Near-optimal interpolation-based time-limited model order reduction

Kasturi Das, Srinivasan Krishnaswamy and Somanath Majhi
Department of EEE, Indian Institute of Technology Guwahati, Guwahati, India

ABSTRACT
This paper presents an interpolatory framework for model order reduction of linear time-invariant (LTI) systems over limited time intervals of the form \([0, \tau]\), with \(\tau < \infty\). We give a new proof for deriving interpolation-based first-order necessary conditions for time-limited \(H_2\) optimality. Based on these optimality conditions, we propose a time-limited rational Krylov framework for time-limited rational interpolation. The interpolatory framework is used to present an iterative algorithm that yields reduced-order models satisfying the optimality conditions approximately. The distance to optimality is quantified in terms of the interpolation errors. The errors depend primarily on the interpolating model’s poles and the time interval size. We test the proposed algorithm in three numerical examples and compare its performance with various time-limited model reduction algorithms available in the literature.

1. Introduction
Linear dynamical models are used to simulate the behaviour of physical systems. For complex systems, large models are required to capture the system dynamics to a high degree of accuracy. Hence, simulating or analysing such systems is computationally expensive. Designing controllers for such large-scale models is also a difficult task. Using model order reduction techniques, one can resolve such issues by replacing the large models with smaller ones based on various performance measures. The references Antoulas et al. (2010), Antoulas et al. (2020), Benner et al. (2021), Benner et al. (2020a) and Benner et al. (2020b) contain a comprehensive discussion on many model reduction techniques and their application areas, such as computational aerodynamics, microelectronics, electromagnetic systems and chemical processes.

In our work, we consider a continuous linear time-invariant (LTI) system \(\Sigma\), with state-space representation:

\[
\begin{align*}
\dot{x}(t) &= Ax(t) + Bu(t), \quad x(0) = 0, \\
y(t) &= Cx(t), \quad t \geq 0,
\end{align*}
\]

where \(A \in \mathbb{R}^{n \times n}\), \(B \in \mathbb{R}^{n \times m}\) and \(C \in \mathbb{R}^{p \times n}\). We consider systems where the number of inputs and outputs are much smaller than the number of states, i.e. \(m, p \ll n\). The impulse response of \(\Sigma\) is the mapping \(g : \mathbb{R}_+ \rightarrow \mathbb{R}^{p \times m}\) given by \(g(t) = Ce^AtB\), where \(\mathbb{R}_+\) denotes the set of non-negative real numbers. We assume that \(\Sigma\) is asymptotically stable. Let us consider a reduced-order system \(\hat{\Sigma}\) as follows:

\[
\begin{align*}
\dot{\hat{x}}(t) &= \hat{A}\hat{x}(t) + \hat{B}\hat{u}(t), \quad \hat{x}(0) = 0, \\
\hat{y}(t) &= \hat{C}\hat{x}(t), \quad t \geq 0,
\end{align*}
\]

where \(\hat{A} \in \mathbb{R}^{r \times r}\), \(\hat{B} \in \mathbb{R}^{r \times m}\) and \(\hat{C} \in \mathbb{R}^{p \times r}\) with \(r \ll n\). The impulse response of \(\hat{\Sigma}\) is the mapping \(\hat{g} : \mathbb{R}_+ \rightarrow \mathbb{R}^{p \times m}\) given by \(\hat{g}(t) = \hat{C}e^{\hat{A}t}\hat{B}\). Let \(G(s)\) and \(\hat{G}(s)\) be the transfer functions of \(\Sigma\) and \(\hat{\Sigma}\), respectively. They are the Laplace transforms of the impulse responses \(g(t)\) and \(\hat{g}(t)\), respectively and are given by \(G(s) = C(sI_n - A)^{-1}B\) and \(\hat{G}(s) = \hat{C}(sI_r - \hat{A})^{-1}\hat{B}\). The \(H_2\) optimal model reduction problem aims to find reduced-order models which minimise the \(H_2\) norm of the error between the full-order and the reduced-order systems (Antoulas et al., 2010; Gugercin et al., 2008).

Numerical algorithms like the Iterative Rational Krylov Algorithm (IRKA) (Gugercin et al., 2008) and the Two-Sided Iteration Algorithm (TSIA) (Xu & Zeng, 2011) obtain \(H_2\) optimal reduced models for large-scale systems. Beattie and Gugercin (2007) and Beattie and Gugercin (2009) also propose iterative numerical algorithms for \(H_2\) optimal model reduction. An iteration-free pseudo optimal rational Krylov (PORK) algorithm is proposed in Wolf (2014). These model reduction methods have been designed for an infinite-time horizon. In specific settings, one may have access to simulation data over a finite-time, or one might be interested in approximating the output trajectory of the original system over a limited time interval. This has inspired the model order reduction problem over a finite-time interval \([0, \tau]\). Accuracy outside this time interval is not essential.

Finite-time model reduction schemes include Proper Orthogonal Decomposition (POD) (Holmes et al., 2012) and time-limited balanced truncation (TL-BT) (Duff & Kürschner, 2021; Gawronski & Juang, 1990; Gugercin & Antoulas, 2003; Kürschner, 2018; Shaker & Shaker, 2013). Error bounds for TL-BT are derived in Redmann and Kürschner (2018) and Redmann (2020). An extensive comparison between TL-BT and modified versions of TL-BT can be found in Werner (2021). In Goyal and Redmann (2019), a time-limited \(H_2\) cost function – defined for a fixed time interval \([0, \tau]\) – is used as an error measure for time-limited model reduction and a model.
reduction algorithm minimising this error is proposed. Using the same cost function, Sinani and Gugercin (2019) proposes a descent-based iterative algorithm for model reduction, valid only for SISO systems. Further, Zulfíqar et al. (2020) proposes an iteration-free, time-limited pseudo-optimal rational Krylov algorithm (TL-PORK).

Interpolatory techniques for model order reduction yield high-fidelity reduced-order models and are numerically efficient. An interpolatory algorithm analogous to IRKA for time-limited $H_2$ optimal model order reduction is not available. In this paper, we propose a projection-based model reduction algorithm inspired by the interpolation-based time-limited $H_2$ optimality conditions derived in Sinani and Gugercin (2019). We also present a different proof for deriving the optimality conditions. Our proof is inspired by Theorem 6 of Breiten et al. (2015), where necessary conditions for optimality of the frequency-weighted $H_2$ model reduction problem are derived. We obtain the projection matrices of the numerical algorithm from a pair of modified rational Krylov spaces. Unlike infinite-time rational interpolation, the reduced models do not precisely satisfy the time-limited rational interpolation conditions. We classify these deviations as interpolation errors and discuss the factors affecting their magnitude. We further show that if the algorithm proposed by us and the iterative projection method for model reduction proposed in Algorithm 1 of Goyal and Redmann (2019) converge, then they converge theoretically to the same reduced model.

The paper is organised as follows. In Section 2, we discuss notations and concepts related to our work. Section 3 presents a new proof for deriving interpolation-based time-limited $H_2$ optimality conditions. We introduce a time-limited rational Krylov framework for time-limited rational interpolation, quantify the interpolation errors and propose a projection-based iterative algorithm in Section 4. Section 5 compares the proposed algorithm with another projection-based iterative algorithm existing in the literature. Three numerical examples are presented in Section 6. The paper is concluded in Section 7.

2. Preliminaries

In this section, we first introduce the notations used in the rest of the paper. Then, we briefly discuss the $H_2$ optimal model reduction problem and introduce the time-limited $H_2$ optimal model reduction problem. The latter is the main focus of our paper.

2.1 Notations

Let $\mathbb{R}$ and $\mathbb{C}$ be the fields of real and complex numbers, respectively. For a matrix $P \in \mathbb{C}^{n \times n}$, $\text{Tr}(P)$ denotes the trace, $P^*$ denotes the conjugate transpose, $\|P\|$ denotes the 2-norm, $\rho(P)$ denotes the spectral radius and $\kappa(P)$ denotes the condition number. $\|P\|_F$ denotes the Frobenius norm of $P$ and is defined as $\|P\|_F = \sqrt{\text{Tr}(P^*P)} = \sqrt{\text{Tr}(PP^*)}$. If $P$ is a diagonal matrix with $n$ diagonal elements, we represent it as $P = \text{diag}(p_1, \ldots, p_n)$. For a function $f : [0, \infty) \rightarrow \mathbb{R}^{p \times m}$, let $F(s)$ denote the Laplace transform of $f(t)$, if it exists. $F(s)$ represents the derivative of $F(s)$ with respect to $s$, i.e. $F'(s) = \frac{d}{ds}F(s)$. For a matrix $V \in \mathbb{R}^{m \times n}$, $\text{Ran}(V)$ denotes the range of the matrix. The time-limited $H_2$ norm is also referred to as the $H_2(r)$ norm. $I_n$ and $I_r$ are the $n \times n$ and $r \times r$ identity matrices, respectively. For $\sigma \in \mathbb{C}$, $\text{Re}(\sigma)$ denotes the real component of $\sigma$, and $|\sigma|$ denotes the absolute value of $\sigma$.

2.2 $H_2$ optimal model reduction over infinite-time and finite-time intervals

In order to discuss the model reduction problems in the time domain, we define the $L_2(I)$ space and the inner product and norm associated with it.

Definition 2.1: The $L_2(I)$ space is a Hilbert space consisting of square-integrable and Lebesgue measurable functions defined on an interval $I \subset \mathbb{R}$ with the inner product defined as

$$\langle f, g \rangle_{L_2(I)} = \int_I f(t)^* g(t) \, dt, \quad (3)$$

where $f, g \in L_2(I)$ are vector or matrix-valued functions. The inner product-induced norm of $f$ is defined as

$$\|f\|_{L_2(I)} = \sqrt{\langle f, f \rangle_{L_2(I)}}. \quad (4)$$

The $H_2$ norm of the system $\Sigma$, given by the state-space representation (1), is defined as $\|\Sigma\|_{H_2} = (\frac{1}{\pi} \int_{-\infty}^{\infty} \|G(i\omega)\|^2_2 d\omega)^{\frac{1}{2}}$. The $H_2$ optimal model reduction problem involves obtaining a reduced system $\hat{\Sigma}$, an $r$th order approximation of the system $\Sigma$, which minimises the $H_2$ error norm defined as $\|\Sigma - \hat{\Sigma}\|_{H_2} = (\frac{1}{\pi} \int_{-\infty}^{\infty} \|G(i\omega) - \hat{G}(i\omega)\|^2_2 d\omega)^{\frac{1}{2}}$. As $\Sigma$ is asymptotically stable, $g \in L_2([0, \infty))$. Hence, $\|g\|_{L_2([0, \infty))} = (\int_0^\infty \|g(t)\|^2_2 dt)^{\frac{1}{2}}$ is defined and by Parseval’s theorem, we can express the $H_2$ norm in the time domain as $\|\Sigma\|_{H_2} = \|g\|_{L_2([0, \infty))}$. Because of this, the $H_2$ optimal model reduction problem in the time domain can be stated as obtaining an asymptotically stable system $\hat{\Sigma}$, of order $r$, which minimises the error norm given by $\|\Sigma - \hat{\Sigma}\|_{H_2} = \|g - \hat{g}\|_{L_2([0, \infty))}$.

For obtaining $H_2$ optimal reduced models, the first-order necessary conditions for $H_2$ optimality are derived. Then, numerical algorithms are proposed, which produce reduced models satisfying the optimality conditions. An example of such an algorithm is the Iterative Rational Krylov Algorithm (IRKA). For a detailed discussion of the optimality conditions and IRKA, one can refer to Gugercin et al. (2008) and Antoulas et al. (2010). Further, it is shown in Flagg et al. (2012) that IRKA is a locally convergent fixed-point iteration to a local minimiser for the $H_2$ optimisation problem for state-space symmetric systems with distinct poles. Numerous experiments have shown that for a fixed reduced-order $r$, IRKA often converges rapidly to an accurate reduced-order approximation.

Our primary objective in this paper is to obtain reduced models which minimise a finite-time $H_2$ error norm between the full-order system $\Sigma$ and the reduced-order system $\hat{\Sigma}$. We now introduce the time-limited impulse response and time-limited transfer function of a system. These terms are used to define the time-limited $H_2$ error norm and the time-limited $H_2$ optimal model reduction problem. They are also used extensively in the later sections.
The time-limited transfer function of the same system is

\[ G(t) = \begin{cases} g(t), & t \in [0, \tau], \\ 0, & t \not\in (\tau, \infty). \end{cases} \]  

(5)

The time-limited transfer function of the same system is

\[ G_\tau(t) = \mathbb{C}(sI_n - A)^{-1}(I_n - e^{-(sI_n - \hat{A})\tau}) \hat{B}. \]

Like \( g_\tau(t) \), the time-limited impulse response matrix for the system \( \hat{\Sigma} \), obtained by restricting the impulse response matrix to the time interval \( [0, \tau] \), is given by \( \hat{g}_\tau(t) \). The corresponding time-limited transfer function matrix is \( \hat{G}_\tau(s) = \mathbb{C}(sI_n - \hat{A})^{-1}(I_n - e^{-(sI_n - \hat{A})\tau}) \hat{B} \). It can be shown that \( G_\tau(t) \) and \( \hat{G}_\tau(t) \) are the Laplace transforms of \( g_\tau(t) \) and \( \hat{g}_\tau(t) \), respectively (Sinani & Gugercin, 2019). As \( g_\tau(t) \) is non-zero over a finite-time interval, it follows that \( g_\tau \in L^2([0, \infty)) \). This brings us to the definition of the time-limited \( H_2 \) norm, which we also refer to as the \( H_2(\tau) \) norm.

**Definition 2.2:** The time-limited \( H_2 \) norm or the \( H_2(\tau) \) norm of the system \( \Sigma \), given by the state space representation (1), is defined as \( \| \Sigma \|_{H_2(\tau)} = \| g_\tau \|_{L^2([0, \infty))} \).

The time-limited \( H_2 \) optimal model reduction problem involves obtaining \( \hat{\Sigma} \), an \( r \)-th order approximation of the full-order model \( \Sigma \), which solves:

\[ \hat{\Sigma} = \text{arg min}_{\dim(\hat{\Sigma}) = r} \| \Sigma - \Sigma_r \|_{H_2(\tau)}^2, \]

(6)

where \( \| \Sigma - \Sigma_r \|_{H_2(\tau)}^2 = \| g_\tau - g_{\Sigma_r} \|_{L^2([0, \infty))}^2 \). Since \( g_\tau(\tau) \in L^2([0, \infty)) \), the reduced-order approximation \( \hat{\Sigma} \) need not be asymptotically stable even if the original system \( \Sigma \) is asymptotically stable. The above time-limited model reduction problem is introduced in Goyal and Redmann (2019). They derive Lyapunov-based first-order necessary conditions for local optimality and use the optimality conditions to develop an iterative algorithm for time-limited model reduction, which we refer to as TL-TSIA. This method is used because the optimisation problem is non-convex, and it is not easy to obtain global minimisers. In a similar way, in the next section, we derive interpolation-based optimality conditions for the time-limited \( H_2 \) optimisation problem. The optimality conditions are then used to propose a time-limited version of the IRKA method in Section 4.

### 3. Interpolation-based time-limited optimality conditions

The interpolation-based optimality conditions for time-limited \( H_2 \) optimal model reduction have already been derived in Sinani and Gugercin (2019). In Theorem 3.2, we derive these conditions using a different method. The following lemma is critical for the proof of the theorem.

**Lemma 3.1:** Consider the impulse response matrix, \( g_1(t) \in \mathbb{R}^{p \times m} \) for \( t \in [0, \infty) \), and the transfer function, \( G_1(s) \in \mathbb{C}^{p \times m} \), of an LTI system. Let \( c \in \mathbb{C}^p \), \( b \in \mathbb{C}^m \) and \( \mu \in \mathbb{C} \). We define \( g_2(t) = cb^*e^{\mu t} \) and \( g_3(t) = cb^*te^{\mu t} \) for \( t > 0 \). Similar to (5), we define \( g_{1,r}(t), g_{2,r}(t) \) and \( g_{3,r}(t) \). Let \( G_{1,r}(s) \) be the Laplace transform of \( g_{1,r}(t) \). Then,

\[ \langle g_{1,r}, g_{2,r} \rangle \mathcal{L}_2([0, \infty)) = c^*G_{1,r}(-\mu)b, \]

(7)

\[ \| g_{2,r} \|_{\mathcal{L}_2([0, \infty))} = \frac{\| b \| \| c \|}{\sqrt{2| \text{Re}(\mu) |}} \sqrt{1 - e^{2\tau \text{Re}(\mu)}}, \]

(8)

\[ \langle g_{1,r}, g_{3,r} \rangle \mathcal{L}_2([0, \infty)) = -c^*G_{1,r}(-\mu)b. \]

(9)

**Proof:** As \( g_1(t), g_2(t), g_3(t) \in \mathcal{L}_2([0, \infty)) \), we obtain their inner product using (3) as follows,

\[ \langle g_1, g_2 \rangle \mathcal{L}_2([0, \infty)) = \int_0^\infty \text{Tr}(g_1(t)b^*e^{\mu t}) \ dt = c^* \left( \int_0^\infty g_1(t)e^{\mu t} \ dt \right) b. \]

By using the definition of the Laplace transform, we obtain (7). Using (4), we can define the norm of \( g_{2,r} \) as follows,

\[ \| g_{2,r} \|_{\mathcal{L}_2([0, \infty))}^2 = \int_0^\infty \text{Tr}(cb^*e^{\mu t}b^*e^{\mu t}) \ dt = \int_0^\tau \| b \| \| c \|^2 e^{2\tau \text{Re}(\mu)} \ dt \]

\[ = \frac{\| b \|^2 \| c \|^2}{2| \text{Re}(\mu) |} |e^{2\tau \text{Re}(\mu)} - 1|. \]

Taking the positive square root of the above expression, we get (8).

It is obvious that \( g_{3,r}(t) \in \mathcal{L}_2([0, \infty)) \), and thus, its inner product with \( g_{1,r}(t) \) is given by

\[ \langle g_1, g_3 \rangle \mathcal{L}_2([0, \infty)) = \int_0^\infty \text{Tr}(g_1(t)b^*te^{\mu t}) \ dt = c^* \left( \int_0^\infty g_1(t)te^{\mu t} \ dt \right) b. \]

Due to the time-multiplication property of the Laplace transform, we obtain (9).}

**Theorem 3.2:** Let \( \hat{\Sigma} \), given by the state-space representation (2), be the best \( r \)-th order approximation of the full-order system \( \Sigma \), given by the state-space representation (1), with respect to the \( H_2(\tau) \) norm. The pole-residue representation of \( \hat{g}(t) \), the impulse response matrix of \( \hat{\Sigma} \), is given by

\[ \hat{g}(t) = \sum_{k=1}^r \hat{c}_k \hat{b}_k^* e^{\lambda_k t}, \]

(10)

with \( \lambda_k \in \mathbb{C}, \hat{c}_k \in \mathbb{C}^p \) and \( \hat{b}_k \in \mathbb{C}^m \) for \( k = 1, \ldots, r \). Let \( g_r(t) \) and \( \hat{g}_r(t) \) be the time-limited impulse response matrices of \( \Sigma \) and \( \hat{\Sigma} \), respectively. Then, for \( k = 1, 2, \ldots, r \),

\[ G_r(-\lambda_k)\hat{b}_k = \hat{c}_k, \]

(11)

\[ \hat{c}_k^* G_r(-\lambda_k) = \hat{c}_k^* \hat{G}_r(-\lambda_k), \] and

(12)

\[ \hat{c}_k^* \hat{G}_r(-\lambda_k) \hat{b}_k = \hat{c}_k^* \hat{c}_k^* \hat{G}_r(-\lambda_k) \hat{b}_k. \]

(13)

where \( G_r(s) \) and \( \hat{G}_r(s) \) are the time-limited transfer functions of \( \Sigma \) and \( \hat{\Sigma} \), respectively.
Proof: Let \( a \in \mathbb{C}^p \) be an arbitrary vector with \( \|a\|_2 = 1 \) and \( k \) be an index with \( 1 \leq k \leq r \). We assume that

\[
(g_t - \hat{g}_t, a^* e^{\lambda_k t})_{L_2([0,\infty))} = \alpha(\neq 0). \tag{14}
\]

By contradiction, we shall prove that \( \alpha = 0 \). Let \( \arg(\alpha) = \theta_0 \) and for some arbitrary \( \varepsilon > 0 \), we perturb \( \hat{g}_t(t) \) as follows,

\[
\hat{g}_t^\varepsilon(t) = (\hat{c}_k + \varepsilon e^{i\theta_0} a) \hat{b}_k^* e^{\lambda_k t} + \sum_{l \neq k} \hat{c}_l \hat{b}_l^* e^{\lambda_l t}. \tag{15}
\]

Using (8), (10), and (15), we get

\[
\| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))} = \left\| -e^{-i\theta_0} a \hat{b}_k^* e^{\lambda_k t} \right\|_{L_2([0,\tau])} = \epsilon \| \hat{b}_k^* \|_2 \sqrt{1 - e^{2\tau Re(\mu)}}.
\]

Hence, \( \| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))} = O(\varepsilon) \) as \( \varepsilon \to 0 \). Since \( \hat{g}_t(t) \) solves the \( H_2(\tau) \) optimisation problem (6), the following holds,

\[
\| g_t - \hat{g}_t \|_{L_2([0,\infty))} \leq \| g_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))} \leq \| (g_t - \hat{g}_t) + (\hat{g}_t - \hat{g}_t^\varepsilon) \|_{L_2([0,\infty))} \leq \| g_t - \hat{g}_t \|_{L_2([0,\infty))} + 2 \Re(g_t - \hat{g}_t, \hat{g}_t - \hat{g}_t^\varepsilon)_{L_2([0,\infty))} + \| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))}^2.
\]

Hence, we have

\[
0 \leq 2 \Re(g_t - \hat{g}_t, \hat{g}_t - \hat{g}_t^\varepsilon)_{L_2([0,\infty))} + \| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))}^2 \leq 2 \Re(e^{-i\theta_0} a \hat{b}_k^* e^{\lambda_k t})_{L_2([0,\infty))} + \| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))}^2.
\]

The above discussion implies that \( 0 \leq -2\varepsilon |\alpha| + O(\varepsilon^2) \). This is possible only if \( \alpha = 0 \), and using (7), we have the following result,

\[
0 = (g_t - \hat{g}_t, a^* e^{\lambda_k t})_{L_2([0,\infty))} = a^*(G_t - \hat{G}_t)(-\lambda_k) \hat{b}_k.
\]

As \( a \) was chosen arbitrarily, we have

\[
(G_t - \hat{G}_t)(-\lambda_k) \hat{b}_k = 0,
\]

which gives us (11). We prove (12) in a similar way by repeating the above analysis with \( \hat{c}_k d^* e^{\lambda_k t} \) instead of \( a^* \hat{b}_k^* e^{\lambda_k t} \) for some arbitrary \( d \in \mathbb{C}^m \) and \( \|d\|_2 = 1 \).

In order to prove (13), we assume that \( (g_t - \hat{g}_t, \hat{c}_k \hat{b}_k^* t e^{\lambda_k t})_{L_2([0,\infty))} = \beta(\neq 0) \). We shall prove that \( \beta = 0 \) by contradiction. Let \( \beta_1 = \arg(\beta) \). For sufficiently small \( \varepsilon > 0 \),

\[
\hat{g}_t^\varepsilon(t) = \hat{c}_k \hat{b}_k^* e^{(\lambda_k + \varepsilon - i\theta_0) t} + \sum_{l \neq k} \hat{c}_l \hat{b}_l^* e^{\lambda_l t}. \tag{16}
\]

For the above expression of \( \hat{g}_t^\varepsilon(t) \), as \( \varepsilon \to 0 \) we get,

\[
\| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))} = \| \hat{c}_k \hat{b}_k^* e^{(\lambda_k + \varepsilon - i\theta_0) t} + \sum_{l \neq k} \hat{c}_l \hat{b}_l^* e^{\lambda_l t} \|_{L_2([0,\tau])} = O(\varepsilon).
\]

Following the same steps as the previous discussion, we get

\[
0 \leq 2 \Re(g_t - \hat{g}_t, \hat{g}_t - \hat{g}_t^\varepsilon)_{L_2([0,\infty))} + \| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))}^2 \tag{17}
\]

The first term of the right-hand side of the above expression can be analysed as

\[
2 \Re(g_t - \hat{g}_t, \hat{c}_k \hat{b}_k^* t e^{\lambda_k t})_{L_2([0,\infty))} = 2 \Re(g_t - \hat{g}_t, \hat{c}_k \hat{b}_k^* e^{\lambda_k t}(-e^{-i\theta_0} t - \cdots))_{L_2([0,\infty))} + 2 \Re(-e^{i\theta_0} (g_t - \hat{g}_t, \hat{c}_k \hat{b}_k^* t e^{\lambda_k t})_{L_2([0,\infty))}) + 2 \Re(-e^{-i\theta_0} (g_t - \hat{g}_t, \hat{c}_k \hat{b}_k^* t e^{\lambda_k t})_{L_2([0,\infty))}) + 2 \| \hat{g}_t - \hat{g}_t^\varepsilon \|_{L_2([0,\infty))}^2
\]

The inequality above is due to the Cauchy-Schwarz inequality. The above discussion implies that (17) can be simplified to

\[
0 \leq -2\varepsilon |\beta| + O(\varepsilon^2),
\]

which is true only if \( \beta = 0 \). For \( \beta = 0 \) and using (9), we get

\[
0 = (g_t - \hat{g}_t, \hat{c}_k \hat{b}_k^* t e^{\lambda_k t})_{L_2([0,\infty))} = -\hat{c}_k^*(G_t - \hat{G}_t)(-\lambda_k) \hat{b}_k.
\]

Thus, we obtain (13).

4. Rational interpolation over a restricted time-interval

This section introduces an interpolation-based projection framework for time-limited model reduction, where we compute a pair of modified rational Krylov subspaces and obtain a reduced model by projecting the full-order model onto them. The subspaces are calculated such that the time-limited transfer function of the reduced system approximates the time-limited transfer function of the full-order system at a given set of interpolation points along given tangential directions. For a set of \( r \) distinct interpolation points \( \{a_1, \ldots, a_r\} \) and corresponding tangential directions \( \{b_1, \ldots, b_r\} \) and \( \{c_1, \ldots, c_r\} \), the modified Krylov subspaces are obtained as follows,

\[
\mathcal{V}_r = \bigcap_{i=1,\ldots,r} \left( (I_n - e^{-a_i^* t} A^T) B_{b_i} \right) \quad \text{and} \quad \mathcal{W}_r = \bigcap_{i=1,\ldots,r} \left( (I_n - e^{-a_i^* t} A^T) C_{c_i} \right). \tag{18}
\]

The complex-valued interpolation points and tangential directions are closed under conjugation. For the subspaces \( \mathcal{V}_r \) and
There exists a neighbourhood of tangent directions, \( b \) and \( c \), respectively. We assume that \( \text{diag} \) functions. The following theorem quantifies the interpolation which are also real. The reduced model obtained above does not result in the exact interpolation of the time-limited transfer functions. The following theorem quantifies the interpolation errors.

**Theorem 4.1:** Let us denote \( \Pi = V_r Z_r^* \). We choose an interpolation point, \( \sigma \in \{\sigma_1, \ldots, \sigma_t\} \) and corresponding right and left tangent directions, \( b \) and \( c \), respectively. We assume that \( \sigma \) is not an eigenvalue of \( A \) or \( \hat{A} \). Let \( A \) and \( \hat{A} \) be diagonalisable matrices with eigenvalue decompositions, \( \hat{A} = Q_A \Lambda Q_A^{-1} \) and \( A = Q_A \Gamma Q_A^{-1} \), respectively, where \( \Lambda = \text{diag}(\lambda_1, \ldots, \lambda_n) \) and \( \Gamma = \text{diag}(\gamma_1, \ldots, \gamma_n) \). Then, we have

\[
G_r(\sigma)b - \hat{G}_r(\sigma)b = e^{-\sigma \tau} CV_r(\sigma I_r - \hat{A})^{-1} \times Z_r^* (e^{A \Pi \tau} - e^{A \tau}) B_b, \\
c^* G_r(\sigma) - c^* \hat{G}_r(\sigma) = e^{-\sigma \tau} c^* C (e^{A \Pi \tau} - e^{A \tau}) \times V_r(\sigma I_r - \hat{A})^{-1} Z_r^* B, \\
c^* G_r'(\sigma)b - c^* \hat{G}_r'(\sigma)b = R_{P_1}(\sigma) + R_{P_2}(\sigma), \quad \text{and} \\
c^* G_r'(\sigma)b - c^* \hat{G}_r'(\sigma)b = R_{Q_1}(\sigma) + R_{Q_2}(\sigma),
\]

where \( R_{P_1}(\sigma), R_{P_2}(\sigma), R_{Q_1}(\sigma) \) and \( R_{Q_2}(\sigma) \) are given by

\[
R_{P_1}(\sigma) = -e^{-\sigma \tau} c^* CV_r(\sigma I_r - \hat{A})^{-2} ((s I_r - \hat{A}) \tau + I_r) \times Z_r^* (e^{A \Pi \tau} - e^{A \tau}) B_b, \\
R_{P_2}(\sigma) = e^{-\sigma \tau} c^* C (s I_r - A) (s I_r - A)^{-2} \times \left( (s I_r - A) \tau + I_r \right) e^{A \tau} B_b, \\
R_{Q_1}(\sigma) = -e^{-\sigma \tau} c^* C (e^{A \Pi \tau} - e^{A \tau}) V_r(\sigma I_r - \hat{A})^{-2} \times (I_r + \tau (s I_r - A) \hat{A}) Z_r^* B, \quad \text{and} \\
R_{Q_2}(\sigma) = e^{-\sigma \tau} c^* C \left( (s I_r - A) \tau + I_r \right) \times (s I_r - A)^{-2} (I_r - Q(\sigma)) B_b.
\]

Here, \( P(s) \) and \( Q(s) \) are matrix-valued functions defined as

\[
P(s) = V_r(s I_r - \hat{A})^{-1} Z_r^* (s I_r - A), \quad \text{and} \\
Q(s) = (s I_r - A) V_r(s I_r - \hat{A})^{-1} Z_r^*.
\]

These functions are projectors for any \( s \in \mathbb{C} \) that is not an eigenvalue of \( A \) or \( \hat{A} \).

**Proof:** As we assume that \( \sigma \) is not an eigenvalue of \( A \) and \( \hat{A} \), there exists a neighbourhood of \( s = \sigma \) where the projectors \( P(s) \) and \( Q(s) \) are analytic. As \( P(s) \) is a projector, \( P^2(s) = P(s) \), and we have

\[
V_r = \text{Ran} P(s) = \text{Ker} (I_r - P(s)).
\]

Similarly, as the matrix \( Q(s) \) is a projector, \( Q^2(s) = Q(s) \), and thus we have

\[
V_r = \text{Ker} Q(s) = \text{Ran} (I_r - Q(s)).
\]

The error in the right tangential interpolation condition is given by

\[
G_r(s)b - \hat{G}_r(s)b = C(s I_r - A)^{-1} (I_r - e^{-s \tau} e^{A \tau}) B_b - \hat{C}(s I_r - \hat{A})^{-1} (I_r - e^{-s \tau} e^{A \tau}) B_b.
\]

Substituting \( \hat{A} \), \( \hat{B} \) and \( \hat{C} \) from (20) in the above expression and using the identity \( e^{A \tau} B = Z_r^* (e^{A \Pi \tau} - e^{A \tau}) B_b \), we get

\[
G_r(s)b - \hat{G}_r(s)b = C(I_r - P(s))(s I_r - A)^{-1} (I_r - e^{-s \tau} e^{A \tau}) B_b + e^{-s \tau} C V_r(s I_r - \hat{A}) Z_r^* (e^{A \Pi \tau} - e^{A \tau}) B_b.
\]

By definition, the vector \( (s I_r - A)^{-1} (I_r - e^{-s \tau} e^{A \tau}) B_b \in V_r \) and, due to (31), also lies in \( (I_r - P(\sigma)) \)’s nullspace. Hence, evaluating the above expression at \( s = \sigma \), we get (21).

Similar to the right tangential interpolation error, for the left tangential interpolation error, using the identity \( \hat{C} e^{A \tau} = C e^{A \Pi \tau} V_r \) and substituting \( \hat{A} \), \( \hat{B} \) and \( \hat{C} \) from (20), we have

\[
c^* G_r(s) - c^* \hat{G}_r(s) = (s I_r - A^*)^{-1} (I_r - e^{-s \tau} e^{A \tau}) C^* c^* \times (I_r - Q(\sigma) B_b + e^{-s \tau} c^* C (e^{A \Pi \tau} - e^{A \tau}) \times V_r(s I_r - \hat{A}) Z_r^* B_b. \]

As \( (s I_r - A^*)^{-1} (I_r - e^{-s \tau} e^{A \tau}) C^* c^* \in V_r \) and (32), evaluating the above expression at \( s = \sigma \), we obtain (22). Left multiplying the left-hand side of (33) by \( c^* \), denoting \( (I_r - e^{-s \tau} e^{A \tau}) B_b \) by \( E(s) \) and differentiating the expression obtained in a neighbourhood of \( s = \sigma \) where \( P(s) \) is analytic, we get

\[
c^* G_r'(s)b - c^* \hat{G}_r'(s)b = c^* C \left( -d \frac{d}{ds} P(s) \right) (s I_r - A)^{-1} E(s) B_b + c^* C (I_r - P(s)) \frac{d}{ds} ((s I_r - A)^{-1} E(s) B_b) + \frac{d}{ds} e^{-s \tau} c^* C V_r(s I_r - \hat{A}) Z_r^* (e^{A \Pi \tau} B - e^{A \tau} B \times (s I_r - A)^{-2} (s I_r - A)^{-1} Z_r^* B_b.
\]

\[
= c^* C (V_r(s I_r - \hat{A})^{-2} Z_r^* \times (s I_r - A) - V_r(s I_r - \hat{A})^{-1} Z_r^* (s I_r - A)^{-1} E(s) B_b + e^{-s \tau} c^* C (I_r - P(s)) \times (s I_r - A)^{-2} (s I_r - A)^{-1} E(s) B_b + e^{-s \tau} c^* C (s I_r - A)^{-1} E(s) B_b\]
\[
+ e^{-s \tau} c^* C V_r(s I_r - \hat{A})^{-2} (s I_r - A) \tau + I_r) Z_r^* \times (e^{A \Pi \tau} - e^{A \tau}) B_b.
\]

(35)
Substituting $s = \sigma$ in the last term above, we note that due to $(\sigma I_n - A)^{-1}E(s)Bb \in V_r$, only the last two terms remain non-zero, and the rest of the terms vanish. We thus obtain (23).

Finally, right multiplying the left-hand side of (34) by $\hat{e}e$ and differentiating the expression in a neighbourhood of $s = \sigma$ and making similar substitutions as above, we get

$$c^*G_r(s)b - c^*\hat{G}_r(s)b = c^*C(sI_n - A)^{-1}((sI_n - A)V_r(sI_n - \hat{A})^{-2}Z_r^*$$
$$- V_r(sI_n - \hat{A})^{-1}Z_r^*Bb$$
$$+ e^{-\tau s}c^*Ce^{\tau r}((\tau (sI_n - A) + I_n) - e^{r(sI_n - A)})$$
$$\times (sI_n - A)^{-2}(I_n - Q(s))Bb$$
$$+ e^{-\tau s}c^*C(e^{\tau r} - e^{\tau rt})V_r(sI_n - \hat{A})^{-2}$$
$$\times (I_n + \tau (sI_n - \hat{A}))Z_r^*Bb. \quad (36)$$

For $s = \sigma$ in the last term above, the vector $(\sigma^* I_n - A^*)^{-1} E(s) C \hat{c}$ lies in $V_r$. Hence, only the last two terms remain while the remaining terms vanish, and we obtain (24). ■

The projection-based interpolation framework introduced at the beginning of this section works well when the interpolation errors are small. We shall now analyze the upper bounds of these errors. From (21), we have

$$\|G_r(\sigma)b - \hat{G}_r(\sigma)b\| \leq |e^{-\sigma r}||\|CV_r\|((\sigma I_n - \hat{A})^{-1}$$
$$\times Z_r^*(e^{\tau r} - e^{\tau rt})Bb\|,$$  

(37)

where the first term of the upper bound equals $|e^{-(\text{Re}(\sigma))r}|$.

For the third term, we have, $(|\|\sigma I_n - \hat{A})^{-1}\| \leq \kappa(Q_2)/\rho(\sigma I_n - \hat{A})$. By applying the projection theorem (Theorem 2 in Section 3.3 of Luenberger (1997)), it can be proven that the vector $(I_n - Z_r Z_r^*) (e^{\tau r} - e^{\tau rt})Bb$ minimises $\|Z_r^*(e^{\tau r} - e^{\tau rt})Bb\|^2$. Thus, the fourth term is, by definition, the distance of $(e^{\tau r} - e^{\tau rt})Bb$ from the kernel of $Z_r^*$.

Similarly, from (22), we have

$$\|e^*G_r(\sigma) - e^*\hat{G}_r(\sigma)\| \leq |e^{-\sigma r}||c^*C(e^{\tau r} - e^{\tau rt})V_r$$
$$\times ((\sigma I_n - \hat{A})^{-1}||Z_r^*B\|, \quad (38)$$

The second term of the upper bound in (38) is the distance of $(e^{\tau r} - e^{\tau rt})C \hat{c}$ from the kernel of $V_r^*$.

Finally, from (23), we have

$$\|e^*G_r(\sigma)b - e^*\hat{G}_r(\sigma)b\|$$

$$\leq \|R_{P_1}(\sigma)\| + \|R_{P_2}(\sigma)\|$$

$$\text{where}$$

$$\|R_{P_1}(\sigma)\|$$

$$\leq |e^{-\sigma r}||\|CV_r\|((\sigma I_n - \hat{A})^{-2}$$

(39)

The third term of the bound in (39) is bounded above by $\kappa(Q_2)/\rho((\sigma I_n - \hat{A})^2)$. Similarly, the third term of the bound in (40) is bounded above by $\kappa(Q_2)/\rho((\sigma I_n - \hat{A})^2)$. If $\rho((\sigma I_n - \hat{A})^2) \ll 1$, $\kappa(Q_2)\rho(e^{\tau r}) \ll 1$, it will contribute significantly to the upper bounds of (37), (38) and (39). If $\rho((\sigma I_n - \hat{A})^2) \ll 1$, it will have a notable contribution to the error bound of the bitangential interpolation error. If the matrix $A$ has unstable eigenvalues ($\gamma_i \gg 0$) such that $\rho(e^{\tau r}) \gg 1$, it will contribute significantly to the upper bound of (40). Also, for sufficiently small values of $\tau$, the error bounds will be negligible.

**Time-Limited Iterative Rational Krylov Algorithm (TL-IRKA).**

The difficulty in constructing a reduced-order model satisfying the time-limited $H_2$ optimality conditions is that one does not know how to choose such interpolation data a priori. Therefore, we use the time-limited rational interpolation framework discussed above to propose an iterative correction method given by Algorithm 1.

The algorithm is stopped when the change in eigenvalues of the reduced state matrix $\hat{A}$ for two consecutive iterations becomes less than a preset tolerance. Though we have not provided a formal proof of the convergence of TL-IRKA, we have tested the algorithm on various examples. We have observed that the algorithm converges after a finite number of iterations for a proper set of initial interpolation points and tangential directions. The reduced model obtained using TL-IRKA will not satisfy the first-order necessary conditions for time-limited $H_2$ optimality exactly. The error expressions derived in Theorem 4.1 can be used to estimate the closeness to optimality of the reduced model.

**Remark 4.1:** Various techniques exist for initialising the interpolation points and tangential directions for TL-IRKA. We can randomly select the initial interpolation points and tangential directions. A second possible method is initialising conventional IRKA randomly and using the reduced-order system to obtain the initial interpolation points and tangential directions. Another possible way of initialising TL-IRKA is effectively computing eigenvalues and corresponding left and right eigenvectors corresponding to the dominant residues using the dominant pole algorithm (Rommes & Martins, 2006).

**Remark 4.2:** TL-IRKA involves using matrix-vector multiplications and linear solvers similar to other Krylov-based model reduction strategies. The main difficulty in implementing the
Algorithm 1: Time-Limited Iterative Rational Krylov Algorithm (TL-IRKA)

Input: The system matrices: $A, B, C$;
- Initial interpolation points: $\{\sigma_1, \ldots, \sigma_r\}$;
- Initial tangential directions: $\tilde{B} = [b_1, \ldots, b_r]$ and $\tilde{C} = [c_1, \ldots, c_r]$;
- A finite-time interval: $[0, \tau]$;

Output: The reduced matrices: $\hat{A}, \hat{B}, \hat{C}$;
1. Compute $V_r$ and $W_r$ using (18) and (19) respectively and let $Z_r = (W_r^*V_r)^{-1}W_r^r$;
2. while (not converged) do
   a. $\hat{A} = Z_r^*AV_r$, $\hat{B} = Z_r^kB$, $\hat{C} = CV_r$;
   b. Compute $\hat{A} = RAR^{-1}$ where $R^{-1}$ and $R$ are the left and right eigenvectors of $\hat{A}$;
   c. Update interpolation points and tangential directions as follows:
      i. $\sigma_i \leftarrow -\lambda_i(A)$ for $i = 1, 2, \ldots, r$;
      ii. $\hat{B} = B^eR^{-e}$ and $\hat{C} = CR$;
   d. Update $V_r$ and $W_r$ using (18) and (19), respectively, and let $Z_r = (W_r^*V_r)^{-1}W_r^r$;
end
3. $\hat{A} = Z_r^*AV_r$, $\hat{B} = Z_r^kB$, $\hat{C} = CV_r$;

Theorem 5.1: We assume that the state matrices of the system $\Sigma$, given by (1), and $\Sigma$, the reduced-order approximation of order $r$ given by (2), are diagonalisable. Reduced models of order $r$ are obtained using TL-IRKA and TL-TSIA. For TL-IRKA, the initial conditions are obtained from the poles and residues of the reduced model $\hat{\Sigma}$ whereas TL-TSIA is initialised with the statespace matrices of $\Sigma$. If both algorithms converge, they should converge to equivalent reduced-order models.

Proof: For the first iteration of TL-IRKA, the projection subspaces $V_r$ and $W_r$ are computed using (18) and (19), respectively. Similarly, for the first iteration of TL-TSIA, the projection matrices $X$ and $Y$ are obtained using (41) and (1), respectively. We prove that the right projection subspace $V_r$, and the right projection subspace defined by the columns of the matrix $X$ are equivalent, i.e., $V_r = \text{Ran}(X)$. Let $(\hat{A}, \hat{B}, \hat{C})$ be the state-space matrices, and $(b_i, c_i, \sigma_i), i = 1, \ldots, r$ be the poles and residues of the reduced-order model $\hat{\Sigma}$. Let $\hat{A} = SDS^{-1}$ and $e^{\hat{A} \tau} = Se^{D \tau}S^{-1}$ be the eigendecompositions for $\hat{A}$ and $e^{A \tau}$, respectively. Substituting these decompositions for $\hat{A}$ and $e^{A \tau}$ in (41), we have

\[
AXS^{-e} + XS^{-e}D^e + BB^eS^{-e} = e^{A \tau}BB^eS^{-e}e^{D^e \tau} = 0
\]

where $\hat{X} = XS^{-e}$ and $B^eS^{-e} = B^e$. Let $b_i \sigma_i$ be the columns of the matrix $\hat{X}$. Since $S$ is a nonsingular matrix, the columns of $X$ and $\hat{X}$ span the same subspace. Let $b_i, c_i$ be the columns of $B$ and let $D = \text{diag}(\lambda_1, \ldots, \lambda_r)$, where $\lambda_i$'s are the eigenvalues of $A$. Using Equation (42), we obtain

\[\hat{X}_i = (-\lambda_i I_n - A)^{-1}(I_n - e^{\sigma_i(\tau)}e^{A \tau})Bb_i, \text{for } i = 1, \ldots, r.\]

This subspace is also the right projection subspace $V_r$ of TL-IRKA, with $b_i$'s as the right tangential directions. Thus, we have shown that $V_r = \text{Ran}(X)$. Similarly, we can show that $W_r = \text{Ran}(Y)$. As the projections obtained in the first iteration of TL-IRKA and TL-TSIA are equivalent, the corresponding reduced models obtained (20) are also equivalent. This will also hold for the subsequent iterations. Hence, if the algorithms converge, they converge to equivalent reduced-order models.

6. Numerical examples

In this section, we investigate the performance of TL-IRKA using three numerical examples. The first two examples are...
SISO models where TL-IRKA is compared to TL-TSIA, TL-BT, IRKA and TL-PORK. The third example is a MIMO model where TL-IRKA is compared to TL-TSIA, TL-BT and IRKA. TL-PORK-1 and TL-PORK-2 are initialised with reduced models obtained from TL-IRKA and IRKA, respectively. The three models are taken from Chahlaoui and Van Dooren (2005). The simulations are done in MATLAB version 8.3.0.532(R2014a) on an Intel(R) Core(TM) i5-6500 CPU running at 3.20GHz with 16 GB RAM and Windows 10, Version 20H2 Operating System. The MATLAB command "impulse" is used to simulate the impulse responses for all three examples. We define the error norm as $\text{AbsErr}(t) = \|\text{vec}((g(t) - \hat{g}(t)))\|_2$ and the relative $H_2(\tau)$ error as $\text{Rel}H_2(\tau) = \frac{\|g(\tau)\|_2 - \|\hat{g}(\tau)\|_2}{\|g(\tau)\|_2}$ for comparing the performance of the different model reduction algorithms, where $g(\tau)$ and $\hat{g}(\tau)$ are the impulse responses of the full-order and the reduced-order system, respectively. In addition, we define the following errors to measure the distance to optimality of the reduced-order models obtained using TL-IRKA and IRKA. For $i = 1, \ldots, r$, we define

1. Right tangential interpolation errors, $\text{RTerr}(i) = \|G_i(\sigma_i) b_i - \hat{G}_i(\sigma_i) b_i\|$ and relative right tangential interpolation error, $\|\text{RTerr}\|_{\text{rel}} = \sum_{i=1}^{r} \text{RTerr}(i)$.
2. Left tangential interpolation errors, $\text{LTerr}(i) = \|c_i^* G_i(\sigma_i) - c_i^* \hat{G}_i(\sigma_i)\|$ and relative left tangential interpolation error, $\|\text{LTerr}\|_{\text{rel}} = \sum_{i=1}^{r} \text{LTerr}(i)$.
3. Bitangential interpolation errors, $\text{derr}(i) = \|c_i^* G_i(\sigma_i) b_i - c_i^* \hat{G}_i(\sigma_i) b_i\|$ and relative bitangential interpolation error, $\|\text{derr}\|_{\text{rel}} = \sum_{i=1}^{r} \text{derr}(i)$.

Here, $G_i(s)$ and $\hat{G}_i(s)$ are the time-limited transfer functions of the full-order and the reduced-order system, respectively. The $\sigma_i$’s are the interpolation points, $b_i$’s are the right tangential directions and $c_i$’s are the left tangential directions. For SISO systems, the tangential directions are scalars and the left and right tangential errors are equal.

**Example 1**

The first example is a beam model of order 348. We obtain reduced models of order $r = 12$ for the time intervals $[0, 0.1]$ s and $[0, 2]$ s. The iterative algorithms are run until the change in the eigenvalues of the reduced state matrix, $\tilde{A}$, becomes less than $10^{-3}$. TL-IRKA and TL-TSIA are initialised randomly and are found to converge for the two time intervals. The number of iterations required for convergence depends on the initialisation. Figure 1(a) compares the impulse responses of the original and the reduced models and the errors $\text{AbsErr}(t)$ for the time interval $[0, 0.1]$ s. Figure 1(b) compares the impulse responses and errors for the time interval $[0, 2]$ s. For both time intervals, the relative $H_2(\tau)$ errors for the different algorithms are compared in Table 1. For the smaller time interval, we observe that the relative $H_2(\tau)$ errors of the reduced systems obtained from TL-IRKA and TL-TSIA are several orders of magnitude less than that of the reduced models due to the other algorithms. TL-IRKA and TL-PORK-1 have smaller $H_2(\tau)$ errors than the other algorithms for $\tau = 2$ s, as evident from Table 1. Also, for the same time interval, we observe that TL-IRKA and TL-TSIA converge to different local optima. Theorem 5.1 shows that if both TL-IRKA and TL-TSIA converge, then theoretically, they should converge to the same local minimum. However, due to finite precision arithmetics, both algorithms converge to different local minima, leading to different $H_2(\tau)$ errors.

The distance to optimality of the reduced-order system obtained by TL-IRKA can be determined from the interpolation errors, a few of which are listed in Table 2. The index $i$ refers to the $i$th interpolation point. We observe that errors corresponding to interpolation points ($-\lambda_i$’s) such that $\text{Re}(\lambda_i) \tau \ll -1$ are negligible.

Finally, Table 3 shows that TL-IRKA performs better than IRKA in satisfying the $H_2(\tau)$ optimality conditions. The relative interpolation errors are negligible for the shorter time interval ($\tau = 0.1$ s). In comparison, the errors are considerably greater for the second time interval ($\tau = 2$ s). The data in Table 2 and Table 3 support the theoretical results from Theorem 4.1.
Table 1. Relative $H_2(\tau)$ Errors for Beam example.

| Algorithm     | TL-BT | TL-IRKA | IRKA | TL-TSIA | TL-PORK-1 | TL-PORK-2 |
|---------------|-------|---------|------|---------|-----------|-----------|
| $\|\text{Err}\|_{H_2}(\tau)$ for $\tau = 0.1$ s | $6.79 \times 10^{-8}$ | $6.55 \times 10^{-11}$ | $0.0580$ | $6.85 \times 10^{-11}$ | $9.25 \times 10^{-9}$ | $0.0123$ |
| $\|\text{Err}\|_{H_2}(\tau)$ for $\tau = 2$ s | $0.0262$ | $0.0115$ | $0.0476$ | $0.0243$ | $0.0114$ | $0.0375$ |

Table 2. Interpolation errors for various time intervals for Beam Example.

| Final-Time | i | $\Re(\lambda_i)$ | $\rho(\lambda_i + \tilde{A})$ | $\|\text{RTErr}\|_{rel}$ | $\|\text{LTErr}\|_{rel}$ | $\|\text{dErr}\|_{rel}$ |
|------------|---|-------------------|------------------------|-----------------|-----------------|-----------------|
| $\tau = 0.1$ s | 1 | $-52.7$ | $538.9$ | $3.8 \times 10^{-17}$ | $3.8 \times 10^{-17}$ | $8.1 \times 10^{-20}$ |
| | 6 | $-7.3$ | $459.1$ | $1.8 \times 10^{-15}$ | $1.8 \times 10^{-15}$ | $6.3 \times 10^{-15}$ |
| | 12 | $1.2$ | $538.9$ | $1.2 \times 10^{-11}$ | $1.2 \times 10^{-11}$ | $1.39 \times 10^{-12}$ |
| $\tau = 2$ s | 1 | $-91.2$ | $74.2$ | $3.5 \times 10^{-15}$ | $3.5 \times 10^{-15}$ | $0.0026$ |
| | 7 | $-1.2$ | $54.3$ | $0.0077$ | $0.0077$ | $0.0022$ |
| | 12 | $-1.9$ | $52.1$ | $0.0026$ | $0.0026$ | $0.0066$ |

Table 3. Relative error in the optimality conditions for Beam example.

| Final-Time | Algorithm | $\|\text{RTErr}\|_{rel}$ | $\|\text{LTErr}\|_{rel}$ | $\|\text{dErr}\|_{rel}$ |
|------------|-----------|-----------------|-----------------|-----------------|
| $\tau = 0.1$ s | TL-IRKA | $4.48 \times 10^{-12}$ | $4.48 \times 10^{-12}$ | $1.32 \times 10^{-11}$ |
| | IRKA | $0.0045$ | $0.0045$ | $5.6221$ |
| $\tau = 2$ s | TL-IRKA | $0.0028$ | $0.0028$ | $0.0187$ |
| | IRKA | $0.0804$ | $0.0804$ | $8.3027$ |

Example 2

For the second case, we have an artificial system of order 1006 called FOM. We obtain reduced models of order $r = 20$ for the time intervals $[0, 0.2]$ s and $[0, 2]$ s. When the change in the eigenvalues of $\tilde{A}$ becomes less than $10^{-5}$, the iterative algorithms are stopped. We randomly initialise the interpolation points for TL-IRKA and TL-TSIA. The algorithms converge for both time intervals. Interpolation errors corresponding to three interpolation points are listed in Table 5 to analyze the distance to optimality of the reduced-order model obtained by TL-IRKA. We see that errors corresponding to interpolation points with $\Re(\lambda_i) \tau \ll -1$ are negligible. As $\Re(\lambda_i) \tau$ increases, the errors also increase. Table 6 compares the relative interpolation errors of TL-IRKA and IRKA for both time intervals. The relative interpolation errors due to TL-IRKA for $[0, 0.2]$ s are smaller than the errors for $[0, 2]$ s. Further, we observe that for reduced models with $r = 20$, TL-IRKA approximates the $H_2(\tau)$ optimality conditions more accurately than IRKA for both the time intervals considered.

Example 3

The final example discussed is a MIMO model of the International Space Station (ISS) with three inputs and three outputs. Reduced models of order $r = 12$ are obtained for three time intervals $[0, 0.01]$ s, $[0, 0.1]$ s and $[0, 1]$ s. The iterative algorithms are stopped when the change in the eigenvalues of $\tilde{A}$ becomes less than $10^{-8}$. The initial interpolation points and tangential directions for TL-IRKA and state matrices for TL-TSIA are randomly chosen, and the algorithms converge for all three time intervals. We compare the error norm, i.e.$\text{AbsErr}(t)$, for the different algorithms in Figure 3.

To compare the performance of the different algorithms, Table 7 lists their relative $H_2(\tau)$ errors for the three time intervals considered. For the smallest time interval, i.e.$[0, 0.01]$ s, TL-IRKA and TL-TSIA perform better than TL-BT and IRKA. For the other time intervals, TL-IRKA yields results comparable with TL-BT and TL-TSIA and performs better than IRKA.
all the time intervals, this term does not contribute much to \( \tau = \text{Final-Time} \).

Table 3. ISS example.

Table 5. Interpolation errors for various time intervals for FOM example.

Table 6. Relative error in the optimality conditions for FOM example.

Table 7. Relative \( H_2(\tau) \) Errors in ISS example.

Table 8. Interpolation errors for various time intervals for ISS example.

The errors corresponding to two interpolation points for each time interval are listed in Table 8. As \(|\text{Re}(\lambda_i)| < 3\) for all the time intervals, this term does not contribute much to the interpolation errors. Instead, the errors depend on the final time \( \tau \). Errors corresponding to \( \tau = 0.01 \) s are negligible and increase by several orders of magnitude for \( \tau = 0.1 \) s and \( \tau = 1 \) s. The relative interpolation errors for TL-IRKA and IRKA are compared in Table 9. For every time interval, TL-IRKA performs better than IRKA. However, for the smallest time interval, the error due to TL-IRKA is negligible compared to IRKA.

It is evident from Tables 1, 4 and 7 that the relative error of TL-IRKA for the shortest time interval is less than that of TL-BT. However, the relative errors of TL-IRKA and TL-BT are comparable for the other time intervals. Further, in all three examples, TL-IRKA and TL-TSIA have similar relative error magnitudes for the various time intervals considered.

7. Conclusion and future work

We have presented an alternative way of deriving the interpolation-based \( H_2(\tau) \) optimality conditions. We have then proposed an algorithm called TL-IRKA, which obtains reduced models satisfying the optimality conditions approximately. The errors in the optimality conditions have also been quantified. We have further compared TL-IRKA to another algorithm.
called TL-TSIA. Finally, through various numerical simulations, we have demonstrated that TL-IRKA performs efficiently.

As mentioned in Remark 4.2, it is possible to improve the computational cost of the proposed model reduction method, TL-IRKA using specific techniques available in the literature. Also, currently, the model reduction method does not preserve stability. This is an important aspect that needs to be studied. Further, the reduced models produced by the method do not exactly satisfy the interpolation-based time-limited $H_2$ optimality conditions. Thus, algorithms which produce reduced models satisfying the interpolation-based optimality conditions exactly, especially for MIMO systems, need to be studied. These are potential areas for future work.

Acknowledgments
The authors would like to thank the anonymous reviewers for their detailed and constructive comments, which have contributed significantly to improving the quality of the paper.

Disclosure statement
No potential conflict of interest was reported by the author(s).

References
Antoulas, A. C., Beattie, C. A., & Gugercin, S. (2010). Interpolatory model reduction of large-scale dynamical systems. In J. Mohammadpour & K. Grigoriadis (Eds.), Efficient Modelling and Control of Large-Scale Systems (pp. 3–58). Springer. https://doi.org/10.1007/978-1-4419-5755-3_1

Antoulas, A. C., Beattie, C. A., & Gugercin, S. (2020). Interpolatory methods for model reduction. SIAM. https://doi.org/10.1137/1.9781611976083

Beattie, C. A., & Gugercin, S. (2007, December 12–14). Krylov based minimization for optimal $H_2$ model reduction. Proceedings of the 46th IEEE Conference on Decision and Control, New Orleans, LA, USA. pp. 4385–4390. https://doi.org/10.1109/CDC.2007.4434939

Beattie, C. A., & Gugercin, S. (2009, December 16–18). A trust region method for optimal $H_2$ model reduction. Joint 48th IEEE Conference on Decision and Control and 28th Chinese Control Conference, Shanghai, P.R.China. pp. 5370–5375. https://doi.org/10.1109/CDC.2009.5400605

Benner, P., Grivet-Talocia, S., Quarteroni, A., Rozza, G., & Silveria, L. M. (2020a). Model order reduction: Volume 2: Snapshot-based methods and algorithms. De Gruyter. https://doi.org/10.1515/9783110671490

Benner, P., Grivet-Talocia, S., Quarteroni, A., Rozza, G., & Silveria, L. M. (2020b). Model order reduction: Volume 3: Applications. De Gruyter. https://doi.org/10.1515/9783110499001

Benner, P., Grivet-Talocia, S., Quarteroni, A., Rozza, G., & Silveria, L. M. (2021). Model order reduction: Volume 1: Systems- and data-driven methods and algorithms. De Gruyter. https://doi.org/10.1515/9783110498967

Benner, P., Köhler, M., & Saak, J. (2011, December). Sparse-dense Sylvester equations in $H_2$-model order reduction (Tech. Rep. No. MPIMD/11-11). Max Planck Institute. https://csc.mpi-magdeburg.mpg.de/preprints/2011/MPIMD11-11.pdf.

Breiten, T., Beattie, C., & Gugercin, S. (2015). Near-optimal frequency-weighted interpolatory model reduction. Systems & Control Letters, 78, 8–18. https://doi.org/10.1016/j.sysconle.2015.01.005

Chahlaoui, Y., & Van Dooren, P. (2005). Benchmark examples for model reduction of linear time-invariant dynamical systems. In P. Benner, D. C. Sorensenand & V. Mehrmann (Eds.), Dimension reduction of large-scale systems (Vol. 45, pp. 379–392). Springer. https://doi.org/10.1007/3-540-27909-1_24

Duff, I. P., & Kürschner, P. (2021). Numerical computation and new output bounds for time-limited balanced truncation of discrete-time systems. Linear Algebra and Its Applications, 623, 367–397. https://doi.org/10.1016/j.laa.2020.09.029

Flagg, G., Beattie, C., & Gugercin, S. (2012). Convergence of the iterative rational Krylov algorithm. Systems & Control Letters, 61(6), 688–691. https://doi.org/10.1016/j.sysconle.2012.03.005

Gawronski, W., & Juang, J. (1990). Model reduction in limited time and frequency intervals. International Journal of Systems Science, 21(2), 349–376. https://doi.org/10.1080/00207729008910366

Goyal, P., & Redmann, M. (2019). Time-limited $H_2$-optimal model order reduction. Applied Mathematics and Computation, 355, 184–197. https://doi.org/10.1016/j.amc.2019.02.065

Gugercin, S., & Antoulas, A. C. (2003, December). A time-limited balanced reduction method. Proceedings of the 42nd IEEE Conference on Decision and Control, Maui, Hawaii USA. pp. 5250–5253. https://doi.org/10.1109/CDC.2003.1272471

Gugercin, S., Antoulas, A. C., & Beattie, C. (2008). $H_2$ model reduction for large-scale linear dynamical systems. SIAM Journal on Matrix Analysis and Applications, 30(2), 609–638. https://doi.org/10.1137/060666123

Holmes, P., Lumley, J., Berkooz, G., & Rowley, C. W. (2012). Turbulence, coherent structures, dynamical systems and symmetry (2nd ed.). Cambridge University Press. https://doi.org/10.1017/CBO9780511919701

Kürschner, P. (2018). Balanced truncation model order reduction in limited time intervals for large systems. Advances in Computational Mathematics, 46(6), 1821–1844. https://doi.org/10.1007/s10444-018-9608-6

Luenberger, D. G. (1997). Optimization by vector space methods. John Wiley & Sons. https://www.wiley.com/en-us/Optimization+by+Vector+Space+Methods-p-9780471181170

Redmann, M. (2020, February). An $L_2$-error bound for time-limited balanced truncation. Systems & Control Letters, 136, 104620. https://doi.org/10.1016/j.sysconle.2019.104620

Redmann, M., & Kürschner, P. (2018). An output error bound for time-limited balanced truncation. Systems & Control Letters, 121, 1–6. https://doi.org/10.1016/j.sysconle.2018.08.004

Rommes, J., & Martins, N. (2006). Efficient computation of transfer function dominant poles using subspace acceleration. IEEE Transactions on Power Systems, 21(3), 1218–1226. https://doi.org/10.1109/TPWRS.2006.876671

Shaker, H. R., & Shaker, F. (2013, June 17–19). Generalized time-limited balanced reduction method. American Control Conference, Washington, DC, USA. pp. 5530–5535. https://doi.org/10.1109/ACC.2013.6580703

Sinani, K., & Gugercin, S. (2019). $H_2(t_f)$ optimality conditions for a finite-time horizon. Automatica, 110, Article 108604. https://doi.org/10.1016/j.automatica.2019.108604

Werner, S. W. (2021). Structure-preserving model reduction for mechanical systems [PhD Dissertation]. Department of Mathematics, Otto von Guericke University, Magdeburg, Germany. https://doi.org/10.25673/38617.

Wolf, T. (2014). $H_2$ pseudo-optimal model order reduction [Doctoral dissertation]. Technische Universität München. https://mediatum.ub.tum.de/dok2/1228083/file.pdf.

Xu, Y., & Zeng, T. (2011). Optimal $H_2$ model reduction for large-scale MIMO systems via tangential interpolation. International Journal of Numerical Analysis & Modeling, 8(1), 174–188. http://www.math.ualbertra.ca/ijnam/Volume-8-2011/No-1-11/2011-01-10.pdf.

Zulfiquar, U., Sreeram, V., & Du, X. (2020). Time-limited pseudo-optimal $H_2$-model order reduction. IET Control Theory & Applications, 14(14), 1995–2007. https://doi.org/10.1049/iet-cta.2019.1105