AN ITERATIVE MODEL ORDER REDUCTION METHOD
FOR LARGE-SCALE DYNAMICAL SYSTEMS

K. MOHAMED[✉1], A. MEHDI[2] and M. ABDELKADER[3]

Abstract

We present a new iterative model order reduction method for large-scale linear time-invariant dynamical systems, based on a combined singular value decomposition–adaptive-order rational Arnoldi (SVD-AORA) approach. This method is an extension of the SVD-rational Krylov method. It is based on two-sided projections: the SVD side depends on the observability Gramian by the resolution of the Lyapunov equation, and the Krylov side is generated by the adaptive-order rational Arnoldi based on moment matching. The use of the SVD provides stability for the reduced system, and the use of the AORA method provides numerical efficiency and a relative lower computation complexity. The reduced model obtained is asymptotically stable and minimizes the error (H_2 and H_∞) between the original and the reduced system. Two examples are given to study the performance of the proposed approach.

2010 *Mathematics subject classification*: primary 93A; secondary 15A, 93B, 65f.

Keywords and phrases: model order reduction, AORA, SVD, Gramian, large scale, Krylov, moment matching, H_∞.

1. Introduction

Modelling is an important tool in the description and characterization of large-scale systems. However, these high order models are difficult to manipulate, the resolution of such models is indeed very demanding in computational resources, storage space, and most importantly the central processing unit (CPU) time, especially when applying a control strategy which becomes very difficult to determine.

[1]Université de Tunis El Manar, École Nationale d'Ingenieurs de Tunis, Laboratoire de Recherche Analyse et Commande des Systèmes, LR-11-ES20, BP 37, Le Belvedere 1002 Tunis, Tunisia;
e-mail: Mohamed.Kouki@isigk.rnu.tn.
[2]Université de Carthage, École Nationale d'Ingénieurs de Carthage, Laboratoire de Recherche Analyse et Commande des Systèmes, LR-11-ES20, BP 37, Le Belvedere 1002 Tunis, Tunisia;
e-mail: mehdi.abbes@enit.rnu.tn.
[3]Université de Tunis El Manar, Faculté des Sciences de Tunis, Laboratoire de Recherche Analyse et Commande des Systèmes, LR-11-ES20, BP 37, Le Belvedere 1002 Tunis, Tunisia;
e-mail: abdelkader.mami@fst.rnu.tn.

1.1. Model order reduction Model order reduction is a good way to overcome the difficulties mentioned above, while the reduced model captures the main characteristics of the original complex model. Hence, the use of model order reduction techniques generates reduced-order systems that capture the essential dynamic behaviour of the original systems (such as stability, passivity, balance), and improves the performance during the numerical simulation procedure. Moreover, the stability of the reduced system and the minimization of the absolute error between the original and reduced systems are important criteria.

According to Antoulas and Serkan, these reduction methods can be divided into two types. First, the Krylov based algorithm is numerically efficient [1, p. 370], [3]; it is focused on an appropriate choice of original system moments which are attributed to the reduced one. Some of these permit a reduced system to be generated around one frequency (for example, Arnoldi [11, 18], Lanczos [1, p. 351] and [8]), around a frequency range (Rational Arnoldi [12, 14], AORA [12], AOGRA [4], AORL [7]). The complexity is $O(n^2r)$ for dense systems, and can be as low as $O(nr^2)$ for sparse systems depending on the sparsity structure, and the storage requirement is $O(nr)$. On the other hand, the SVD based algorithms are based in the resolution of the Lyapunov equation. We cite the balanced truncation [1, p. 234], the least squares and optimal Hankel norm [1, p. 276]. The complexity of these methods is $O(n^3)$, and they are expensive to implement in large-scale systems.

1.2. Problem description The model reduction problem we are interested in is the stability of the reduced system which is not always guaranteed [1, 7, 12]. However, Antoulas [1, p. 364] gives a Krylov based multiple restart algorithm which gives a reduced stable system. The choice of the appropriate moments was also studied by Druskin and Simoncini [5]. On the other hand, an SVD-rational algorithm generates a stable reduced system but the original model is limited in the order ($n < 200$) due to the complexity of the resolution of Lyapunov equation [1, 10], which requires the determination of both Gramian matrices P and Q of controllability [1, p. 200] and observability [9], respectively. A combined SVD-rational Krylov algorithm is proposed by Serkan et al. [9, 16], which takes advantage of the two preceding methods [1, p. 34], [10].

In this paper, we propose a new iterative SVD-AORA algorithm that combines the advantages of the SVD and AORA algorithms. This new algorithm presents better results than the SVD-Krylov algorithm.

The model reduction problem that we are addressing can be stated as follows.

We are given a linear dynamical system in state space form [1, p. 34], [13, 14, 17],

$$\zeta = \begin{cases} E\dfrac{dx(t)}{dt} = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases} \tag{1.1}$$

where $E, A \in \mathbb{R}^{n \times n}$ with E as an identity matrix, $B \in \mathbb{R}^n$, $C \in \mathbb{R}^n$; for simplicity, we take $D = 0$, and $x(t) \in \mathbb{R}^n$ is the state vector; $u(t) \in \mathbb{R}$ and $y(t) \in \mathbb{R}$ are called the input and output vectors, respectively. We assume that $u(t)$ is an impulse function. We also assume that the full-order model is asymptotically stable (that is, all eigenvalues λ_i of A have positive real parts $(\text{Real}(\lambda_i(A)) > 0$ for $i = 1, \ldots, n))$ and is minimal, that is, reachable and observable.

The transfer function $f(s)$ (where s is a complex variable) of the original system is obtained by applying the Laplace transform to equation (1.1) [14, 17]:

$$f(s) = C^T (sE - A)^{-1} B = C^T X(s), \tag{1.2}$$

where $X(s) = (sE - A)^{-1} B$ and $Y(s) = C^T X(s)$ are the state variables and the output, respectively.

The problem consists of approximating the reduced system matrices $\hat{E}$ and $\hat{A} \in \mathbb{R}^{r \times r}$, the reduced input matrix $\hat{B} \in \mathbb{R}^r$ and the reduced output matrix $\hat{C} \in \mathbb{R}^r$, where $r \ll n$. The state space description of the reduced model is as follows [1, p. 371], [14, 16, 17]:

$$\hat{\zeta} = \begin{cases} \hat{E}\dfrac{d\hat{x}(t)}{dt} = \hat{A}\hat{x}(t) + \hat{B}u(t), \\ \hat{y}(t) = \hat{C}x(t) + \hat{D}u(t). \end{cases}$$

The transfer function of the reduced system is given by [14, 16, 17]

$$\hat{f}(s) = \hat{C}^T (s\hat{E} - \hat{A})^{-1} \hat{B}. \tag{1.3}$$

This paper is organized as follows. In Section 2 the basic tools are given. In Section 3, the iterative SVD-AORA method is presented with applications on numerical examples. In Section 4 we give a comparison between the proposed method and other method, and we conclude the paper in Section 5.

2. Basic tools

2.1. Krylov subspace   Given a linear time-invariant system in the state space form (1.1), we assume that $u(t)$ is an impulse function [12, 14]. Let us define two matrices, $\psi_i = -(s_i E - A)^{-1} E$ and $\xi_i = (s_i E - A)^{-1} B$, where $(s_i E - A)$ is assumed to be nonsingular, and an expansion frequency $s_i \in S$ with S as a set of the predetermined expansion points. Applying the Taylor expansion of $X(s)$ at expansion points s_i, we obtain [1, p. 37], [12, 14]

$$X(s) = \sum_{j=0}^{\infty} X^j(s_i)(s - s_i)^i,$$

where $X^j(s_i) = \psi_i^j \xi_i$ and $Y^j(s_i) = C^T X^j(s_i)$ are the jth-order system moment and the jth-order output moment at s_i, respectively. We use the modified Gram–Schmidt orthogonalization technique to generate the Krylov subspace $K_r(\psi, \xi) = \text{span}\{\xi_i, \psi_i \xi_i, \ldots, \psi_i^{j-1} \xi_i\}$ [6, 8, 12, 14, 15]. Let $V_r \in \mathbb{R}^{n \times r}$ be the orthonormal basis

generated from the successive Krylov subspaces. Our goal is to find the reduced system of matrices $\hat{E}$, $\hat{A}$, $\hat{B}$ and $\hat{C}$ of order $r \ll n$, which can be defined by the congruence transformation

$$\hat{E} = V_r^T E V_r, \quad \hat{A} = V_r^T A V_r, \quad \hat{B} = V_r^T B, \quad \hat{C} = C^T V_r.$$

The modified Gram–Schmidt orthogonalization technique is numerically efficient to generate the matching moments of each Krylov subspace.

2.2. Singular value decomposition   We apply the SVD method to a rectangular matrix of dimension (m, n) [1, 14, 16, 17]. There exist an orthogonal matrix $\mathcal{U}_m$ of dimension (m, m), an orthogonal matrix $\mathcal{V}_n$ of dimension (n, n) and a diagonal matrix $\Sigma = \text{diag}(\sigma_1, \sigma_2, \ldots, \sigma_p, \sigma_{p+1}, \ldots, \sigma_n)$, where p is the maximum number of singular values different from zero. Note that these values are all positive and $\sigma_1 \geq \sigma_2 \geq \cdots \geq \sigma_p$ [1, p. 58], [14, 17].

The matrix A is written as $A = \mathcal{U}_m \Sigma \mathcal{V}_n^T$ [16].

Take a linear time-invariant system as in Equation (1.1), which is asymptotically stable and minimal; therefore, matrix A is stable. The solutions of the following Lyapunov equations allow the determination of g_c and g_o:

$$A g_c + g_c A^T + B B^T = 0, \quad A^T g_o + g_o A + C C^T = 0,$$

where g_c and $g_o \in \mathbb{R}^n$ are the unique symmetric positive definite matrices, called the controllability and the observability Gramians, respectively. These two matrices are written in a factored form as $g_c = U U^T$ and $g_o = L L^T$, where U and L are upper and lower triangular matrices, respectively. They are also called the Cholesky factors [16] or the square roots of g_c and g_o, respectively. By applying the singular value decomposition on the matrix $U^T L$, we obtain three matrices $\mathbb{W}$, $\mathbb{Y}$ and χ [14], from which we determine two projection matrices μ and θ, that are used in the calculation of the reduced one:

$$\mu = L \mathbb{Y}_r \chi_r^{-1/2}, \quad \theta = U \mathbb{W}_r \chi_r^{-1/2},$$

where $\mathbb{Y}_r$ and $\mathbb{W}_r$ present the first r columns of $\mathbb{Y}$ and $\mathbb{W}$, respectively, and χ_r contains the r first Hankel singular values [1, 9] of χ.

The goal is to find the reduced system matrices $\hat{E}$, $\hat{A}$, $\hat{B}$ and $\hat{C}$ of order $r \ll n$, via balanced truncation. They can be defined by the congruence transformation:

$$\hat{E} = \mu^T E \theta, \quad \hat{A} = \mu^T A \theta, \quad \hat{B} = \mu^T B, \quad \hat{C} = C^T \theta.$$

The SVD technique is much used in the average-scale systems.

2.3. H_∞ of a dynamical system   To evaluate the performance of the proposed method, we need to determine the H_∞ norm. In this section, we give an overview.

DEFINITION 2.1. In the general case, taking a stable system in the form of equation (1.1), in which the eigenvalues of A have negative real parts, the equation of the H_∞ norm of

the stable system is as follows [1, p. 55], [9, 16]:

$$\|\zeta\|_2 = \|\mathcal{S}\|_{2-\text{induced}} = \|H\|_{H_\infty} = \sup_w \sigma_{\max}(f(iw))$$

where $\sigma_{\max}$ is the maximum Hankel singular value of the dynamical system, $w \in \mathfrak{R}$ and $\mathcal{S}$ is a convolution operator. From this, we treat a special case: take an asymptotically stable dynamical system as in equation (1.1), then the H_∞ norm is defined as

$$\|\zeta\|_{H_\infty} = \sup \|f(iw)\|_2.$$

2.4. Adaptive-order rational Arnoldi algorithm   The adaptive-order rational Arnoldi (AORA) algorithm is an improvement of the Arnoldi and rational Arnoldi algorithms [4, 12]. The AORA algorithm gives a reduced model around a frequency range with an automatic choice of matching moments. However, the expansion frequency s_i and the number of matched moment $\hat{j}_i$ must be given. no

Given a fixed set of expansion points $S = [s_1, s_2, \ldots, s_i]$ and the number of matched moments $r = \sum \hat{j}_i$, where r is the order of the reduced system, the method can generate an orthonormal matrix V_r from the successive Krylov subspaces K_r satisfying the orthogonality relation $V_r^T V_r = E_r$. The reduced-order model is defined by the relationships

$$\hat{E} = V_r^T E V_r, \hat{A} = V_r^T A V_r, \quad \hat{B} = V_r^T B, \quad \hat{C} = C^T V_r.$$

The main steps of the AORA algorithm are the following.

Step 1.   Initialize the first two vectors $k^{(0)}(s_i)$ and $R^{(0)}(s_i)$ of the Krylov sequence for each expansion point s_i where $i \in \{1, 2, \ldots, \hat{i}\}$, knowing that $k^{(0)}(s_i) = R^{(0)}(s_i)$ and the normalization coefficient $h_\Pi(s_i) = 1$ for all s_i.

Step 2.   Choose the expansion frequency s_i, knowing that s_i gives the greatest difference between the j_ith-order moments of the original system $Y(s)$ and the reduced system $\hat{Y}(s)$; the chosen expansion frequency is called $s_{i_j^*}$.

Step 2.1.   Select the expansion point $s_{i_j^*}$ and apply the Arnoldi algorithm around this point. The new orthonormal vector v_j is incorporated into the orthonormal matrix V_{j-1}. The normalization coefficient $h_\Pi(s_i)$ is updated according to $s_{i_j^*}$.

Step 2.2.   Update the residue vector $R^{(j)}(s_i)$ and generate the Hessenberg matrix H, using the modified Gram–Schmidt orthogonalization technique.

Step 3.   Generate the real orthogonal matrix V_r using the reduced QR factorization, if there exist any complex expansion points.

The details of the adaptive-order rational Arnoldi method can be found in Algorithm 1 [12]:

Algorithm 1: AORA (**Inputs:** E, A, B, C, D, S; **Output:**V_r)

---

(1) Initialize;

for each $s_i \in S$ **do**

$\quad | \quad k^{(0)}(s_i) := (s_i E - A)^{-1} B$, $R^{(0)}(s_i) = k^{(0)}(s_i)$ and $h_\Pi(s_i) := 1$

end

(2) Begin AORA iterations;

for $j = 1, 2, \ldots, r$ **do**

$\quad$ (2.1) Select the expansion frequency with the maximum output moment error;

$\quad$ Choose $s_i \in S$ as the i giving $\max_i(|h_\Pi(s_i C^T R^{(j-1)}(s_i)|))$. Set $s_{i_j^*}$ as the expansion frequency in the jth iteration;

$\quad$ (2.2) Generate the orthonormal vector at $s_{i_j^*}$;

$\quad h_{j,j-1}(s_{i_j^*}) := \|R^{(j-1)}(s_{i_j^*})\|$;

$\quad v_j = \|R^{(j-1)}(s_{i_j^*})\|/h_{j,j-1}(s_{i_j^*})$;

$\quad h_\Pi(s_{i_j^*}) := h_\Pi(s_{i_j^*})h_{j,j-1}(s_{i_j^*})$;

$\quad$ (2.3) Update the residue $R^{(j)}(s_i)$ for the next iteration;

$\quad$ **for** each $s_i \in S$ **do**

$\quad\quad$ **if** ($s_i == s_{i_j^*}$) **then**

$\quad\quad\quad | \quad k^{(j)}(s_{i_j^*}) := -(s_i E - A)^{-1} E v_j$;

$\quad\quad$ **else**

$\quad\quad\quad | \quad k^{(j)}(s_{i_j^*}) := k^{(j)}(s_i)$;

$\quad\quad$ **end**

$\quad\quad R^{(j)}(s_i) := k^{(j)}(s_i)$;

$\quad\quad$ **for** $t = 1, 2, \ldots, j$ **do**

$\quad\quad\quad | \quad h_{t,j}(s_i) := v_t^H R^j(s_i)$, $R^{(j)}(s_i) := R^{(j)}(s_i) - h_{t,j}(s_i)$;

$\quad\quad$ **end**

$\quad$ **end**

end

$V_r = [v_1, v_2, \ldots, v_r]$;

(3) Generate real V_r for complex expansion points;

if there exists any $s_i \in S$ such that s_i is not real number **then**

$\quad | \quad V_{\text{real}} := \text{real}(V_r)$, $V_{\text{imag}} = \text{imag}(V_r)$, $[V_r, rr] = qr([V_{\text{real}} V_{\text{imag}}])$;

end

2.5. SVD-rational Krylov algorithm While the AORA algorithm does not always guarantee stability of the reduced system, the SVD-rational Krylov algorithm gives a reduced model with guaranteed stability [1, p. 372], [9]. This is possible with the resolution of the Lyapunov equation in the SVD part of the algorithm. Hence, the SVD-rational Krylov algorithm [9] combines the advantages of the rational Krylov based method and the singular value decomposition based method, and the use of SVD provides stability for the reduced system. This method can yield two matrices.

The first matrix V_r is generated by the rational Krylov technique, and it depends on the observability Gramian. The second matrix Z_r is generated by the singular value decomposition technique. The two matrices Z_r and V_r satisfy the orthogonality relation $Z_r^T V_r = E_r$. The reduced-order model is defined by the relationships

$$\hat{E} = Z_r^T E V_r, \quad \hat{A} = Z_r^T A V_r, \quad \hat{B} = Z_r^T B, \quad \hat{C} = C^T V_r.$$

The main steps of the SVD-rational Krylov algorithm are the following.

Step 1. Given an expansion frequency range, the number of frequencies must be equal to the order of reduced system. The initial choice of these points is arbitrary.

Step 2. Use the basis rational Krylov method to construct the orthonormal basis V_r.

Step 3. Calculate the Gramian matrix of observability g_o.

Step 4. Construct the second projection matrix Z_r using the Gramian matrix and the first projection matrix V_r.

Step 5. Calculate the new range of expansion frequency and recalculate the two new projection matrices V_r and Z_r

Step 5.1. Compute the reduced state matrix $\hat{A}$.

Step 5.2. Determine the eigenvalues of the matrix $\hat{A}$.

Step 5.3. Recalculate the two projection matrices V_r and Z_r.

The different steps of the SVD-rational Krylov method can be found in Algorithm 2 [1]:

Algorithm 2: SVD-rational Krylov (**Inputs:** E, A, B, C, S; **Outputs:** V_r, Z_r)

---

(1) Initialization of frequency expansion;
    s_i for $i = 1, 2, \ldots, r$;
(2) Construction of the matrix V_r by the rational Krylov based method, knowing that
    $\text{Ran}(V_r) = \text{vect}\{(s_1 E - A)^{-1} B, \ldots, (s_r E - A)^{-1} B\}$, with $V_r^T V_r = E_r$;
(3) Calculate the Gramian matrix of observabilty;
    $g_o = \int_0^\infty e^{tA^T} C^T C e^{tA} dt$;
(4) Construction of the matrix Z_r;
    $Z_r = g_0 V_r (V_r^T g_0 V_r)^{-1}$;
(5) Calculate the new expansion points;
    (a) $\hat{A} = Z_r^T A V_r$;
    (b) $s_i = -\lambda(\hat{A})$ for $i = 1, \ldots, r$;
    (c) recalculate the matrix V_r by the AORA method, knowing that
    $\text{Ran}(V_r) = \text{vect}\{(s_i E_n - A)^{-1} B, (s_i E_n - A)^{-1} E * (s_i E_n - A)^{-1} B, \ldots, (s_i E_n - A)^{j-1} E * (s_i E_n - A)^{-1} B\}$, with $V_r^T V_r = E_r$;
    (d) $Z_r = g_0 V_r (V_r^T g_0 V_r)^{-1}$;
(6) Parameters of reduced model;
    $\hat{E} = Z_r^T E V_r, \hat{A} = Z_r^T A V_r, \hat{B} = Z_r^T B, \hat{C} = C^T V_r$.

---

3. Iterative SVD-AORA

3.1. Presentation　Rather than choosing a matching moment in the reduced model of the SVD-rational Krylov algorithm, the use of an AORA algorithm allows an automatic choice of matching moments, which gives better results than the rational Krylov algorithm. This method preserves the stability of the reduced model and minimizes the error [6, 12]. The minimization of error is due to the use of an AORA algorithm to generate the first subspace V_r, and the preservation of stability is obtained by the use of a method based on singular value decomposition to generate the second subspace Z_r.

3.1.1 *Generating the subspace* V_r. Initially setting the expansion points $S = [s_1, s_2, \ldots, s_r]$ for $i = 1, \ldots, r$ arbitrarily or experimentally such that $V_r^T V_r = E_r$, we try several times to find the best point by only using the AORA algorithm, before iteratively applying the SVD-AORA algorithm.

The basic concept of the AORA method [12] is to select an expansion point $s_{i_j^*}$ among all expansion points S in the $(r + 1)$th iteration such that

$$Y^{(\hat{j}_{r+1}^*)}(s_{i_{r+1}^*}) = \hat{Y}_{r+1}^{(\hat{j}_{r+1}^*)}(s_{i_{r+1}^*}) \quad \text{or} \quad \mathcal{E}_{r+1}^{(\hat{j}_{r+1}^*)}(s_{i_{r+1}^*}) = 0.$$

From the new expansion point $s_{i_{r+1}^*}$ and the new orthonormal vector v_{r+1}, the new transfer function error in the $(r + 1)$th iteration can be expressed as [1, p. 392], [9, 12]

$$\begin{aligned}
\mathcal{E}_{r+1}(s) &= Y(s) - \hat{Y}_{r+1}(s) \\
&= \mathcal{E}_{r+1}^{\hat{j}_{r+1}^*}(s_{i_{r+1}^*})(s - s_{i_{r+1}^*})^{\hat{j}_{r+1}^*} \\
&\quad + \mathcal{E}_{r+1}^{(\hat{j}_{r+1}^*+1)}(s_{i_{r+1}^*})(s - s_{i_{r+1}^*})^{(\hat{j}_{r+1}^*+1)} + O((s - s_{i_{r+1}^*})^{(\hat{j}_{r+1}^*+2)}) \\
&= 0.(s - s_{i_{r+1}^*})^{\hat{j}_{r+1}^*} + \mathcal{E}_{r+1}^{(\hat{j}_{r+1}^*+1)}(s_{i_{r+1}^*})(s - s_{i_{r+1}^*})^{(\hat{j}_{r+1}^*+1)} + O((s - s_{i_{r+1}^*})^{(\hat{j}_{r+1}^*+2)}).
\end{aligned}$$

The selection of the expansion point $s_{i_{r+1}^*}$ is based on the criterion

$$\max_{s_i \in S} |Y^{\hat{j}_i}(s_i) - \hat{Y}^{\hat{j}_i}(s_i)| = \max_{s_i \in S} |h_\pi(s_i)C^T R^{(j-1)}(s_i)|,$$

where h_π is the normalization coefficient, and R is the residue vector [12]. Theorems 3.1 and 3.2 summarize this result.

THEOREM 3.1. *Assume that an asymptotically stable linear system in the form of equation* (1.1) *fixes the expansion point s_i for $j = 0, 1, \ldots, r$ and $i = 1, \ldots, r$, and let $\psi_i = -(s_i E - A)^{-1}E$ and $\xi_i = (s_i E - A)^{-1}B$. Then after r steps, the adaptive-order rational Arnoldi can iteratively generate an orthonormal basis $V_r \in \mathbb{R}^{n \times r}$ from the successive Krylov subspaces $K_r(\psi, \xi) = \text{span}\{\xi_i, \psi_i\xi_i, \ldots, \psi_i^{j-1}\xi_i\} = \text{span}\{v_1, v_2, \ldots, v_r\}$. Since $X^{(j)}(s_i) \in \text{colspan}\{V_r\}$ for $j = 0, 1, \ldots, r$ and $i = 1, \ldots, r$, we have*

$$X^{(j)}(s_i) = V_r \hat{X}_r^{(j)}(s_i) \quad and \quad Y^{(j)}(s_i) = \hat{Y}_r^{(j)}(s_i).$$

PROOF. Let V_r be the orthonormal basis generated by using the adaptive-order rational Arnoldi [12] with r iteration and S expansion points, such that the expansion points are arbitrarily fixed in the first iteration of Algorithm 3. To demonstrate this theorem, we consider two cases.

First. if $s_{i_{r+1}}^* \notin \{s_{i_1^*}, \ldots, s_{i_r^*}\}$ and $s_{i_{r+1}} \in \{s_{i_1^*}, \ldots, s_{i_r^*}\}$, then from the AORA algorithm it is known that [12]

$$k_j = (s_{i_{r+1}}^* E - A)^{-1} B - V_r[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}]^T = h_{r+1,r} V_{r+1},$$

which can be rewritten as

$$V_{r+1}[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}]^T = (s_{i_{r+1}^*} E - A)^{-1} B$$
$$= (s_{i_{r+1}^*} E - A)^{-1}(s_{i_1^*} E - A)k_0$$
$$= h_{1,0}^{i_{r+1}^*}(s_{i_{r+1}^*} E - A)^{-1}(s_{i_1^*} E - A)V_{r+1}e_1, \quad (3.1)$$

where e_1 is the first column of the identity matrix. Now, multiplying both sides of equation (3.1) by $(s_{i_{r+1}^*} E - A)$ and rearranging each term, we obtain

$$(s_{i_{r+1}^*} E - A)V_{r+1}[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}]^T = (s_{i_{r+1}^*} E - A)h_{1,0}^{i_{r+1}^*}(s_{i_{r+1}^*} E - A)^{-1}(s_{i_1^*} E - A)V_{r+1}e_1$$
$$= (s_{i_{r+1}^*} E - A)V_r[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}]^T h_{1,0}^{i_{r+1}^*}(s_{i_1^*} E - A)V_{r+1}e_1$$
$$- AV_{r+1}([h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}] - h_{1,0}e_1)$$
$$= EV_{r+1}(s_{i_1^*} h_{1,0}^{i_{r+1}^*} e_1 - s_{i_{r+1}^*}[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}]^T). \quad (3.2)$$

Adding $s_i EV_{r+1}([h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T - h_{1,0}^{i_{r+1}^*} e_1)$ to both sides of equation (3.2) yields

$$s_i EV_{r+1}([h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T - h_{1,0}^{i_{r+1}^*} e_1) - AV_{r+1}([h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}] - h_{1,0}e_1)$$
$$= s_i EV_{r+1}([h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T - h_{1,0}^{i_{r+1}^*} e_1) + EV_{r+1}(s_{i_1^*} h_{1,0}^{i_{r+1}^*} e_1 - s_{i_{r+1}^*}[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r,r}^{i_{r+1}^*}]^T),$$

and after rearranging, we obtain

$$(s_i E - A)V_{r+1}([h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T - h_{1,0}^{i_{r+1}^*} e_1)$$
$$= EV_{r+1}((s_{i_1^*}) - s_i)h_{1,0}^{i_{r+1}^*} e_1 + (s_i - s_{i_{r+1}^*})[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T). \quad (3.3)$$

We assume that

$$f_r = [h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T - h_{1,0}^{i_{r+1}^*} e_1 \quad (3.4)$$

and

$$g_r = (s_{i_1^*}) - s_i)h_{1,0}^{i_{r+1}^*} e_1 + (s_i - s_{i_{r+1}^*})[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T, \quad (3.5)$$

where f_r and g_r are the rth columns of the Hessenberg matrices [12] F_r and G_r, respectively.

Next, if $s_{i_{r+1}}^* \in \{s_{i_1^*}, \ldots, s_{i_r^*}\}$ and $s_{i_{r+1}} \in \{s_{i_1^*}, \ldots, s_{i_r^*}\}$, then we obtain

$$V_{r+1}[h_{1,r}^{i_{r+1}^*}, \ldots, h_{r+1,r}^{i_{r+1}^*}]^T = -(s_i E - A)^{-1} V_{r+1}(e_k) \quad (3.6)$$

and

$$(s_iE - A)V_{r+1}[h_{1,r}^{i^*_{r+1}}, \ldots, h_{r+1,r}^{i^*_{r+1}}]^T = EV_{r+1}(-e_r), \tag{3.7}$$

where

$$f_r = [h_{1,r}^{i^*_{r+1}}, \ldots, h_{r+1,r}^{i^*_{r+1}}] \tag{3.8}$$

and

$$g_r = -e_r. \tag{3.9}$$

Combining equations (3.3)–(3.9) for all r steps of AORA algorithm process yields

$$(s_iE - A)V_{r+1}f_r = EV_{r+1}g_r,$$
$$(s_iE - A)V_{r+1}f_r = EV_{r+1}g_r,$$
$$-(s_iE - A)EV_r = V_r(-F_{r,r}G_{r,r}^{-1}),$$

where $F_{r,r}$ and $G_{r,r}$ present two Hessenberg matrices. For $j = 0, 1, \ldots, r$ and $i = 1, \ldots, r$, we know that $X^{(j)}(s_{i^*}) \in V_r$; then $X^{(j)}(s_{i^*}) = V_r V_r^T X^{(j)}(s_{i^*})$. By using the recursion formula we obtain

$$X^{(j)}(s_i^*) = (-(s_{i^*}E - A)^{-1}E)^{-j} V_r V_r^T X^{(0)}(s_{i^*})$$
$$= V_r(-F_{r,r}G_{r,r}^{-1})^j V_r^T X^{(0)}(s_{i^*}).$$

Then

$$\hat{X}_r^{(j)}(s_{i^*}) = V_r^T X^{(j)}(s_i^*)$$
$$= V_r(-F_{r,r}G_{r,r}^{-1})^j V_r^T X^{(0)}(s_{i^*}),$$

which is equivalent to proving that

$$\hat{X}_r^{(0)}(s_{i^*}) = V_r^T X^{(0)}(s_{i^*}) \quad \text{and} \quad -(s_{i^*}\hat{E} - \hat{A})^{-1}\hat{E} = -F_{r,r}G_{r,r}^{-1}.$$

In this case,

$$(s_{i^*}E - A)^{-1}b = V_r V_r^T(s_{i^*}E - A)^{-1}B,$$
$$V_r^T B = V_r^T(s_{i^*}E - A)V_r V_r^T(s_{i^*}E - A)^{-1}B,$$
$$(s_{i^*}V_r^T EV_r - V_T^T AV_r)^{-1}V_r^T = V_r^T(s_{i^*}E - A)^{-1}B.$$

Moreover,

$$-(s_{i^*}E - A)^{-1}EV_r = V_r(-F_{r,r}G_{r,r}^{-1}).$$

Therefore, $Y^j(s_{i^*}) = CX^{(j)}(s_i^*) = CV_r\hat{X}^{(j)}(s_{i^*}) = \hat{C}\hat{X}_r^{(j)}(s_i).$ □

THEOREM 3.2 (Maximum output moment errors). *We assume that both output moments of the original system and those of the reduced-order system are matched, so that $Y^{(j)}(s_i) = \hat{Y}_r^{(j)}(s_i)$ for $j = 0, 1, \ldots, \hat{j}_i$ and $i = 1, \ldots, r$. Select the expansion point $s_i \in S$ in the $(j + 1)$th iteration, using the AORA algorithm; then s_i yields the greatest error*

between the j_ith output moments of the original system and the reduced one (that is, $Y^{(\hat{j}_i)}(s_i)$ and $\hat{Y}_r^{(\hat{j}_i)}(s_i)$, respectively), which is given as

$$|\mathcal{E}_r^{(\hat{j}_i)}(s_i)| = \max_{s_i \in S} |Y^{\hat{j}_i}(s_i) - \hat{Y}^{\hat{j}_i}(s_i)| = \max_{s_i \in S} |h_\pi(s_i)C^T k^{(r)}(s_i)|,$$

where $h_\pi(s_i) = \prod_j \|R^{(j-1)}(s_i)\|$ is the normalization coefficient, and $r = \sum_{i=1}^{\hat{i}} \hat{j}_i$.

PROOF. The proof was given by Lee et al. [12].   □

3.1.2 *Generating the subspace Z_r.* We must calculate the Gramian matrix of observability from resolution of the following linear Lyapunov equation:

$$A^T g_o + g_o A + C^T C = 0. \tag{3.10}$$

This equation admits unique symmetric solutions if and only if $\lambda_i(A) + \bar{\lambda}_j(A) \neq 0$ for all i, j [9], where $\lambda_i(A)$ and $\bar{\lambda}_j(A)$ are the eigenvalues and conjugate eigenvalues of A, respectively. We use the solution of the large Lyapunov equation to obtain a good low-rank approximation. The Gramian matrix g_o is the solution to equation (3.10). This matrix (g_o) and the subspace V_r are used to generate the subspace Z_r, which is defined as $Z_r = g_o V_r (V_r^T g_o V_r)^{-1}$ with $Z_r^T V_r = E_r$. Then $\|f(iw) - f_r(iw)\|_{H_\infty}$ is minimized. Theorem 3.3 summarizes this result.

THEOREM 3.3. *We take an asymptotically stable system ζ with transfer function $f(s)$ as in equation (1.2), and assume that the V_r and Z_r are two projection matrices obtained by the SVD-AORA process. Let $\hat{\zeta}$ be a reduced system generated by the congruence transformation with transfer function $f_r(s)$ as in equation (1.3), knowing that the system has fixed and stable reduced eigenvalues $\{\lambda_1, \ldots, \lambda_r\}$. Then $\|f(iw) - f_r(iw)\|_{H_\infty}$ is minimized if and only if*

$$f(s) = f_r(s) \quad \text{for } s = -\lambda_1, -\lambda_2 \ldots, -\lambda_r.$$

PROOF. The proof has been given by Gugercin [9] and Lee et al. [12].   □

The following are the main steps of the proposed Algorithm 3.

Step 1.   Set the expansion point in an arbitrary manner.

Step 2.   Compute the V_r base with AORA knowing that the orthogonality condition is satisfied (that is, $V_r^T V_r = E_r$). All vectors of the orthonormal basis $V_r \in \mathbb{R}^{n \times r}$ are generated iteratively by the use of the Arnoldi algorithm around a single frequency, which is selected by the maximum error criterion. An upper Hessemberg matrix $H_r \in R^{r*r}$ is generated, which satisfies

$$\psi V_r = V_r H_r + h_{r+1,r} v_{r+1} e_r^T \quad \text{and} \quad v_1 = \xi/\|\xi\|.$$

Step 3.   Calculate the Gramian matrix of observability g_o.

Step 4.   Calculate the second projection matrix Z_r, using the Gramian matrix g_o and the projection matrix V_r.

Step 5. Calculate the reduced state matrix $\hat{A}$ and the corresponding eigenvalues. Using these eigenvalues as points of expansion to recalculate the base V_r, apply the AORA method again. Repeat until a stopping criterion in the expansion frequency is satisfied. The stopping criterion is a tolerance value already set. It is the relative change between two successive expansion points. Calculate again the Z_r matrix with respect to new base V_r, the same Gramian matrix of observability calculated in the third step with the satisfying orthogonality condition $Z_r^T V_r = E_r$.

Step 6. The reduced-order model is defined as

$$\hat{E} = Z_r^T E V_r, \quad \hat{A} = Z_r^T A V_r, \quad \hat{B} = Z_r^T B, \quad \hat{C} = C^T V_r.$$

The iterative SVD-AORA method is illustrated in Algorithm 3.

Algorithm 3: Iterative SVD-AORA (**Inputs:** E, A, B, C, S, tol; **Outputs:** V_r, Z_r)

(1) Initialization of the frequency expansion;
    s_i for $i = j = 1, 2, \ldots, r$;
(2) Construction of the matrix V_r by the AORA method knowing that;
    $\text{Ran}(V_r) = \text{vect}\{(s_i E - A)^{-1}B, (s_i E - A)^{-1}E * (s_i E - A)^{-1}B, \ldots, (s_i E - A)^{j-1}E * (s_i E - A)^{-1}B\}$, with $V_r^T V_r = E_r$;
(3) Calculate the Gramian matrix of observability;
    $g_o = \int_0^\infty e^{tA^T} C^T C e^{tA} dt$;
(4) Construction of Z_r;
    $Z_r = g_o V_r (V_r^T g_o V_r)^{-1}$;
(5) Compute the final V_r and Z_r matrices:
while *the relative changes in s_i : $((s_{i+1} - s_i)/s_i) \geq tol$* **do**
        (a) $\hat{A} = V_r^T A V_r$;
        (b) $s_i = -\lambda(\hat{A})$ for $i = 1, \ldots, r$;
        (c) Construction the matrix V_r by the AORA method knowing that:
        $\text{Ran}(V_r = \text{vect}\{(s_i E - A)^{-1}B, (s_i E - A)^{-1}E * (s_i E - A)^{-1}B, \ldots, (s_i E - A)^{j-1}E * (s_i E - A)^{-1}B\})$, with $V_r^T V_r = E_r$;
        (d): $Z_r = g_o V_r (V_r^T g_o V_r)^{-1}$;
end
(6) The reduced system parameters can be defined by the congruence
    transformation: $\hat{E} = Z_r^T E V_r, \hat{A} = Z_r^T A V_r, \hat{B} = Z_r^T B, \hat{C} = C^T V_r$.

3.2. Application To evaluate the proposed Algorithm 3, we take two SISO models of order 1006 (full order model (FOM) 1006) and 348 (Clamped beam 348) [2, 3], present the largest singular values of the frequency response of the original system and the reduced one, and give the absolute error and the pole distribution of the reduced system.
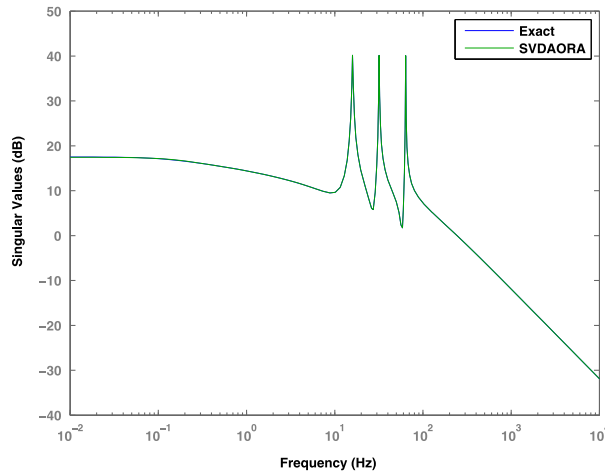
FIGURE 1. Largest singular value of the frequency response of the original system of (exact) order 1006 and the reduced one (SVD-AORA) of order 12 to a frequency range with the iterative SVD-AORA method (colour available online).
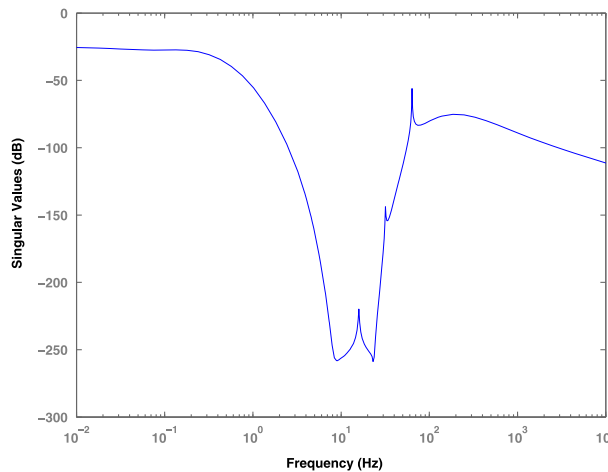

FIGURE 2. Absolute error between the original system of order 1006 and the reduced one of order 12 with the iterative SVD-AORA method.

EXAMPLE 3.4 (FOM model). Figure 1 shows the largest singular value of the frequency response of the original system of order 1006 and the reduced one of order 12; we see a good correlation between the two systems over the entire frequency range of the original system. Figure 2 shows the variation of the absolute error between the original system and the reduced one. Figure 3 depicts the pole distribution of the reduced system in the complex plane. We note that all poles have negative real parts. This shows the stability of the reduced system.

FIGURE 3. Pole distribution of FOM reduced system (order 12) with the iterative SVD-AORA method.
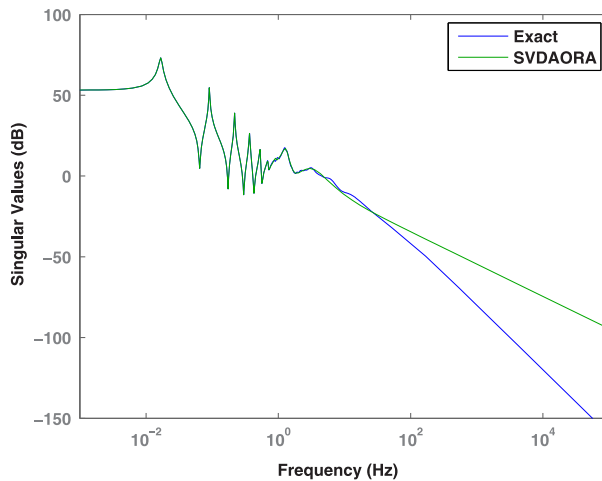

FIGURE 4. Largest singular value of the frequency response of the original system of (exact) order 348 and the reduced one (SVD-AORA) of order 22 to a frequency range with the iterative SVD-AORA method (colour available online).

EXAMPLE 3.5 (Clamped Beam model). Figure 4 shows the largest singular value of the frequency response of the original system of order 348 and the reduced one of order 22; we see a good correlation between the two systems over the entire frequency range of the original system. The variation of the absolute error between the original system and the reduced one is presented in the Figure 5. Figure 6 shows the pole distribution
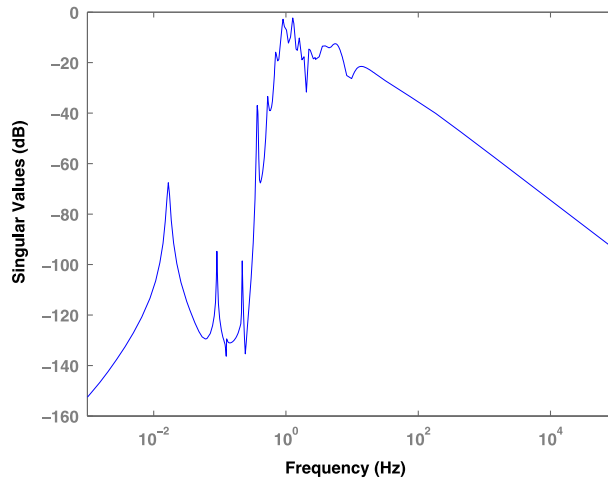
FIGURE 5. Absolute error between original system of order 348 and the reduced one of order 22 with iterative SVD-AORA method.
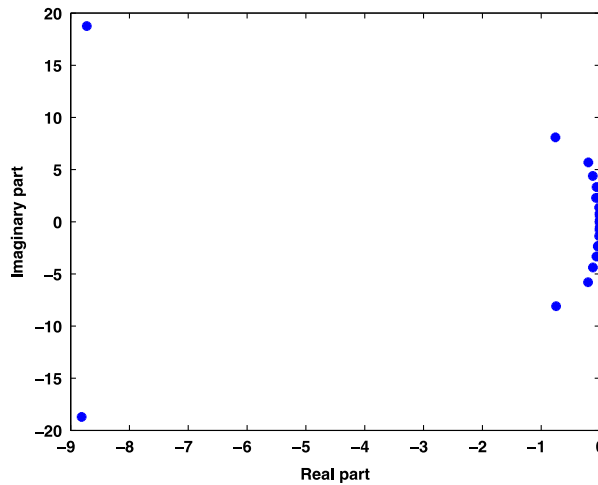

FIGURE 6. Pole distribution of Clamped Beam reduced system (order 22) with iterative SVD-AORA method.

of the reduced system in the complex plane. We note that all poles have negative real parts. This explains the stability of the reduced system.

4. Comparison with AORA and SVD-rational Krylov algorithms

A comparison is given in Figures 7 and 9 of three reduced models obtained by the AORA (Algorithm 1), SVD-rational Krylov (Algorithm 2) and iterative SVD-AORA
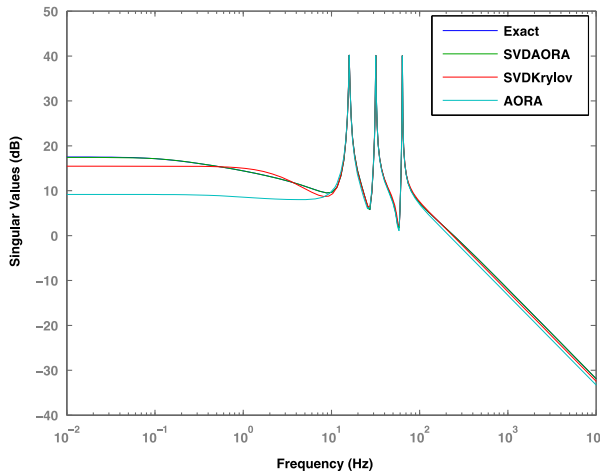
FIGURE 7. Largest singular value of the frequency response of the original system whose (exact) order is 1006 and the three reduced systems whose order is 12, with three methods (colour available online).
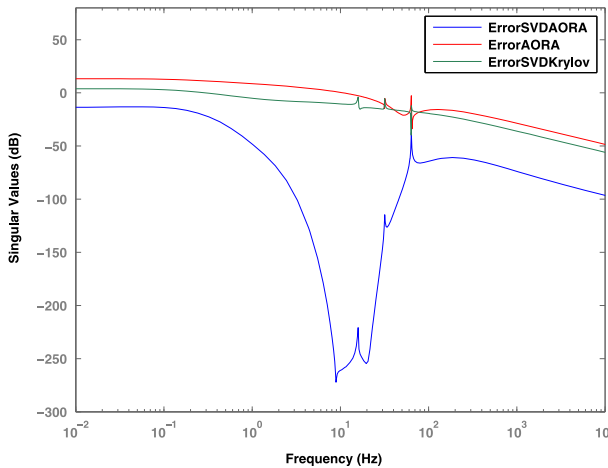

FIGURE 8. Absolute error systems of FOM model (colour available online).

(Algorithm 3) methods for the FOM and Clamped Beam models, respectively. Table 1 presents the H_∞-norms of the error systems and the CPU time for each method. We see from the table that the iterative SVD-AORA has the lowest H_∞-error for both models. However, in terms of computation time, the proposed method is outperformed by the competitive methods.

We observe that the best result is obtained by the iterative SVD-AORA method. We also note the previous difference in performance between the three methods from the
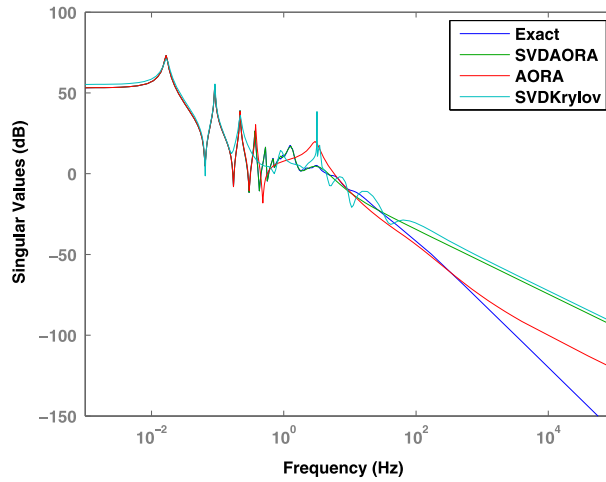
FIGURE 9. Largest singular value of the frequency response of the original system whose (exact) order is 348 and the three reduced systems whose order is 22, with three methods (colour available online).

TABLE 1. H_∞ norms of the errors systems and the CPU-time for each method.

Models	FOM		Clamped beam	
Methods	CPU-time	H_∞-error	CPU-Time	H_∞-error
AORA	1.311×10^2	1.810×10^{-3}	70.95	0.0686
SVD-rational Krylov	3.874×10^2	1.877×10^{-4}	1.231×10^2	5.970×10^{-4}
Iterative SVD-AORA	5.114×10^2	6.705×10^{-6}	1.410×10^2	4.934×10^{-5}

correspondence between the original system and the reduced one in Figures 8 and 10 with the absolute error variation; the best results are obtained by the iterative SVD-AORA method.

5. Conclusion

We have proposed a new iterative SVD-AORA model reduction method for large-scale systems, which takes advantages of both SVD and AORA methods. First, the stability of the reduced system is guaranteed by the use of the Lyapunov equation, which determines the Gramian observability matrix. Secondly, the application of the AORA method to generate an orthonormal basis is based on the criterion of maximum error. We conclude that this method is very efficient in the generation of orthonormal bases and highly capable in the numerical calculation. It combines the two methods to guarantee the digital stability and the efficiency of the algorithm. It also guarantees the
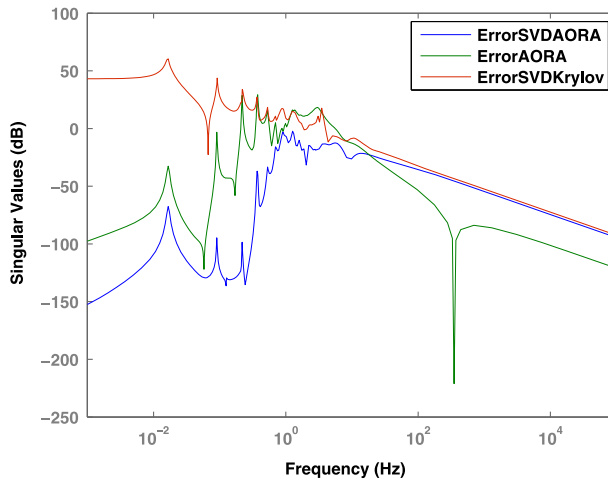
FIGURE 10. Absolute error systems of Clamped beam model (colour available online).

stability of the reduced system, and minimizes the relative error between the original system and the reduced one.

References

[1] A. C. Antoulas, *Approximation of large-scale dynamical systems* (Society for Industrial and Applied Mathematics, Philadelphia).

[2] Y. Chahlaoui and P. V. Dooren, "A collection of benchmark examples for model reduction of linear time invariant dynamical systems", *Niconet* (2002) http://www.icm.tu-bs.de/NICONET/benchmodred.html.

[3] Y. Chahlaoui and P. V. Dooren, "Benchmark examples for model reduction of linear time-invariant dynamical systems", *Dimens. Reduction of Large-Scale Systems* **45** (2005) 379–392; doi:10.1007/3-540-27909-1_24.

[4] C. C. Chu, M. H. Lai and W. S. Feng, "Model-order reductions for MIMO systems using global Krylov subspace methods", *Math. Comput. Simulation* **79** (2008) 1153–1164; doi:10.1016/j.matcom.2007.09.007.

[5] V. Druskin and V. Simoncini, "Adaptive rational Krylov subspaces for large-scale dynamical systems", *Systems Control Lett.* **60** (2011) 546–560; doi:10.1016/j.sysconle.2011.04.013.

[6] M. Frangos and I. M. Jaimoukha, "Adaptive rational interpolation: Arnoldi and Lanczos-like equations", *Eur. J. Control* **14** (2008) 342–354; doi:10.3166/ejc.14.342-354.

[7] K. Gallivan, E. Grimme and P. V. Dooren, "A rational Lanczos algorithm for model reduction", *Numer. Algorithms* **12** (1996) 33–63; doi:10.1007/BF02141740.

[8] E. J. Grimme, D. C. Sorensen and P. V. Dooren, "Model reduction of state space systems via an implicitly restarted Lanczos method", *Numer. Algorithms* **12** (1995) 1–31; doi:10.1007/BF02141739.

[9] S. Gugercin, "An iterative SVD-Krylov based method for model reduction of large-scale dynamical systems", *Linear Algebra Appl.* **428** (2008) 1964–1986; doi:10.1016/j.laa.2007.10.041.

[10] S. Gugercin, D. C. Sorensen and A. C. Antoulas, "A modified low-rank Smith method for large-scale Lyapunov equations", *Numer. Algorithms* **32** (2003) 27–55; doi:10.1023/A:1022205420182.

[11] Y. Hashimoto and T. Nodera, "Inexact shift-invert Arnoldi method for evolution equations", *ANZIAM J.* **58** (2016) E1–E27; doi:10.21914/anziamj.v58i0.10766.

[12]   H. J. Lee, C. C. Chu and W. S. Feng, "An adaptive-order rational Arnoldi method for model-order reductions of linear time-invariant systems", *Linear Algebra Appl.* **415** (2006) 235–261; doi:10.1016/j.laa.2004.10.011.

[13]   K. Mohamed, A. Mehdi and M. Abdelkader, "Arnoldi model reduction for switched linear systems", in: *5th International Conference Modeling, Simulation and Applied Optimization (Hammamet, Tunisia)* (IEEE, Piscataway, NJ, 2013) 1–6; doi:10.1109/ICMSAO.2013.6552568.

[14]   K. Mohamed, A. Mehdi and M. Abdelkader, "A survey of linear invariant time model reduction", *ICIC Express Letters, An International Journal of Research and Surveys* **7** (2013) 909–916; http://www.ijicic.org/el-7(3)b.htm.

[15]   K. Mohamed, A. Mehdi and M. Abdelkader, "Rational Arnoldi & adaptive order rational Arnoldi for switched linear systems", *Neural Parallel Sci. Comput.* **22** (2014) 75–88; http://cat.inist.fr/?aModele=afficheN&cpsidt=28604453.

[16]   K. Mohamed, A. Mehdi and M. Abdelkader, "Iterative dual rational Krylov and iterative SVD-dual rational Krylov model reduction for switched linear systems", in: *Complex system modelling and control through intelligent soft computations*, *Studies in Fuzziness and Soft Computing*, Volume 319 (eds Q. Zhu and A. T. Azar), (Springer International Publishing, Switzerland, 2015) 407–435.

[17]   K. Mohamed, A. Mehdi and M. Abdelkader, "SVD-AORA method for dynamic linear time invariant model order reduction", *IFAC-PapersOnLine* **48**(1) (2015) 695–696; doi:10.1016/j.ifacol.2015.05.175.

[18]   K. Mohamed, A. Mehdi and M. Abdelkader, "Arnoldi model reduction for switched linear systems", *Int. J. Oper. Res.* **27** (2016) 85–112; doi:10.1504/IJOR.2016.078456.