathbf{E}\mathbf{V})$. This involves the Observability Gramian, which indicates the SVD-side of the method. By such an approach, stability is preserved in the reduced system yet the computation of the Gramian usually requires high numerical effort. Anyhow, the ISRK algorithm aims at stability preservation (as it is an SVD-Krylov method) in combination with $\mathcal{H}_2$ optimality, due to its analogy to the IRKA algorithm. However, the resulting reduced system $G_r(s)$ interpolates $G(s)$ only at the mirror images of the reduced poles and not its first derivative. Nevertheless, upon convergence, $G_r(s)$ is the best approximation in the $\mathcal{H}_2$ sense among all reduced systems sharing the same set of poles.

2.5 Approximation error

Due to the approximation of $G(s)$ by $G_r(s)$, an error $y_e(t)$ is induced, defined as

$$\begin{aligned} y_e(t) = &\quad y(t) - y_r(t) &= \\ = &\quad \mathbf{c}\mathbf{x}(t) - \mathbf{c}_r\mathbf{x}_r(t) &= \mathbf{c}\mathbf{e}(t), \end{aligned} \tag{15}$$

where the state error $\mathbf{e}(t)$ is defined as:

$$\mathbf{e}(t) = \mathbf{x}(t) - \mathbf{V}\mathbf{x}_r(t). \tag{16}$$

The error $y_e(t)$ can be formulated as the output of the state-space descriptor system

$$\begin{aligned} \begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_q \end{bmatrix} \begin{bmatrix} \dot{\mathbf{x}}(t) \\ \dot{\mathbf{x}}_r(t) \end{bmatrix} &= \begin{bmatrix} \mathbf{A} & \mathbf{0} \\ \mathbf{0} & \mathbf{A}_r \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_r(t) \end{bmatrix} + \begin{bmatrix} \mathbf{b} \\ \mathbf{b}_r \end{bmatrix} u(t), \\ y_e(t) &= \begin{bmatrix} \mathbf{c} & -\mathbf{c}_r \end{bmatrix} \begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_r(t) \end{bmatrix}, \end{aligned} \tag{17}$$

which is referred to as the *error system* $G_e(s)$ in the following. Therefore, in order to investigate the error in model order reduction, a system of order $n+q$ has to be studied, which becomes demanding in large-scale settings. In contrast to the Truncated Balanced Realization, there exists no general error-bound in Krylov subspace methods, up to the knowledge of the authors. Please note that methods like IRKA and ISRK result in locally $\mathcal{H}_2$ optimal reduced models, but the value of the $\mathcal{H}_2$ norm remains unknown. As a first result, in [Meier and Luenberger (1965)] it was shown that in the $\mathcal{H}_2$ optimal case (both IRKA and ISRK) the error norm is given as:

$$\|G_e(s)\|_2^2 = \|G(s)\|_2^2 - \|G_r(s)\|_2^2. \tag{18}$$

This theoretically allows to calculate the $\mathcal{H}_2$ error for the ISRK algorithm by $\|G_e(s)\|_2^2 = \mathbf{b}^T\mathbf{Q}\mathbf{b} - \mathbf{b}_r^T\mathbf{Q}_r\mathbf{b}_r$. However, please note that equation (18) holds true exclusively for an optimal reduced system. But then, in real applications the reduced system is iteratively generated and thus, due to approximate convergence, evaluating the error by equation (18) might become inaccurate.

2.6 The problem

To sum up, no general error-bound is known for model order reduction by Krylov subspace methods. Even for $\mathcal{H}_2$ optimal methods the value remains unknown. The theoretical possibility to evaluate the $\mathcal{H}_2$ error by equation (18) suffers from the prerequisite of full convergence. In this article a factorization of the error system is introduced, directing to facilitated investigation of the error system. As a first result, a numerically efficient $\mathcal{H}_2$ error bound for the ISRK algorithm in every iteration is derived that provides additional information on the convergence.

3. MAIN RESULTS

For the aim of this paper it is useful to introduce a state transformation to the error system $G_e(s)$ as

$$\begin{bmatrix} \mathbf{x}(t) \\ \mathbf{x}_r(t) \end{bmatrix} = \begin{bmatrix} \mathbf{I}_n & \mathbf{V} \\ \mathbf{0} & \mathbf{I}_q \end{bmatrix} \begin{bmatrix} \mathbf{e}(t) \\ \mathbf{x}_r(t) \end{bmatrix}, \tag{19}$$

resulting from definition (16). In addition, premultiplying equation (17) by the matrix

$$\mathbf{M} = \begin{bmatrix} \mathbf{I}_n & -\mathbf{E}\mathbf{V} \\ \mathbf{0} & \mathbf{I}_q \end{bmatrix} \tag{20}$$

from the left hand side, leads to

$$\begin{aligned} \overbrace{\begin{bmatrix} \mathbf{E} & \mathbf{0} \\ \mathbf{0} & \mathbf{I}_q \end{bmatrix}}^{\mathbf{E}_e} \begin{bmatrix} \dot{\mathbf{e}} \\ \dot{\mathbf{x}}_r \end{bmatrix} &= \overbrace{\begin{bmatrix} \mathbf{A} & (\mathbf{I}-\boldsymbol{\mathcal{P}})\,\mathbf{A}\mathbf{V} \\ \mathbf{0} & \mathbf{A}_r \end{bmatrix}}^{\mathbf{A}_e} \begin{bmatrix} \mathbf{e} \\ \mathbf{x}_r \end{bmatrix} + \overbrace{\begin{bmatrix} \mathbf{b}_\perp \\ \mathbf{b}_r \end{bmatrix}}^{\mathbf{b}_e} u(t), \\ y_e(t) &= \underbrace{\begin{bmatrix} \mathbf{c} & \mathbf{0} \end{bmatrix}}_{\mathbf{c}_e} \begin{bmatrix} \mathbf{e} \\ \mathbf{x}_r \end{bmatrix}, \end{aligned} \tag{21}$$

where $\mathbf{b}_\perp = (\mathbf{I}-\boldsymbol{\mathcal{P}})\,\mathbf{b}$. The benefit from this transformation is the decoupling of the states $\mathbf{x}_r(t)$, since they are not affected by $\mathbf{e}(t)$ (zero in the bottom left entry of $\mathbf{A}_e$ and $\mathbf{E}_e$) and do not directly operate on the output (zero in output-vector $\mathbf{c}_e$). They only affect the error $\mathbf{e}(t)$ by the top right entry of $\mathbf{A}_e$:

$$(\mathbf{I}-\boldsymbol{\mathcal{P}})\,\mathbf{A}\mathbf{V} = \mathbf{A}\mathbf{V} - \mathbf{E}\mathbf{V}\mathbf{A}_r. \tag{22}$$

The main contribution of this paper will be based on representation (21) of the error system. Towards this goal, the properties of (22) are further investigated.

Lemma 2. Let the matrix $\mathbf{V}$ form a basis of an input Krylov subspace (13). Then $(\mathbf{I}-\boldsymbol{\mathcal{P}})\,\mathbf{A}\mathbf{V} \in \mathbb{R}^{n\times q}$ is of rank one and $\mathrm{sp}\{(\mathbf{I}-\boldsymbol{\mathcal{P}})\,\mathbf{A}\mathbf{V}\} = \mathrm{sp}\{\mathbf{b}_\perp\}$.

Proof. The proof is first given for a Krylov sequence (11) involving one expansion point s_0. The proof for the general Krylov subspace (13) then follows by concatenation of Krylov sequences. Note that by the definition of the Krylov vectors $\mathbf{v}_K^1 = (\mathbf{A}-s_0\mathbf{E})^{-1}\,\mathbf{b}$ and $\mathbf{v}_K^i = (\mathbf{A}-s_0\mathbf{E})^{-1}\,\mathbf{E}\mathbf{v}_K^{i-1}$, which leads to:

$$\mathbf{A}\mathbf{v}_K^1 = \mathbf{b} + s_0\mathbf{E}\mathbf{v}_K^1, \tag{23}$$

$$\mathbf{A}\mathbf{v}_K^i = \mathbf{E}\mathbf{v}_K^{i-1} + s_0\mathbf{E}\mathbf{v}_K^i. \tag{24}$$

For a Krylov sequence (11), equation (23) holds true, together with $q-1$ equations (24). Collecting these q equations into a matrix-vector product yields:

$$\begin{aligned} \mathbf{A}\underbrace{\begin{bmatrix} \mathbf{v}_K^1 \dots \mathbf{v}_K^q \end{bmatrix}}_{\mathbf{V}_K} &= \mathbf{b}\begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix} + \\ &+ \mathbf{E}\underbrace{\begin{bmatrix} \mathbf{v}_K^1 \dots \mathbf{v}_K^q \end{bmatrix}}_{\mathbf{V}_K} \underbrace{\begin{bmatrix} s_0 & 1 & & \\ & s_0 & \ddots & \\ & & \ddots & 1 \\ & & & s_0 \end{bmatrix}}_{\mathbf{S}}. \end{aligned} \tag{25}$$

In general, $\mathbf{V}$ does not explicitly contain the sequence of Krylov vectors $\mathbf{v}_K^i$, but is connected to $\mathbf{V}_K$ by a full rank matrix $\mathbf{T} \in \mathbb{R}^{q\times q}$:

$$\mathbf{V} = \mathbf{V}_K\mathbf{T}, \qquad \mathbf{V}_K = \mathbf{V}\mathbf{T}^{-1}. \tag{26}$$

Inserting (26) into equation (25) and multiplying from the right by $\mathbf{T}$, leads to:

$$\mathbf{A}\mathbf{V} = \mathbf{b}\widetilde{\mathbf{c}_r} + \mathbf{E}\mathbf{V}\mathbf{T}^{-1}\mathbf{S}\mathbf{T}, \tag{27}$$

where $\widetilde{\mathbf{c}_r} := \begin{bmatrix} 1 & 0 & \dots & 0 \end{bmatrix}\mathbf{T} \in \mathbb{R}^{1\times q}$. Due to the projection onto $\mathrm{sp}(\mathbf{E}\mathbf{V})$, $\boldsymbol{\mathcal{P}}\mathbf{E}\mathbf{V} = \mathbf{E}\mathbf{V}$ holds true, and multiplying (27) from the left by $(\mathbf{I}-\boldsymbol{\mathcal{P}})$ results in

$$(\mathbf{I}-\boldsymbol{\mathcal{P}})\,\mathbf{A}\mathbf{V} = (\mathbf{I}-\boldsymbol{\mathcal{P}})\,\mathbf{b}\widetilde{\mathbf{c}_r} + \mathbf{0} = \mathbf{b}_\perp\widetilde{\mathbf{c}_r}, \tag{28}$$

and completes the proof for the Krylov sequence (11). Note that equation (25) holds true for any Krylov sequence $\mathbf{V}_{K,i}$ of order q_i associated with the expansion point s_i. Thus,

concatenating the equations (25) for $i = 1 \dots L$ different Krylov sequences leads to:

$$\begin{aligned}\mathbf{A}\left[\mathbf{V}_{K,1} \dots \mathbf{V}_{K,L}\right] = \mathbf{b}\left[\mathbf{e}_1^{q_1} \dots \mathbf{e}_1^{q_L}\right] + \\ + \mathbf{E}\left[\mathbf{V}_{K,1} \dots \mathbf{V}_{K,L}\right] \operatorname{diag}\left(\mathbf{S}_1 \dots \mathbf{S}_L\right),\end{aligned} \tag{29}$$

where $\mathbf{e}_1^{q_i} \in \mathbb{R}^{1 \times q_i}$ has "1" as the first entry and zeros elsewhere. The proof for a general Krylov subspace (13) is completed similarly by multiplying from the right a suitably chosen matrix $\mathbf{T}$ and from the left $(\mathbf{I} - \boldsymbol{\mathcal{P}})$.

Remark 3. Note that the matrix $\mathbf{T}$ does not have to be known for the calculation of $\widetilde{\mathbf{c}_r}$. Instead, applying the pseudo inverse of $\mathbf{b}_\perp$ to (28) yields:

$$\widetilde{\mathbf{c}_r} = \frac{1}{\mathbf{b}_\perp^T \mathbf{b}_\perp} \mathbf{b}_\perp^T \left(\mathbf{I} - \boldsymbol{\mathcal{P}}\right) \mathbf{A}\mathbf{V}. \tag{30}$$

3.1 Factorization of the error system

Using the fact that the columns of $(\mathbf{I} - \boldsymbol{\mathcal{P}})\,\mathbf{A}\mathbf{V}$ are multiples of $\mathbf{b}_\perp$, it is possible to state the main result of this paper in the following theorem.

Theorem 4. Let the matrix $\mathbf{V}$ form a basis of an input Krylov subspace as defined in (13). The error system $G_e(s)$ can then be factorized into a product of two systems of order n and q, respectively:

$$G_e(s) = G_{\mathbf{b}_\perp}(s) \cdot G_{\widetilde{\mathbf{c}_r}}(s), \tag{31}$$

where

$$G_{\mathbf{b}_\perp}(s) := \mathbf{c}\left(s\mathbf{E} - \mathbf{A}\right)^{-1} \mathbf{b}_\perp, \tag{32}$$

$$G_{\widetilde{\mathbf{c}_r}}(s) := \widetilde{\mathbf{c}_r}\left(s\mathbf{I}_r - \mathbf{A}_r\right)^{-1} \mathbf{b}_r + 1. \tag{33}$$

Proof. The proof follows from the state space representation (21) of the error system. The second row of the state equation defines the dynamics of the reduced states by $\mathcal{X}_r(s) = \left(s\mathbf{I}_r - \mathbf{A}_r\right)^{-1} \mathbf{b}_r \mathcal{U}(s)$, where $\mathcal{X}_r(s)$ denotes the Laplace-transform of $\mathbf{x}_r(t)$. Using Lemma 2, the dynamics of the error $\mathbf{e}(t)$ is given by the first row:

$$\mathcal{E}(s) = \left(s\mathbf{E} - \mathbf{A}\right)^{-1} \cdot \left[\mathbf{b}_\perp \widetilde{\mathbf{c}_r} \mathcal{X}_r(s) + \mathbf{b}_\perp \mathcal{U}(s)\right] \tag{34}$$

$$= \left(s\mathbf{E} - \mathbf{A}\right)^{-1} \mathbf{b}_\perp \cdot \underbrace{\left[\widetilde{\mathbf{c}_r} \frac{\mathcal{X}_r(s)}{\mathcal{U}(s)} + 1\right]}_{G_{\widetilde{\mathbf{c}_r}}} \cdot \mathcal{U}(s) \tag{35}$$

Using the definition of the output of the error system $\mathcal{Y}_e(s) = \mathbf{c}\,\mathcal{E}(s)$ completes the proof.

It is worth to mention that the system $G_{\widetilde{\mathbf{c}_r}}(s)$ is of order q and thus its dynamics can be easily investigated. Its output is the input to the system $G_{\mathbf{b}_\perp}(s)$ which shares $\mathbf{E}$, $\mathbf{A}$ and $\mathbf{c}$ with the original system $G(s)$ and only differs from it in its input vector. The output of $G_{\mathbf{b}_\perp}(s)$ then defines the overall output error $y_e(t)$ resulting from the approximation. Some properties of the factorized error system are investigated in the following theorem.

Theorem 5. Let the matrices $\mathbf{V}$ and $\mathbf{W}$ form bases of input and output Krylov subspaces as defined in (13) and (14). The expansion points used to construct $\mathbf{V}$ are the transfer zeros of the system $G_{\widetilde{\mathbf{c}_r}}(s)$ and the expansion points used to construct $\mathbf{W}$ are a subset of the transfer zeros of $G_{\mathbf{b}_\perp}(s)$.

Proof. For the proof of the first part, recall equation (26), which implies by the help of the pseudo inverse:

$$\mathbf{T} = \left(\mathbf{V}_K^T \mathbf{V}_K\right)^{-1} \mathbf{V}_K^T \mathbf{V}. \tag{36}$$

From Lemma 3.2 in [Grimme (1997)] it is true that

$$\begin{aligned}\mathbf{V}\left[\left(\mathbf{A}_r - s_0\mathbf{E}_r\right)^{-1}\mathbf{E}_r\right]^{i-1}\left(\mathbf{A}_r - s_0\mathbf{E}_r\right)^{-1}\mathbf{b}_r = \\ = \underbrace{\left[\left(\mathbf{A} - s_0\mathbf{E}\right)^{-1}\mathbf{E}\right]^{i-1}\left(\mathbf{A} - s_0\mathbf{E}\right)^{-1}\mathbf{b}}_{\mathbf{v}_K^i = \mathbf{V}_K\mathbf{e}_i},\end{aligned} \tag{37}$$

for any expansion point s_0 of $\mathbf{V}$. This means that the left hand side of equation (37) equals a certain Krylov vector $\mathbf{v}_K^i$, which can be written in terms of the full Krylov matrix $\mathbf{V}_K$ by $\mathbf{v}_K^i = \mathbf{V}_K\mathbf{e}_i$, where $\mathbf{e}_i$ denotes the i-th unit column vector of appropriate length. Using $\widetilde{\mathbf{c}_r} := \left[1\ 0\ \dots\ 0\right]\mathbf{T}$, the transfer function $G_{\widetilde{\mathbf{c}_r}}(s_0)$ is given by:

$$G_{\widetilde{\mathbf{c}_r}}(s_0) = 1 + \widetilde{\mathbf{c}_r}\left(s_0\mathbf{I}_r - \mathbf{A}_r\right)^{-1}\mathbf{b}_r \tag{38}$$

$$= 1 - \left[1\ 0\ \dots\ 0\right]\left(\mathbf{V}_K^T\mathbf{V}_K\right)^{-1} \cdot \cdot \mathbf{V}_K^T \underbrace{\mathbf{V}\left(\mathbf{A}_r - s_0\mathbf{I}_r\right)^{-1}\mathbf{b}_r}_{\mathbf{v}_K^1 = \mathbf{V}_K\mathbf{e}_1}, \tag{39}$$

$$= 1 - \left[1\ 0\ \dots\ 0\right]\mathbf{e}_1 = 0. \tag{40}$$

Thus, the system $G_{\widetilde{\mathbf{c}_r}}(s)$ has a transfer zero at the expansion point s_0. In order to prove that $G_{\widetilde{\mathbf{c}_r}}(s)$ has as many transfer zeros as moments are matched around s_0, the i-th derivative $G_{\widetilde{\mathbf{c}_r}}^{(i)}(s_0)$ for $i = 2 \dots (q-1)$ at s_0 is computed as:

$$G_{\widetilde{\mathbf{c}_r}}^{(i)}(s_0) = -s_0^i\left[1\ 0\ \dots\ 0\right]\left(\mathbf{V}_K^T\mathbf{V}_K\right)^{-1}\mathbf{V}_K^T \cdot \cdot \underbrace{\mathbf{V}\left[\left(s_0\mathbf{I}_r - \mathbf{A}_r\right)^{-1}\mathbf{E}_r\right]^i\left(s_0\mathbf{I}_r - \mathbf{A}_r\right)^{-1}\mathbf{b}_r}_{\mathbf{v}_K^{i+1} = \mathbf{V}_K\mathbf{e}_{i+1}}, \tag{41}$$

$$= -s_0^i\left[1\ 0\ \dots\ 0\right]\mathbf{e}_{i+1} = 0. \tag{42}$$

This completes the proof of the first part for a Krylov sequence (11). The general case (13) follows accordingly to the proof of Lemma 2 and hence is omitted. For the proof of the second part, note that due to moment matching, the original and the reduced system share the same value and $q_i - 1$ derivatives of their transfer function at the expansion points s_i. Thus, every expansion point used to construct $\mathbf{V}$ or $\mathbf{W}$, respectively, results in a transfer zero of the error system $G_e(s)$ with multiplicity q_i. As already proven, the expansion points of $\mathbf{V}$ are in a one-to-one relation to the transfer zeros of the system $G_{\widetilde{\mathbf{c}_r}}(s)$. Hence, the system $G_{\mathbf{b}_\perp}(s)$ has to have transfer zeros at the expansion points of $\mathbf{W}$ with associated multiplicity.

4. APPLICATIONS

In this section a general error bound for model order reduction by Krylov subspace methods is introduced, based on the main results from section 3. It is stated in which settings this bound is reasonable to compute, along with the special case of $\mathcal{H}_2$ optimality, when the upper bound coincides with the exact value.

4.1 General Error Bound

The following Lemma gives an upper bound on the $\mathcal{H}_2$ error resulting from Krylov subspace methods.

Lemma 6. Let the error-system be factorized as introduced in Theorem 4. An upper bound on the $\mathcal{H}_2$ norm of the error system is given by:

$$\|G_{\mathbf{e}}(s)\|_2 \leq \|G_{\mathbf{b}_\perp}(s)\|_2 \cdot \left\|G_{\widetilde{\mathbf{c}_r}}(s)\right\|_\infty . \tag{43}$$

Proof. From Theorem 4 it is true that $\|G_{\mathbf{e}}(s)\|_2 = \left\|G_{\mathbf{b}_\perp}(s) \cdot G_{\widetilde{\mathbf{c}_r}}(s)\right\|_2$. Then,

$$\|G_{\mathbf{e}}(s)\|_2^2 = \frac{1}{2\pi}\int_{-\infty}^{\infty} |G_{\mathbf{b}_\perp}(j\omega)|^2 \cdot \left|G_{\widetilde{\mathbf{c}_r}}(j\omega)\right|^2 \mathrm{d}\omega \tag{44}$$

$$\leq \frac{1}{2\pi}\int_{-\infty}^{\infty} |G_{\mathbf{b}_\perp}(j\omega)|^2 \,\mathrm{d}\omega \cdot \left\|G_{\widetilde{\mathbf{c}_r}}(s)\right\|_\infty^2 \tag{45}$$

$$= \|G_{\mathbf{b}_\perp}(s)\|_2^2 \cdot \left\|G_{\widetilde{\mathbf{c}_r}}(s)\right\|_\infty^2 , \tag{46}$$

which completes the proof.

Using equation (6), an upper bound on the relative $\mathcal{H}_2$ error can be given as:

$$\frac{\|G_{\mathbf{e}}(s)\|_2}{\|G(s)\|_2} \leq \sqrt{\frac{\mathbf{b}_\perp^T \mathbf{Q} \mathbf{b}_\perp}{\mathbf{b}^T \mathbf{Q} \mathbf{b}}} \cdot \left\|G_{\widetilde{\mathbf{c}_r}}(s)\right\|_\infty . \tag{47}$$

However, in order to calculate this newly introduced error bound (47), one Lyapunov equation (7) of order n has to be solved, which might not be reasonable due to its numerical effort. Nevertheless, taking into account that the solution $\mathbf{Q}$ to the Lyapunov equation (7) is independent from the reduced model, its computation becomes more justified: Calculating $\mathbf{Q}$ *once* allows to give an error-bound for *any* reduction, irrespective of the choice of expansion points or the reduced order. Only simple matrix-vector products and the computation of the $\mathcal{H}_\infty$ norm of a system of reduced order q is then required. All in all, in iterative methods or in settings when multiple reductions are performed, the error-bound (47) is well applicable, since the costs for calculating $\mathbf{Q}$ are shared by the different reductions. One possible scenario would be to reduce the system to very low order, and then increment the order of the reduced system until the $\mathcal{H}_2$ error bound falls below a desired value. Due to efficient error estimation in every iteration, this can be seen as a first step towards an automatic algorithm for model reduction by Krylov subspace methods.

4.2 SVD-Krylov methods

The new bound is particularly suitable for model reduction by SVD-Krylov methods, since it is assumed that with these methods the Observability Gramian $\mathbf{Q}$ is already available.

4.3 $\mathcal{H}_2$ optimal case

Theorem 7. Let $G_r(s)$ be the $\mathcal{H}_2$ optimum in the set of reduced models sharing the same reduced poles. Then the upper bound introduced in Lemma 6 matches the exact value of the error, given by:

$$\|G_{\mathbf{e}}(s)\|_2 = \|G_{\mathbf{b}_\perp}(s)\|_2 = \sqrt{\mathbf{b}_\perp^T \mathbf{Q} \mathbf{b}_\perp}. \tag{48}$$

Proof. A necessary condition for $\mathcal{H}_2$ optimality in a set of reduced system poles is that these poles are equal to the mirror images of the expansion points of $\mathbf{V}$ [Gugercin et al. (2008); Gugercin (2008)]. Thus, due to Theorem 5, the transfer zeros of the system $G_{\widetilde{\mathbf{c}_r}}(s)$ are the mirror images of its poles, and hence the system becomes an all-pass filter with a magnitude of 0 dB:

$$\left|G_{\widetilde{\mathbf{c}_r}}(j\omega)\right| = 1, \quad \forall \omega \in \mathbb{R}. \tag{49}$$

Therefore, the inequality in the proof of Lemma 6 becomes an equality and $\left\|G_{\widetilde{\mathbf{c}_r}}(s)\right\|_\infty = 1$ completes the proof.

Remark 8. Please note that the smallest possible value for $\left\|G_{\widetilde{\mathbf{c}_r}}(s)\right\|_\infty$ is 1, due to the feedthrough of the system. Simulations have shown that even for the non-optimal case, the system $G_{\widetilde{\mathbf{c}_r}}(s)$ is often close to an all-pass filter. Thus, evaluating the bode plot of $G_{\widetilde{\mathbf{c}_r}}(s)$ gives an indication on the overestimation by the upper bound (47). The closer $G_{\widetilde{\mathbf{c}_r}}(s)$ is to an all-pass filter with a magnitude of 0 dB, the tighter the bound is.

Taking the facts above into account, the newly introduced error bound is ideally applicable to the ISRK algorithm. On the one hand, the Observability Gramian $\mathbf{Q}$ of the original system is calculated anyway. On the other hand, upon convergence, the relative $\mathcal{H}_2$ error is exactly given by:

$$\frac{\|G_{\mathbf{e}}(s)\|_2}{\|G(s)\|_2} = \sqrt{\frac{\mathbf{b}_\perp^T \mathbf{Q} \mathbf{b}_\perp}{\mathbf{b}^T \mathbf{Q} \mathbf{b}}}. \tag{50}$$

Remark 9. Note that in this case the system $G_{\widetilde{\mathbf{c}_r}}(s)$ still can be investigated. By verifying its all-pass character, information on the convergence of the ISRK algorithm is delivered, which would not be available by direct evaluation of the error from equation (18).

5. TECHNICAL EXAMPLE

The performance and practical value of the newly introduced error bound is demonstrated at a model from the Oberwolfach Benchmark Collection [Oberwolfach (2003)]. It describes the heat flow within a steel profile and can be used for the optimization of an industrial cooling process as different control laws and cooling strategies can be precalculated to predict the temperature distribution in the material. The system consists of $n = 5177$ ordinary differential equations (1) with 35185 and 35241 non-zero elements in $\mathbf{A}$ and $\mathbf{E}$, respectively.

Consider the following task: An approximation to the original system shall be found while guaranteeing a relative $\mathcal{H}_2$ error below 1e-4. Therefore, two SVD-Krylov algorithms are employed: i) ISRK, which iteratively choses the expansion points in $\mathbf{V}$ at the mirror images of the reduced poles, and ii) ICOP, which matches q moments about a single expansion point resulting from a Laguerre-based optimality criterion [Eid (2009)]. Either way, $\mathbf{Q}$ is firstly calculated for the computation of $\mathbf{W}$ which allows for a fast evaluation of the new error bound. Then, starting from $q = 1$, the order of the reduced system is incremented iteratively until the relative upper bound on the $\mathcal{H}_2$ norm falls below 1e-4. This procedure is expressed in the following algorithm.

Algorithm 1: Error Bound by Inexpensive Factorization (EB-IFAC)

1. Solve the generalized Lyapunov equation (7) for the Observability Gramian $\mathbf{Q}$. Set $q := 1$.
2. Choose initial expansion points.

3.a) Compute the Krylov Subspace $\mathbf{V}$.
3.b) Compute $\mathbf{W} := \mathbf{QEV}$, perform reduction.
3.c) Determine new expansion points and return to 3.a) until ISRK/ICOP converges.

4. Compute $\mathbf{b}_\perp = \mathbf{b} - \mathbf{EVb}_r$ and $\widetilde{\mathbf{c}_r}$ by equation (30). Evaluate error bound.
5. If error bound $<$ 1e-4: Stop.
   Else: Increment q and return to step 2.

Simulation results can be seen in Table 1. As both ISRK and ICOP reach the objective for the first time by $q = 7$, all values are given for that reduced order, only. In each case, the computation of the Observability Gramian takes about $201sec$ - much longer than the iterative calculation of the Krylov subspaces, the reduction processes and the numerous evaluations of the error bound together.

The authors would like to emphasize that calculating the error norm using the classical representation of $G_e(s)$ (17) would have taken more than $200sec$ in *each* iteration!

For ISRK, $G_{\widetilde{c_r}}(s)$ is an all-pass filter and the error bound in fact delivers the exact $\mathcal{H}_2$ error. In contrast, the error resulting from SVD-ICOP is slightly overestimated by the bound, as $G_{\widetilde{c_r}}(s)$ is not of all-pass type; however, the real error is actually smaller than for ISRK. In addition, ICOP outperforms ISRK by means of simulation time, as it requires only one LU-factorization and exhibits faster convergence per iteration in this example.

For comparison, reduction results from the purely Krylov based IRKA algorithm and from TBR are additionally given in Table 1. Both methods deliver comparably good approximations for $q = 7$. Please note, however, that no error bound is given for IRKA, since no Gramian is computed by the algorithm; this is in fact the reason for the comparably short reduction time. Finally, note that no $\mathcal{H}_2$ error bound is available for TBR, as the presented factorization is only applicable to Krylov subspace methods.

	relative $\mathcal{H}_2$ bound	real relative $\mathcal{H}_2$ error	$\|G_{\widetilde{c_r}}\|_\infty$	time $[sec]$
ISRK	2.13e-5	2.13e-5	1.00	201+11.4
ICOP	6.82e-5	4.15e-6	1.00	201+3.0
IRKA	-	2.12e-5	1.00	5.21
TBR	-	2.16e-5	-	1440

Table 1. Simulation results

6. CONCLUSIONS AND FUTURE WORK

A new factorization of the error system in model order reduction for SISO LTI systems by means of Krylov subspace methods has been presented in this paper. The partitioning allows to give a general upper bound on the $\mathcal{H}_2$ error and can drastically reduce the effort of its computation. In case an $\mathcal{H}_2$ optimal reduced system is found, the bound actually matches the exact error. Assuming that the Observability Gramian of the original system is known, the evaluation of the $\mathcal{H}_2$ error bound is numerically inexpensive. Due to these properties, this contribution is particularly well-suited for the ISRK algorithm.

The generalization to the MIMO case as well as to matching the *Markov* parameters instead of moments turns out to be feasible and will be discussed in a separate work. In addition, the existence of a dual factorization based on an output Krylov Subspace will be presented.

ACKNOWLEDGEMENTS

The second author of this work has been supported by the Cusanuswerk Scholarship Award Programme, which is gratefully acknowledged.

REFERENCES

Bai, Z., Slone, R.D., Smith, W.T., and Ye, Q. (1999). Error Bound for Reduced System Model by Padé Approximation via the Lanczos Process. *IEEE trans. Coputer-Aided Design of ICs and systems*, 18(2), 133–141.

Bechtold, T., Rudnyi, E.B., and Korvink, J.G. (2005). Error indicators for fully automatic extraction of heat-transfer macromodels for mems. *Journal of Micromechanics and Microengineering*, 15, 430–440.

Doyle, J.C., Francis, B., and Tannembaum, A. (1992). *Feedback Control Theory*. Maxwell/Macmillan.

Eid, R. (2009). *Time domain model reduction by moment matching*. PhD thesis, Institute of Automatic Control, Technische Universität München.

Grimme, E.J. (1997). *Krylov Projection Methods for Model Reduction*. PhD thesis, Dep. of Electrical Eng., Uni. Illinois at Urbana Champaign.

Gugercin, S., Antoulas, A.C., and Beattie, C. (2008). H2 model reduction for large-scale dynamical systems. *SIAM Journal on Matrix Analysis and Applications,*, 30(2), 609–638.

Gugercin, S. (2008). An iterative svd-krylov based method for model reduction of large-scale dynamical systems. *Lineal Algebra and its Applications*, 428, 1964–1986.

Konkel, Y., Farle, O., and Dyczij-Edlinger, R. (2008). Ein Fehlerschätzer für die Krylov-Unterraum basierte Ordnungsreduktion zeitharmonischer Anregungsprobleme. In *Modellbildung, Identifikation und Simulation in der Automatisierungstechnik*. Tagungsband GMA-Fachausschuss 1.30.

Meier, L. and Luenberger, D.G. (1965). Approximation of linear constant systems. *Preprints 1966 Joint Automatic Control Conference*, 728734. Seattle, August 1966.

Moore, B.C. (1981). Principal Component Analysis in Linear Systems: Controllability, Observability and Model Reduction. *IEEE Trans. Aut. Control*, AC-26, 17–32.

Oberwolfach (2003). *Oberwolfach Model Reduction Benchmark Collection*. available online at http://www.imtek.uni-freiburg.de/simulation/benchmark/.

Villemagne, C.D. and Skelton, R.E. (1987). Model Reduction using a Projection Formulation. *Int. J. Control*, 46, 2141–2169.