



MAX-PLANCK-GESELLSCHAFT

**Max Planck Institute Magdeburg
Preprints**

Lihong Feng Athanasios C. Antoulas Peter Benner

**Some a posteriori error bounds for reduced
order modelling of (non-)parametrized
linear systems**



Abstract

We propose a posteriori error bounds for reduced-order models of non-parametrized linear time invariant (LTI) systems and parametrized LTI systems. The error bounds estimate the errors of the transfer functions of the reduced models, and are independent of the model reduction methods used. It is shown that for some special non-parametrized LTI systems, particularly efficiently computable error bounds can be derived. According to the error bounds, reduced models of both non-parametric and parametric systems, computed by Krylov subspace based model reduction methods, can be obtained automatically and reliably. Simulations for several examples from engineering applications have demonstrated the robustness of the error bounds.

Imprint:

Max Planck Institute for Dynamics of Complex Technical Systems, Magdeburg

Publisher:

Max Planck Institute for
Dynamics of Complex Technical Systems

Address:

Max Planck Institute for
Dynamics of Complex Technical Systems
Sandtorstr. 1
39106 Magdeburg

<http://www.mpi-magdeburg.mpg.de/preprints/>

1 Introduction

The technique of model order reduction (MOR) has been successfully applied in many fields e.g. mechanical engineering, fluid dynamics, control, circuit simulation, microelectromechanical systems (MEMS) simulation etc.. The robustness of MOR has been revealed in all the above applications areas.

The purpose of MOR is to reduce the number of degrees of freedom in the original large-scale systems described by algebraic equations, ordinary differential equations (ODEs), or differential algebraic equations (DAEs) while attaining good accuracy. These systems usually come from (time-)spatial discretization of partial differential equations describing the underlying process, devices, structure or dynamics, etc.. Sometimes, the mathematical models are described directly by ODEs/DAEs, for example, the many models obtained based on modified nodal analysis (MNA) in circuit or MEMS simulation.

Parametric model order reduction (PMOR) is an advanced MOR technique for more complex mathematical models, where some variables, called parameters, are entries of the system matrices that are allowed to vary, such that the systems are parametrized. For a parametrized system, PMOR methods aim to preserve the parameters as symbolic quantities in the reduced models, such that a single reduced model is sufficiently accurate for all possible variations of the parameters.

Krylov subspace based moment-matching MOR and Gramian based MOR are popular MOR methods for non-parametrized LTI systems. The very basic method of Gramian based MOR is balanced truncation, which is well-known for its global error bound. Recent algorithmic progress has made this method applicable to truly large-scale systems, see, e.g., [7]. As these advances include certain approximations, the global error bound is not computable exactly, and therefore should be confirmed by an a posteriori error bound. Moment-matching MOR methods are computationally efficient, and are widely used in large-scale problems arising from circuit or MEMS simulation. However, they suffer from the lack of a global error bound, which leads to the fact that the reduced model cannot be generated automatically and reliably.

Some attempts have been made for getting error estimation for moment-matching MOR methods applied to non-parametrized LTI systems [8, 13, 20, 31, 27, 41]. While showing the efficiency of their error estimators, these are more or less heuristics [8, 13, 20, 27, 31]. Based on system theory, an error bound is derived in [41], but faces the high computational complexity. The residual of the state vector is simply used in [27] as the error estimator of the reduced-order model. All these error estimators are limited to non-parametrized systems. An a posteriori error bound for parametrized LTI systems is proposed in time domain in [28]. Although it is stated that it can be seen as a posteriori error bound for the Krylov subspace based method (e.g. moment-matching MOR), it is hardly computable

In recent years, numerous model order reduction methods for parametrized LTI systems have been developed, for example, the Krylov subspace based (multi-moment matching) PMOR methods [16, 18, 19, 21], the interpolation based PMOR methods [3, 5, 6, 35], the Loewner approach to parametric model reduction [32], and the reduced basis methods [34, 15]. A survey of PMOR methods can be found in [10]. Among these methods, only for the reduced basis method a posteriori error bounds are known. These enable automatic generation of a reliable reduced parametrized model.

Error bounds/estimators have been intensively studied on the reduced basis method for parametrized systems. Many error estimators developed for the reduced basis methods estimate the error in the state vectors (field variables) [38, 29, 24, 30], not for the outputs of the systems. In many applications, the output or the transfer function (output in the frequency domain) of the system are of interest. The error estimations for the state vectors often tends to overestimate the output errors. Nevertheless, output error estimators for reduced basis methods are proposed in [37, 40], which are only applicable to steady state systems. In [25], output error bound for linear parabolic equations is proposed, which estimates the output error of the reduced-order model in time domain. Output error bounds in time domain are also introduced in [45, 46, 47] based on space-time variational formulation of the original system. However, direct application of those time-domain error bounds to the frequency-domain PMOR methods, such as the Krylov subspace based multi-moment matching PMOR methods [16, 18, 19, 21], is unclear. Typically, almost all the error bounds for the reduced basis methods necessitate the bilinear forms of the PDE models [25, 24, 37, 40, 34, 38, 29, 30, 45, 46, 47].

The above observations motivate us to derive output error bounds for the dynamic systems in the discretized vector space. We propose several a posteriori output error bounds for the reduced models of both, non-parametrized LTI systems and parametrized LTI systems. The error bounds are the bounds for the difference between the transfer function of the original system and that of the reduced model, and are applicable to any MOR or PMOR methods based on approximation/interpolation of the transfer function, including the Krylov subspace based (multi-)moment matching methods [16, 18, 19, 21].

The basic idea originates from the output error bounds proposed for the reduced basis methods [37, 15, 40]. The

main theoretical contributions of the newly derived error bounds are firstly, the error bounds are independent of the discretization method (finite difference, finite element, finite volume) applied to the original PDEs. Secondly, the error bounds can be directly used in the discretized vector space, without going back to the PDEs, and especially to the bilinear form (weak formulation) associated with the finite element discretization. This is typically useful when only discretized systems are available in some situations. In particular, most of the dynamic models in circuit and MEMS simulation are derived using commercial software, where the usable mathematical models appear directly as ordinary differential equations (ODEs), or differential algebraic equations (DAEs). The bilinear form of the PDE models are usually unknown.

Technically, the output error bounds provide a way of automatically generating reliable reduced models computed by the Krylov subspace based (P)MOR methods, which is desired in design automation for circuits and MEMS. Although Krylov subspace based (P)MOR methods have been integrated into some simulation tools [39], the reduced model cannot be guaranteed to satisfy the required accuracy due to the lack of an robust error bound. We are making the design automation reliable by proposing some a posteriori output error bounds for both non-parametrized and parametrized linear systems.

The paper is organized as follows. In the next section, we review the general structure of projection based MOR and the transfer functions of the LTI systems. In Section 3, we introduce an a posteriori error bound for non-parametrized LTI systems with symmetric system matrices. A posteriori error bounds for general non-parametrized LTI systems and parametrized LTI systems are proposed in Section 4 and Section 5. How to efficiently compute the error bounds is discussed in Section 6. Section 7 relates the error bound analysis in Section 4 and Section 5 to a new reduction method. In the section that follows, the basic idea of Krylov subspace based MOR methods, also called moment-matching MOR/multi-moment matching PMOR, are reviewed. One will see that automatic generation of reduced models relies on adaptive selection of expansion points. Algorithms for automatic selection of expansion points according to the a posteriori error bounds are proposed. Simulation results are presented in Section 9. Conclusions and future work are given in the end.

2 Preliminaries

2.1 MOR for non-parametrized LTI systems

In general, a non-parametrized LTI system is described by

$$\begin{aligned} E \frac{d}{dt} x(t) &= Ax(t) + Bu(t), \\ y(t) &= Cx(t). \end{aligned} \quad (1)$$

Here, $x(t) \in \mathbb{R}^n$ is the state vector, $u(t) \in \mathbb{R}^{m_1}$ is the input signal and $y(t) \in \mathbb{R}^{m_2}$ is the output response. $E \in \mathbb{R}^{n \times n}$, $A \in \mathbb{R}^{n \times n}$, $B \in \mathbb{R}^{n \times m_1}$, $C \in \mathbb{R}^{m_2 \times n}$ are the system matrices. We consider general multiple-input and multiple-output (MIMO) systems, i.e. $m_1 \geq 1$ and $m_2 \geq 1$.

Almost all MOR methods are based on the idea of projection, i.e. a basis of a subspace \mathcal{V} which approximates the manifold in which the state vector $x(t)$ (approximately) resides is first computed, then the reduced order model is obtained by Petrov-Galerkin projection. If we use system (1) as an example, usually a matrix $V \in \mathbb{R}^{n \times r}$ is computed, whose columns span the subspace \mathcal{V} . $x(t)$ is approximated by its projection onto the subspace, i.e., $x(t) \approx Vz(t)$,

$$\begin{aligned} EV \frac{d}{dt} z(t) &\approx AVz(t) + Bu(t), \\ y(t) &\approx CVz(t). \end{aligned} \quad (2)$$

The reduced model is derived by forcing the residual $re = EV \frac{d}{dt} z(t) - AVz(t) - Bu(t)$ to be zero in a test subspace \mathcal{W} spanned by the columns of a matrix $W \in \mathbb{R}^{n \times r}$ i.e.,

$$\begin{aligned} W^T EV \frac{d}{dt} z(t) &= W^T AVz(t) + W^T Bu(t), \\ y(t) &= CVz(t). \end{aligned} \quad (3)$$

The above process in (3) is the so-called Petrov-Galerkin projection. The variable $r \ll n$ indicates the size of the system in (3), which is also called the order of the reduced model. Model reduction methods based on projection differ in the computation of the matrices W and V , see [7]. Methods related to balanced truncation [33, 9] compute W, V from the controllability Gramian and observability Gramian of the system, whereas methods based on moment-matching [36, 23] compute W, V according to the series expansion of the transfer function $H(s)$,

$$H(s) = C(sE - A)^{-1}B. \quad (4)$$

$H(s)$ is derived from the input-output relation in the frequency domain, through the Laplace transform of (1) (with zero initial condition $x(0) = 0$),

$$y(s) = H(s)u(s) = C(sE - A)^{-1}Bu(s). \quad (5)$$

Here s is the Laplace variable, and is related to the frequency f by $s = 2\pi jf$, where $j = -1$ is the imaginary unit. Similarly, the transfer function $\hat{H}(s)$ of the reduced model is

$$\hat{H}(s) = \hat{C}(s\hat{E} - \hat{A})^{-1}\hat{B},$$

where $\hat{C} = CV$, $\hat{E} = W^T EV$, $\hat{A} = W^T AV$, $\hat{B} = W^T B$.

2.2 PMOR for parametrized LTI systems

The parametrized LTI systems are usually in the following forms,

$$\begin{aligned} E(\tilde{\mu}) \frac{dx}{dt} &= A(\tilde{\mu})x + B(\tilde{\mu})\bar{u}(t), \\ y(t, \tilde{\mu}) &= C(\tilde{\mu})x, \end{aligned} \quad (6)$$

or

$$\begin{aligned} M(\tilde{\mu}) \frac{d^2x}{dt^2} + K(\tilde{\mu}) \frac{dx}{dt} + A(\tilde{\mu})x &= B(\tilde{\mu})\bar{u}(t), \\ y(t, \tilde{\mu}) &= C(\tilde{\mu})x. \end{aligned} \quad (7)$$

Here $\tilde{\mu} = (\tilde{\mu}_1, \dots, \tilde{\mu}_{\tilde{p}})$ is a vector of the parameters $\tilde{\mu}_1, \dots, \tilde{\mu}_{\tilde{p}}$. After Laplace transform (with zero initial condition), both of the systems in (6) and (7) can be generally written as

$$\begin{aligned} G(\mu)x(\mu) &= B(\mu)u(\mu_p), \\ y(\mu) &= C(\mu)x, \end{aligned} \quad (8)$$

where the entries in $\mu = (\mu_1, \mu_2, \dots, \mu_p)$ could be certain functions of the parameters $\tilde{\mu}_1, \dots, \tilde{\mu}_{\tilde{p}}$ and the Laplace variable s .

Similar to the definition in (5), the transfer function of the parametrized system in (6) or (7) is defined through the input-output relation in (8) as below:

$$H(\mu) = C(\mu)G^{-1}(\mu)B(\mu) \quad (9)$$

The reduced model of the system in (6) or (7) can be derived by Petrov-Galerkin projection with a pair of projection matrices W and V ,

$$\begin{aligned} W^T E(\tilde{\mu})V \frac{dz}{dt} &= W^T AV(\tilde{\mu})z + W^T B(\tilde{\mu})\bar{u}(t), \\ \hat{y}(t, \tilde{\mu}) &= C(\tilde{\mu})Vz, \end{aligned} \quad (10)$$

or

$$\begin{aligned} W^T M(\tilde{\mu})V \frac{d^2z}{dt^2} + W^T K(\tilde{\mu})V \frac{dz}{dt} + W^T A(\tilde{\mu})Vz &= W^T B(\tilde{\mu})\bar{u}(t), \\ \hat{y}(t, \tilde{\mu}) &= C(\tilde{\mu})Vz. \end{aligned} \quad (11)$$

While many interpolation based PMOR methods use Petrov-Galerkin projection [10], the Galerkin projection $W = V$ is often used by the Krylov subspace based PMOR methods. The transfer function $\hat{H}(\mu)$ of the reduced models can also be obtained by applying Laplace transform to the reduced models in (10) or in (11),

$$\hat{H}(\mu) = \hat{C}(\mu)\hat{G}^{-1}(\mu)\hat{B}(\mu),$$

where $\hat{C}(\mu) = C(\mu)V$, $\hat{B}(\mu) = W^T B(\mu)$, and $\hat{G}(\mu) = W^T G(\mu)V$.

3 Output error bound for special non-parametrized LTI systems

The technique of deriving the output error bound in this section is motivated by the method in [37], where an output error bound for the reduced model is derived based on the weak formulation of the PDEs. The error bound in [37] is derived in functional space, and is only valid for parametrized systems with real parameters. Here, we consider estimating the output error of the reduced model directly in the vector space. Further extensions have been made, so that the derived error bound is applicable to systems with complex parameters, and finally, it can be used for adaptive selection of the multiple expansion points discussed in subsection 8.1.1.

In this section, we assume that E and A are symmetric matrices. Although the assumption is not a request for the derivation of the error bound, they are needed to make the error bound computable, see Section 6.

3.1 Derivation of an error bound for SISO systems

We only consider single-input and single-output (SISO) system in this subsection. The results will be used to get a posteriori output error bound for MIMO systems in the next subsection.

To derive the error bound, the norm $\|\cdot\|_{\tilde{A}}: \mathbb{C}^n \rightarrow \mathbb{R}$ for a complex vector x is defined as

$$\|x\|_{\tilde{A}} = (x^* \tilde{A} x)^{1/2},$$

Here, the matrix \tilde{A} is assumed to be symmetric, positive definite. It can be simply taken as the identity matrix, then the norm reduces to the standard 2-norm. x^* is the conjugate transpose of x . The norm $\|\cdot\|_{\tilde{A}}$ is actually associated with the inner product: $\langle \cdot, \cdot \rangle: \langle x_1, x_2 \rangle = x_2^* \tilde{A} x_1, \forall x_1, x_2 \in \mathbb{C}^n$.

We also assume that the matrix-valued function $G(s) := sE - A: \mathbb{C} \mapsto \mathbb{C}^{n \times n}$ satisfies

$$\operatorname{Re}(x^* G(s) x) \geq \alpha(s) (x^* \tilde{A} x), \forall x \in \mathbb{C}^n, x \neq 0, \quad (12)$$

and

$$\operatorname{Im}(x^* G(s) x) \geq \gamma(s) (x^* \tilde{A} x), \forall x \in \mathbb{C}^n, x \neq 0, \quad (13)$$

where $\operatorname{Re}(\cdot)$ means the real part of $x^* G(s) x$, and $\operatorname{Im}(\cdot)$ is the imaginary part. $\alpha(s) > 0, \gamma(s) > 0$ may depend on the parameter s . Our goal is to derive an error bound for the error $|H(s) - \hat{H}(s)|$, where $|\cdot|$ means the absolute value or modulus of a complex number.

Notice that $H(s)$ is actually the output $y(s)$ in (5) corresponding to the impulse input $u(t) = \delta(t)$, where $\delta(t)$ is the δ function, so that $u(s) = \mathcal{L}(\delta(t)) = 1$ (the Laplace transform of $\delta(t)$). Therefore the error bound for the reduced transfer function $\hat{H}(s)$ is also the output error bound for the reduced model in the frequency domain.

We first define the primal system in the frequency domain, which is the Laplace transform of the original system (1) with $u(t) = \delta(t)$,

$$\begin{aligned} G(s)x(s) &= B, \\ y(s) &= Cx(s). \end{aligned} \quad (14)$$

It is easy to see that the output $y(s)$ in (14) equals to the transfer function $H(s)$ of the original system (1). The reduced model for the primal system is defined as,

$$\begin{aligned} W^T G(s) V z(s) &= W^T B, \\ \hat{y}(s) &= C V z(s), \end{aligned} \quad (15)$$

where W, V are those used in (3). Here $\hat{x}(s) := V z(s)$ is the approximation of $x(s)$ in (14). Analogously, $\hat{y}(s)$ equals to the transfer function $\hat{H}(s)$.

To assist the derivation of the error bound, we need a dual system in the frequency domain,

$$\begin{aligned} G^*(s) x^{du}(s) &= -C^T, \\ y^{du}(s) &= B^T x^{du}(s). \end{aligned} \quad (16)$$

The reduced model for the dual system is defined as,

$$\begin{aligned} V^T G^*(s) W z^{du}(s) &= -V^T C^T, \\ \hat{y}^{du}(s) &= B^T W z^{du}(s), \end{aligned} \quad (17)$$

where $G^*(s) = \bar{s}E^T - A^T$ is the conjugate transpose of $G(s)$, and \bar{s} is the conjugate of s . $\hat{x}^{du}(s) := W z^{du}(s)$ is the approximate solution to the dual system, $x^{du}(s) \approx \hat{x}^{du}(s)$.

Let $r^{pr}(s) = B - G(s)\hat{x}(s)$ be the residual of the primal system in (14), and $r^{du}(s) = -C^T - G^*(s)\hat{x}^{du}(s)$ be the residual of the dual system. In order to make the final description of the error bound as simple as possible, we first show that by the Ritz representation theorem, $r^{pr}(s)$ can be represented through a vector $\hat{\epsilon}^{pr} \in \mathbb{C}^n$, and $r^{du}(s)$ can be represented through a vector $\hat{\epsilon}^{du} \in \mathbb{C}^n$.

Define a function $f^{pr}(\xi) = (r^{pr})^* \xi: \mathbb{C}^n \mapsto \mathbb{C}$. From the Ritz representation theorem, there exists a unique vector $\hat{\epsilon}^{pr} \in \mathbb{C}^n$, such that

$$f^{pr}(\xi) = \langle \xi, \hat{\epsilon}^{pr} \rangle = (\hat{\epsilon}^{pr})^* \tilde{A} \xi. \quad (18)$$

We also define a function $f^{du}(\xi) = (r^{du})^* \xi: \mathbb{C}^n \mapsto \mathbb{C}$. Similarly, there exists a unique vector $\hat{\epsilon}^{du} \in \mathbb{C}^n$, such that

$$f^{du}(\xi) = \langle \xi, \hat{\epsilon}^{du} \rangle = (\hat{\epsilon}^{du})^* \tilde{A} \xi. \quad (19)$$

In the following, we use $x, x^{du}, \hat{x}, \hat{x}^{du}, z, z^{du}$ to represent $x(s), x^{du}(s), \hat{x}(s), \hat{x}^{du}(s), z(s), z^{du}(s)$ in (14-17) for the purpose of simplicity. We first propose a relation between $H(s) - \hat{H}(s)$ and the errors of the state vectors \hat{x} and \hat{x}^{du} of the reduced primal and dual systems. This relation will be repeatedly used to derive the error bound.

Proposition 1 If the reduced model (15) of the primal system and that of the dual system (17) are obtained by the same pair W and V , then

$$H(s) - \hat{H}(s) = Cx - C\hat{x} = -(\epsilon^{du})^* G(s) \epsilon^{pr},$$

where $\epsilon^{du} = x^{du} - \hat{x}^{du}$ and $\epsilon^{pr} = x - \hat{x}$.

Proof From (15), we have $W^T B - W^T G(s) \hat{x} = 0$, i.e.

$$\begin{aligned} W^T G(s) x - W^T G(s) \hat{x} &= 0 \\ \Leftrightarrow W^T G(s) (x - \hat{x}) &= 0 \\ \Rightarrow (z^{du})^* W^T G(s) (x - \hat{x}) &= 0 \\ \Leftrightarrow (\hat{x}^{du})^* G(s) (x - \hat{x}) &= 0 \\ \Leftrightarrow (\hat{x}^{du})^* G(s) \epsilon^{pr} &= 0. \end{aligned} \quad (20)$$

From (16) and (20), we get

$$\begin{aligned} Cx - C\hat{x} &= C\epsilon^{pr} \\ &= -(x^{du})^* (G^*(s))^* \epsilon^{pr} \\ &= -(x^{du})^* G(s) \epsilon^{pr} + (\hat{x}^{du})^* G(s) \epsilon^{pr}. \\ &= -(\epsilon^{du})^* G(s) \epsilon^{pr}. \end{aligned} \quad (21)$$

■

Since the computation of ϵ^{du} and ϵ^{pr} involves computation of x^{du} and x , the solutions of the full dual and the full primal systems, $|(\epsilon^{du})^* G(s) \epsilon^{pr}|$ cannot act as a computable error bound for $|H(s) - \hat{H}(s)|$. Next, we will use Proposition 1 and the assumptions on $G(s)$ to derive computable error bounds for the real part, and imaginary part of $H(s) - \hat{H}(s)$ separately. The final error bound can be obtained from the error bounds for the real and imaginary parts.

Proposition 2 If the reduced model (15) of the primal system and that of the dual system (17) are obtained by the same pair W and V , and $G(s)$ satisfies (12), then

$$-S_R - \beta_R \leq \text{Re}(H(s) - \hat{H}(s)) \leq S_R - \beta_R.$$

Here,

$$\beta_R = \frac{1}{4\alpha(s)} (\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{du} + \frac{1}{4\alpha(s)} (\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{pr}, \quad S_R = \frac{\kappa_0}{4\alpha(s)} (\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{pr} + \frac{1}{4\kappa_0\alpha(s)} (\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{du},$$

and

$$\kappa_0 = \left(\frac{(\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{du}}{(\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{pr}} \right)^{1/2}.$$

Proof We begin by defining a new vector $\hat{\epsilon}^- = \frac{1}{\alpha(s)} \hat{\epsilon}^{pr} - \frac{1}{\kappa\alpha(s)} \hat{\epsilon}^{du}$. Here and below $\kappa > 0$ is a variable to be specified. We first derive an upper bound for $\text{Re}(H(s) - \hat{H}(s))$. Since $\alpha(s) > 0$,

$$\begin{aligned} &\kappa\alpha(s) \langle \epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^-, \epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^- \rangle \\ &= \kappa\alpha(s) (\epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^-)^* \tilde{A} (\epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^-) \geq 0 \\ \Leftrightarrow &\kappa\alpha(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\alpha(s)}{4} (\hat{\epsilon}^-)^* \tilde{A} \hat{\epsilon}^- - \frac{\kappa\alpha(s)}{2} (\hat{\epsilon}^-)^* \tilde{A} \epsilon^{pr} - \frac{\kappa\alpha(s)}{2} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^- \geq 0 \\ \Leftrightarrow &\frac{\kappa\alpha(s)}{2} (\hat{\epsilon}^-)^* \tilde{A} \epsilon^{pr} + \frac{\kappa\alpha(s)}{2} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^- \leq \kappa\alpha(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\alpha(s)}{4} (\hat{\epsilon}^-)^* \tilde{A} \hat{\epsilon}^-. \end{aligned} \quad (22)$$

From the property of inner product

$$(\hat{\epsilon}^-)^* \tilde{A} \epsilon^{pr} = \overline{(\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^-}, \quad (23)$$

we only have to estimate $(\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^-$ in the last inequality of (22). From the definition of $\hat{\epsilon}^-$, and (18) (19) (21), we get

$$\begin{aligned} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^- &= \frac{1}{\alpha(s)} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^{pr} - \frac{1}{\kappa\alpha(s)} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^{du} \\ &= \frac{1}{\alpha(s)} (\epsilon^{pr})^* r^{pr}(s) - \frac{1}{\kappa\alpha(s)} (r^{du}(s))^* \epsilon^{pr} \\ &= \frac{1}{\alpha(s)} (\epsilon^{pr})^* (B - G(s) \hat{x}) - \frac{1}{\kappa\alpha(s)} (-C^T - G^*(s) \hat{x}^{du})^* \epsilon^{pr} \\ &= \frac{1}{\alpha(s)} (\epsilon^{pr})^* (G(s)x - G(s)\hat{x}) - \frac{1}{\kappa\alpha(s)} \overline{(G^*(s)x^{du} - G^*(s)\hat{x}^{du})^* \epsilon^{pr}} \\ &= \frac{1}{\alpha(s)} (\epsilon^{pr})^* G(s) \epsilon^{pr} - \frac{1}{\kappa\alpha(s)} \overline{(G^*(s)\epsilon^{du})^* \epsilon^{pr}} \\ &= \frac{1}{\alpha(s)} (\epsilon^{pr})^* G(s) \epsilon^{pr} - \frac{1}{\kappa\alpha(s)} (\epsilon^{du})^* G(s) \epsilon^{pr} \\ &= \frac{1}{\alpha(s)} (\epsilon^{pr})^* G(s) \epsilon^{pr} + \frac{1}{\kappa\alpha(s)} (Cx - C\hat{x}). \end{aligned} \quad (24)$$

Using the relation in (23), and substituting (24) into the last inequality of (22) yield

$$\begin{aligned}
& \frac{\kappa\alpha(s)}{2}(\hat{\epsilon}^-)^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{2}(\epsilon^{pr})^* \tilde{A}\hat{\epsilon}^- \leq \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^-)^* \tilde{A}\hat{\epsilon}^- \\
& \Leftrightarrow \frac{\kappa}{2} \left[2\operatorname{Re}((\epsilon^{pr})^* G(s)\epsilon^{pr}) + \frac{2}{\kappa}\operatorname{Re}(Cx - C\hat{x}) \right] \leq \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^-)^* \tilde{A}\hat{\epsilon}^- \\
& \Leftrightarrow \kappa\operatorname{Re}((\epsilon^{pr})^* G(s)\epsilon^{pr}) + \operatorname{Re}(Cx - C\hat{x}) \leq \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^-)^* \tilde{A}\hat{\epsilon}^- \\
& \Leftrightarrow \operatorname{Re}(Cx - C\hat{x}) \leq -\kappa\operatorname{Re}((\epsilon^{pr})^* G(s)\epsilon^{pr}) + \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^-)^* \tilde{A}\hat{\epsilon}^- \\
& \Leftrightarrow \operatorname{Re}(Cx - C\hat{x}) \leq \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^-)^* \tilde{A}\hat{\epsilon}^-,
\end{aligned} \tag{25}$$

where the last inequality holds due to (12). Substituting $\hat{\epsilon}^- = \frac{1}{\alpha(s)}\hat{\epsilon}^{pr} - \frac{1}{\kappa\alpha(s)}\hat{\epsilon}^{du}$ into the last inequality of (25) gives

$$\begin{aligned}
& \operatorname{Re}(Cx - C\hat{x}) \leq \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^-)^* \tilde{A}\hat{\epsilon}^- \\
& \Leftrightarrow \operatorname{Re}(Cx - C\hat{x}) \leq \frac{\kappa}{4\alpha(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{pr} + \frac{1}{4\kappa\alpha(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{du} - \frac{1}{4\alpha(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{du} - \frac{1}{4\alpha(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{pr} \\
& \Leftrightarrow \operatorname{Re}(Cx - C\hat{x}) \leq f_R(\kappa) - \beta_R \\
& \Leftrightarrow \operatorname{Re}(H(s) - \hat{H}(s)) \leq f_R(\kappa) - \beta_R,
\end{aligned} \tag{26}$$

where $f_R(\kappa) = \frac{\kappa}{4\alpha(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{pr} + \frac{1}{4\kappa\alpha(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{du}$, $\beta_R = \frac{1}{4\alpha(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{du} + \frac{1}{4\alpha(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{pr}$.

Next we derive a lower bound, $\operatorname{Re}(H(s) - \hat{H}(s)) \geq -f_R(\kappa) - \beta_R$. We now define a second new vector $\hat{\epsilon}^+ = \frac{1}{\alpha(s)}\hat{\epsilon}^{pr} + \frac{1}{\kappa\alpha(s)}\hat{\epsilon}^{du}$, and estimate

$$\begin{aligned}
& \kappa\alpha(s)(\epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^+, \epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^+) \\
& = \kappa\alpha(s)(\epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^+)^* \tilde{A}(\epsilon^{pr} - \frac{1}{2}\hat{\epsilon}^+) \geq 0 \\
& \Leftrightarrow \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+ - \frac{\kappa\alpha(s)}{2}(\hat{\epsilon}^+)^* \tilde{A}\epsilon^{pr} - \frac{\kappa\alpha(s)}{2}(\epsilon^{pr})^* \tilde{A}\hat{\epsilon}^+ \geq 0 \\
& \Leftrightarrow \frac{\kappa\alpha(s)}{2}(\hat{\epsilon}^+)^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{2}(\epsilon^{pr})^* \tilde{A}\hat{\epsilon}^+ \leq \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+.
\end{aligned} \tag{27}$$

It is not difficult to see from (24) that

$$(\epsilon^{pr})^* \tilde{A}\hat{\epsilon}^+ = \frac{1}{\alpha(s)}(\epsilon^{pr})^* G(s)\epsilon^{pr} - \frac{1}{\kappa\alpha(s)}\overline{(Cx - C\hat{x})}. \tag{28}$$

Using the relation $(\hat{\epsilon}^+)^* \tilde{A}\epsilon^{pr} = \overline{(\epsilon^{pr})^* \tilde{A}\hat{\epsilon}^+}$, and substituting (28) into the last inequality of (27) yield

$$\begin{aligned}
& \frac{\kappa\alpha(s)}{2}(\hat{\epsilon}^+)^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{2}(\epsilon^{pr})^* \tilde{A}\hat{\epsilon}^+ \leq \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+ \\
& \Leftrightarrow \frac{\kappa}{2} \left[2\operatorname{Re}((\epsilon^{pr})^* G(s)\epsilon^{pr}) - \frac{2}{\kappa}\operatorname{Re}(Cx - C\hat{x}) \right] \leq \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+ \\
& \Leftrightarrow \kappa\operatorname{Re}((\epsilon^{pr})^* G(s)\epsilon^{pr}) - \operatorname{Re}(Cx - C\hat{x}) \leq \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+ \\
& \Leftrightarrow -\operatorname{Re}(Cx - C\hat{x}) \leq -\kappa\operatorname{Re}((\epsilon^{pr})^* G(s)\epsilon^{pr}) + \kappa\alpha(s)(\epsilon^{pr})^* \tilde{A}\epsilon^{pr} + \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+ \\
& \Leftrightarrow -\operatorname{Re}(Cx - C\hat{x}) \leq \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+,
\end{aligned} \tag{29}$$

where the last inequality holds due to the relation in (12). Substituting $\hat{\epsilon}^+ = \frac{1}{\alpha(s)}\hat{\epsilon}^{pr} + \frac{1}{\kappa\alpha(s)}\hat{\epsilon}^{du}$ into the last inequality of (29), we can assert that

$$\begin{aligned}
& -\operatorname{Re}(Cx - C\hat{x}) \leq \frac{\kappa\alpha(s)}{4}(\hat{\epsilon}^+)^* \tilde{A}\hat{\epsilon}^+ \\
& \Leftrightarrow -\operatorname{Re}(Cx - C\hat{x}) \leq \frac{\kappa}{4\alpha(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{pr} + \frac{1}{4\kappa\alpha(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{du} + \frac{1}{4\alpha(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{du} + \frac{1}{4\alpha(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{pr} \\
& \Leftrightarrow -\operatorname{Re}(Cx - C\hat{x}) \leq f_R(\kappa) + \beta_R \\
& \Leftrightarrow \operatorname{Re}(Cx - C\hat{x}) \geq -f_R(\kappa) - \beta_R \\
& \Leftrightarrow \operatorname{Re}(H(s) - \hat{H}(s)) \geq -f_R(\kappa) - \beta_R.
\end{aligned} \tag{30}$$

Combining the last equality of (26) with that of (30), we have $|\operatorname{Re}(H(s) - \hat{H}(s)) + \beta_R| \leq f_R(\kappa)$. It is not difficult to check that when $\kappa = \kappa_0$, $f_R(\kappa)$ reaches the minimum, and $f_R(\kappa_0) = S_R$. ■

Proposition 3 *If the reduced model (15) of the primal system and that of the dual system (17) are obtained by the same pair W , V , and $G(s)$ satisfies (13), then*

$$-S_I + \beta_I \leq \operatorname{Im}(H(s) - \hat{H}(s)) \leq S_I + \beta_I.$$

Here,

$$\beta_I = \frac{1}{4\gamma(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{du} + \frac{1}{4\gamma(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{pr}, S_I = \frac{\kappa_0}{4\gamma(s)}(\hat{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{pr} + \frac{1}{4\kappa_0\gamma(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{du},$$

and

$$\kappa_0 = \left(\frac{(\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{du}}{(\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{pr}} \right)^{1/2}.$$

Proof We first derive an upper bound for $Im(H(s) - \hat{H}(s))$. Let us define a new vector $\tilde{\epsilon}^- = \frac{1}{\gamma(s)} \hat{\epsilon}^{pr} - \frac{1}{\kappa\gamma(s)} \hat{\epsilon}^{du}$. Since $\gamma(s) > 0$,

$$\begin{aligned} & \kappa\gamma(s) \langle \epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^-, \epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^- \rangle \\ &= \kappa\gamma(s) (\epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^-)^* \tilde{A} (\epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^-) \geq 0 \\ &\Leftrightarrow \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^- - \frac{\mathcal{J}\kappa\gamma(s)}{2} (\tilde{\epsilon}^-)^* \tilde{A} \epsilon^{pr} + \frac{\mathcal{J}\kappa\gamma(s)}{2} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^- \geq 0 \\ &\Leftrightarrow \frac{\mathcal{J}\kappa\gamma(s)}{2} (\tilde{\epsilon}^-)^* \tilde{A} \epsilon^{pr} - \frac{\mathcal{J}\kappa\gamma(s)}{2} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^- \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^-. \end{aligned} \quad (31)$$

Since the only difference between $\hat{\epsilon}^-$ and $\tilde{\epsilon}^-$ is the denominator, from (24) we see

$$\begin{aligned} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^- &= \frac{1}{\gamma(s)} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^{pr} - \frac{1}{\kappa\gamma(s)} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^{du} \\ &= \frac{1}{\gamma(s)} (\epsilon^{pr})^* G(s) \epsilon^{pr} + \frac{1}{\kappa\gamma(s)} (Cx - C\hat{x}). \end{aligned} \quad (32)$$

We proceed analogously to the proof of Proposition 2. Using the relation $(\tilde{\epsilon}^-)^* \tilde{A} \epsilon^{pr} = \overline{(\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^-}$, and substituting (32) into the last inequality of (31) yield

$$\begin{aligned} & \frac{\mathcal{J}\kappa\gamma(s)}{2} (\tilde{\epsilon}^-)^* \tilde{A} \epsilon^{pr} - \frac{\mathcal{J}\kappa\gamma(s)}{2} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^- \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^- \\ &\Leftrightarrow \frac{\mathcal{J}^2\kappa}{2} [-2Im((\epsilon^{pr})^* G(s) \epsilon^{pr}) + \frac{2}{\kappa} Im(Cx - C\hat{x})] \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^- \\ &\Leftrightarrow \kappa Im((\epsilon^{pr})^* G(s) \epsilon^{pr}) - Im(Cx - C\hat{x}) \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^- \\ &\Leftrightarrow -Im(Cx - C\hat{x}) \leq -\kappa Im(\epsilon^{pr})^* G(s) \epsilon^{pr} + \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^- \\ &\Leftrightarrow -Im(Cx - C\hat{x}) \leq \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^-, \end{aligned} \quad (33)$$

where the last inequality holds because of (13). Substituting $\tilde{\epsilon}^- = \frac{1}{\gamma(s)} \hat{\epsilon}^{pr} - \frac{1}{\kappa\gamma(s)} \hat{\epsilon}^{du}$ into the last inequality of (33) gives

$$\begin{aligned} & -Im(Cx - C\hat{x}) \leq \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^-)^* \tilde{A} \tilde{\epsilon}^- \\ &\Leftrightarrow -Im(Cx - C\hat{x}) \leq \frac{\kappa}{4\gamma(s)} (\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{pr} + \frac{1}{4\kappa\gamma(s)} (\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{du} - \frac{1}{4\gamma(s)} (\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{du} - \frac{1}{4\gamma(s)} (\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{pr} \\ &\Leftrightarrow -Im(Cx - C\hat{x}) \leq f_I(\kappa) - \beta_I \\ &\Leftrightarrow Im(Cx - C\hat{x}) \geq -f_I(\kappa) + \beta_I \\ &\Leftrightarrow Im(H(s) - \hat{H}(s)) \geq -f_I(\kappa) + \beta_I, \end{aligned} \quad (34)$$

where $f_I(\kappa) = \frac{\kappa}{4\gamma(s)} (\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{pr} + \frac{1}{4\kappa\gamma(s)} (\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{du}$, $\beta_I = \frac{1}{4\gamma(s)} (\hat{\epsilon}^{pr})^* \tilde{A} \hat{\epsilon}^{du} + \frac{1}{4\gamma(s)} (\hat{\epsilon}^{du})^* \tilde{A} \hat{\epsilon}^{pr}$.

We continue to prove the next claim $Im(H(s) - \hat{H}(s)) \leq f_I(\kappa) + \beta_I$. Defining a second new vector $\tilde{\epsilon}^+ = \frac{1}{\gamma(s)} \hat{\epsilon}^{pr} + \frac{1}{\kappa\gamma(s)} \hat{\epsilon}^{du}$ yields

$$\begin{aligned} & \kappa\gamma(s) \langle \epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^+, \epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^+ \rangle \\ &= \kappa\gamma(s) (\epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^+)^* \tilde{A} (\epsilon^{pr} + \frac{1}{2} \tilde{\epsilon}^+) \geq 0 \\ &\Leftrightarrow \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^+)^* \tilde{A} \tilde{\epsilon}^+ - \frac{\mathcal{J}\kappa\gamma(s)}{2} (\tilde{\epsilon}^+)^* \tilde{A} \epsilon^{pr} + \frac{\mathcal{J}\kappa\gamma(s)}{2} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^+ \geq 0 \\ &\Leftrightarrow \frac{\mathcal{J}\kappa\gamma(s)}{2} (\tilde{\epsilon}^+)^* \tilde{A} \epsilon^{pr} - \frac{\mathcal{J}\kappa\gamma(s)}{2} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^+ \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^+)^* \tilde{A} \tilde{\epsilon}^+. \end{aligned} \quad (35)$$

From (32), we know

$$\begin{aligned} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^+ &= \frac{1}{\gamma(s)} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^{pr} + \frac{1}{\kappa\gamma(s)} (\epsilon^{pr})^* \tilde{A} \hat{\epsilon}^{du} \\ &= \frac{1}{\gamma(s)} (\epsilon^{pr})^* G(s) \epsilon^{pr} - \frac{1}{\kappa\gamma(s)} (Cx - C\hat{x}). \end{aligned} \quad (36)$$

Using the relation $(\tilde{\epsilon}^+)^* \tilde{A} \epsilon^{pr} = \overline{(\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^+}$, and substituting (36) into (35), we obtain

$$\begin{aligned} & \frac{\mathcal{J}\kappa\gamma(s)}{2} (\tilde{\epsilon}^+)^* \tilde{A} \epsilon^{pr} - \frac{\mathcal{J}\kappa\gamma(s)}{2} (\epsilon^{pr})^* \tilde{A} \tilde{\epsilon}^+ \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^+)^* \tilde{A} \tilde{\epsilon}^+ \\ &\Leftrightarrow \frac{\mathcal{J}^2\kappa}{2} [-2Im((\epsilon^{pr})^* G(s) \epsilon^{pr}) - \frac{2}{\kappa} Im(Cx - C\hat{x})] \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^+)^* \tilde{A} \tilde{\epsilon}^+ \\ &\Leftrightarrow \kappa Im((\epsilon^{pr})^* G(s) \epsilon^{pr}) + Im(Cx - C\hat{x}) \leq \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^+)^* \tilde{A} \tilde{\epsilon}^+ \\ &\Leftrightarrow Im(Cx - C\hat{x}) \leq -\kappa Im(\epsilon^{pr})^* G(s) \epsilon^{pr} + \kappa\gamma(s) (\epsilon^{pr})^* \tilde{A} \epsilon^{pr} + \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^+)^* \tilde{A} \tilde{\epsilon}^+ \\ &\Leftrightarrow Im(Cx - C\hat{x}) \leq \frac{\kappa\gamma(s)}{4} (\tilde{\epsilon}^+)^* \tilde{A} \tilde{\epsilon}^+, \end{aligned} \quad (37)$$

where the last inequality holds because of (13). Substituting $\tilde{\epsilon}^+ = \frac{1}{\gamma(s)}\tilde{\epsilon}^{pr} + \frac{1}{\kappa\gamma(s)}\hat{\epsilon}^{du}$ into the last inequality of (37), it follows immediately

$$\begin{aligned} \text{Im}(Cx - C\hat{x}) &\leq \frac{\kappa\gamma(s)}{4}(\tilde{\epsilon}^+)^* \tilde{A}\tilde{\epsilon}^+ \\ \Leftrightarrow \text{Im}(Cx - C\hat{x}) &\leq \frac{\kappa}{4\gamma(s)}(\tilde{\epsilon}^{pr})^* \tilde{A}\tilde{\epsilon}^{pr} + \frac{1}{4\kappa\gamma(s)}(\hat{\epsilon}^{du})^* \tilde{A}\hat{\epsilon}^{du} + \frac{1}{4\gamma(s)}(\tilde{\epsilon}^{pr})^* \tilde{A}\hat{\epsilon}^{du} + \frac{1}{4\gamma(s)}(\hat{\epsilon}^{du})^* \tilde{A}\tilde{\epsilon}^{pr} \\ \Leftrightarrow \text{Im}(Cx - C\hat{x}) &\leq f_I(\kappa) + \beta_I \\ \Leftrightarrow \text{Im}(H(s) - \hat{H}(s)) &\leq f_I(\kappa) + \beta_I. \end{aligned} \quad (38)$$

From (34) and (38), we have $|\text{Im}(H(s) - \hat{H}(s)) - \beta_I| \leq f_I(\kappa)$. When $\kappa = \kappa_0$, $\min_{\kappa} f_I(\kappa) = f(\kappa_0) = S_I$. ■

Based on Proposition 2 and 3, we can immediately get the error bound for $|H(s) - \hat{H}(s)|$ given in the following theorem.

Theorem 1 *The error of $\hat{H}(s)$ is bounded by $\Delta(s)$ defined as below:*

$$|H(s) - \hat{H}(s)| = \sqrt{|\text{Re}(H(s) - \hat{H}(s))|^2 + |\text{Im}(H(s) - \hat{H}(s))|^2} \leq \sqrt{B_R^2 + B_I^2} := \Delta(s). \quad (39)$$

Here $B_R = \max\{|S_R - \beta_R|, |-S_R - \beta_R|\}$ and $B_I = \max\{|-S_I + \beta_I|, |S_I + \beta_I|\}$.

3.2 Error bound for MIMO systems

For MIMO systems, the transfer function $H(s)$ is a matrix. The ik -th entry $H_{ik}(s)$ corresponds to a SISO system, whose input is $u_k(t)$, the k -th entry of the input vector $u(t)$, and whose output is $y_i(t) = C(i, :)x(t)$, the i -th entry of the output response $y(t)$. The transfer function $\hat{H}(s)$ of the reduced model in (3) is also a matrix. The ik -th entry $\hat{H}_{ik}(s)$ is an approximation of $H_{ik}(s)$ for the corresponding SISO system. Therefore, the error between $\hat{H}_{ik}(s)$ and $H_{ik}(s)$ can be measured by $\Delta_{ik}(s)$, which can be computed from (39) as

$$|H_{ik}(s) - \hat{H}_{ik}(s)| \leq \sqrt{B_R^2 + B_I^2} =: \Delta_{ik}(s),$$

$B_R = \max\{|S_R - \beta_R|, |-S_R - \beta_R|\}$ and $B_I = \max\{|-S_I + \beta_I|, |S_I + \beta_I|\}$. To compute B_R, B_I corresponding to $\Delta_{ik}(s)$, the output matrix C should be replaced by $C(i, :)$, the i th row in C . The input matrix B should be replaced by $B(:, k)$, the k th column in B , $\forall 1 \leq i \leq m_1, 1 \leq k \leq m_2$.

Once all $\Delta_{ik}(s)$, $1 \leq i \leq m_1, 1 \leq k \leq m_2$ are computed by (39), the final error bound can be taken as the maximum of them, i.e.

$$\|H(s) - \hat{H}(s)\|_{\max} = \max_{ik} |H_{ik}(s) - \hat{H}_{ik}(s)| \leq \max_{ik} \Delta_{ik}(s). \quad (40)$$

In Section 6, it will be shown that the assumptions (12) and (13) on $G(s)$ imply that the matrix E must be symmetric positive definite, which is a strict limit on E . In the next section, we derive another error bound for general non-parametrized LTI systems, where E and A are allowed to be nonsymmetric, and E can be singular.

4 Output error bound for general non-parametrized LTI systems

In this section, we derive an output error bound for more general LTI systems, where the matrices E, A do not have to be symmetric, and E could be singular. In stead of (12) and (13), the matrix-valued function $G(s)$ is assumed to satisfy

$$\inf_{\substack{w \in \mathbb{C}^n \\ w \neq 0}} \sup_{\substack{v \in \mathbb{C}^n \\ v \neq 0}} \frac{w^* G(s) v}{\|w\|_2 \|v\|_2} = \beta(s) > 0. \quad (41)$$

We still rely on the the primal system in (14) and the dual system in (16).

The reduced model for the primal system is of the same form as the system in (15), where W, V are also used to get the reduced model (3) for the original system. The reduced model for the dual system can be constructed more flexibly as,

$$\begin{aligned} (W^{du})^T G^*(s) V^{du} z^{du}(s) &= -(W^{du})^T C^T, \\ \hat{y}^{du}(s) &= B^* V^{du} z^{du}(s), \end{aligned} \quad (42)$$

where W^{du} and V^{du} can be different from V and W for the primal system, that means it is allowed that $W^{du} \neq V$ and $V^{du} \neq W$. Here $\hat{x}^{du} = V^{du} z^{du}$ is the approximate solution to (16).

Define two new variables $e(s) = (\hat{x}^{du})^* r^{pr}(s)$ and $\tilde{y}(s) = C\hat{x} - e(s)$.

Theorem 2 For a SISO LTI system, if $G(s)$ satisfies (41), then $|y(s) - \tilde{y}(s)| \leq \tilde{\Delta}_g(s)$, $\tilde{\Delta}_g(s) := \frac{\|r^{du}(s)\|_2 \|r^{pr}(s)\|_2}{\beta(s)}$. As a result, $|H(s) - \hat{H}(s)| = |Cx - C\hat{x}| \leq \Delta_g(s)$, $\Delta_g(s) := \tilde{\Delta}_g(s) + |e(s)|$.

Proof The dual system $G^*(s)x^{du} = -C^T$ implies that

$$(x - \hat{x})^* G^*(s)x^{du} = -(x - \hat{x})^* C^T. \quad (43)$$

From the definition of the residual $r^{pr} = B - G(s)\hat{x} = G(s)(x - \hat{x})$ for the primal system, we get

$$(x^{du})^* r^{pr} = (x^{du})^* G(s)(x - \hat{x}). \quad (44)$$

Combining (43) with (44), it is obvious

$$-C(x - \hat{x}) = (x^{du})^* r^{pr}.$$

Then

$$\begin{aligned} |y(s) - \tilde{y}(s)| &= |Cx - C\hat{x} + (\hat{x}^{du})^* r^{pr}(s)| \\ &= |-(x^{du})^* r^{pr} + (\hat{x}^{du})^* r^{pr}(s)| \\ &= |-(x^{du} - \hat{x}^{du})^* r^{pr}(s)| \\ &\leq \|(x^{du} - \hat{x}^{du})\|_2 \|r^{pr}(s)\|_2. \end{aligned} \quad (45)$$

Replacing w in (41) with $x^{du} - \hat{x}^{du}$ yields

$$\sup_{\forall v \in \mathbb{C}^n} \frac{(x^{du} - \hat{x}^{du})^* G(s)v}{\|(x^{du} - \hat{x}^{du})\|_2 \|v\|_2} \geq \beta(s),$$

i.e.

$$\sup_{\forall v \in \mathbb{C}^n} \frac{(x^{du} - \hat{x}^{du})^* G(s)v}{\|v\|_2} \geq \beta(s) \|x^{du} - \hat{x}^{du}\|_2. \quad (46)$$

Since

$$\frac{(x^{du} - \hat{x}^{du})^* G(s)v}{\|v\|_2} \leq \frac{\|G^*(s)(x^{du} - \hat{x}^{du})\|_2 \|v\|_2}{\|v\|_2} = \|G^*(s)(x^{du} - \hat{x}^{du})\|_2,$$

and when $v = v_0 = G^*(s)(x^{du} - \hat{x}^{du})$,

$$\frac{(x^{du} - \hat{x}^{du})^* G(s)v_0}{\|v_0\|_2} = \|G^*(s)(x^{du} - \hat{x}^{du})\|_2,$$

it suffices to make the following observation

$$\sup_{\forall v \in \mathbb{C}^n} \frac{(x^{du} - \hat{x}^{du})^* G^*(s)v}{\|v\|_2} = \|G^*(s)(x^{du} - \hat{x}^{du})\|_2. \quad (47)$$

Combining (46) with (47) shows that

$$\|G^*(s)(x^{du} - \hat{x}^{du})\|_2 \geq \beta(s) \|x^{du} - \hat{x}^{du}\|_2.$$

From the definition of $r^{du}(s) = -C^T - G^*(s)\hat{x}^{du} = G^*(s)x^{du} - G^*(s)\hat{x}^{du}$, we get

$$\|r^{du}(s)\|_2 \geq \beta(s) \|x^{du} - \hat{x}^{du}\|_2. \quad (48)$$

Substituting (48) into (45) we obtain,

$$|y(s) - \tilde{y}(s)| \leq \frac{\|r^{du}(s)\|_2 \|r^{pr}(s)\|_2}{\beta(s)}.$$

Since $\tilde{y}(s) = C\hat{x} - e(s)$,

$$|Cx - C\hat{x}| - |e(s)| \leq |y(s) - \tilde{y}(s)| \leq \frac{\|r^{du}(s)\|_2 \|r^{pr}(s)\|_2}{\beta(s)}.$$

Finally,

$$|H(s) - \hat{H}(s)| = |Cx - C\hat{x}| \leq \frac{\|r^{du}(s)\|_2 \|r^{pr}(s)\|_2}{\beta(s)} + |e(s)| =: \Delta_g(s).$$

■

Remark For MIMO systems, we can use the similar technique in subsection 3.2 to get the final error bound for the reduced transfer matrix.

Remark If $W^{du} = V$, and $V^{du} = W$, and $W^T G(s)V$ is invertible for any s , then $e(s) = 0$. Since

$$\begin{aligned}
e(s) &= (\hat{x}^{du})^* r^{pr} \\
&= (V^{du} z^{du})^* (B - G(s)Vz) \\
&= (W z^{du})^* (B - G(s)Vz) \\
&= (z^{du})^* W^T [B - G(s)V(W^T G(s)V)^{-1} W^T B] \\
&= (z^{du})^* [W^T B - W^T G(s)V(W^T G(s)V)^{-1} W^T B] \\
&= 0.
\end{aligned} \tag{49}$$

Although the matrix pair $W^{du} = V$, $V^{du} = W$ make $|e(s)| = 0$, they are not always the optimal choice for the dual system in the sense of making the residual r^{du} as small as possible. As a result, the error bound which is also influenced by r^{du} decreases probably much slower than using an optimal W^{du} , V^{du} which are different from W , V . In this situation, although no contribution of $|e(s)| = 0$ to the error bound, the contribution of r^{du} is much bigger. Taking the dual system in (16) for a non-parametric LTI system as an example. If only Galerkin projection is used to get the reduced model i.e. $W = V$, then based on the idea of model order reduction and moment-matching MOR (see subsection 8.1.1), the subspace \mathcal{V}^{du} which includes the trajectory of x^{du} in the frequency domain should be $\mathcal{V}^{du} = \mathcal{K}_{q+1}(\tilde{E}_c(s_0), \tilde{C}(s_0))$ defined in (61), so that $\text{range}(V^{du}) = \mathcal{K}_{q+1}(\tilde{E}_c(s_0), \tilde{C}(s_0))$ is the proper choice for the projection matrix V^{du} , rather than $V^{du} = W = V = \mathcal{K}_{q+1}(\tilde{E}_b(s_0), \tilde{B}(s_0))$ defined in (60). Simulation results in Section 9 also support our analysis. However if Petrov-Galerkin projection is used to obtain the reduced model, the choice $W^{du} = V$, and $V^{du} = W$ is already optimal for moment-matching MOR.

5 Output error bound for parametrized LTI systems

In order to derive the error bound for the transfer function $\hat{H}(\mu)$ of the reduced model (10) or (11), we need the parametrized primal system and the parametrized dual system defined in the frequency domain. Similar to the definitions for the non-parametrized system, the primal system is defined as,

$$\begin{aligned}
G(\mu)x(\mu) &= B(\mu), \\
y(\mu) &= C(\mu)x(\mu),
\end{aligned} \tag{50}$$

where the output $y(\mu)$ is exactly the transfer function $H(\mu)$ in (9).

The parametrized dual system is defined as,

$$\begin{aligned}
G^*(\mu)x^{du}(\mu) &= -C(\mu)^T, \\
y^{du}(\mu) &= B(\mu)^T x^{du}(\mu).
\end{aligned} \tag{51}$$

Recall the discussion in Section 4, we also need the residuals caused by the reduced models for the primal and the dual systems. The reduced model of the primal system is obtained by the pair W , V used in (10) or (11).

$$\begin{aligned}
W^T G(\mu)Vz(\mu) &= W^T B(\mu), \\
\hat{y}(\mu) &= C(\mu)Vz(\mu),
\end{aligned} \tag{52}$$

where $\hat{x}(\mu) = Vz(\mu)$ approximates $x(\mu)$. It can be easily seen that $\hat{H}(\mu) = \hat{y}(\mu)$. The reduced model of the dual system is

$$\begin{aligned}
(W^{du})^T G^*(\mu)V^{du} z^{du}(\mu) &= -(W^{du})^T C(\mu)^T, \\
\hat{y}^{du}(\mu) &= B(\mu)^T V^{du} z^{du}(\mu),
\end{aligned} \tag{53}$$

where $\hat{x}^{du}(\mu) = V^{du} z^{du}(\mu)$ is the approximation of $x^{du}(\mu)$. The two residuals are $r^{pr}(\mu) = B(\mu) - G(\mu)\hat{x}(\mu)$ and $r^{du}(\mu) = -C(\mu)^T - G^*(\mu)\hat{x}^{du}(\mu)$. Here the matrix pairs W , V and W^{du} , V^{du} should be computed by PMOR methods (see Section 8.1.2).

Comparing the systems in (50) and (51) with the systems in (14) and (16), we notice that they are in the same form, and only s is changed into μ . It is also true for the corresponding reduced models in (52) and (53) versus the reduced systems in (15) and (42). Therefore, we have a similar theorem for the error bound of $\hat{H}(\mu)$.

Defining two new variables $e(\mu) = (\hat{x}^{du}(\mu))^* r^{pr}(\mu)$ and $\tilde{y}(\mu) = C(\mu)\hat{x}(\mu) - e(\mu)$ and assuming that $G(\mu)$ satisfies

$$\inf_{\substack{w \in \mathbb{C}^n \\ w \neq 0}} \sup_{\substack{v \in \mathbb{C}^n \\ v \neq 0}} \frac{w^* G(\mu)v}{\|w\|_2 \|v\|_2} = \beta(\mu) > 0, \tag{54}$$

we have the following theorem.

Theorem 3 For a SISO linear parametrized system in (6) or in (7), if $G(\mu)$ satisfies (54), then $|y(\mu) - \tilde{y}(\mu)| \leq \tilde{\Delta}_p(\mu)$, $\tilde{\Delta}_p(\mu) := \frac{\|r^{du}(\mu)\|_2 \|r^{pr}(\mu)\|_2}{\beta(\mu)}$. As a result, $|H(\mu) - \hat{H}(\mu)| = |C(\mu)x(\mu) - C(\mu)\hat{x}(\mu)| \leq \Delta_p(\mu)$, $\Delta_p(\mu) := \tilde{\Delta}_p(\mu) + |e(\mu)|$.

Remark The proof is straight forward by following the proof in Section 4. Similar to the analysis in (49), when $W^{du} = V$, $V^{du} = W$, and $W^T G(\mu)V$ is invertible for any μ , $e(\mu) = 0$. However, analogous to the analysis in Section 4, V, W are possibly not the optimal projection matrices for the dual system.

6 Computation of the error bounds

6.1 Computation of $\Delta(s)$

In subsection 3.1, the error bound $\Delta(s)$ is derived for special LTI systems, where E, A are required to be symmetric. $\Delta(s)$ is decided by the two Ritz representation vectors $\hat{\epsilon}^{pr}$ and $\hat{\epsilon}^{du}$, and the two variables $\alpha(s)$ and $\gamma(s)$.

computation of $\alpha(s)$ and $\gamma(s)$

From the two assumptions (12)(13) on $G(s)$, we have for $s = \sigma_0 + j\omega$,

$$\begin{aligned} \frac{\operatorname{Re}(x^* G(s)x)}{x^* \tilde{A}x} &= \frac{\operatorname{Re}(x^*(j\omega E + \sigma_0 E - A)x)}{x^* \tilde{A}x} \\ &= \frac{\operatorname{Re}(x^*(j\omega \tilde{A} + \tilde{A})x)}{x^* \tilde{A}x} \\ &= \frac{x^* \tilde{A}x}{\tilde{x}^* (\tilde{R}^{-1})^* \tilde{A} \tilde{R}^{-1} \tilde{x}} \\ &= \frac{x^* \tilde{A}x}{\tilde{x}^* (\tilde{R}^{-1})^* \tilde{A} \tilde{R}^{-1} \tilde{x}} \\ &\geq \lambda_{\min}((\tilde{R}^{-1})^* \tilde{A} \tilde{R}^{-1}) := \alpha(s), \end{aligned}$$

where $\tilde{A} = \sigma_0 E - A$, $\tilde{R}^* \tilde{R} = \tilde{A}$ is the Cholesky factorization of \tilde{A} , and $\tilde{x} = R x$. $\lambda_{\min}((\tilde{R}^{-1})^* \tilde{A} \tilde{R}^{-1})$ is the smallest eigenvalue of $(\tilde{R}^{-1})^* \tilde{A} \tilde{R}^{-1}$. If E, A are symmetric, $x^* \tilde{A}x$ is the real part of $x^* G(s)x$. Therefore, from the property of Rayleigh quotient, we can take $\alpha(s)$ as the smallest eigenvalue of the symmetric matrix $(\tilde{R}^{-1})^* \tilde{A} \tilde{R}^{-1}$.

Remark If $\sigma_0 = s_0$, then we have a simple case $\alpha(s) = 1$. Usually, $s = j\omega$, and $\sigma_0 = 0$. If A is symmetric negative definite, we can use $\tilde{A} = -A$ to define the norm, then $\alpha = 1$ when $s = j\omega$. If $\tilde{A} = I$, the identity matrix, then $\alpha(s)$ simplifies to the minimal eigenvalue of the matrix \tilde{A} .

Remark For a system with symmetric negative definite matrix A , symmetric positive matrix E , and if $C = B^T$, then the system is passive [22]. Passivity is an important property for linear systems arising from circuit design. An essential reason is that the interconnection of stable subcircuits may not stable, but the interconnection of passive subcircuits is passive. Moreover, passivity implies stability of the system, so that a passive interconnection of subcircuits guarantees stability of the whole circuit.

For the estimation of $\gamma(s)$, we have

$$\begin{aligned} \frac{\operatorname{Im}(x^* G(s)x)}{x^* \tilde{A}x} &= \frac{\omega x^* E x}{x^* \tilde{A}x} \\ &= \frac{\omega \tilde{x}^* (\tilde{R}^{-1})^* E \tilde{R}^{-1} \tilde{x}}{\tilde{x}^* \tilde{x}} \\ &\geq \omega \lambda_{\min}((\tilde{R}^{-1})^* E \tilde{R}^{-1}) := \gamma(s), \end{aligned}$$

$\lambda_{\min}((\tilde{R}^{-1})^* E \tilde{R}^{-1})$ is the smallest eigenvalue of $(\tilde{R}^{-1})^* E \tilde{R}^{-1}$. Since it is assumed that $\gamma(s) > 0$, E must be symmetric positive definite.

Remark Once we have computed the smallest eigenvalue of $(\tilde{R}^{-1})^* \tilde{A} \tilde{R}^{-1}$, we have obtained $\alpha(s)$. Furthermore, $\lambda_{\min}((\tilde{R}^{-1})^* E \tilde{R}^{-1})$ can be computed a priori, then for each value of s , $\gamma(s)$ is $\lambda_{\min}((\tilde{R}^{-1})^* E \tilde{R}^{-1})$ multiplied by ω . That means, the two eigenvalue problems are solved only once, and can be computed independent of s .

computation of $\hat{\epsilon}^{pr}$ and $\hat{\epsilon}^{du}$

Next, we show how to efficiently compute the two Ritz representation vectors $\hat{\epsilon}^{pr}$ and $\hat{\epsilon}^{du}$. From (18), we see that

$$\xi^* r^{pr}(s) = \xi^* \tilde{A} \hat{\epsilon}^{pr}, \quad \forall \xi \in \mathbb{C}^n.$$

If we take $\xi = e_i, e_i = (0, \dots, 0, 1, 0, \dots, 0), i = 1, 2, \dots, n$, where 1 is at the i -th entry of e_i , then

$$\begin{aligned} r_1^{pr}(s) &= \tilde{a}_{11}\hat{\epsilon}_1^{pr} + \dots + \tilde{a}_{1n}\hat{\epsilon}_n^{pr}, \\ r_2^{pr}(s) &= \tilde{a}_{21}\hat{\epsilon}_1^{pr} + \dots + \tilde{a}_{2n}\hat{\epsilon}_n^{pr}, \\ &\vdots \\ r_n^{pr}(s) &= \tilde{a}_{n1}\hat{\epsilon}_1^{pr} + \dots + \tilde{a}_{nn}\hat{\epsilon}_n^{pr}. \end{aligned}$$

Here $r^{pr}(s) = (r_1^{pr}(s), \dots, r_n^{pr}(s))^T$, $(\tilde{A})_{ik} = \tilde{a}_{ik}, 1 \leq i, k \leq n$, and $\hat{\epsilon}^{pr} = (\hat{\epsilon}_1^{pr}, \dots, \hat{\epsilon}_n^{pr})$. In matrix form, it is

$$\tilde{A}\hat{\epsilon}^{pr} = r^{pr}(s).$$

Since \tilde{A} is assumed to be positive definite, $\hat{\epsilon}^{pr}$ is uniquely decided by (6.1): $\hat{\epsilon}^{pr} = \tilde{A}^{-1}r^{pr}(s)$. Similarly from (19), $\hat{\epsilon}_n^{du}$ can be uniquely determined by $\hat{\epsilon}^{du} = \tilde{A}^{-1}r^{du}(s)$.

The two Ritz vectors $\hat{\epsilon}^{pr}$ and $\hat{\epsilon}^{du}$ are functions of s . For each value of s , two full-size linear systems must be solved to get $\hat{\epsilon}^{pr}$, and $\hat{\epsilon}^{du}$, which looks expensive. However, from the formulation of $r^{pr}(s)$, we get

$$\begin{aligned} \hat{\epsilon}^{pr}(s) &= \tilde{A}^{-1}r^{pr}(s), \\ &= \tilde{A}^{-1}(B - sEVz + AVz), \\ &= \tilde{A}^{-1}[B - sEV(sW^T EV - W^T AV)^{-1}W^T B + \\ &\quad AV(sW^T EV - W^T AV)^{-1}W^T B], \\ &= \tilde{A}^{-1}B - s\tilde{A}^{-1}EV(sW^T EV - W^T AV)^{-1}W^T B \\ &\quad + \tilde{A}^{-1}AV(sW^T EV - W^T AV)^{-1}W^T B. \end{aligned}$$

The terms $W^T EV, W^T AV, W^T B$ needs only one-time computation, and can be repeatedly used for any value of s . Although the inverse of $(sW^T EV - W^T AV)$ needs to be computed for each possible value of s , they are of the reduced size, and can be implemented very fast. The matrix V usually has few columns, therefore, the terms $\tilde{A}^{-1}EV, \tilde{A}^{-1}AV, \tilde{A}^{-1}B$ can be computed by solving $2r + m_1$ linear systems. Here r is the number of the columns in V , which is also the size of the reduced model. m_1 is the number of the columns in B . As a result, the estimation of $\hat{\epsilon}^{pr}$ at any fixed value of s can be done efficiently. Likewise, $\hat{\epsilon}^{du}(s)$ can be computed by following the similar implementations. Furthermore, when the standard 2-norm is used, \tilde{A} reduces to the identity matrix. There is no need to solve linear systems.

6.2 Computation of $\Delta_g(s)$ and $\Delta_p(\mu)$

The key for computing $\Delta_g(s)$ or $\Delta_p(\mu)$ is how to compute $\beta(s)$ or $\beta(\mu)$. The condition (41) is equivalent to

$$\inf_{\forall w \in \mathbb{C}^n} \frac{1}{\|w\|_2} \sup_{\forall v \in \mathbb{C}^n} \frac{w^* G(s)v}{\|v\|_2} = \beta(s). \quad (55)$$

On the one hand,

$$\frac{w^* G(s)v}{\|v\|_2} \leq \frac{\|G^*(s)w\|_2 \|v\|_2}{\|v\|_2} = \|G^*(s)w\|_2.$$

On the other hand, when $v = G^*(s)w$,

$$\frac{w^* G(s)v}{\|v\|_2} = \|G^*(s)w\|_2.$$

Therefore,

$$\sup_{\forall v \in \mathbb{C}^n} \frac{w^* G(s)v}{\|v\|_2} = \|G^*(s)w\|_2.$$

Substitute it into (55), we get

$$\inf_{\forall w \in \mathbb{C}^n} \frac{\|G^*(s)w\|_2}{\|w\|_2} = \beta(s).$$

From Rayleigh quotient,

$$\min_{\forall w \in \mathbb{C}^n} \frac{w^* G(s)G^*(s)w}{w^* w} = \lambda_{min}(G(s)G^*(s)).$$

Therefore $\beta(s) = \sqrt{\lambda_{\min}(G(s)G^*(s))}$. Analogously, $\beta(\mu)$ is the square root of the minimal eigenvalue of $G(\mu)G^*(\mu)$. The error bound $\Delta_g(s)$ includes the two residuals $r^{pr}(s)$ and $r^{du}(s)$. For the computation of $r^{pr}(\mu)$ and $r^{du}(\mu)$, we need the affine assumption on $G(\mu)$. That is

$$G(\mu) = E_0 + E_1\mu_1 + \dots + E_p\mu_p. \quad (56)$$

With the affine form, it is not difficult to see that $r^{pr}(\mu)$ and $r^{du}(\mu)$ can be efficiently computed [34].

Remark Clearly, for each value of s , the minimal eigenvalue of a large matrix $G(s)G^*(s)$, or equivalently, the minimal singular value of $G(s)$, must be computed to get $\beta(s)$. It is impractical if $\beta(s)$ must be estimated at many samples of s . The method proposed in [29] can be used to compute a lower bound $\beta_{LB}(s)$ of $\beta(s)$, such that $\beta_{LB}(s)$ could be computed efficiently through solving a sequence of small optimization problems. Without solving any large-scale problems, $\beta_{LB}(s)$ is expected to be quickly available for each sample of s . However, since it is a lower bound of $\beta(s)$, $\Delta_g(s)$ computed by $\beta_{LB}(s)$ would over estimate the real error of the reduced transfer function $\tilde{H}(s)$. It is observed that when the range of s is very large, β_{LB} is not close to $\beta(s)$ at all. Furthermore, the accuracy of $\beta_{LB}(s)$ highly relies on the optimization solvers, which cannot guarantee to converge to an optimal solution for each sample of s . For parametrized systems, the method in [29] will become more complicated, which may easily lead to a meaningless lower bound $\beta_{LB}(\mu)$. More efficient methods for computing or estimating $\beta(s)$ will be the future work.

7 Reformulated reduced system

It is discussed in Section 5 that except for some special cases, the value of $|e(\mu)|$ in the error bound is nonzero in general. Motivated by the analysis in a new research [2], we show in this section that a different reduced model can be constructed from $\tilde{y}(s)$ in Theorem 2 or $\tilde{y}(\mu)$ in Theorem 3, so that $e(s)$ or $e(\mu)$ in the error bound disappear.

The non-parametrized LTI systems discussed in Section 4 can be considered as a special case of the parametrized LTI systems in Section 5, with $\mu = s$. Therefore we generally consider parametrized LTI systems in the following.

From the definition of $e(\mu) = (\hat{x}^{du})^* r^{pr}(\mu)$, and $r^{pr} = B(\mu) - G(\mu)V(W^T G(\mu)V)^{-1}W^T B(\mu)$, $\hat{x}^{du} = V^{du} z^{du} = -V^{du}[(W^{du})^T G^*(\mu)V^{du}]^{-1}(W^{du})^T C(\mu)^*$, we observe that,

$$\begin{aligned} & y^{pr}(\mu) - \tilde{y}^{pr}(\mu) \\ &= C(\mu)x - C(\mu)\hat{x} + (\hat{x}^{du})^* r^{pr} \\ &= C(\mu)x - C(\mu)\hat{x} - C(\mu)(W^{du})[(V^{du})^T G(\mu)W^{du}]^{-1}(V^{du})^T [B(\mu) - G(\mu)V(W^T G(\mu)V)^{-1}W^T B(\mu)] \\ &= C(\mu)G^{-1}(\mu)B(\mu) - C(\mu)V(W^T G(\mu)V)^{-1}W^T B(\mu) - \\ & C(\mu)W^{du}[(V^{du})^T G(\mu)W^{du}]^{-1}(V^{du})^T [B(\mu) - G(\mu)V(W^T G(\mu)V)^{-1}W^T B(\mu)] \\ &= C(\mu)G^{-1}(\mu)B(\mu) - C(\mu)V(W^T G(\mu)V)^{-1}W^T B(\mu) - C(\mu)W^{du}[(V^{du})^T G(\mu)W^{du}]^{-1}(V^{du})^T B(\mu) - \\ & C(\mu)W^{du}[(V^{du})^T G(\mu)W^{du}]^{-1}(V^{du})^T G(\mu)V(W^T G(\mu)V)^{-1}W^T B(\mu). \end{aligned} \quad (57)$$

The right-hand side of the last equality in (57) can be written into the matrix form as below,

$$\underbrace{C(\mu)G^{-1}(\mu)B(\mu)}_{H(\mu)} - \underbrace{\begin{bmatrix} C(\mu)V & C(\mu)W^{du} \end{bmatrix} \begin{bmatrix} W^T G(\mu)V & 0 \\ (V^{du})^T G(\mu)V & (V^{du})^T G(\mu)W^{du} \end{bmatrix}^{-1} \begin{bmatrix} W^T B(\mu) \\ (V^{du})^T B(\mu) \end{bmatrix}}_{\tilde{H}(\mu)}. \quad (58)$$

Clearly, $y^{pr}(\mu) = H(\mu)$ and $\tilde{y}(\mu) = \tilde{H}(\mu)$. If using $\tilde{H}(\mu)$ to approximate $H(\mu)$, the error bound for $\tilde{H}(\mu)$ is $\tilde{\Delta}_p(\mu)$, i.e. $|H(\mu) - \tilde{H}(\mu)| \leq \tilde{\Delta}_p(\mu)$. There is no additional term $|e(\mu)|$ in the error bound.

Next we consider constructing a reduced system whose transfer function is $\tilde{H}(\mu)$. From $\tilde{H}(\mu)$, the corresponding system with zero initial condition¹, can be written as

$$\begin{aligned} & \begin{bmatrix} W^T G(\mu)V & 0 \\ (V^{du})^T G(\mu)V & (V^{du})^T G(\mu)W^{du} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} W^T B(\mu) \\ (V^{du})^T B(\mu) \end{bmatrix}, \\ & y = \begin{bmatrix} C(\mu)V & C(\mu)W^{du} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} \end{aligned}$$

¹Here, we assume that the original system (e.g. (6) or (7)) has zero initial condition. For a system with nonzero initial condition, one can use coordinate transformation $\tilde{x} = x - x(0)$ to get a transformed system with state vector \tilde{x} , and with zero initial condition $\tilde{x}(0) = 0$. A reduced system with zero initial condition can be obtained by applying MOR to the transformed system [17].

in Laplace domain.

If the parametrized LTI system in (6) or (7) is considered, then $G(\mu)$ can be written as $G(\mu) = sE(\tilde{\mu}) - A(\tilde{\mu})$ or $G(\mu) = s^2M(\tilde{\mu}) + sK(\tilde{\mu}) + A(\tilde{\mu})$. Inserting, e.g. $G(\mu) = sE(\tilde{\mu}) - A(\tilde{\mu})$ into (58), we get

$$\begin{bmatrix} s\hat{E}_{11}(\tilde{\mu}) - \hat{A}_{11}(\tilde{\mu}) & 0 \\ s\hat{E}_{21}(\tilde{\mu}) - \hat{A}_{21}(\tilde{\mu}) & s\hat{E}_{22}(\tilde{\mu}) - \hat{A}_{22}(\tilde{\mu}) \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} W^T B(\mu) \\ (V^{du})^T B(\mu) \end{bmatrix},$$

$$y = \begin{bmatrix} C(\mu)V & C(\mu)W^{du} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix},$$

where $\hat{E}_{11}(\tilde{\mu}) = W^T E(\tilde{\mu})V$, $\hat{A}_{11}(\tilde{\mu}) = W^T A(\tilde{\mu})V$, $\hat{E}_{21} = (V^{du})^T E(\tilde{\mu})V$, $\hat{A}_{21}(\tilde{\mu}) = (V^{du})^T A(\tilde{\mu})V$, $\hat{E}_{22}(\tilde{\mu}) = (V^{du})^T E(\tilde{\mu})W^{du}$, $\hat{A}_{22}(\tilde{\mu}) = (V^{du})^T A(\tilde{\mu})W^{du}$.

Using inverse Laplace transform, the reduced system in time domain is

$$\begin{bmatrix} \hat{E}_{11}(\tilde{\mu}) & 0 \\ \hat{E}_{21}(\tilde{\mu}) & \hat{E}_{22}(\tilde{\mu}) \end{bmatrix} \begin{bmatrix} \dot{z}_1(t) \\ \dot{z}_2(t) \end{bmatrix} = \begin{bmatrix} \hat{A}_{11}(\tilde{\mu}) & 0 \\ \hat{A}_{21}(\tilde{\mu}) & \hat{A}_{22}(\tilde{\mu}) \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix} + \begin{bmatrix} W^T B(\tilde{\mu}) \\ (V^{du})^T B(\tilde{\mu}) \end{bmatrix} u(t), \quad (59)$$

$$y = \begin{bmatrix} C(\tilde{\mu})V & C(\tilde{\mu})W^{du} \end{bmatrix} \begin{bmatrix} z_1(t) \\ z_2(t) \end{bmatrix}.$$

It shows that for the original system in (6), a reduced system in (59) can be derived, whose transfer function is $\tilde{H}(s)$, satisfying $|H(\mu) - \tilde{H}(\mu)| \leq \tilde{\Delta}_p(\mu)$. For the second order parametrized system in (7), the corresponding reduced system can also be obtained in a similar way. Notice also that the reduced system (59) cannot be obtained by means of an explicit Petrov-Galerkin projection applied to the original system in (6). Instead the projection

$$\tilde{W} = \begin{bmatrix} W & \\ & V^{du} \end{bmatrix} \quad \tilde{V} = \begin{bmatrix} V & \\ & W^{du} \end{bmatrix}$$

is applied to a non-minimal realization of the original system, namely

$$\begin{bmatrix} E(\tilde{\mu}) & 0 \\ E(\tilde{\mu}) & E(\tilde{\mu}) \end{bmatrix} \begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} A(\tilde{\mu}) & 0 \\ A(\tilde{\mu}) & A(\tilde{\mu}) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} + \begin{bmatrix} B(\tilde{\mu}) \\ B(\tilde{\mu}) \end{bmatrix} u(t),$$

$$y = \begin{bmatrix} C(\tilde{\mu}) & C(\tilde{\mu}) \end{bmatrix} \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix}.$$

In summary, there are two reduced systems available for the original system in (6), one is the reduced system in (10) constructed directly from the original system; the other is the reformulated reduced system in (59). On the one hand, if using the same stopping criteria ϵ_{tol} in Algorithm 2, the reduced system in (10) is usually less accurate than the one in (59) (though both satisfy the error tolerance ϵ_{tol}), because the error bound $\Delta_p(\mu)$ for $\tilde{H}(\mu)$ is rougher than $\tilde{\Delta}_p(\mu)$ for $\tilde{H}(\mu)$. On the other hand, the reduced system in (59) could be of much bigger size than the one in (10). From this point of view, the reduced system in (10) is practically preferable, since it is of much smaller size and also satisfies the acceptable error tolerance. The analysis is aided by an example in Section 9.

8 Automatic generation of the reduced models

We explore algorithms of automatic constructing reliable reduced-order models in this section. In particular, we show that the Krylov subspace based MOR methods can be adaptively implemented using the proposed error bounds. To this end, we first present a brief review of the MOR methods, and point out the necessity of adaptive implementation of the methods.

8.1 Review of Krylov subspace based MOR methods

8.1.1 Moment-matching MOR for non-parametrized LTI systems

For moment-matching methods, the matrices W , V are constructed from the transfer function (matrix) $H(s)$ in (4). If we expand $H(s)$ into series, around an expansion point s_0 as

$$\begin{aligned} H(s) &= C[(s - s_0 + s_0)E - A]^{-1}B \\ &= C[(s - s_0)E + (s_0E - A)]^{-1}B \\ &= C[I + (s_0E - A)^{-1}E(s - s_0)]^{-1}(s_0E - A)^{-1}B \\ &= \sum_{i=0}^{\infty} \underbrace{C[-(s_0E - A)^{-1}E]^i (s_0E - A)^{-1}B}_{:=m_i(s_0)} (s - s_0)^i, \end{aligned}$$

then $m_i(s_0)$, $i = 0, 1, 2, \dots$ are called the i th order moments of the transfer function. The columns of V span the Krylov subspace

$$\text{range}\{V\} = \mathcal{K}_{q+1}(\tilde{E}_b(s_0), \tilde{B}(s_0)) := \text{span}\{\tilde{B}(s_0), \dots, (\tilde{E}_b(s_0))^q \tilde{B}(s_0)\}, \quad (60)$$

where $\tilde{E}_b(s_0) = (s_0 E - A)^{-1} E$, $\tilde{B}(s_0) = (s_0 E - A)^{-1} B$. The columns of W span the Krylov subspace

$$\text{range}\{W\} = \mathcal{K}_{q+1}(\tilde{E}_c(s_0), \tilde{C}(s_0)) := \text{span}\{\tilde{C}(s_0), \dots, (\tilde{E}_c(s_0))^q \tilde{C}(s_0)\}, \quad (61)$$

where $\tilde{C}(s_0) = (s_0 E - A)^{-T} C^T$, $\tilde{E}_c(s_0) = (s_0 E - A)^{-T} E^T$. Obviously, W and V span two Krylov subspaces. It is proved in [23], that the transfer function of the reduced model $\hat{H}(s)$ matches the first $2q + 1$ moments of the original transfer function $H(s)$. It is obvious that the accuracy of the reduced model depends on the expansion point s_0 . In many cases, only a single expansion point is insufficient to attain the required accuracy. Multiple-point expansion is preferred such that the large error caused at frequencies far away from the expansion point can be reduced. A reduced model of better accuracy and smaller order can be obtained by multiple-point expansion. Therefore proper selection of multiple expansion points is important. Previous studies on multiple-point expansion are found in [1, 26, 20, 12, 14, 27]. In [26], the expansion points are chosen such that the reduced model is locally optimal. Binary principle is used in [1, 20, 12] for adaptive, but heuristic selection of the expansion points. In subsection 8.2, we readdress the problem of selecting multiple expansion points by using the global a posteriori error bounds proposed in Section 3 and Section 4.

8.1.2 Review of multi-moment matching PMOR methods

Multi-moment matching PMOR methods can be found in [42, 16, 18, 43, 19]. In this section, a robust PMOR method in [19] is reviewed. All these methods are based on Galerkin projection, i.e. $W = V$. Assume that $G(\mu)$ has the affine form defined in (56). To compute the matrix V , the series expansion of the state x in (8) is needed. Given an expansion point $\mu^0 = [\mu_1^0, \mu_2^0, \dots, \mu_p^0]$, x can be expanded as

$$\begin{aligned} x &= [I - (\sigma_1 M_1 + \dots + \sigma_p M_p)]^{-1} \tilde{B}_M u(\mu_p) \\ &= \sum_{i=0}^{\infty} (\sigma_1 M_1 + \dots + \sigma_p M_p)^i \tilde{B}_M u(\mu_p), \end{aligned}$$

where $\tilde{B}_M(\mu) = [G(\mu^0)]^{-1} B(\mu)$, $M_i = -[G(\mu^0)]^{-1} E_i$, $i = 1, 2, \dots, p$, and $\sigma_i = \mu_i - \mu_i^0$, $i = 1, 2, \dots, p$. We call the coefficients in the above series expansion the moment matrices of the parametrized system. The corresponding multi-moments of the transfer function are those moment matrices multiplied by C from the left.

To get the projection matrix V , instead of directly computing the moment matrices [16], a numerically robust method is proposed in [19] (the detailed algorithm is described in [11]). The method combines the recursions in (62) below, with a repeated modified Gram-Schmidt process so that the moment matrices are computed implicitly.

$$\begin{aligned} R_0 &= B_M, \quad R_1 = [M_1 R_0, \dots, M_p R_0], \\ R_2 &= [M_1 R_1, \dots, M_p R_1], \\ &\vdots \\ R_q &= [M_1 R_{q-1}, \dots, M_p R_{q-1}], \\ &\vdots \end{aligned} \quad (62)$$

where $B_M = \tilde{B}_M$, if $B(\mu)$ dose not depend on μ . Otherwise, $B_M = [\tilde{B}_{M1}, \dots, \tilde{B}_{Mp}]$, if $B(\mu)$ can be affinely approximated by $B(\mu) \approx B_1 \mu_1 + \dots + B_p \mu_p$. Here $\tilde{B}_{Mi} = [G(\mu^0)]^{-1} B_i$, $i = 1, \dots, p$. The computed $V = V_{\mu^0}$ is an orthonormal basis of the subspace spanned by the moment matrices,

$$\text{range}\{V_{\mu^0}\} = \text{span}\{R_0, R_1, \dots, R_q\}_{\mu^0}, \quad (63)$$

and depends on the expansion point μ^0 . It is proved in [11] that the multi-moments of the transfer function of the original system can be matched by those of the transfer function of the reduced model. The accuracy of the reduced model can be improved by increasing the number of the terms in (63), whereby more multi-moments can be matched.

It is noticed that the dimension of R_j increases exponentially. If the number of the parameters p in a parametrized system is larger than 2, it is advantageous to use multiple point expansion, such that only the low order moment matrices, e.g. R_j , $j \leq 2$ have to be computed for each expansion point. As a result, the order of the reduced model

can be kept small. Given a group of expansion points $\mu^i, i = 0, \dots, exp$, a matrix V_{μ^i} can be computed from (63) for each μ^i as

$$\text{range}\{V_{\mu^i}\} = \text{span}\{R_0, R_1, \dots, R_q\}_{\mu^i}.$$

The final projection matrix V is a combination (orthogonalization) of all the matrices V_{μ^i} ,

$$V = \text{orth}\{V_{\mu^0}, \dots, V_{\mu^{exp}}\}. \quad (64)$$

Here, selecting the expansion points μ^i is unavoidable. Algorithm 2 is proposed in subsection 8.2 for adaptively selecting the expansion points μ^i using the a posteriori error bound $\Delta_p(\mu)$ in Section 5.

8.2 Algorithms of automatic generation of reduced-order models

The algorithms in this section follow the idea of the greedy algorithm widely used in the reduced basis community. A large sample space Ξ_{train} of the variable s or the vector of parameters μ , covering the whole interesting frequency/parameter domain, must be initially given. During each step of the algorithm, a point \hat{s} or $\hat{\mu}$ in Ξ_{train} , which causes the largest error (indicated by the error bound $\Delta(s)$, $\Delta_g(s)$ or $\Delta_p(\mu)$), is chosen as the next expansion point. The process continues until the largest error among all the samples in Ξ_{train} is smaller than an acceptable error tolerance $\epsilon_{tol} (< 1)$ for the reduced model.

Algorithm 1 Automatic generation of the reduced model by adaptively selecting expansion points \hat{s} for non-parametrized LTI systems

- 1: $W = \emptyset; V = \emptyset;$
 - 2: $\epsilon = 1;$
 - 3: **Initial expansion point:** $\hat{s};$
 - 4: Ξ_{train} : a large set of samples of s , taken over the interesting range of the frequency.
 - 5: **while** $\epsilon > \epsilon_{tol}$ **do**
 - 6: $\text{range}(V_{\hat{s}}) = \mathcal{K}_{q+1}(\tilde{E}_b(\hat{s}), \tilde{B}(\hat{s}));$
 - 7: $\text{range}(W_{\hat{s}}) = \mathcal{K}_{q+1}(\tilde{E}_c(\hat{s}), \tilde{C}(\hat{s}));$
 - 8: $V = \text{orth}\{V, V_{\hat{s}}\}; W^{du} = V;$
 - 9: $W = \text{orth}\{W, W_{\hat{s}}\}; V^{du} = W;$
 - 10: $\hat{s} = \arg \max_{s \in \Xi_{train}} \Delta(s)$ (or $\Delta_g(s)$);
 - 11: $\epsilon = \Delta(\hat{s})$ (or $\Delta_g(\hat{s})$);
 - 12: **end while.**
-

Remark Petrov-Galerkin projection with $W \neq V$ is used in the algorithm. One can certainly use only V to get the reduced model, which reduces to the Galerkin projection method in [36]. It is discussed in [36] that by using the Galerkin projection, the reduced model preserves the passivity of the original system, which is an important property in circuit simulation. To compute the error bound $\Delta_g(s)$, W^{du} , V^{du} are needed. In the algorithm, we use $W^{du} = V$ and $V^{du} = W$, such that the second part $|e(s)|$ in the error bound reduces to zero. However, as is discussed at the end of Section 4, if for Galerkin projection where only one projection matrix V^{du} is needed for the dual system, it is preferred to use

$$\text{range}(V_{\hat{s}}^{du}) = \mathcal{K}_{q+1}((\tilde{E}_c(\hat{s})), \tilde{C}(\hat{s})), \quad (65)$$

for a chosen expansion point \hat{s} , rather than $V^{du} = V$, i.e. V^{du} should be computed based on the trajectory of the state x^{du} of the dual system.

In the following, we consider an algorithm for parametrized LTI systems. Since the multi-moment matching PMOR method in subsection 8.1.2 is a Galerkin projection method. We use also Galerkin projection in the algorithm, though the algorithm can be straight forwardly extended to any other Petrov-Galerkin methods. From (64), we see that the reduced model depends on the expansion points $\mu_i, i = 0, \dots, exp$. Algorithm 2 adaptively chooses the multiple expansion points, which cause the largest errors at the subsequent iteration steps. The next expansion point $\hat{\mu} = (\hat{\mu}_1, \dots, \hat{\mu}_p)$ is chosen as the point at which the current reduced transfer function $\hat{H}(\mu)$ has the biggest error measured by the error bound $\Delta_p(\mu)$. The projection matrices for the dual system at a particular expansion point $\hat{\mu}$ are chosen as $W_{\hat{\mu}}^{du} = V_{\hat{\mu}}^{du} = \text{span}\{R_0^{du}, \dots, R_q^{du}\}_{\hat{\mu}}$, where $R_0^{du}, \dots, R_q^{du}$ are composed of the moment matrices of the dual system, and are defined analogously as for R_0, \dots, R_q in (62). In particular,

$R_j^{du} = [M_1^{du} R_{j-1}, \dots, M_p^{du} R_{j-1}]$, $j = 1, \dots, q, \dots$. $R_0^{du} = [G^*(\hat{\mu})]^{-1}(-C^T)$, and $M_i^{du} = [G^*(\hat{\mu})]^{-1}E_i^T$, $i = 1, \dots, p$.

Algorithm 2 Automatic generation of the reduced model by adaptively selecting expansion points $\hat{\mu}$ for parametrized LTI systems

```

1:  $W = \emptyset; V = \emptyset;$ 
2:  $\epsilon = 1;$ 
3: Initial expansion point:  $\hat{\mu};$ 
4:  $\Xi_{train}$ : a large set of the samples of  $\mu$ , taken over the interesting range of all the parameters  $\mu_1, \dots, \mu_p.$ 
5: while  $\epsilon > \epsilon_{tol}$  do
6:    $V_{\hat{\mu}} = \text{span}\{R_0, \dots, R_q\}_{\hat{\mu}};$ 
7:    $V = \text{orth}\{V, V_{\hat{\mu}}\}; W = V;$ 
8:    $V_{\hat{\mu}}^{du} = \text{span}\{R_0^{du}, \dots, R_q^{du}\}_{\hat{\mu}};$ 
9:    $V^{du} = \text{orth}\{V^{du}, V_{\hat{\mu}}^{du}\}; W^{du} = V^{du};$ 
10:   $\hat{\mu} = \text{arg max}_{\mu \in \Xi_{train}} \Delta_p(\mu);$ 
11:   $\epsilon = \Delta_p(\hat{\mu});$ 
12: end while.

```

9 Simulation results

In this section, we use some examples to show the performance of the error bounds. There are four examples. Three of them are non-parametrized LTI systems, two of which have symmetric E and A . The last example is a parametrized LTI system with four parameters. We use Galerkin projection $W = V$ to get the reduced models for all the examples. When computing $\Delta_g(s)$, the projection matrix V^{du} for the dual system is computed by (65).

When computing $\Delta(s)$, we take $\tilde{A} = I$, such that the norm $\|\cdot\|_{\tilde{A}}$ reduces to the standard 2-norm. In this case, the two Ritz representation vectors $\tilde{\epsilon}^{pr}$, $\tilde{\epsilon}^{du}$ equal to the two residuals, therefore, there is no need to solve the two linear systems in subsection 6.1.

The error bounds derived in the above sections are designed for the absolute error, e.g. $\epsilon^{ab}(\mu) = |H(\mu) - \hat{H}(\mu)|$ for a SISO system. In the following results, we also show the performance of the error bound for the relative error defined as $\epsilon^{re}(\mu) = \epsilon^{ab}/|H(\mu)|$. Accordingly, $\Delta^{re}(\mu) = \Delta(\mu)/|\hat{H}(\mu)|$, $\Delta_g^{re}(\mu) = \Delta_g(\mu)/|\hat{H}(\mu)|$, $\Delta_p^{re}(\mu) = \Delta(\mu)_p/|\hat{H}(\mu)|$ are used as the error bounds for the relative errors, since $H(\mu)$ is never computed in practice.

For a MIMO system, the true error is firstly defined entry-wisely, then the maximum is taken, so that $\epsilon^{ab}(\mu) = \max_{ij} |H_{ij}(\mu) - \hat{H}_{ij}(\mu)|$ is the absolute true error, and $\epsilon^{re}(\mu) = \max_{ij} \frac{|H_{ij}(\mu) - \hat{H}_{ij}(\mu)|}{|H_{ij}(\mu)|}$ is the relative true error, $i = 1, \dots, m_1, j = 1, \dots, m_2$. The error bound for the absolute error is already defined in (40), $\Delta(\mu) = \max_{ij} \Delta_{ij}(\mu)$.

The error bound for the relative error is defined as $\Delta^{re}(\mu) = \max_{ij} \frac{\Delta_{ij}(\mu)}{|\hat{H}_{ij}(\mu)|}$. The same definitions also apply to $\Delta_g(\mu)$ and $\Delta_p(\mu)$. For both SISO and MIMO systems, when there are no parameters, $\mu = s$ in the error bounds as well as in the true errors.

At each iteration step of Algorithm 1 and Algorithm 2, the maximal error bound in Ξ_{train} , is computed, and is used as the error bound for the reduced model. Therefore, the maximal true error $\epsilon_{max}^{ab} = \max_{\mu_i \in \Xi} \epsilon^{ab}(\mu_i)$ or $\epsilon_{max}^{re} = \max_{\mu_i \in \Xi} \epsilon^{re}(\mu_i)$ is used for a comparison. In the following, Ξ in the definitions of ϵ_{max}^{ab} or ϵ_{max}^{re} may refer to Ξ_{train}^j or $\Xi_{ver}^j, j = 1, 2, 3, 4$. When the error bound for $\tilde{H}(\mu)$ of the reformulated reduced system in Section 7 is studied, the corresponding true errors are denoted by $\tilde{\epsilon}_{max}^{ab}$ and $\tilde{\epsilon}_{max}^{re}$ respectively.

9.1 Results for $\Delta(s)$

The two examples of LTI systems involved in this subsection have symmetric positive definite matrix E and symmetric matrix A . One example is a SISO system, the model of a peek inductor. The other is a single-input, multiple-output (SIMO) system, the model of an optical filter. They can be found in the Oberwolfach MOR benchmark Collection². Both examples are in the form of (1). The model of the peek inductor is of size $n = 1434$, and

²URL: <http://portal.uni-freiburg.de/imteksimulation/downloads/benchmark>

the size of the optical filter model is $n = 1668$. There are 5 outputs for the filter model. The working frequency range for the peek inductor is $f \in [0, 10GHz]$, numerically $f \in [0, 10^{10}]$. The optical filter is assumed to work in the range of $f \in [0, 1KHz]$, numerically $f \in [0, 10^3]$. For each example, the variable s is sampled in the above frequency interval to form the sample space Ξ_{train} in Algorithm 1.

Example 1: the peek inductor

For this example, we take the sample space as:

$$\Xi_{train}^1 : \{s_i = 2\pi j f_i, f_i = 10^{(i/100)}, i = 1, \dots, 1000\}$$

where s_i are the samples in Ξ_{train}^1 . The first 6 moments ($q = 5$ in Algorithm 1) are matched for each chosen expansion point \hat{s} . The initial expansion point is taken as $\hat{s} = 2\pi j \hat{f} = 2\pi j$, with $\hat{f} = 1$. Three more expansion points are adaptively selected by Algorithm 1. Finally, a reduced model of order $r = 24$, and with sufficient accuracy is derived. The results are listed in Table 1. $\Delta^{re}(\hat{s})$ in the table is the relative error bound at the selected

Table 1: Peek inductor $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n = 1434$, $r = 24$, Ξ_{train}^1

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{re}	$\Delta^{re}(\hat{s})$
1	1	0.23	1.86×10^4
2	1×10^{10}	0.04	2.85×10^3
3	4×10^7	6.6×10^{-5}	0.3
4	3.89×10^8	4×10^{-8}	3.5×10^{-4}

expansion point \hat{s} , which is also the maximal error of the reduced model in Ξ_{train}^1 estimated by $\Delta^{re}(s)$. ϵ_{max}^{re} is the true maximal relative error of the reduced model in Ξ_{train}^1 , at the current iteration step. The final reduced model is obtained at the last iteration step, and this is also the case for the results in all the Tables (Figures) below. The data shows that $\Delta^{re}(s)$ is a bound for the true error of the reduced model at all the samples in Ξ_{train}^1 . The results for the absolute error is listed in Table 2, where $\Delta(\hat{s})$ has been demonstrated to be a rigorous bound for ϵ_{max}^{ab} , the maximal absolute error of the reduced model in Ξ_{train}^1 .

Table 2: Peek inductor $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n = 1434$, $r = 24$, Ξ_{train}^1

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{ab}	$\Delta(\hat{s})$
1	1	0.02	252.8
2	1.4×10^7	1.9×10^{-4}	2.42
3	1×10^{10}	3.6×10^{-6}	2.5×10^{-2}
4	1.2×10^8	7.5×10^{-9}	1×10^{-4}

In Figure 1, we further show that $\Delta^{re}(s)$ can actually bound the true error $\epsilon^{re}(s)$ of the final reduced model in the whole frequency range. That means, if we use more densely distributed samples: e.g. 2000 exponentially distributed samples,

$$\Xi_{ver}^1 : \{s_i = 2\pi j f_i, i = 1, \dots, 2000, f_i = 10^{(i/200)}\}$$

to represent the whole interesting frequency interval $[0, 10^{10}]$, the errors of the reduced model at those sample points are still smaller than $\Delta^{re}(\hat{s}) = 3.5 \times 10^{-4}$ at the last iteration step in Table 1, which is the error bound for the final reduced model.

Similar results can be given by the absolute error bound $\Delta(s)$, and will not be repeated.

It is interesting to see that the training sample space Ξ_{train} doesn't have to be too rich. If 2000 samples are taken to form the training space (instead of the previously used 1000 samples in Ξ_{train}^1):

$$\Xi_{train}^{rich_1} : \{f_i = 10^{(i/200)}, s_i = 2\pi j f_i, i = 1, \dots, 2000\},$$

the results in Table 3 and 4 are more or less the same as those in Table 1 and 2.

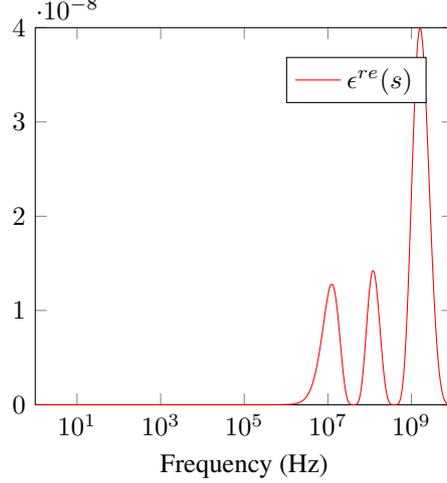


Figure 1: Peek inductor, the true error of the reduced model over Ξ_{ver}^1 .

Table 3: Peek inductor $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n = 1434$, $r = 24$, $\Xi_{train}^{rich_1}$

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{re}	$\Delta^{re}(\hat{s})$
1	1	0.23	1.86×10^4
2	1×10^{10}	0.04	2.85×10^3
3	4×10^7	6.6×10^{-5}	0.3
4	3.89×10^8	4×10^{-8}	3.5×10^{-4}

Example 2: the tunable optical filter³

The second example is a SIMO system, so the definitions for true errors and the error bounds of the MIMO systems are used.

In this example, we choose 600 sample points in the interesting frequency interval $f \in [0, 10^3]$ to form the sample space,

$$\Xi_{train}^2 : \{f_i = 10^{(i/200)}, s_i = 2\pi j f_i, i = 1, \dots, 600\}.$$

Finally, 42 samples are selected as the expansion points. The reduced model is of order $r = 12$. In Algorithm 1, we take $q = 1$, i.e. the first 2 moments are matched for each expansion point.

The dashed line in Figure 2(a) shows the absolute error bound $\Delta(\hat{s})$ at each of the selected expansion point \hat{s} , which bounds the true maximal absolute error ϵ_{max}^{ab} , plotted by the solid line.

In order to validate the error bound, the true errors $\epsilon^{ab}(s)$ of the final reduced model at more dense sample points are plotted in Figure 2(b). There are 2100 exponentially distributed sample points taken in $[0, 10^3]$:

$$\Xi_{ver}^2 : \{f_i = 10^{(i/700)}, s_i = 2\pi\sqrt{-1}f_i, i = 1, \dots, 2100\}.$$

The true error at each sample point is also below the error bound at the final iteration step in Figure 2(a).

The results of the relative error bound $\Delta^{re}(s)$ is presented in Table 5. Here the final reduced model with good accuracy is obtained within 5 iteration steps, and 5 expansion points \hat{s} are used.

9.2 Results for $\Delta_g(s)$

We use the model of an interconnect⁴ to demonstrate the behavior of $\Delta_g(s)$. The model is of size $n = 6134$. Since the matrix E is singular, the error bound $\Delta(s)$ is not valid anymore. The interesting frequency range is $f \in [0, 3GHz]$, i.e. $f \in [0, 3 \times 10^9]$.

A training sample space containing 900 samples:

$$\Xi_{train}^3 : \{f_i = 3 \times 10^{(i/100)}, s_i = 2\pi j f_i, i = 1, \dots, 900\}$$

³From the Oberwolfach model reduction benchmark collection <http://simulation.uni-freiburg.de/downloads/benchmark>

⁴The detailed description for the example can be found in [20]

Table 4: Peek inductor $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n = 1434$, $r = 24$, $\Xi_{train}^{rich_1}$

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{ab}	$\Delta(\hat{s})$
1	1	0.02	252.8
2	1.4×10^7	1.9×10^{-4}	2.42
3	1×10^{10}	3.6×10^{-6}	2.5×10^{-2}
4	1.2×10^8	7.5×10^{-9}	9.2×10^{-5}

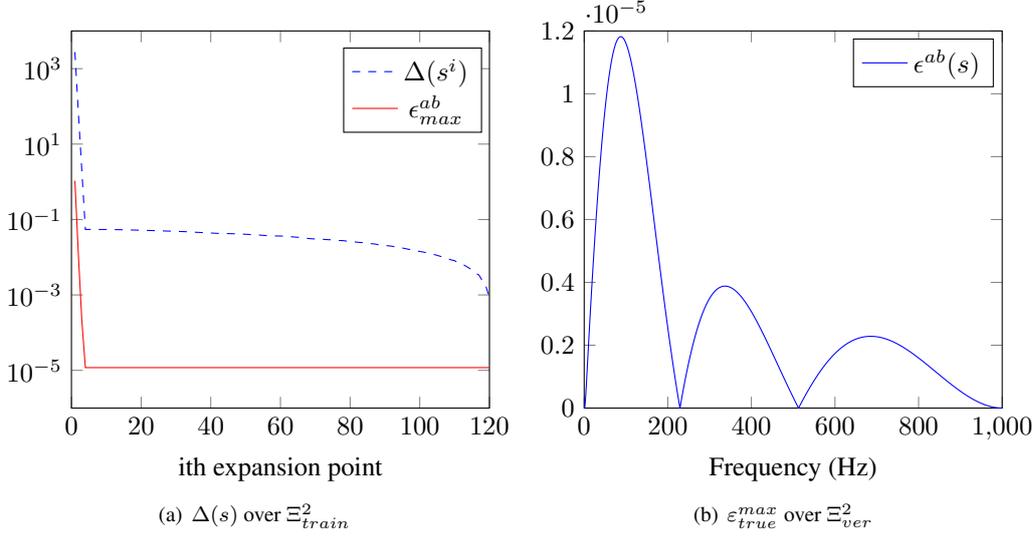


Figure 2: Optical filter, $q = 1$, $\epsilon_{tol} = 10^{-3}$, $n = 1668$, $r = 12$.

is used to compute the expansion points. The simulation results of $\Delta_g(s)$ and of $\Delta_g^{re}(s)$ are listed in Table 6 and Table 7. A reduced model is successfully found by using either $\Delta_g(s)$ or $\Delta_g^{re}(s)$ after 7 iteration steps.

To test the rigorousness of the error bound $\Delta_g(s)$, we plot in Figure 3 the true absolute errors of the final reduced model at 2700 exponentially distributed samples in $f \in [0, 3 \times 10^9]$:

$$\Xi_{ver}^3 : \{f_i = 3 \times 10^{(i/300)}, s_i = 2\pi\sqrt{-1}f_i, i = 1, \dots, 2700\}.$$

The errors are all below the error bound $\Delta_g(\hat{s}) = 6.64 \times 10^{-8}$ in Table 6, the error bound for the final reduced model derived at the last iteration step.

It is analyzed at the end of Section 4 that for Galerkin projection, the error bound $\Delta_g(s)$ using V^{du} computed from (65) should perform better than the error bound, say $\Delta_g^0(s)$ using $V^{du} = V$. Here we show the behavior of $\Delta_g^0(s)$ in Figure 4. With the same inputs for Algorithm 1, we compare $\Delta_g^0(s)$ in Figure 4 to $\Delta_g(s)$ in Table 6. It is obvious that $\Delta_g^0(s)$ decreases much slower than $\Delta_g(s)$. Using $\Delta_g(s)$, the algorithm constructs the final reduced model within 7 iteration steps; while using $\Delta_g^0(s)$, it takes 84 iterations. It is observed that the residual of the dual system $\|r^{du}\|_2$ decreases at least as fast as $\|r^{pr}\|_2$ if $V^{du} \neq V$. However, $\|r^{du}\|_2$ stagnates at around $O(1)$ for the case $V^{du} = V$, while $\|r^{pr}\|_2$ keeps decreasing.

9.3 Results for $\Delta_p(\mu)$ and $\tilde{\Delta}_p(\mu)$

In this subsection, we show the behavior of the error bounds $\Delta_p(\mu)$ and $\tilde{\Delta}_p(\mu)$ in Theorem 3 by using the parameterized system for a MEMS model⁵ as an example. It is of the following formulation

$$\begin{aligned} M(d)\ddot{x} + D(\theta, \alpha, \beta, d)\dot{x} + T(d)x &= Bu(t), \\ y &= Cx. \end{aligned}$$

Here, $M(d) = (M_1 + dM_2)$, $T(d) = (T_1 + \frac{1}{d}T_2 + dT_3)$, $D(\theta, \alpha, \beta, d) = \theta(D_1 + dD_2) + \alpha M(d) + \beta T(d) \in R^{n \times n}$, $n=17,913$. The parameters are d, θ, α, β .

⁵Benchmark available at <http://modlreduction.org>.

Table 5: The tunable optical filter $q = 1$, $\epsilon_{tol} = 10^{-3}$, $n = 1668$, $r = 12$

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{re}	$\Delta^{re}(\hat{s})$
1	1	2.5	3.2×10^3
2	966	3.8×10^{-3}	21.1
3	462.4	3.4×10^{-5}	0.17
4	676	4.8×10^{-6}	7.9×10^{-2}
5	188.4	9.3×10^{-8}	8.7×10^{-4}

Table 6: RLC tree $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n = 6134$, $r = 24$, Ξ_{train}^3

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{ab}	$\Delta_g(\hat{s})$
1	0	0.3916	2.93×10^6
2	3.0000×10^9	5.4495×10^{-4}	3.93×10^5
3	1.7665×10^9	4.1075×10^{-8}	31.54
4	1.1146×10^9	4.1076×10^{-8}	3.11
5	2.1733×10^9	4.1073×10^{-8}	0.23
6	2.4385×10^9	4.1070×10^{-8}	3.3×10^{-3}
7	0.3525×10^9	4.1077×10^{-8}	6.64×10^{-8}

After Laplace transform, the system in frequency domain is

$$\begin{aligned} s^2 M(d)x + sD(\theta, \alpha, \beta, d)x + T(d)x &= Bu(s), \\ y &= Cx. \end{aligned}$$

The above system can be rewritten into the affine form,

$$\begin{aligned} G(\mu)x &= Bu(\mu), \\ y &= Cx, \end{aligned}$$

where $G(\mu) = T_1 + \mu_1 M_1 + \mu_2 M_2 + \mu_3 D_1 + \mu_4 D_2 + \mu_5 M_1 + \mu_6 M_2 + \mu_7 T_1 + \mu_8 T_2 + \mu_9 T_3 + \mu_{10} T_2 + \mu_{11} T_3$. Here $\mu = (\mu_1, \dots, \mu_{11})^T$ includes the newly generated parameters, $\mu_1 = s^2$, $\mu_2 = s^2 d$, $\mu_3 = s\theta$, $\mu_4 = s\theta d$, $\mu_5 = s\alpha$, $\mu_6 = s\alpha d$, $\mu_7 = s\beta$, $\mu_8 = s\beta/d$, $\mu_9 = s\beta d$, $\mu_{10} = 1/d$, $\mu_{11} = d$.

The transfer function of this system is of very small magnitude, which is in the interval $[10^{-7}, 10^{-4}]$ [21]. Therefore, the tolerance ϵ_{tol} for the absolute error of the reduced model is assigned a small value $\epsilon_{tol} = 10^{-7}$. The tolerance ϵ_{tol} for the relative error is taken as $\epsilon_{tol} = 10^{-2}$.

For this example, the training sample space is taken as Ξ_{train}^4 :

$$\{3 \text{ random } \theta \in [10^{-7}, 10^{-5}], 10 \text{ random } s, 5 \text{ random } d \in [1, 2], \text{ and } \alpha = 0, \beta = 0\}.$$

The frequency range for $s = 2\pi j f$ is $f \in [0.025, 0.25] \text{ KHz}$, numerically $f \in [0.025, 0.25]$. There are totally 150 samples of $\mu = (\mu_1, \dots, \mu_{11})$. It is indicated in [44], that $\alpha = 0$ and $\beta = 0$ do not affect the accuracy of the reduced model, therefore they are taken as zeros in Ξ_{train}^4 .

Behavior of $\Delta_p(\mu)$

Figure 5 shows the error bound $\Delta_p(\mu)$ and the true absolute error ϵ_{max}^{ab} at each iteration step of Algorithm 2. The plot on the right is the effectivity $\frac{\Delta_p(\mu)}{\epsilon_{max}^{ab}}$, which shows the sharpness of the error bound. It is already below 10 at the final iterations in Algorithm 2, showing the error bound is close to the true error. Here R_0, R_1, R_2 in (62) are used for each expansion point μ^i to generate the projection matrix V , the resulting reduced model is of size 804, where 33 expansion points have been selected.

To further reduce the size of the reduced model, one may use only R_0, R_1 for each μ^i . The case is shown in Figure 6. The computed reduced model is of a much smaller size 210, and 36 expansion points have been selected.

To verify the reduced model obtained by the above two cases,

Table 7: RLC tree $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n = 6134$, $r = 26$, Ξ_{train}^3

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{re}	$\Delta_g^{re}(\hat{s})$
1	0	1.25	6.07×10^7
2	2.67×10^9	1.4660×10^{-4}	6.08×10^4
3	1.77×10^9	3.3174×10^{-8}	22.69
4	9.27×10^8	3.3161×10^{-8}	8.23
5	3.00×10^9	8.7115×10^{-10}	1.47
6	1.34×10^9	8.7809×10^{-10}	1.32
7	3.44×10^8	8.6021×10^{-10}	9.70×10^{-6}

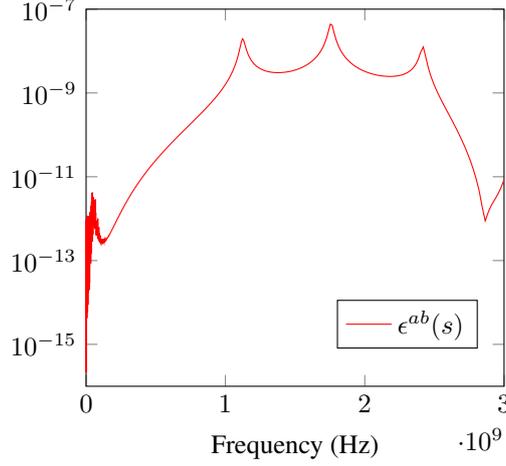


Figure 3: The true error of the reduced model (with size $r = 24$) over Ξ_{ver}^3 .

- Case 1: $V_{\mu^i} = \text{span}\{B_M, R_1, R_2\}_{\mu^i}$,
- Case 2: $V_{\mu^i} = \text{span}\{B_M, R_1\}_{\mu^i}$,

a much denser sample space is taken as

$$\Xi_{ver}^4 : \{5 \text{ random } \theta, 50 \text{ random } s, 10 \text{ random } d, \text{ and } \alpha = 0, \beta = 0\}.$$

There are totally 2500 samples of $\mu = (\mu_1, \dots, \mu_{11})$. The data of the two reduced models are listed in Table 8. In the table, $\Delta_p(\mu^{final})$ is the value of the error bound $\Delta_p(\mu)$ at the expansion point μ^{final} selected by Algorithm 2 at the final iteration step, which is the error bound for the final reduced model. The true error of the reduced model is very close to but below $\Delta(\mu^{final})$ in each case, showing that the error bound is both rigorous and sharp. The number of the iterations indicates the total iteration steps implemented in the greedy algorithm. To evaluate the transfer function over Ξ_{ver}^4 , one needs 1295 seconds if the reduced model of size 804 is used. In stead, only 29 seconds is needed, if the reduced model in the second case is used.

Table 8: Verification of the final ROMs over Ξ_{ver}^4 .

Cases	$\Delta_p(\mu^{final})$	ϵ_{max}^{ab}	iterations	ROM size	time
Case 1	7.4×10^{-8}	1.77×10^{-9}	33	804	1295s
Case 2	7.1×10^{-8}	1.4×10^{-9}	36	210	29s

The relative error bound $\Delta_p^{re}(\mu)$ behaves as well as the absolute error bound. The results in Figure 7 are computed using B_M, R_1 for each expansion point. The reduced model is of order $r = 216$, and 39 expansion points have been selected after 39 iterations. The figure again shows the robustness of the error bound.

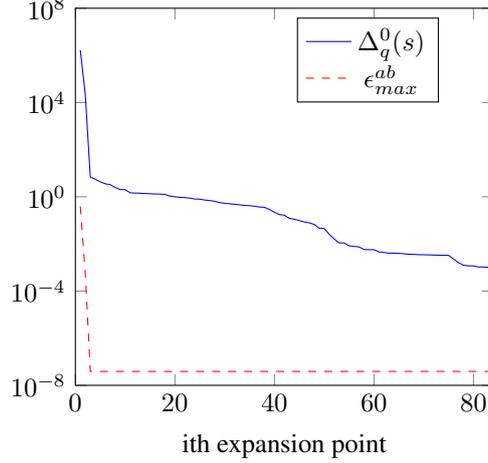


Figure 4: The true absolute error vs the error bound $\Delta_g^0(s)$ for the RLC tree model, over Ξ_{train}^3 , $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n=6134$, $r=24$.

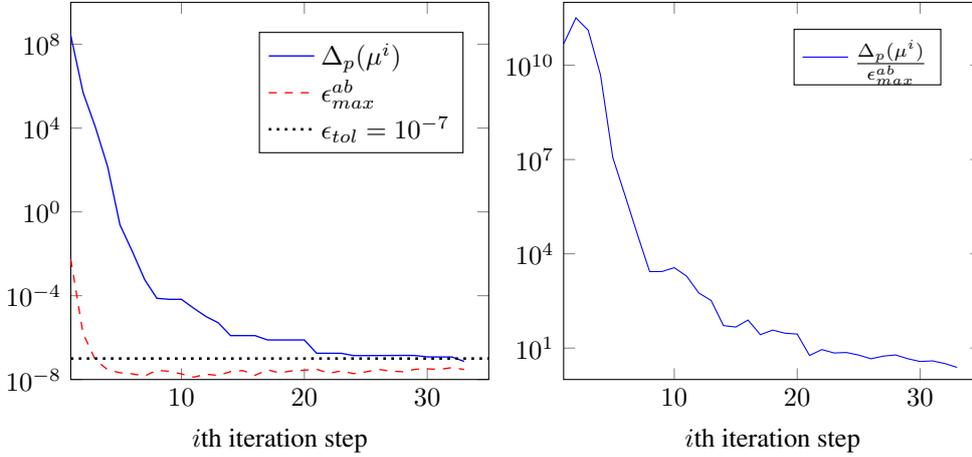


Figure 5: $V_{\mu^i} = \text{span}\{R_0, R_1, R_2\}_{\mu^i}$, $i = 1, \dots, 33$. $\epsilon_{tol} = 10^{-7}$, ROM size=804.

Behavior of $\tilde{\Delta}_p(\mu)$

In Section 7, we propose a reformulated reduced system, whose transfer function is exactly $\tilde{y}(\mu)$, so that the error bound for the transfer function of the reformulated reduced system is $\tilde{\Delta}_p(\mu)$. Figure 8 shows the decay of $\tilde{\Delta}_p(\mu)$ with the iterations in the greedy algorithm, Algorithm 2. When compared to the error bound $\Delta_p(\mu)$ for the reduced system in (10), the error bound $\tilde{\Delta}_p(\mu)$ for the reformulated reduced system is sharper. When $\tilde{\Delta}_p(\mu)$ is used in the greedy algorithm, there are 34 iterations used, instead of 36 iterations for $\Delta_p(\mu)$, shown in Figure 6. However, the difference is not that much. The big difference is the size of the reduced models. The reformulated reduced system is of size $r = 429$, while the reduced system obtained using $\Delta_p(\mu)$ is of size $r = 216$.

9.4 Sharpness of the error bounds

It can be seen that for the examples studied, the error bounds $\Delta(s)$ and $\Delta_g(s)$ are not sharp at most iterations of the greedy algorithm. One key reason might be the values $\alpha(s)$, $\gamma(s)$ on the denominator of $\Delta(s)$, and the value $\beta(s)$ on the denominator of $\Delta_g(s)$ are too small. In Subsection 9.1, $\alpha(s)$ is around $O(10^{-9})$ and $\gamma(s)$ is at $O(10^{-13})$ for the first example optical filter. For the second example peek inductor, $\gamma(s)$ is at $O(10^{-13})$, though $\alpha(s)$ is around $O(10^{-3})$.

In Subsection 9.2, $\beta(s)$ in $\Delta_g(s)$ is around 1×10^{-4} . If we use another example to check $\Delta_g(s)$, it behaves much better. The example is a model of a CDplayer, of size $n = 120$. It is observed that the values of $\beta(s)$ at all

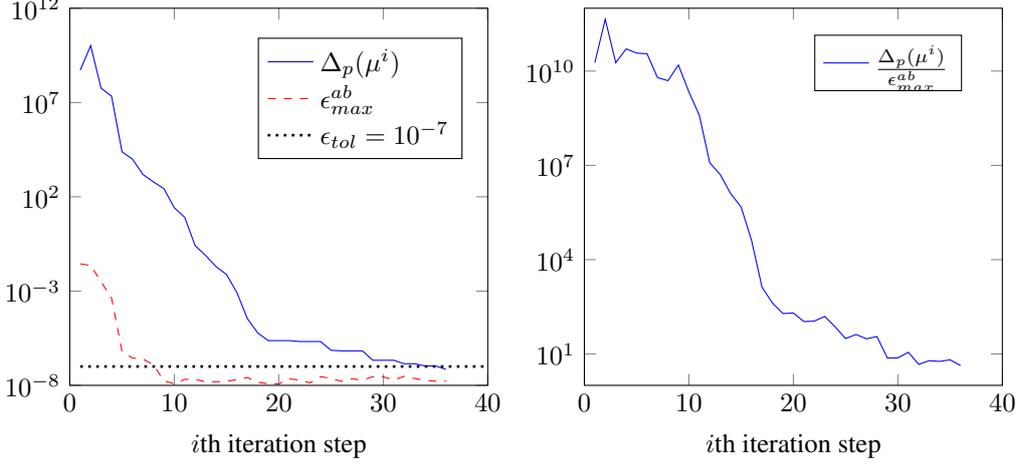


Figure 6: $V_{\mu^i} = \text{span}\{B_M, R_1\}_{\mu^i}$, $i = 1, \dots, 36$, $\epsilon_{tol} = 10^{-7}$, ROM size=210.

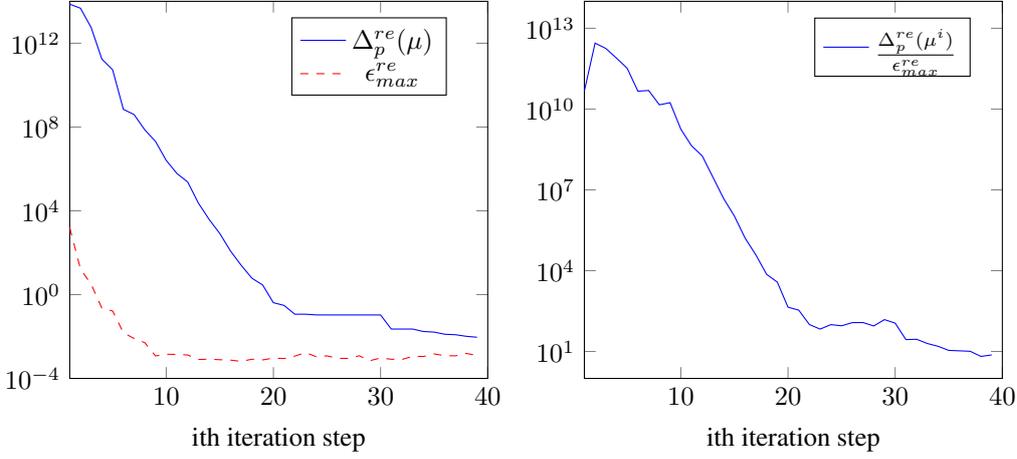


Figure 7: $V_{\mu^i} = \text{span}\{B_M, R_1\}_{\mu^i}$, $i = 1, \dots, 39$, $\epsilon_{tol} = 10^{-2}$, ROM size=216.

the samples of s are between 0.2 and 700. The decay of the error bound $\Delta_g(s)$ with the iterations in the greedy algorithm is listed in Table 9. Compared to the results of $\Delta_g(s)$ in Table 6, the error bound is much sharper.

Notice that in the beginning, either the error bound $\Delta_p(\mu)$ or $\tilde{\Delta}_p(\mu)$ is actually not sharp at all, this is also because the smallest singular value $\beta(\mu)$ of the matrix $G(\mu)$, which appears in the denominator of the error bound, is very small, around $O(10^{-8})$. With the iterations in the greedy algorithm going on, the two residuals on the numerator decrease very fast, so that the error bound quickly becomes much sharper.

It should be pointed out that the greedy algorithm is used to construct the reduced model, hence the final reduced model is only available at the last iteration step of the algorithm. The reduced models at the intermediate iterations are less important than the reduced model at the final iteration step. What does matter is that the error of the finally derived reduced model is not only below the acceptable tolerance ϵ_{tol} , but also closely estimated by the error bound. Therefore, the sharpness of the error bound at the last iteration step is of most importance.

10 Conclusions

In this work, we proposed some a posteriori error bounds for reduced order modelling of linear parametrized systems. The error bound $\Delta(s)$ for linear systems with symmetric A and symmetric positive definite E can be cheaply computed, so that the reduced model can be constructed efficiently. The computation of the error bound $\Delta_g(s)$ or $\Delta_p(\mu)$ requires solving an eigenvalue problem for each sample in the training sample space. More efficient and robust methods for computing or estimating $\Delta_g(s)$ and $\Delta_p(\mu)$ will be studied in the future. The error

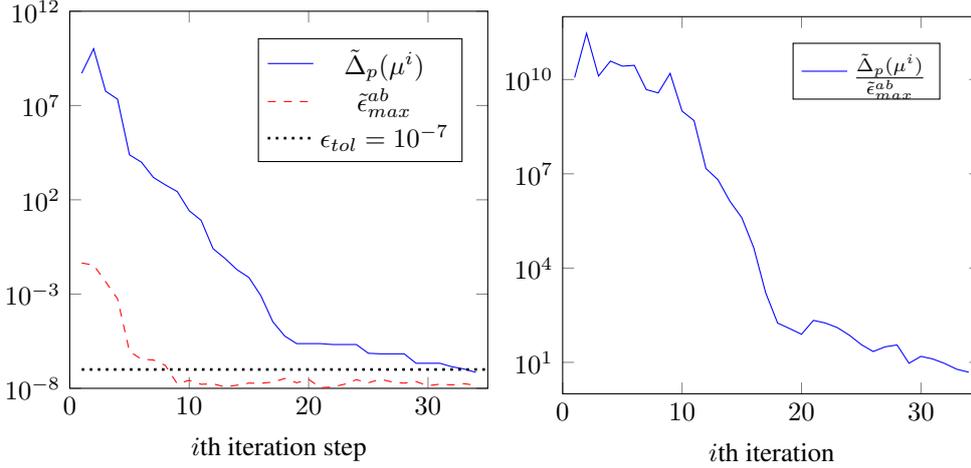


Figure 8: $V_{\mu^i} = \text{span}\{B_M, R_1\}_{\mu^i}$, $i = 1, \dots, 34$, $\epsilon_{tol} = 10^{-7}$, ROM size=429.

Table 9: RLC tree $q = 5$, $\epsilon_{tol} = 10^{-3}$, $n = 120$, $r = 60$, Ξ_{train}^3

iteration	$\hat{s}/(2\pi j)$	ϵ_{max}^{re}	$\Delta_q^{re}(\hat{s})$
1	0	61.02	8.3×10^3
2	3.61	28.43	2.10×10^3
3	48.8	8.88	3.86×10^3
4	11.8	0.74	513.7
5	94.4	0.73	199.6
6	615	0.0019	0.27
7	482	9×10^{-4}	0.02
8	1000	4.4×10^{-5}	1.27×10^{-4}

bounds are rigorous. The sharpness of the error bounds depends, nevertheless, on the properties of the system matrices E , A or $G(\mu)$. It is demonstrated that with the guidance of any of the proposed error bounds, the reduced models computed with Krylov subspace based MOR methods can be generated automatically and reliably. It is possible that the error bound $\Delta_g(s)$ or $\Delta_p(\mu)$ may realize automatic implementation of other MOR methods which are based on approximation of the transfer function.

References

- [1] R. Achar and M. S. Nakhla. Simulation of high-speed interconnects. *Proceedings of the IEEE*, 89(5):693-728, 2001.
- [2] A. C. Antoulas, P. Benner, L. Feng. Model reduction by iterative error-system approximation. *Technical report*, 2014.
- [3] D. Amsallem and C. Farhat. An online method for interpolating linear parametric reduced-order models. *SIAM J. Sci. Comput.*, 33(5):2169-2198, 2011.
- [4] Z. Bai, R.D. Slone, W.T. Smith, Q. Ye. Error bound for reduced system model by Padé approximation via the Lanczos process. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 18(2):133-141, 1999.
- [5] U. Baur, P. Benner. Model reduction for parametric systems using balanced truncation and interpolation. *at-Automatisierungstechnik* 57(8):411-420, 2009.
- [6] U. Baur, P. Benner, C. Beattie, S. Gugercin. Interpolatory projection methods for parameterized model reduction. *SIAM Journal Scientific Computing*, 33:2489-2518, 2011.

- [7] U. Baur, P. Benner, L. Feng. Model order reduction for linear and nonlinear systems: a system-theoretic perspective. *Archives of Computational Methods in Engineering*, 21: 331–358, 2014.
- [8] T. Bechtold, E. B. Rudnyi, J. G. Korvink. Error indicators for fully automatic extraction of heat-transfer macromodels for MEMS. *Journal of Micromechanics and Microengineering*, 15(3):430-440, 2005.
- [9] P. Benner. System-theoretic methods for model reduction of large-scale systems: simulation, control, and inverse Problems. In *Proceedings of MATHMOD 2009, 6th Vienna International Conference on Mathematical Modelling*, ARGESIM Report 35:126-145, 2009.
- [10] P. Benner, S. Gugercin, K. Willcox. A survey of model reduction methods for parametric systems. *SIAM Review* (to appear).
- [11] P. Benner and L. Feng. A robust algorithm for parametric model order reduction based on implicit moment matching. In *Springer MS &A series, Vol. 8: Reduced Order Methods for modeling and computational reduction* (A. Quarteroni, G. Rozza, Eds), 159-185, 2014.
- [12] M. Bollhöfer and A. Bodendiek. Adaptive-order rational Arnoldi method for Maxwell’s equations. In *Scientific Computing in Electrical Engineering (Conference abstracts)*, 77-78, 2012.
- [13] A. Bodendiek and M. Bollhöfer. Adaptive-order rational Arnoldi-type methods in computational electromagnetism. *BIT Numerical Mathematics*, 54:357380, 2014.
- [14] T. Bonin, H. Fassbender, A. Soppa and M. Zaeh. A fully adaptive rational global Arnoldi method for the model-order reduction of second-order MIMO systems with proportional damping. *Preprint, TU Braunschweig, Germany*, 2014.
- [15] S. Boyabal. Mathematical modelling and numerical simulation in materials science. *PhD thesis*, Université Paris Est, 2009.
- [16] L. Daniel, O.C. Siong, L.S. Chay, K.H. Lee, J. White. A multiparameter moment-matching model-reduction approach for generating geometrically parameterized interconnect performance models. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 22(5):678-693, 2004.
- [17] Feng, L., Koziol, D., Rudnyi, E., Korvink, J.: Model order reduction for scanning electrochemical microscope: The treatment of nonzero initial condition. In *Proceedings of Sensors* pp. 1236-1239 (2004)
- [18] L. Feng, E.B. Rudnyi, J.G. Korvink. Preserving the film coefficient as a parameter in the compact thermal model for fast electro-thermal simulation. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 24(12):1838-1847, 2005.
- [19] L. Feng and P. Benner. A robust algorithm for parametric model order reduction. In *Proceedings of Applied Mathematics and Mechanics*, 7(1):1021501-1021502, 2008.
- [20] L. Feng, J.G. Korvink, P. Benner. A fully adaptive scheme for model order reduction based on moment-matching. *Max Planck Institute preprint*, 2012. Available from: <http://www.mpi-magdeburg.mpg.de/preprints/>. Accepted by IEEE Transactions on Components, Packaging and Manufacturing Technology, 2015.
- [21] L. Feng, P. Benner, J.G. Korvink. Subspace recycling accelerates the parametric macro-modeling of MEMS. *International Journal for Numerical Methods in Engineering*, 94(1):84-110, 2013.
- [22] R.W. Freund. Model reduction methods based on Krylov subspaces. *Acta Numerica*, 12:267-319, 2003.
- [23] E. J. Grimme. Krylov projection methods for model reduction. *PhD thesis*, Univ. Illinois, Urbana-Champaign, 1997.
- [24] M. Grepl. Reduced–basis approximations and posteriori error estimation for parabolic partial differential equations. *PhD thesis*, Massachusetts Institute of Technology, 2005.
- [25] M. A. Grepl and A. T. Patera. A posteriori error bounds for reduced-basis approximations of parametrized parabolic partial differential equations. *M2AN Math. Model. Numer. Anal.*, 39:157-181, 2005.

- [26] S. Gugercin, A. C. Antoulas, C. A. Beattie. \mathcal{H}_2 model reduction for large-scale linear dynamical systems. *SIAM Journal on Matrix Analysis and Applications*, 30(2): 609-638, 2008.
- [27] U. Hetmaniuk, R. Tezaur and C. Farhat. An adaptive scheme for a class of interpolatory model reduction methods for frequency response problems. *International Journal for Numerical Methods in Engineering*, 93:11091124, 2013.
- [28] B. Haasdonk and M. Ohlberger. Efficient reduced models for parametrized dynamical systems by off-line/online decomposition. In *Proceedings of MATHMOD 2009, 6th Vienna International Conference on Mathematical Modelling*, 2009.
- [29] D.B.P. Huynh, G. Rozza, S. Sen, A.T. Patera. A successive constraint linear optimization method for lower bounds of parametric coercivity and inf-sup stability constants. *Comptes Rendus Mathematique*, 345(8):473-478, 2007.
- [30] N. Jung, A. T. Patera, B. Haasdonk, B. Lohmann. Model order reduction and error estimation with an application to the parameter-dependent eddy current equation. *Mathematical and Computer Modelling of Dynamical Systems*, 17(6):561-582, 2011.
- [31] H-J. Lee, Ch-Ch. Chu, W-Sh. Feng. An adaptive-order rational Arnoldi method for model-order reductions of linear time-invariant systems. *Linear Algebra and its Applications*, 235261, 2006.
- [32] S. Lefteriu, A. C. Antoulas, A. C. Ionita. Parametric model reduction in the Loewner framework, In *Proceedings of 18th IFAC World Congress*, 12752-12756, 2011.
- [33] B. C. Moore. Principal component analysis in linear systems: controllability, observability, and model reduction. *IEEE Transactions on Automatic Control*, AC-26:17-32, 1981.
- [34] A. T. Patera and G. Rozza. Reduced basis approximation and a posteriori error estimation for parameterized partial differential equations. *Version 1.0, Copyright MIT 2006, to appear in (tentative rubric) MIT Pappalardo Graduate Monographs in Mechanical Engineering*. Available at URL <http://augustine.mit.edu>.
- [35] H. Panzer, J. Mohring, R. Eid, and B. Lohmann. Parametric model order reduction by matrix interpolation. *at-Automatisierungstechnik*, 58(8):475484, 2010.
- [36] A. Odabasioglu, M. Celik, L. T. Pileggi. PRIMA: passive reduced-order interconnect macromodeling algorithm. *IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems*, 17(8):645-654, 1998.
- [37] D. V. Rovas. Reduced-basis output bound methods for parametrized partial differential equations. *PhD thesis*, Massachusetts Institute of Technology, 2003.
- [38] G. Rozza, D.B.P. Huynh, A.T. Patera. Reduced basis approximation and a posteriori error estimation for affinely parametrized elliptic coercive partial differential equations. *Archives of Computational Methods in Engineering*, 15:229-275, 2008.
- [39] E. B. Rudnyi and J. G. Korvink. Model order reduction for large scale engineering models developed in ANSYS. *Lecture Notes in Computer Science*, Springer-Verlag, 3732:349-356, 2006.
- [40] S. Sen. Reduced basis approximation and a *posteriori* error estimation for non-coercive elliptic problems. *PhD thesis*, Massachusetts Institute of Technology, 2007.
- [41] Thomas Wolf, Heiko Panzer, Boris Lohmann. Gramian-based error bound in model reduction by Krylov subspace methods. In *Proceedings of IFAC World Congress*, pages 3587-3591, 2011.
- [42] P. Gunupudi, R. Khazaka, M. Nakhla, T. Smy, D. Celo. Passive parameterized time-domain macromodels for high-speed transmission-line networks. *IEEE Transactions on Microwave Theory and Techniques*, 51(12):2347- 2354, 2003.
- [43] Y.T. Li, Z. Bai, Y.-F. Su, X. Zeng. Model order reduction of parameterized interconnect networks via a two-directional Arnoldi process. In *Proceedings of International Conference on Computer-Aided Design*, 868-873, 2007.

- [44] B. Salimbahrami, R. Eid, B. Lohmann. Model reduction by second order Krylov subspaces: extensions, stability and proportional damping. *In Proceedings of IEEE International Symposium on Intelligent Control*, 2997-3002, 2006.
- [45] K. URBAN AND A. T. PATERA, *An improved error bound for reduced basis approximation of linear parabolic problems*, *Math. Comp.*, 83 (2014), pp. 1599–1615.
- [46] M. YANO, *A space-time Petrov-Galerkin certified reduced basis method: Application to the Boussinesq equations*, *SIAM J. Sci. Comput.*, 36 (2014), pp. 232–266.
- [47] M. YANO, A. T. PATERA, AND K. URBAN, *A space-time hp-interpolation-based certified reduced basis method for Burgers' equation*, *Math. Models Methods Appl. Sci.*, 24 (2014), pp. 1903–1935.

