Universität Augsburg

Institut für
Mathematik

Nguyen Thanh Son, Tatjana Stykel

**Solving Parameter-Dependent Lyapunov Equations
Using Reduced Basis Method With Application to
Parametric Model Order Reduction**

SOLVING PARAMETER-DEPENDENT LYAPUNOV EQUATIONS USING REDUCED BASIS METHOD WITH APPLICATION TO PARAMETRIC MODEL ORDER REDUCTION

NGUYEN THANH SON* AND TATJANA STYKEL[†]

Abstract. Our aim is to numerically solve parameter-dependent Lyapunov equations using reduced basis method. Such equations arise in parametric model order reduction. We restrict ourselves to systems that affinely depend on the parameter as our main ingredient is the min-Θ approach. In those cases, we derive various *a posteriori* error estimates. Based on these estimates, Greedy algorithms for constructing reduced basis are formulated. Thanks to the derived results, a novel so-called parametric balanced truncation model reduction method is developed. Numerical examples are presented.

Key words. Parameter-dependent Lyapunov equations, reduced basis method, affine dependence, min-Θ approach, error estimates, Greedy algorithm, model order reduction, parametric balanced truncation.

1. Introduction. In this paper, we consider the following parametric algebraic Lyapunov equations (PALE)

$$A(\mu)X(\mu)E^T(\mu) + E(\mu)X(\mu)A^T(\mu) = -B(\mu)B^T(\mu), \qquad (1.1)$$

where $A(\mu), E(\mu) \in \mathbb{R}^{N \times N}$ and $B(\mu) \in \mathbb{R}^{N \times m}$ with $m \ll N$ are given. The coefficient matrices and the right-hand side depend on parameter μ in a compact domain $\mathcal{D} \subset \mathbb{R}^d$. For the rest of the paper, we assume that $E(\mu)$ is nonsingular and all eigenvalues of a pencil $\lambda E(\mu) - A(\mu)$ have negative real part for all $\mu \in \mathcal{D}$. With these assumptions, equation (1.1) has a unique symmetric positive semidefinite solution $X(\mu)$ for all $\mu \in \mathcal{D}$, see, e.g., [18]. Solving Lyapunov equations is of great importance in many control problems including stability analysis, stabilization problem, model reduction by balanced truncation and optimal control [2, 10]. The PALE (1.1) arises naturally in parametric model reduction [4] or in the design of low gain feedback [34].

Because of their role in the control theory, various works have been devoted to the numerical solution of Lyapunov equations. First of all, it is worth to mention the direct methods by Bartels and Stewart [3] and by Hammarling [13]. To avoid expensive computations in the direct methods, iterative ones such as the sign function method [6], the alternating directions implicit method [20, 33], and Krylov subspace methods [16, 25] have been developed. Exploiting the fact that the right-hand sides of Lyapunov equations in most applications have low rank, low-rank versions of the mentioned iterative methods have been formulated, [9, 19, 24, 26], just to name a few, see also [5, 27] for the recent surveys on the state-of-the-art algorithms.

Although a lot of attention has been paid to Lyapunov equations, only very few publications dedicated to solving PALEs can be found. To our knowledge, the method proposed in [17] is the only released printed work on this subject. Nevertheless, the purpose of this work, computing solutions for many different parameter samples, is not our goal. Unlike [17], we would like to compute the solution $X(\mu)$ for any $\mu \in \mathcal{D}$.

For dealing with parameter-dependent problems, reduced basis method [23] is an effective tool. This method was initially proposed for coercive elliptic partial differ-

---

* Department of Mathematics and Informatics, Thai Nguyen University of Sciences, 23000 Thai Nguyen, Vietnam (`ntson@tnus.edu.vn`).

† Institute of Mathematics, University of Augsburg, Universitätsstr. 14, D-86159 Augsburg, Germany (`stykel@math.uni-augsburg.de`).

ential equations and then extended to non-coercive equations [32], Burgers equations [31], and Navier-Stokes equations [8]. Applying this method to PALEs is most probably the first time performed in the present paper, while an extension to parametric Riccati equation can be found in a very recent paper [12].

To use the reduced basis method, first we have to convert the Lyapunov equation to a linear system using a Kronecker product. The size of the resulting linear system is N^2, where N is itself already large. This fact results in very expensive computations as well as huge storage requirement. The key point to avoid these difficulties is to keep all computations within matrices and vectors of dimension N. The norm used to measure errors must also be carefully chosen in order to make all computations feasible. Similar to [23], an *a posteriori* error estimate will be constructed, based on which a Greedy algorithm is utilized for determining the reduced basis.

To this end, the rest of the paper is organized as follows. In Section 2, we repeat how the Lyapunov equation can be converted to a linear equation by Kronecker product. We also provide some formulae that allow us to replace N^2-sized operations with N-sized operations. Section 3 introduces a reduced basis method for parameter-dependent linear systems. Important components of this method are the Greedy algorithm and *a posteriori* error estimates which will be discussed in Sections 3.1 and 3.2, respectively. We also explain how to efficiently compute the residual norm, which is the main ingredient for the error estimate in this section. In Section 4, we present an extension of the reduced basis framework to parametric symmetric Lyapunov equations. Nonsymmetric case is treated in Section 5. Based on these results, we develop in Section 6 a new approach for parametric model order reduction. Two numerical examples will be presented in Section 7. Finally, in Section 8, we conclude the issue as well as pose some problems for the future work.

Throughout this paper, we will use bold font for matrices and vectors of dimension N^2 and the standard font for that of dimension N. Given a matrix $A \in \mathbb{R}^{N \times N}$ and a vector $\mathbf{v} \in \mathbb{R}^{N^2}$, $\text{vec}(A)$ will denote the column vector in $\mathbb{R}^{N^2}$ generated from A by stacking all columns of A, and $\text{mat}(\mathbf{v})$ is the $N \times N$-matrix such that $\text{vec}(\text{mat}(\mathbf{v})) = \mathbf{v}$. The trace of the matrix A is denoted by $\text{trace}(A)$, $\ker(A) = \{v \in \mathbb{R}^N : Av = 0\}$, $\text{span}(A)$ is the subspace spanned by columns of A, A^T stands for the transpose of A, and $A > 0$ ($A \geq 0$) means that A is positive definite (semidefinite), i.e., $v^T A v > 0$ ($v^T A v \geq 0$) for all $v \neq 0$. The smallest and largest singular values of A are denoted by $\sigma_{\min}(A)$ and $\sigma_{\max}(A)$, respectively, whereas the smallest and largest eigenvalues of symmetric A are denoted by $\lambda_{\min}(A)$ and $\lambda_{\max}(A)$, respectively. Similarly, $\lambda_{\min}(E, A)$ and $\lambda_{\max}(E, A)$ will denote the smallest and the largest eigenvalue of the matrix pencil $\lambda E - A$, respectively. We denote by $\|v\| = \sqrt{(v^T v)}$ the Euclidean vector norm of $v \in \mathbb{R}^N$, by $\|A\|_2 = \sigma_{\max}(A)$ the spectral norm of $A \in \mathbb{R}^{N \times M}$, and by $\text{cond}(A) = \|A\|_2 \|A^{-1}\|_2$ the condition number of invertible A. Furthermore, the inner product of two matrices $A, B \in \mathbb{R}^{N \times M}$ is defined as $\langle A, B \rangle = \text{trace}(B^T A)$, and $\|A\|_F = \sqrt{\langle A, A \rangle}$ is the Frobenius matrix norm.

2. Lyapunov equations and linear systems. We begin this section with a definition of the Kronecker product. For $A = [a_{ij}] \in \mathbb{R}^{k \times l}$ and $B \in \mathbb{R}^{m \times n}$, the Kronecker product $A \otimes B$ is the $km \times ln$ block matrix

$$A \otimes B = \begin{bmatrix} a_{11}B & \cdots & a_{1l}B \\ \vdots & \ddots & \vdots \\ a_{k1}B & \cdots & a_{kl}B \end{bmatrix}.$$

The following lemmas provide some useful properties of the Kronecker product, vec- and mat-operators.

LEMMA 2.1. *Let* $E, A, X, Y \in \mathbb{R}^{N \times N}$, $\mathbf{x} = \mathrm{vec}(X)$ *and* $\mathbf{y} = \mathrm{vec}(Y)$. *Then*

1. $\mathbf{y}^T \mathbf{x} = \langle X, Y \rangle$,
2. $\mathrm{vec}(AXE^T + EXA^T) = \mathbf{L}\mathbf{x}$ *with* $\mathbf{L} = E \otimes A + A \otimes E$.

Proof. See [15]. □

LEMMA 2.2. *Let* $\mathbf{V} = [\mathbf{v}_1, \dots, \mathbf{v}_k] \in \mathbb{R}^{N^2 \times k}$, $V_j = \mathrm{mat}(\mathbf{v}_j)$ *for* $j = 1, \dots, k$, $\mathbf{U} = [\mathbf{u}_1, \dots, \mathbf{u}_l] \in \mathbb{R}^{N^2 \times l}$, $U_j = \mathrm{mat}(\mathbf{u}_j)$ *for* $j = 1, \dots, l$, *and* $\mathbf{L} = E \otimes A + A \otimes E$.

1. $\mathbf{L}\mathbf{V} = [\, \mathrm{vec}(AV_1E^T + EV_1A^T), \dots, \mathrm{vec}(AV_kE^T + EV_kA^T)\,]$.
2. *The entries of* $\mathbf{U}^T\mathbf{L}\mathbf{V} \in \mathbb{R}^{l \times k}$ *are given by* $(\mathbf{U}^T\mathbf{L}\mathbf{V})_{ij} = \langle AV_jE^T + EV_jA^T, U_i \rangle$.
3. *Let* $\mathbf{L}_s = E_s \otimes A_s + A_s \otimes E_s$, $s = 1, 2$. *Then the entries of* $\mathbf{U}^T\mathbf{L}_1^T\mathbf{L}_2\mathbf{V} \in \mathbb{R}^{l \times k}$ *are given by* $(\mathbf{U}^T\mathbf{L}_1^T\mathbf{L}_2\mathbf{V})_{ij} = \langle A_2V_jE_2^T + E_2V_jA_2^T, A_1U_iE_1^T + E_1U_iA_1^T \rangle$.

Proof. The equalities can be easily verified by calculation. □

Let $\mathbb{S}$ denote a space of $N \times N$ symmetric matrices. Consider a Lyapunov operator $\mathcal{L}_\mu : \mathbb{S} \to \mathbb{S}$ given by

$$\mathcal{L}_\mu(X) = -A(\mu)XE^T(\mu) - E(\mu)XA^T(\mu).$$

Then the PALE (1.1) takes the form $\mathcal{L}_\mu(X(\mu)) = B(\mu)B^T(\mu)$. Using Lemma 2.1, this equation can also be written as a linear system

$$\mathbf{L}(\mu)\mathbf{x}(\mu) = \mathbf{b}(\mu), \tag{2.1}$$

where $\mathbf{x}(\mu) = \mathrm{vec}(X(\mu))$, $\mathbf{b}(\mu) = \mathrm{vec}(B(\mu)B^T(\mu))$, and

$$\mathbf{L}(\mu) = -E(\mu) \otimes A(\mu) - A(\mu) \otimes E(\mu) \tag{2.2}$$

is the matrix representation of the linear Lyapunov operator $\mathcal{L}_\mu$. The following theorem establishes some properties of $\mathcal{L}_\mu$ and $\mathbf{L}(\mu)$.

THEOREM 2.3. *Let* $-A(\mu)$ *and* $E(\mu)$ *be symmetric, positive definite for all* $\mu \in \mathcal{D}$.

1. *The matrix* $\mathbf{L}(\mu)$ *in* (2.2) *is symmetric and positive definite for all* $\mu \in \mathcal{D}$ *and its smallest and largest eigenvalues are bounded as*

$$\lambda_{\min}(\mathbf{L}(\mu)) \geq 2\,\lambda_{\min}(-A(\mu))\,\lambda_{\min}(E(\mu)), \tag{2.3}$$
$$\lambda_{\max}(\mathbf{L}(\mu)) \leq 2\,\lambda_{\max}(-A(\mu))\,\lambda_{\max}(E(\mu)) \tag{2.4}$$

for all $\mu \in \mathcal{D}$.

2. *The Lyapunov operator* $\mathcal{L}_\mu$ *is uniformly coercive, i.e., it holds*

$$\alpha(\mu) := \inf_{V \in \mathbb{R}^{N \times N}\setminus\{0\}} \frac{\langle \mathcal{L}_\mu(V), V \rangle}{\|V\|_F^2} > 0 \tag{2.5}$$

for all $\mu \in \mathcal{D}$.

3. *The Lyapunov operator* $\mathcal{L}_\mu$ *is uniformly continuous, i.e., it holds*

$$\gamma(\mu) := \sup_{W,V \in \mathbb{R}^{N \times N}\setminus\{0\}} \frac{\langle \mathcal{L}_\mu(V), W \rangle}{\|W\|_F \|V\|_F} < \infty \tag{2.6}$$

for all $\mu \in \mathcal{D}$.

Proof. 1. Since $-A(\mu)$ and $E(\mu)$ are both symmetric, the matrix $\mathbf{L}(\mu)$ in (2.2) is also symmetric for all $\mu \in \mathcal{D}$. Using Weyl's theorem [14, Theorem 4.3.1] and the

multiplicativity property of eigenvalues of the Kronecker product [15, Theorem 4.2.12], we have

$$\begin{aligned}\lambda_{\min}(\mathbf{L}(\mu)) &\geq \lambda_{\min}(-E(\mu) \otimes A(\mu)) + \lambda_{\min}(-A(\mu) \otimes E(\mu)) \\ &= 2\lambda_{\min}(-A(\mu))\lambda_{\min}(E(\mu)) > 0.\end{aligned}$$

The last inequality follows from the positive definiteness of $-A(\mu)$ and $E(\mu)$. Thus, the bound (2.3) holds and $\mathbf{L}(\mu)$ is positive definite. The bound (2.4) can be proved analogously.

2. We obtain from Lemma 2.1 and the Courant-Fischer theorem [14, Theorem 4.2.6] that

$$\alpha(\mu) = \inf_{V \in \mathbb{R}^{N \times N} \setminus \{0\}} \frac{\langle \mathcal{L}_\mu(V), V \rangle}{\|V\|_F^2} = \inf_{\mathbf{v} \in \mathbb{R}^{N^2} \setminus \{0\}} \frac{\mathbf{v}^T \mathbf{L}(\mu)\mathbf{v}}{\|\mathbf{v}\|^2} = \lambda_{\min}(\mathbf{L}(\mu)) > 0$$

for all $\mu \in \mathcal{D}$, and, hence, $\mathcal{L}_\mu$ is uniformly coercive.

3. Using again Lemma 2.1 and the Courant-Fischer theorem, we have

$$\gamma(\mu) = \sup_{W, V \in \mathbb{R}^{N \times N} \setminus \{0\}} \frac{\langle \mathcal{L}_\mu(V), W \rangle}{\|W\|_F \|V\|_F} = \sup_{\mathbf{w}, \mathbf{v} \in \mathbb{R}^{N^2} \setminus \{0\}} \frac{\mathbf{w}^T \mathbf{L}(\mu)\mathbf{v}}{\|\mathbf{w}\| \|\mathbf{v}\|} = \lambda_{\max}(\mathbf{L}(\mu)) < \infty$$

for all $\mu \in \mathcal{D}$. Thus, $\mathcal{L}_\mu$ is uniformly continuous. $\square$

The parameter-dependent quantities $\alpha(\mu)$ and $\gamma(\mu)$ are called *coercivity constant* and *continuity constant* of the Lyapunov operator $\mathcal{L}_\mu$, respectively. From the proof of Theorem 2.3 we obtain the important relations

$$\alpha(\mu) = \lambda_{\min}(\mathbf{L}(\mu)), \qquad \gamma(\mu) = \lambda_{\max}(\mathbf{L}(\mu)), \tag{2.7}$$

which together with the estimates (2.3) and (2.4) will be very useful in the following.

3. Reduced basis method. In this section, we consider the application of the reduced basis method to the linear system (2.1) with $\mathbf{L}(\mu)$ as in (2.2), where $-A(\mu)$ and $E(\mu)$ are assumed to be symmetric and positive definite. This method consists of the following steps. For selected parameters $\mu_1, \ldots, \mu_k \in \mathcal{D}$, we construct first a reduced basis matrix $\mathbf{V}_k = [\mathbf{x}(\mu_1), \ldots, \mathbf{x}(\mu_k)]$, where $\mathbf{x}(\mu_j)$ is the solution of system (2.1) at $\mu = \mu_j$ for $j = 1, \ldots, k$. Then, for any $\mu \in \mathcal{D}$, an approximate solution can be computed by Galerkin projection $\mathbf{x}(\mu) \approx \mathbf{V}_k \hat{\mathbf{x}}(\mu)$, where $\hat{\mathbf{x}}(\mu)$ solves the reduced linear system

$$\hat{\mathbf{L}}(\mu)\hat{\mathbf{x}}(\mu) = \hat{\mathbf{b}}(\mu) \tag{3.1}$$

with $\hat{\mathbf{L}}(\mu) = \mathbf{V}_k^T \mathbf{L}(\mu)\mathbf{V}_k$ and $\hat{\mathbf{b}}(\mu) = \mathbf{V}_k^T \mathbf{b}(\mu)$. This seemingly simple procedure raises several issues: optimal choice of the sample parameters $\mu_1, \ldots, \mu_k$ providing good reduced basis subspaces that guarantee a rapid convergence of the reduced basis approximation $\mathbf{V}_k \hat{\mathbf{x}}(\mu)$ to $\mathbf{x}(\mu)$ over the entire parameter domain $\mathcal{D}$; rigorous error estimates for the approximate solution and efficient computations.

3.1. Greedy algorithm. The key point to the success of the reduced basis method is the construction of an appropriate basis. It should be done in such a way that the error of the approximation is smaller than a given tolerance while the dimension of the reduced basis is kept as small as possible. One way to do this is to employ a Greedy algorithm, e.g., [23], which successively determines snapshots depending on the error magnitude.

Suppose that we have already computed a basis matrix $\mathbf{V}_k = [\,\mathbf{x}(\mu_1), \ldots, \mathbf{x}(\mu_k)\,]$. Then the error and the residual of the approximate solution $\mathbf{V}_k \hat{\mathbf{x}}(\mu)$ are given by

$$\begin{aligned}
\mathbf{e}_k(\mu) &= \mathbf{x}(\mu) - \mathbf{V}_k \hat{\mathbf{x}}(\mu), \\
\mathbf{r}_k(\mu) &= \mathbf{b}(\mu) - \mathbf{L}(\mu)\mathbf{V}_k \hat{\mathbf{x}}(\mu),
\end{aligned}$$

respectively. They satisfy the equation

$$\mathbf{L}(\mu)\mathbf{e}_k(\mu) = \mathbf{r}_k(\mu), \tag{3.2}$$

which immediately implies the error estimate

$$\|\mathbf{e}_k(\mu)\| = \|\mathbf{L}^{-1}(\mu)\mathbf{r}_k(\mu)\| \leq \frac{\|\mathbf{r}_k(\mu)\|}{\alpha(\mu)} \leq \frac{\|\mathbf{r}_k(\mu)\|}{\alpha_{\mathrm{LB}}(\mu)} =: \Delta_k(\mu). \tag{3.3}$$

Here, $\alpha_{\mathrm{LB}}(\mu)$ is a positive lower bound for the coercivity constant $\alpha(\mu)$, and $\Delta_k(\mu)$ is the resulting error estimator. To enlarge the basis in order to somewhat reduce the approximation error, we will find the next value μ_{k+1} such that $\Delta_k(\mu_{k+1})$ is the largest in $\mathcal{D}$. Of course, we cannot pursue the search on $\mathcal{D}$, which is a continuous set. Instead, one usually does it on a discrete subset of $\mathcal{D}$. To ensure that we do not miss any good candidate, this training set, denoted by $\mathcal{D}_{\mathrm{train}}$, should be rather dense and, therefore, large. In practice, we choose $\mathcal{D}_{\mathrm{train}}$ first and pick μ_1 arbitrarily in $\mathcal{D}_{\mathrm{train}}$. We also need to specify a tolerance *tol* for the approximation. The Greedy algorithm is then given as follows.

Algorithm 1 Greedy algorithm for linear systems

Input: tolerance $tol_{\mathrm{rb}} < 1$, training set $\mathcal{D}_{\mathrm{train}}$, initial parameter $\mu_1 \in \mathcal{D}_{\mathrm{train}}$.
Output: a basis matrix $\mathbf{V}_k$.
1: Solve $\mathbf{L}(\mu_1)\mathbf{x}(\mu_1) = \mathbf{b}(\mu_1)$.
2: $\Delta_1^{\max} = 1$, $\mathcal{M}_1 = \{\mu_1\}$ and $\mathbf{V}_1 = \mathbf{x}(\mu_1)$, $k = 2$.
3: **while** $\Delta_{k-1}^{\max} \geq tol_{\mathrm{rb}}$ **do**
4: $\quad \mu_k = \underset{\mu \in \mathcal{D}_{\mathrm{train}} \backslash \mathcal{M}_{k-1}}{\arg\max} \; \Delta_{k-1}(\mu)$
5: $\quad \Delta_k^{\max} = \Delta_{k-1}(\mu_k)$
6: $\quad \mathcal{M}_k = \mathcal{M}_{k-1} \cup \{\mu_k\}$
7: $\quad$ solve $\mathbf{L}(\mu_k)\mathbf{x}(\mu_k) = \mathbf{b}(\mu_k)$
8: $\quad \mathbf{V}_k = [\,\mathbf{V}_{k-1}, \; \mathbf{x}(\mu_k)\,]$
9: $\quad k \leftarrow k + 1$
10: **end while**

For the success of this algorithm, an efficient, sharp and rigorous error estimate is required. This issue will be addressed in the next subsection.

3.2. Error estimation. First, we impose some further restrictions on the problems treated in this paper. To wit, we assume that the matrices $A(\mu)$, $E(\mu)$ and $B(\mu)$ are *affine* in the parameter μ, i.e.,

$$\textbf{(A1)} \qquad A(\mu) = \sum_{j=1}^{n_A} \theta_j^A(\mu) A_j, \quad E(\mu) = \sum_{j=1}^{n_E} \theta_j^E(\mu) E_j, \quad B(\mu) = \sum_{j=1}^{n_B} \theta_j^B(\mu) B_j,$$

where A_j, E_j and B_j are independent of μ, n_A, n_E and n_B are very small compared to N, $\theta_j^A(\mu)$, $\theta_j^E(\mu)$ and $\theta_j^B(\mu)$ are continuous in $\mathcal{D}$ and their evaluations at each $\mu \in \mathcal{D}$ are cheap. This assumption permits us to decompose the computation of the

solution of the linear system (2.1) into an offline stage (computationally expensive), in which the reduced basis matrix $\mathbf{V}_k$ is constructed and all parameter-independent matrices are computed and stored, and an online stage (computationally inexpensive), in which the reduced system (3.1) is solved for any $\mu \in \mathcal{D}$ to get the approximate solution $\mathbf{x}(\mu) \approx \mathbf{V}_k \hat{\mathbf{x}}(\mu)$.

Furthermore, we require that the matrices $-A(\mu)$ and $E(\mu)$ are *parametrically coercive*, i.e.,

(**A2**) $E_j = E_j^T \geq 0$ and $\theta_j^E(\mu) > 0$ for all $\mu \in \mathcal{D}$ and $j = 1, \ldots, n_E$,

(**A3**) $-A_j = -A_j^T \geq 0$ and $\theta_j^A(\mu) > 0$ for all $\mu \in \mathcal{D}$ and $j = 1, \ldots, n_A$.

The following lemma shows that under assumptions (A1)–(A3) the linear system (2.1) maintains the affine dependence and parametric coercivity.

LEMMA 3.1. *Let $A(\mu)$, $E(\mu)$ and $B(\mu)$ satisfy* (A1)–(A3). *Then the matrix $\mathbf{L}(\mu)$ and the vector $\mathbf{b}(\mu)$ in the linear system* (2.1) *are affine in the parameter μ. Moreover, $\mathbf{L}(\mu)$ is parametrically coercive.*

Proof. The proof of affine dependence is straightforward and based on the properties of the Kronecker product [15]. Nevertheless, since we will need the explicit form of the affine dependence later on, it is briefly written down here. One can easily verify that the coefficient matrix and the right-hand side in (2.1) take the form

$$\mathbf{L}(\mu) = \sum_{i=1}^{n_E} \sum_{j=1}^{n_A} \theta_{ij}^{\mathbf{L}}(\mu) \mathbf{L}_{ij}, \qquad \mathbf{b}(\mu) = \sum_{i=1}^{n_B} \sum_{j=1}^{n_B} \theta_{ij}^{\mathbf{b}}(\mu) \mathbf{b}_{ij}, \qquad (3.4)$$

where

$$\begin{aligned} \theta_{ij}^{\mathbf{L}}(\mu) &= \theta_i^E(\mu)\theta_j^A(\mu) > 0, & \mathbf{L}_{ij} &= -E_i \otimes A_j - A_j \otimes E_i, \\ \theta_{ij}^{\mathbf{b}}(\mu) &= \theta_i^B(\mu)\theta_j^B(\mu), & \mathbf{b}_{ij} &= \mathrm{vec}(B_i B_j^T). \end{aligned} \qquad (3.5)$$

Obviously, $\mathbf{L}_{ij}$ is symmetric and positive semidefinite. This fact together with the positiveness of $\theta_{ij}^{\mathbf{L}}$ completes the proof. $\square$

REMARK 3.1. *In general, the parametric coercivity of $\mathbf{L}(\mu)$ does not imply its coercivity, i.e., the positivity of $\alpha(\mu) = \lambda_{\min}(\mathbf{L}(\mu))$. If, however, in addition to the parametric coercivity, we assume that*

(**A4**)   *there exist at least one pair (A_j, E_i) such that $-A_j > 0$ and $E_i > 0$,*

then the coercivity is satisfied. Assumption (A4) *can also be replaced by a condition that*

$$\bigcap_{j=1,\ldots,n_A} \ker(A_j) = \emptyset \quad and \quad \bigcap_{j=1,\ldots,n_E} \ker(E_j) = \emptyset.$$

To derive *a posteriori* error estimates, we first have to find positive lower bounds for the coercivity constant $\alpha(\mu)$. We will mainly employ a min-Θ approach [23] based on the affine decomposition (A1). To this end, for a fixed value $\bar{\mu} \in \mathcal{D}$, we define the following functions

$$\theta_{\min}^{\mathbf{L},\bar{\mu}}(\mu) = \min_{\substack{i=1,\ldots,n_E \\ j=1,\ldots,n_A}} \frac{\theta_{ij}^{\mathbf{L}}(\mu)}{\theta_{ij}^{\mathbf{L}}(\bar{\mu})}, \qquad \theta_{\max}^{\mathbf{L},\bar{\mu}}(\mu) = \max_{\substack{i=1,\ldots,n_E \\ j=1,\ldots,n_A}} \frac{\theta_{ij}^{\mathbf{L}}(\mu)}{\theta_{ij}^{\mathbf{L}}(\bar{\mu})}, \qquad \theta^{\mathbf{L},\bar{\mu}}(\mu) = \frac{\theta_{\max}^{\mathbf{L},\bar{\mu}}(\mu)}{\theta_{\min}^{\mathbf{L},\bar{\mu}}(\mu)}.$$

For convenience, an upper bound for the continuity constant $\gamma(\mu)$ is also included in the following lemma.

LEMMA 3.2. *Let $E(\mu)$ and $A(\mu)$ satisfy* (A1)–(A4) *and let $\bar{\mu}, \bar{\mu}_1, \bar{\mu}_2 \in \mathcal{D}$.*

1. *For all $\mu \in \mathcal{D}$, the coercivity constant $\alpha(\mu)$ in (2.5) is bounded from below as*

$$\alpha(\mu) \geq \alpha_{\mathrm{LB}}(\mu) := \max\left(\alpha_{\mathrm{LB}}^{\mathbf{L},\bar{\mu}}(\mu), \ \alpha_{\mathrm{LB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu), \ \alpha_{\mathrm{LB}}^{\mathbf{L}}(\mu)\right) > 0, \qquad (3.6)$$

where

$$\begin{aligned}
\alpha_{\mathrm{LB}}^{\mathbf{L},\bar{\mu}}(\mu) &= 2\,\theta_{\min}^{\mathbf{L},\bar{\mu}}(\mu)\,\lambda_{\min}\big(-A(\bar{\mu})\big)\,\lambda_{\min}\big(E(\bar{\mu})\big), \\
\alpha_{\mathrm{LB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu) &= 2\,\theta_{\min}^{A,\bar{\mu}_1}(\mu)\,\theta_{\min}^{E,\bar{\mu}_2}(\mu)\,\lambda_{\min}\big(-A(\bar{\mu}_1)\big)\,\lambda_{\min}\big(E(\bar{\mu}_2)\big), \\
\alpha_{\mathrm{LB}}^{\mathbf{L}}(\mu) &= 2\sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\theta_{ij}^{\mathbf{L}}(\mu)\,\lambda_{\min}(-A_j)\,\lambda_{\min}(E_i).
\end{aligned}$$

2. *For all $\mu \in \mathcal{D}$, the continuity constant $\gamma(\mu)$ in (2.6) is bounded from above as*

$$\gamma(\mu) \leq \gamma_{\mathrm{UB}}(\mu) := \min\left(\gamma_{\mathrm{UB}}^{\mathbf{L},\bar{\mu}}(\mu), \ \gamma_{\mathrm{UB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu), \ \gamma_{\mathrm{UB}}^{\mathbf{L}}(\mu)\right), \qquad (3.7)$$

where

$$\begin{aligned}
\gamma_{\mathrm{UB}}^{\mathbf{L},\bar{\mu}}(\mu) &= 2\,\theta_{\max}^{\mathbf{L},\bar{\mu}}(\mu)\,\lambda_{\max}\big(-A(\bar{\mu})\big)\,\lambda_{\max}\big(E(\bar{\mu})\big), \\
\gamma_{\mathrm{UB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu) &= 2\,\theta_{\max}^{A,\bar{\mu}_1}(\mu)\,\theta_{\max}^{E,\bar{\mu}_2}(\mu)\,\lambda_{\max}\big(-A(\bar{\mu}_1)\big)\,\lambda_{\max}\big(E(\bar{\mu}_2)\big), \\
\gamma_{\mathrm{UB}}^{\mathbf{L}}(\mu) &= 2\sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\theta_{ij}^{\mathbf{L}}(\mu)\,\lambda_{\max}(-A_j)\,\lambda_{\max}(E_i).
\end{aligned}$$

Proof. 1. From the Courant-Fischer theorem and the estimate (2.3) we obtain that

$$\begin{aligned}
\lambda_{\min}\big(\mathbf{L}(\mu)\big) &= \min_{\|v\|=1}\sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\theta_{ij}^{\mathbf{L}}(\mu)\,\mathbf{v}^T\mathbf{L}_{ij}\mathbf{v} = \min_{\|v\|=1}\sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\frac{\theta_{ij}^{\mathbf{L}}(\mu)}{\theta_{ij}^{\mathbf{L}}(\bar{\mu})}\,\theta_{ij}^{\mathbf{L}}(\bar{\mu})\,\mathbf{v}^T\mathbf{L}_{ij}\mathbf{v} \\
&\geq \theta_{\min}^{\mathbf{L},\bar{\mu}}(\mu)\,\lambda_{\min}\big(\mathbf{L}(\bar{\mu})\big) \geq 2\,\theta_{\min}^{\mathbf{L},\bar{\mu}}(\mu)\,\lambda_{\min}\big(-A(\bar{\mu})\big)\,\lambda_{\min}\big(E(\bar{\mu})\big).
\end{aligned}$$

On the other hand, using (2.3) again and applying the min-Θ approach to the matrices $A(\mu)$ and $E(\mu)$, we have

$$\begin{aligned}
\lambda_{\min}\big(\mathbf{L}(\mu)\big) &\geq 2\,\lambda_{\min}\big(-A(\mu)\big)\,\lambda_{\min}\big(E(\mu)\big) \\
&\geq 2\,\theta_{\min}^{A,\bar{\mu}_1}(\mu)\,\theta_{\min}^{E,\bar{\mu}_2}(\mu)\lambda_{\min}\big(-A(\bar{\mu}_1)\big)\,\lambda_{\min}\big(E(\bar{\mu}_2)\big).
\end{aligned}$$

Finally, Weyl's theorem [14, Theorem 4.3.1] implies that

$$\lambda_{\min}\big(\mathbf{L}(\mu)\big) \geq \sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\theta_{ij}^{\mathbf{L}}(\mu)\,\lambda_{\min}(\mathbf{L}_{ij}) \geq 2\sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\theta_{ij}^{\mathbf{L}}(\mu)\,\lambda_{\min}(-A_j)\,\lambda_{\min}(E_i).$$

Thus, the bound (3.6) holds.

2. The bound (3.7) can be proved similarly using (2.4). $\square$

REMARK 3.2. *One can observe that if $E(\mu) = E$ is constant and $\bar{\mu} = \bar{\mu}_1 = \bar{\mu}_2$ then $\alpha_{\mathrm{LB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu)$ and $\gamma_{\mathrm{UB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu)$ become $\alpha_{\mathrm{LB}}^{\mathbf{L},\bar{\mu}}(\mu)$ and $\gamma_{\mathrm{UB}}^{\mathbf{L},\bar{\mu}}(\mu)$, respectively. Note also that the quantities $\alpha_{\mathrm{UB}}^{\mathbf{L}}(\mu)$ and $\gamma_{\mathrm{UB}}^{\mathbf{L}}(\mu)$ in (3.6) and (3.7), respectively, have no relation to min-Θ approach. We put them into the bounds in order to possibly have tighter bounds.*

The following theorem provides an *a posteriori* error estimate for the reduced basis solution $\mathbf{V}_k\hat{\mathbf{x}}(\mu)$.

THEOREM 3.3. *Let assumptions* (A1)–(A4) *be fulfilled and let* $\alpha_{\mathrm{LB}}(\mu)$ *and* $\gamma_{\mathrm{UB}}(\mu)$ *be as in* (3.6) *and* (3.7), *respectively. Then the error* $\mathbf{e}_k(\mu) = \mathbf{x}(\mu) - \mathbf{V}_k\hat{\mathbf{x}}(\mu)$ *satisfies the bounds*

$$\|\mathbf{e}_k(\mu)\| \leq \Delta_k(\mu) \leq \frac{\gamma_{\mathrm{UB}}(\mu)}{\alpha_{\mathrm{LB}}(\mu)} \|\mathbf{e}_k(\mu)\|, \tag{3.8}$$

where the error estimator $\Delta_k(\mu)$ *is given by*

$$\Delta_k(\mu) = \frac{\|\mathbf{r}_k(\mu)\|}{\alpha_{\mathrm{LB}}(\mu)}. \tag{3.9}$$

Proof. The error estimate $\|\mathbf{e}_k(\mu)\| \leq \Delta_k(\mu)$ immediately follows from (3.3) and (3.6). Furthermore, using (2.7), (3.2) and (3.7), we have

$$\Delta_k(\mu) = \frac{\|\mathbf{r}_k(\mu)\|}{\alpha_{\mathrm{LB}}(\mu)} \leq \frac{\gamma_{\mathrm{UB}}(\mu)}{\alpha_{\mathrm{LB}}(\mu)} \|\mathbf{e}_k(\mu)\|.$$

This completes the proof. $\square$

The effectivity of the error estimator $\Delta_k(\mu)$ is measured by the quantity

$$\eta_k(\mu) = \frac{\Delta_k(\mu)}{\|\mathbf{e}_k(\mu)\|}.$$

The error estimate is tight if $\eta_k(\mu)$ is very close to 1. It follows from (3.8) that

$$1 \leq \eta_k(\mu) \leq \frac{\gamma_{\mathrm{UB}}(\mu)}{\alpha_{\mathrm{LB}}(\mu)}. \tag{3.10}$$

Therefore, to sharpen the error estimate, one could choose the parameter $\bar{\mu} \in \mathcal{D}$ such that the quotient $\gamma_{\mathrm{UB}}(\mu)/\alpha_{\mathrm{LB}}(\mu)$ is as small as possible. The solution to this optimization problem, however, goes beyond the purpose of this paper.

Recall that in the Greedy algorithm, finding the maximizer of the error estimator $\Delta_k(\mu)$ on a large discrete set $\mathcal{D}_{\mathrm{train}}$ is required. This involves the repeated computation of the residual norms $\|\mathbf{r}_k(\mu)\|$ in the vector space of huge dimension N^2 for all $\mu \in \mathcal{D}_{\mathrm{train}}$. Only a good strategy for computing the residuals can avoid the unfeasibly expensive computations. This can be done thanks to the choice of norm, the affine dependence in $\mathbf{L}(\mu)$ and $\mathbf{b}(\mu)$ and a suitable arrangement.

To simplify the notation, we re-index the expressions for the coefficient matrix and the right-hand side in (3.4) by replacing the two-index system by the one-index system, say

$$\mathbf{L}(\mu) = \sum_{p=1}^{n_A n_E} \theta_p^{\mathbf{L}}(\mu)\mathbf{L}_p, \qquad \mathbf{b}(\mu) = \sum_{p=1}^{n_B^2} \theta_p^{\mathbf{b}}(\mu)\mathbf{b}_p, \tag{3.11}$$

where

$$\begin{aligned}
\mathbf{L}_p &= \mathbf{L}_{ij}, & \theta_p^{\mathbf{L}}(\mu) &= \theta_{ij}^{\mathbf{L}}(\mu) & \text{for} \quad p &= (i-1)n_A + j, \\
\mathbf{b}_p &= \mathbf{b}_{ij}, & \theta_p^{\mathbf{b}}(\mu) &= \theta_{ij}^{\mathbf{b}}(\mu) & \text{for} \quad p &= (i-1)n_B + j.
\end{aligned} \tag{3.12}$$

Then the residual norm can be represented as

$$\|\mathbf{r}_k(\mu)\|^2 = \big(\mathbf{b}(\mu) - \mathbf{L}(\mu)\mathbf{V}_k\hat{\mathbf{x}}(\mu)\big)^T \big(\mathbf{b}(\mu) - \mathbf{L}(\mu)\mathbf{V}_k\hat{\mathbf{x}}(\mu)\big)$$

$$= \sum_{p,q=1}^{n_B^2} \theta_p^{\mathbf{b}}(\mu)\theta_q^{\mathbf{b}}(\mu)\mathbf{b}_p^T\mathbf{b}_q - 2\sum_{p=1}^{n_B^2}\sum_{q=1}^{n_A n_E} \theta_p^{\mathbf{b}}(\mu)\theta_q^{\mathbf{L}}(\mu)\mathbf{b}_p^T\mathbf{L}_q\mathbf{V}_k\hat{\mathbf{x}}(\mu) \qquad (3.13)$$

$$+ \sum_{p,q=1}^{n_A n_E} \theta_p^{\mathbf{L}}(\mu)\theta_q^{\mathbf{L}}(\mu)\hat{\mathbf{x}}^T(\mu)\mathbf{V}_k^T\mathbf{L}_p^T\mathbf{L}_q\mathbf{V}_k\hat{\mathbf{x}}(\mu),$$

where $\hat{\mathbf{x}}(\mu)$ is the solution of the reduced linear system (3.1). Note that the parameter-dependent coefficient matrix and the right-hand side in (3.1) also admit the affine decomposition

$$\hat{\mathbf{L}}(\mu) = \sum_{p=1}^{n_A n_E} \theta_p^{\mathbf{L}}(\mu)\mathbf{V}_k^T\mathbf{L}_p\mathbf{V}_k, \qquad \hat{\mathbf{b}}(\mu) = \sum_{p=1}^{n_B^2} \theta_p^{\mathbf{b}}(\mu)\mathbf{V}_k^T\mathbf{b}_p. \qquad (3.14)$$

One can easily realize that several matrix-matrix and matrix-vector products in (3.13) and (3.14), namely $\mathbf{b}_p^T\mathbf{b}_q$, $\mathbf{b}_p^T\mathbf{L}_q\mathbf{V}_k$, $\mathbf{V}_k^T\mathbf{L}_p^T\mathbf{L}_q\mathbf{V}_k$, $\mathbf{V}_k^T\mathbf{L}_p\mathbf{V}_k$ and $\mathbf{V}_k^T\mathbf{b}_p$, are independent of μ. They are expensive to compute but cheap to store. Another worthwhile advantage is that all these parameter-independent quantities can be computed and stored hierarchically with respect to k when running the Greedy algorithm. In other words, at each Greedy step only one entry of the vectors, and one row and one column of the matrices must be computed and added to their previous versions. Once all parameter-independent quantities are available, for each $\mu \in \mathcal{D}_{\text{train}}$, one first computes (3.14), solves (3.1) for $\hat{\mathbf{x}}(\mu)$ and, finally, computes (3.13). These three steps are cheap since their computational complexity depends only on k, n_E, n_A and n_B, which are very small. Therefore, the search on $\mathcal{D}_{\text{train}}$ is quite fast, which makes the Greedy algorithm feasible and efficient even for large-scale problems and the large training set $\mathcal{D}_{\text{train}}$.

4. Low-rank reduced basis method for Lyapunov equation. A major drawback of the reduced basis method described above is that it operates with matrices and vectors of huge dimension N^2 and, as a consequence, suffers from high computational complexity and storage requirements. Fortunately, thanks to Lemma 2.2, the reduced basis method can directly be applied to the PALE (1.1) keeping all operations in terms of $N \times N$-matrices and N-vectors. Moreover, assuming that the solution of the PALE (1.1) with the low-rank right-hand side $B(\mu)B^T(\mu)$ is well approximated by a low-rank matrix $X(\mu) \approx Z(\mu)Z^T(\mu)$, we can further reduce computational cost and memory requirements both in offline and online stages.

4.1. Offline phase. In the offline phase, instead of solving the linear system (2.1), we compute the solutions of the PALE (1.1) for selected parameters $\mu_1, \ldots, \mu_k$. Let $Z_j \in \mathbb{R}^{N \times n_j}$ be the low-rank Cholesky factor of the solution $X(\mu_j) \approx Z_j Z_j^T$ of (1.1) at $\mu = \mu_j$, $j = 1, \ldots, k$. These factors can efficiently be computed by the low-rank alternating directions implicit (LR-ADI) method [5, 19, 24], the rational Krylov method [9, 26] or the Riemannian method [30]. Then for a reduced basis matrix $V_k = [\, Z_1, \ldots, Z_k \,]$ and any $\mu \in \mathcal{D}$, we determine an approximate solution of (1.1) as

$$X(\mu) \approx \text{mat}(\mathbf{V}_k\hat{\mathbf{x}}(\mu)) =: X_{\text{RB}}(\mu),$$

where $\mathbf{V}_k = [\text{vec}(Z_1 Z_1^T), \ldots, \text{vec}(Z_k Z_k^T)]$ and $\hat{x}(\mu)$ solves the reduced linear system (3.1). The solution $X_{\mathrm{RB}}(\mu)$ can also be written as

$$X_{\mathrm{RB}}(\mu) = \sum_{j=1}^{k} \hat{x}_j(\mu) Z_j Z_j^T = V_k \begin{bmatrix} \hat{x}_1(\mu)I_{n_1} & & \\ & \ddots & \\ & & \hat{x}_k(\mu)I_{n_k} \end{bmatrix} V_k^T, \qquad (4.1)$$

where $\hat{\mathbf{x}}(\mu) = [\hat{x}_1(\mu), \ldots, \hat{x}_1(\mu)]^T$. Note that we never form the matrix $\mathbf{V}_k$ explicitly to construct the reduced linear system (3.1). Instead, we exploit the affine decomposition (3.14) for the coefficient matrix $\hat{\mathbf{L}}(\mu)$ and the right-hand side $\hat{\mathbf{b}}(\mu)$ and compute the entries of the parameter-independent matrices $\mathbf{V}_k^T \mathbf{L}_p \mathbf{V}_k$ for $p = (i-1)n_A + j$ with $i = 1, \ldots, n_E$, $j = 1, \ldots, n_A$, and the vectors $\mathbf{V}_k^T \mathbf{b}_p$ for $p = (i-1)n_B + j$ with $i, j = 1, \ldots, n_B$, using Lemmas 2.1, 2.2 and relations (3.5), (3.12) as follows

$$\begin{aligned}
(\mathbf{V}_k^T \mathbf{L}_p \mathbf{V}_k)_{rl} &= \langle -E_i Z_l Z_l^T A_j^T - A_j Z_l Z_l^T E_i^T, Z_r Z_r^T \rangle \\
&= -2\text{trace}\big(Z_r^T (E_i Z_l)(A_j Z_l)^T Z_r \big), \\
(\mathbf{V}_k^T \mathbf{b}_p)_r &= \langle B_i B_j^T, Z_r Z_r^T \rangle = \text{trace}\big((B_j^T Z_r)(Z_r^T B_i) \big)
\end{aligned}$$

for $r, l = 1, \ldots, k$. Thus, taking advantage of the structure of $\mathbf{L}_p$, $\mathbf{b}_p$ and $\mathbf{V}_k$ reduces the computational cost, for example, for $\mathbf{V}_k^T \mathbf{L}_p \mathbf{V}_k$ from $\mathcal{O}(N^4 k)$ to $\mathcal{O}(N^2 n)$ with $n = n_1 + \ldots + n_k$. In counting, we did not exploit the sparsity of E_i and A_j.

REMARK 4.1. *Note that even E_i and A_j are assumed to be symmetric, we always write E_i^T and A_j^T if the transpose matrices are needed. This will simplify the extension of the reduced basis method to nonsymmetric problems, see Section 5.*

For the approximate solution $X_{\mathrm{RB}}(\mu)$, we obtain from Theorem 3.3 the error estimate

$$\|X(\mu) - X_{\mathrm{RB}}(\mu)\|_F = \|\mathbf{e}_k(\mu)\| \leq \Delta_k(\mu) \qquad (4.2)$$

with $\Delta_k(\mu)$ as in (3.9). This error estimator can now be utilized in the Greedy parameter sampling procedure presented in Algorithm 2.

Algorithm 2 Greedy algorithm for Lyapunov equations

Input: tolerance $tol_{\mathrm{rb}} < 1$, training set $\mathcal{D}_{\mathrm{train}}$, initial parameter $\mu_1 \in \mathcal{D}_{\mathrm{train}}$.
Output: a basis matrix V_k.
 1: Solve the PALE (1.1) at $\mu = \mu_1$ for $X(\mu_1) \approx Z_1 Z_1^T$.
 2: $\Delta_1^{\max} = 1$, $\mathcal{M}_1 = \{\mu_1\}$, $V_1 = Z_1$, $k = 2$.
 3: **while** $\Delta_{k-1}^{\max} \geq tol_{\mathrm{rb}}$ **do**
 4:     $\mu_k = \underset{\mu \in \mathcal{D}_{\mathrm{train}} \backslash \mathcal{M}_{k-1}}{\arg\max} \Delta_{k-1}(\mu)$
 5:     $\Delta_k^{\max} = \Delta_{k-1}(\mu_k)$
 6:     $\mathcal{M}_k = \mathcal{M}_{k-1} \cup \{\mu_k\}$
 7:     solve the PALE (1.1) at $\mu = \mu_k$ for $X(\mu_k) \approx Z_k Z_k^T$
 8:     $V_k = [\, V_{k-1}, \; Z_k \,]$
 9:     $k \leftarrow k + 1$
10: **end while**

For the efficient computation of the residuals $\|\mathbf{r}_k(\mu)\|$, we use again the affine representation (3.13), where computing the parameter-independent quantities demands

a more detailed discussion. Consider first $\mathbf{b}_p^T\mathbf{b}_q$ for $p = (i-1)n_B+j$, $q = (f-1)n_B+g$ and $i, j, f, g = 1, \ldots, n_B$. Using (3.5) and (3.12) we obtain

$$\mathbf{b}_p^T\mathbf{b}_q = \langle B_f B_g^T, B_i B_j^T \rangle = \text{trace}\big((B_i^T B_f)(B_g^T B_j)\big).$$

The components of the vector $\mathbf{b}_p^T\mathbf{L}_q\mathbf{V}_k$ for $p = (i-1)n_B + j$, $q = (f-1)n_A + g$, $i, j = 1, \ldots, n_B$, $f = 1, \ldots, n_E$ and $g = 1, \ldots, n_A$ can be expressed as

$$\begin{aligned}(\mathbf{b}_p^T\mathbf{L}_q\mathbf{V}_k)_l &= \langle -A_g Z_l Z_l^T E_f^T - E_f Z_l Z_l^T A_g^T, B_i B_j^T \rangle \\ &= -\text{trace}\big(B_i^T(E_f Z_l)(A_g Z_l)^T B_j + B_i^T(A_g Z_l)(E_f Z_l)^T B_j\big), \quad l = 1, \ldots, k.\end{aligned}$$

Finally, the matrix $\mathbf{V}_k^T\mathbf{L}_p^T\mathbf{L}_q\mathbf{V}_k$ for $p = (i - 1)n_A + j$, $q = (f - 1)n_A + g$ and $i, f = 1, \ldots, n_E$, $j, g = 1, \ldots, n_A$ can be determined element-wise

$$\begin{aligned}(\mathbf{V}_k^T\mathbf{L}_p^T\mathbf{L}_q\mathbf{V}_k)_{rl} &= \langle -A_g Z_l Z_l^T E_f^T - E_f Z_l Z_l^T A_g^T, -A_j Z_r Z_r^T E_i^T - E_i Z_r Z_r^T A_j^T \rangle \\ &= 2\,\text{trace}\big((E_i Z_r)^T(E_f Z_l)(A_g Z_l)^T(A_j Z_r) + (E_i Z_r)^T(A_g Z_l)(E_f Z_l)^T(A_j Z_r)\big)\end{aligned}$$

for $r, l = 1, \ldots, k$. Here, we used Lemma 2.2 and relations (3.5) and (3.12).

4.2. Online phase. Once the reduced basis matrix V_k is constructed such that the error estimator does not exceed a given tolerance, the solution of the PALE (1.1) at any $\mu \in \mathcal{D}$ can be obtained in the online phase as in (4.1). Nevertheless, a serious disadvantage of this approach is that the resulting solution $X_{\text{RB}}(\mu)$ is not necessarily positive semidefinite since the solution $\hat{\mathbf{x}}(\mu)$ of (1.1) may have negative entries. This difficulty can be circumvented by computing the approximate solution in the form $X(\mu) \approx V_k\hat{X}(\mu)V_k^T =: \hat{X}_{\text{RB}}(\mu)$, where $\hat{X}(\mu)$ solves the reduced Lyapunov equation

$$\hat{A}(\mu)\hat{X}(\mu)\hat{E}^T(\mu) + \hat{E}(\mu)\hat{X}(\mu)\hat{A}^T(\mu) = -\hat{B}(\mu)\hat{B}^T(\mu) \tag{4.3}$$

with $\hat{E}(\mu) = V_k^T E(\mu)V_k$, $\hat{A}(\mu) = V_k^T A(\mu)V_k$ and $\hat{B}(\mu) = V_k^T B(\mu)$. Since $-A(\mu)$ and $E(\mu)$ are symmetric and positive definite, this equation has a unique symmetric positive semidefinite solution $\hat{X}(\mu) = \hat{Z}(\mu)\hat{Z}^T(\mu)$. Then $\hat{X}_{\text{RB}}(\mu)$ can be written in the factorized form $\hat{X}_{\text{RB}}(\mu) = Z_{\text{RB}}(\mu)Z_{\text{RB}}^T(\mu)$ with $Z_{\text{RB}}(\mu) = V_k\hat{Z}(\mu)$.

Let

$$\hat{R}_k(\mu) = A(\mu)\hat{X}_{\text{RB}}(\mu)E^T(\mu) + E(\mu)\hat{X}_{\text{RB}}(\mu)A^T(\mu) + B(\mu)B^T(\mu) \tag{4.4}$$

be the residual associated with the approximate solution $\hat{X}_{\text{RB}}(\mu)$. Then the error $X(\mu) - \hat{X}_{\text{RB}}(\mu)$ can be estimated similarly to the linear system case as

$$\|X(\mu) - \hat{X}_{\text{RB}}(\mu)\|_F \leq \frac{\|\hat{R}_k(\mu)\|_F}{\alpha(\mu)} \leq \frac{\|\hat{R}_k(\mu)\|_F}{\alpha_{\text{LB}}(\mu)} =: \hat{\Delta}_k(\mu) \tag{4.5}$$

with $\alpha_{\text{LB}}(\mu)$ as in (3.6). Replacing the approximate solution $\mathbf{V}_k\hat{\mathbf{x}}(\mu)$ in (3.13) by

$$\text{vec}(\hat{X}_{\text{RB}}(\mu)) = (V_k \otimes V_k)\text{vec}(\hat{X}(\mu)),$$

we obtain the following expression for the residual

$$
\begin{aligned}
\|\hat{R}_k(\mu)\|_F^2 &= \|\mathbf{b}(\mu) - \mathbf{L}(\mu)(V_k \otimes V_k)\mathrm{vec}(\hat{X}(\mu))\|^2 \\
&= \sum_{i,j=1}^{n_B} \sum_{f,g=1}^{n_B} \theta_{ijfg}^B(\mu)\mathrm{trace}\big((B_i^T B_f)(B_g^T B_j)\big) \\
&\quad + 2\sum_{i,j=1}^{n_B} \sum_{f=1}^{n_E} \sum_{g=1}^{n_A} \theta_{ijfg}^{AEB}(\mu)\mathrm{trace}\big(B_i^T(E_f V_k)\hat{X}(\mu)(A_g V_k)^T B_j\big) \\
&\quad + 2\sum_{i,j=1}^{n_B} \sum_{f=1}^{n_E} \sum_{g=1}^{n_A} \theta_{ijfg}^{AEB}(\mu)\mathrm{trace}\big(B_i^T(A_g V_k)\hat{X}(\mu)(E_f V_k)^T B_j\big) \\
&\quad + 2\sum_{i,f=1}^{n_E} \sum_{j,g=1}^{n_A} \theta_{ijfg}^{AE}(\mu)\mathrm{trace}\big((E_f V_k)^T(E_i V_k)\hat{X}(\mu)(A_j V_k)^T(A_g V_k)\hat{X}(\mu)\big) \\
&\quad + 2\sum_{i,f=1}^{n_E} \sum_{j,g=1}^{n_A} \theta_{ijfg}^{AE}(\mu)\mathrm{trace}\big((E_f V_k)^T(A_j V_k)\hat{X}(\mu)(E_i V_k)^T(A_g V_k)\hat{X}(\mu)\big),
\end{aligned}
$$

where $\theta_{ijfg}^B(\mu) = \theta_i^B(\mu)\theta_j^B(\mu)\theta_f^B(\mu)\theta_g^B(\mu)$, $\theta_{ijfg}^{AEB}(\mu) = \theta_i^B(\mu)\theta_j^B(\mu)\theta_f^E(\mu)\theta_g^A(\mu)$ and $\theta_{ijfg}^{AE}(\mu) = \theta_i^E(\mu)\theta_j^A(\mu)\theta_f^E(\mu)\theta_g^A(\mu)$. Again, all parameter-independent matrices can be precomputed and stored in the offline stage. Then the error estimator $\hat{\Delta}_k(\mu)$ can be calculated in the online stage at low computational cost which is independent of the large dimension N.

Note that in the Greedy algorithm, instead of $\Delta_k(\mu)$ one can also use the estimator $\hat{\Delta}_k(\mu)$. It should, however, be emphasized that the computation of the reduced basis solution $X_{\mathrm{RB}}(\mu)$ in (4.1) is less expensive than that of $\hat{X}_{\mathrm{RB}}(\mu) = V_k\hat{X}(\mu)V_k^T$. This is caused by the fact that solving the linear system (3.1) with $\hat{\mathbf{L}}(\mu) \in \mathbb{R}^{k \times k}$ is cheaper than solving the Lyapunov equation (4.3) with $\hat{E}(\mu)$, $\hat{A}(\mu) \in \mathbb{R}^{n \times n}$, where $n = n_1 + \ldots + n_k$ may be significantly larger than k. The dimension n depends on m (number of columns of $B(\mu)$), the convergence rate of the iterative method used for solving the PALE (1.1) and the number of Greedy steps. To keep the offline computational cost low, we compute the error estimator based on $X_{\mathrm{RB}}(\mu)$, whereas in the online phase, where the positive semidefiniteness of the solution is most essential, we calculate $\hat{X}_{\mathrm{RB}}(\mu)$. Moreover, due to the fact that $\mathrm{span}(\mathbf{V}_k) \subset \mathrm{span}(V_k \otimes V_k)$, the approximate solution $\hat{X}_{\mathrm{RB}}(\mu)$ is to predicted to have smaller error. It is reasonably compliant with the fact that computing $\hat{X}_{\mathrm{RB}}(\mu)$ is more expensive. In order to retain a low-rank structure in $\hat{X}_{\mathrm{RB}}(\mu)$ and to reduce the online computational cost, the column compression in V_k should be performed with a prescribed tolerance tol_{cc}. This can be done by computing a rank-revealing QR decomposition or a singular value decomposition (SVD) of V_k.

4.3. Energy norm based error estimates. Instead of the Euclidean vector norm and the Frobenius matrix norm, one can also quantify the approximation error in the energy norm which is frequently used in the reduced basis method [23]. In our setting, the energy vector norm is defined by $\|\mathbf{x}\|_\mu = \sqrt{\mathbf{x}^T\mathbf{L}(\mu)\mathbf{x}}$. A matrix counterpart can be introduced as

$$
\|X\|_{F,\mu} = \sqrt{\langle \mathcal{L}_\mu(X), X\rangle}.
$$

It is easy to verify that $\|X\|_{F,\mu} = \|\mathrm{vec}(X)\|_\mu$ and

$$
\sqrt{\alpha(\mu)}\|X\|_F \leq \|X\|_{F,\mu} \leq \sqrt{\gamma(\mu)}\|X\|_F. \tag{4.6}
$$

with $\alpha(\mu)$ and $\gamma(\mu)$ as in (2.5) and (2.6), respectively. Furthermore, for a symmetric matrix X, we have $\|X\|_{F,\mu} = \sqrt{2\operatorname{trace}(-A(\mu)XE^T(\mu)X)}$. The first inequality in (4.6) implies the energy based error estimates

$$\|X(\mu) - X_{\mathrm{RB}}(\mu)\|_{F,\mu} \leq \frac{\|\mathbf{r}_k(\mu)\|}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}} =: \Delta_k^{en}(\mu), \tag{4.7}$$

$$\|X(\mu) - \hat{X}_{\mathrm{RB}}(\mu)\|_{F,\mu} \leq \frac{\|\hat{R}_k(\mu)\|_F}{\sqrt{\alpha_{\mathrm{LB}}(\mu)}} =: \hat{\Delta}_k^{en}(\mu). \tag{4.8}$$

The corresponding effectivity constants satisfy

$$1 \leq \eta_k^{en}(\mu) = \frac{\Delta_k^{en}(\mu)}{\|X(\mu) - X_{\mathrm{RB}}(\mu)\|_\mu} \leq \sqrt{\frac{\gamma_{\mathrm{UB}}(\mu)}{\alpha_{\mathrm{LB}}(\mu)}},$$

$$1 \leq \hat{\eta}_k^{en}(\mu) = \frac{\hat{\Delta}_k^{en}(\mu)}{\|X(\mu) - \hat{X}_{\mathrm{RB}}(\mu)\|_\mu} \leq \sqrt{\frac{\gamma_{\mathrm{UB}}(\mu)}{\alpha_{\mathrm{LB}}(\mu)}}.$$

We see that the upper bound for $\eta_k^{en}(\mu)$ and $\hat{\eta}_k^{en}(\mu)$ is smaller than that for $\eta_k(\mu)$ in (3.10). This implies that the estimates (4.7) and (4.8) are sharper than (4.2) and (4.5), respectively.

Similarly to [23, Lemma 4A and Proposition 4C], one can get even a sharper error estimate

$$\|\mathbf{e}_k(\mu)\|_\mu \leq \frac{\|\mathbf{L}^{-1}(\bar{\mu})\mathbf{r}_k(\mu)\|_{\bar{\mu}}}{\sqrt{\theta_{\min}^{\mathbf{L},\bar{\mu}}(\mu)}} =: \Delta_k^{en,\bar{\mu}}(\mu) \tag{4.9}$$

with the effectivity constant

$$1 \leq \eta_k^{en,\bar{\mu}}(\mu) := \frac{\Delta_k^{en,\bar{\mu}}(\mu)}{\|\mathbf{e}_k(\mu)\|_\mu} \leq \sqrt{\theta^{\mathbf{L},\bar{\mu}}(\mu)}.$$

Let us, however, consider one term in the affine decomposition of

$$\|\mathbf{L}^{-1}(\bar{\mu})\mathbf{r}_k(\mu)\|_{\bar{\mu}}^2 = \mathbf{r}_k^T(\mu)\mathbf{L}^{-1}(\bar{\mu})\mathbf{r}_k(\mu)$$
$$= \big(\mathbf{b}(\mu) - \mathbf{L}(\mu)\mathbf{V}_k\hat{\mathbf{x}}(\mu)\big)^T \mathbf{L}^{-1}(\bar{\mu})\big(\mathbf{b}(\mu) - \mathbf{L}(\mu)\mathbf{V}_k\hat{\mathbf{x}}(\mu)\big),$$

say the last one. The parameter-independent matrix products there have the form $\mathbf{V}_k^T \mathbf{L}_p^T \mathbf{L}^{-1}(\bar{\mu})\mathbf{L}_q\mathbf{V}_k$ with $p = (i-1)n_A + j$ and $q = (f-1)n_A + g$. To compute the (r,l)-th entry of this matrix, we need to
- compute $K_q = -A_g Z_l Z_l^T E_f^T - E_f Z_l Z_l^T A_g^T$,
- solve the Lyapunov equation $-A(\bar{\mu})XE^T(\bar{\mu}) - E(\bar{\mu})XA^T(\bar{\mu}) = K_q$,
- compute the inner product $\langle X, -A_j Z_r Z_r^T E_i^T - E_i Z_r Z_r^T A_j^T \rangle$.

Thus, the computation of the error estimator $\Delta_k^{en,\bar{\mu}}(\mu)$ requires solving several Lyapunov equations most probably without low-rank right-hand sides which makes the error estimate (4.9) inefficient.

5. Nonsymmetric case. The reduced basis method as described above can also be applied to the PALE (1.1), where $A(\mu)$ is nonsymmetric, with some adjustments in formulating the error estimates.

Assume that the pencil $\lambda E(\mu) - A(\mu)$ is strictly dissipative, i.e.,

$$E(\mu) = E^T(\mu) > 0, \qquad A(\mu) + A^T(\mu) < 0 \tag{5.1}$$

for all $\mu \in \mathcal{D}$. These conditions guarantee the solvability of the reduced Lyapunov equation (4.3) for any projection matrix V_k. They are fulfilled if assumptions (A1), (A2) together with

(A3´) $A_j + A_j^T \leq 0$ and $\theta_j^A(\mu) > 0$ for all $\mu \in \mathcal{D}$ and $j = 1, \ldots, n_A$;

(A4´) there exist at least one pair (A_j, E_i) such that $A_j + A_j^T < 0$ and $E_i > 0$

hold. Taking into account (3.2) and (3.3), we just need to find upper bounds for $\|\mathbf{L}(\mu)\|_2 = \sigma_{\max}\big(\mathbf{L}(\mu)\big)$ and $\|\mathbf{L}^{-1}(\mu)\|_2 = 1/\sigma_{\min}\big(\mathbf{L}(\mu)\big)$. For this purpose, we introduce

$$S(\mu) = \frac{1}{2}\big(A(\mu) + A^T(\mu)\big), \tag{5.2}$$

$$\mathbf{L}_S(\mu) = \frac{1}{2}\big(\mathbf{L}(\mu) + \mathbf{L}^T(\mu)\big) = -E(\mu) \otimes S(\mu) - S(\mu) \otimes E(\mu).$$

Obviously, these matrices inherit the affine structure. Moreover, $S(\mu) = S^T(\mu) < 0$ and $\mathbf{L}_S(\mu) = \mathbf{L}_S^T(\mu) > 0$ for all $\mu \in \mathcal{D}$. The following lemma establishes the bounds for the smallest and largest singular values of $\mathbf{L}(\mu)$.

LEMMA 5.1. *Let $E(\mu)$ and $A(\mu)$ satisfy (A1), (A2), (A3´) and (A4´), and let $\bar{\mu}, \bar{\mu}_1, \bar{\mu}_2 \in \mathcal{D}$.*

1. *For all $\mu \in \mathcal{D}$, the smallest singular value of $\mathbf{L}(\mu)$ is bounded from below as*

$$\sigma_{\min}\big(\mathbf{L}(\mu)\big) \geq \tilde{\alpha}_{\mathrm{LB}}(\mu) := \max\Big(\tilde{\alpha}_{\mathrm{LB}}^{\mathbf{L},\bar{\mu}}(\mu),\ \tilde{\alpha}_{\mathrm{LB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu),\ \tilde{\alpha}_{\mathrm{LB}}^{\mathbf{L}}(\mu)\Big) > 0, \tag{5.3}$$

where

$$\begin{aligned}
\tilde{\alpha}_{\mathrm{LB}}^{\mathbf{L},\bar{\mu}}(\mu) &= 2\,\theta_{\min}^{\mathbf{L},\bar{\mu}}(\mu)\,\lambda_{\min}\big(-S(\bar{\mu})\big)\,\lambda_{\min}\big(E(\bar{\mu})\big), \\
\tilde{\alpha}_{\mathrm{LB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu) &= 2\,\theta_{\min}^{A,\bar{\mu}_1}(\mu)\,\theta_{\min}^{E,\bar{\mu}_2}(\mu)\,\lambda_{\min}\big(-S(\bar{\mu}_1)\big)\,\lambda_{\min}\big(E(\bar{\mu}_2)\big), \\
\tilde{\alpha}_{\mathrm{LB}}^{\mathbf{L}}(\mu) &= 2\sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\theta_{ij}^{\mathbf{L}}(\mu)\,\lambda_{\min}(-S_j)\,\lambda_{\min}(E_i)
\end{aligned}$$

with $S(\mu)$ as in (5.2) and $S_j = (A_j + A_j^T)/2$.

2. *For all $\mu \in \mathcal{D}$, the largest singular value of $\mathbf{L}(\mu)$ is bounded from above as*

$$\sigma_{\max}\big(\mathbf{L}(\mu)\big) \leq \tilde{\gamma}_{\mathrm{UB}}(\mu) := \min\Big(\tilde{\gamma}_{\mathrm{UB}}^{\mathbf{L},\bar{\mu}}(\mu),\ \tilde{\gamma}_{\mathrm{UB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu),\ \tilde{\gamma}_{\mathrm{UB}}^{\mathbf{L}}(\mu)\Big), \tag{5.4}$$

where

$$\begin{aligned}
\tilde{\gamma}_{\mathrm{UB}}^{\mathbf{L},\bar{\mu}}(\mu) &= 2\,\theta_{\max}^{\mathbf{L},\bar{\mu}}(\mu)\,\sigma_{\max}\big(A(\bar{\mu})\big)\,\lambda_{\max}\big(E(\bar{\mu})\big), \\
\tilde{\gamma}_{\mathrm{UB}}^{A,\bar{\mu}_1;E,\bar{\mu}_2}(\mu) &= 2\,\theta_{\max}^{A,\bar{\mu}_1}(\mu)\,\theta_{\max}^{E,\bar{\mu}_2}(\mu)\,\sigma_{\max}\big(A(\bar{\mu}_1)\big)\,\lambda_{\max}\big(E(\bar{\mu}_2)\big), \\
\tilde{\gamma}_{\mathrm{UB}}^{\mathbf{L}}(\mu) &= 2\sum_{i=1}^{n_E}\sum_{j=1}^{n_A}\theta_{ij}^{\mathbf{L}}(\mu)\,\sigma_{\max}(A_j)\,\lambda_{\max}(E_i).
\end{aligned}$$

Proof. 1. Based on an important inequality between singular values of a matrix and eigenvalues of its symmetric part [15, Corollary 3.1.5], we get

$$\sigma_{\min}\big(\mathbf{L}(\mu)\big) \geq \lambda_{\min}\big(\mathbf{L}_S(\mu)\big).$$

Then the bound (5.3) immediately follows from Lemma 3.2, part 1.

2. Using the multiplicativity property of the singular values of the Kronecker product [15, Theorem 4.2.15], we obtain that

$$\sigma_{\max}\big(\mathbf{L}(\mu)\big) \leq 2\,\sigma_{\max}\big(A(\mu)\big)\,\lambda_{\max}\big(E(\mu)\big).$$

Then the bound (5.4) can be proved analogously to Lemma 3.2, part 2. □

We use now the bounds (5.3) and (5.4) to derive the error estimates for the reduced basis solutions of the PALE (1.1).

THEOREM 5.2. *Let $E(\mu)$ and $A(\mu)$ satisfy (A1), (A2), (A3´) and (A4´), and let $X_{\mathrm{RB}}(\mu)$ and $\hat{X}_{\mathrm{RB}}(\mu)$ be the reduced basis solutions of the PALE (1.1). Then the errors $X(\mu) - X_{\mathrm{RB}}(\mu)$ and $X(\mu) - \hat{X}_{\mathrm{RB}}(\mu)$ can be estimated as*

$$\|X(\mu) - X_{\mathrm{RB}}(\mu)\|_F \leq \frac{\|\mathbf{r}_k(\mu)\|}{\tilde{\alpha}_{\mathrm{LB}}(\mu)} =: \Delta_k^{ns}(\mu) \leq \frac{\tilde{\gamma}_{\mathrm{UB}}(\mu)}{\tilde{\alpha}_{\mathrm{LB}}(\mu)}\|X(\mu) - X_{\mathrm{RB}}(\mu)\|_F,$$

$$\|X(\mu) - \hat{X}_{\mathrm{RB}}(\mu)\|_F \leq \frac{\|\hat{R}_k(\mu)\|_F}{\tilde{\alpha}_{\mathrm{LB}}(\mu)} =: \hat{\Delta}_k^{ns}(\mu) \leq \frac{\tilde{\gamma}_{\mathrm{UB}}(\mu)}{\tilde{\alpha}_{\mathrm{LB}}(\mu)}\|X(\mu) - \hat{X}_{\mathrm{RB}}(\mu)\|_F,$$

where $\tilde{\alpha}_{\mathrm{LB}}(\mu)$ and $\tilde{\gamma}_{\mathrm{UB}}(\mu)$ are as in (5.3) and (5.4), respectively.

Proof. The result follows from (3.2), (3.3) and Lemma 5.1. □

Alternative error estimates can be derived using a 2-logarithmic norm of the pencil $\lambda E(\mu) - A(\mu)$ defined as

$$\ell\big(E(\mu), A(\mu)\big) = \lambda_{\max}\big(E(\mu), S(\mu)\big).$$

If $E(\mu) \equiv I$, then $\ell\big(I, A(\mu)\big) = \ell\big(A(\mu)\big) = \lambda_{\max}\big(S(\mu)\big)$ is the 2-logarithmic matrix norm which is frequently used in differential equations and numerical analysis [28]. Conditions (5.1) imply that $\ell\big(E(\mu), A(\mu)\big) < 0$ for all $\mu \in \mathcal{D}$, so it is not a norm in the usual sense. Define a weighted matrix norm

$$\|X\|_{E(\mu)} = \|E(\mu)X\|_F = \|G^T(\mu)XG(\mu)\|_F,$$

where $G(\mu)$ is a Cholesky factor of $E(\mu) = G(\mu)G^T(\mu)$. The following theorem establishes an error estimate for the reduced basis solution $\hat{X}_{RB}(\mu) = V_k\hat{X}(\mu)V_k^T$.

THEOREM 5.3. *Let $X(\mu)$ and $\hat{X}_{RB}(\mu)$ be the exact and approximate solutions of the PALE (1.1). Then the error $X(\mu) - \hat{X}_{RB}(\mu)$ can be estimated as*

$$\|X(\mu) - \hat{X}_{RB}(\mu)\|_{E(\mu)} \leq \frac{\|\hat{R}_k(\mu)\|_F}{\alpha_{\mathrm{LB}}^{E,A,\bar{\mu}}(\mu)} =: \hat{\Delta}_k^{E,A,\bar{\mu}}(\mu) \leq \frac{\gamma_{\mathrm{UB}}^{E,A,\bar{\mu}}(\mu)}{\alpha_{\mathrm{LB}}^{E,A,\bar{\mu}}(\mu)}\|X(\mu) - \hat{X}_{RB}(\mu)\|_{E(\mu)},$$

where

$$\alpha_{\mathrm{LB}}^{E,A,\bar{\mu}}(\mu) = 2\,\frac{\theta_{\min}^{A,\bar{\mu}}(\mu)}{\theta^{E,\bar{\mu}}(\mu)}\lambda_{\min}\big(E(\bar{\mu})\big)\,\lambda_{\min}\big(E(\bar{\mu}), -S(\bar{\mu})\big),$$

$$\gamma_{\mathrm{UB}}^{E,A,\bar{\mu}}(\mu) = 2\,\theta_{\max}^{A,\bar{\mu}}(\mu)\sqrt{\theta^{E,\bar{\mu}}(\mu)}\sigma_{\max}\big(A(\bar{\mu})\big)\sqrt{\frac{\lambda_{\max}\big(E(\bar{\mu})\big)}{\lambda_{\min}\big(E(\bar{\mu})\big)}},$$

and $\hat{R}_k(\mu)$ is the residual given in (4.4).

Proof. Let $\Xi(\mu) = X(\mu) - \hat{X}_{RB}(\mu)$. It follows from

$$G^{-1}(\mu)\hat{R}_k(\mu)G^{-T}(\mu) = -A_G(\mu)G^T(\mu)\Xi(\mu)G(\mu) - G^T(\mu)\Xi(\mu)G(\mu)A_G^T(\mu)$$

with $A_G(\mu) = G^{-1}(\mu)A(\mu)G^{-T}(\mu)$ that

$$-G^T(\mu)\Xi(\mu)G(\mu) = \int_0^\infty e^{A_G(\mu)t}G^{-1}(\mu)\hat{R}_k(\mu)G^{-T}(\mu)e^{A_G^T(\mu)t}dt.$$

The matrix exponential is estimated as

$$\|e^{A_G(\mu)t}\|_2 \le e^{\ell\left(G^{-1}(\mu)A(\mu)G^{-T}(\mu)\right)t} = e^{\ell\left(E(\mu),A(\mu)\right)t},$$

see [7, 22]. Therefore,

$$\|\Xi(\mu)\|_{E(\mu)} \le \|G^{-1}(\mu)\hat{R}_k(\mu)G^{-T}(\mu)\|_F \int_0^\infty e^{2\ell(E(\mu),A(\mu))t}dt \le \frac{\|\hat{R}_k(\mu)\|_F \|E^{-1}(\mu)\|_2}{-2\,\ell\left(E(\mu),A(\mu)\right)}.$$

It remains to find a lower bound for $-\ell\left(E(\mu), A(\mu)\right)$. We again employ the min-Θ approach. For

$$S(\mu) = \sum_{j=1}^{n_A} \theta_j^A(\mu)S_j$$

with $S_j = (A_j + A_j^T)/2$, we obtain similarly to the proof of Lemma 3.2 that

$$
\begin{aligned}
-\ell\left(E(\mu), A(\mu)\right) = \lambda_{\min}\left(E(\mu), -S(\mu)\right) &= \min_{v\in\mathbb{R}^n\backslash\{0\}} \frac{v^T(-S(\mu))v}{v^T E(\mu)v} \\
&= \min_{v\in\mathbb{R}^n\backslash\{0\}} \frac{\sum_{j=1}^{n_A}\theta_j^A(\mu)v^T(-S_j)v}{\sum_{j=1}^{n_E}\theta_j^E(\mu)v^T E_j v} \ge \min_{v\in\mathbb{R}^n\backslash\{0\}} \frac{\theta_{\min}^{A,\bar{\mu}}(\mu)v^T(-S(\bar{\mu}))v}{\theta_{\max}^{E,\bar{\mu}}(\mu)v^T E(\bar{\mu})v} \\
&= \frac{\theta_{\min}^{A,\bar{\mu}}(\mu)}{\theta_{\max}^{E,\bar{\mu}}(\mu)}\lambda_{\min}\left(E(\bar{\mu}), -S(\bar{\mu})\right).
\end{aligned}
$$

Hence, we have the error estimate

$$\|X(\mu) - \hat{X}_{RB}(\mu)\|_{E(\mu)} \le \frac{\|\hat{R}_k(\mu)\|_F}{\alpha_{\mathrm{LB}}^{E,A,\bar{\mu}}(\mu)} =: \hat{\Delta}_k^{E,A,\bar{\mu}}(\mu)$$

with

$$\alpha_{\mathrm{LB}}^{E,A,\bar{\mu}}(\mu) = 2\,\frac{\theta_{\min}^{A,\bar{\mu}}(\mu)}{\theta^{E,\bar{\mu}}(\mu)}\lambda_{\min}\left(E(\bar{\mu})\right)\lambda_{\min}\left(E(\bar{\mu}), -S(\bar{\mu})\right).$$

Taking into account that

$$
\begin{aligned}
\|\hat{R}_k(\mu)\|_F &= \|A(\mu)(\hat{X}_{\mathrm{RB}}(\mu) - X(\mu))E^T(\mu) + E(\mu)(\hat{X}_{\mathrm{RB}}(\mu) - X(\mu))A^T(\mu)\|_F \\
&\le 2\,\|A(\mu)\|_2\sqrt{\|E(\mu)\|_2\|E^{-1}(\mu)\|_2}\|X(\mu) - \hat{X}_{RB}(\mu)\|_{E(\mu)} \\
&\le 2\,\theta_{\max}^{A,\bar{\mu}}(\mu)\sigma_{\max}\left(A(\bar{\mu})\right)\sqrt{\theta^{E,\bar{\mu}}(\mu)\frac{\lambda_{\max}\left(E(\bar{\mu})\right)}{\lambda_{\min}\left(E(\bar{\mu})\right)}}\|X(\mu) - \hat{X}_{RB}(\mu)\|_{E(\mu)},
\end{aligned}
$$

the effectivity constant can be estimated as

$$\eta_k^{E,A,\bar{\mu}}(\mu) := \frac{\hat{\Delta}_k^{E,A,\bar{\mu}}(\mu)}{\|X(\mu) - \hat{X}_{RB}(\mu)\|_{E(\mu)}} \le \frac{2\theta_{\max}^{A,\bar{\mu}}(\mu)\sqrt{\theta^{E,\bar{\mu}}(\mu)}\sigma_{\max}\left(A(\bar{\mu})\right)\lambda_{\max}^{1/2}\left(E(\bar{\mu})\right)}{\alpha_{\mathrm{LB}}^{E,A,\bar{\mu}}(\mu)\lambda_{\min}^{1/2}\left(E(\bar{\mu})\right)}.$$

This completes the proof. $\qquad\square$

6. Application to parametric model order reduction. Given a parametric linear dynamical system

$$E(\mu)\dot{x}(t,\mu) = A(\mu)x(t,\mu) + B(\mu)u(t,\mu),$$
$$y(t,\mu) = C(\mu)x(t,\mu), \tag{6.1}$$

where $A(\mu)$, $E(\mu)$ and $B(\mu)$ satisfy (A1), and the output matrix $C(\mu) \in \mathbb{R}^{l \times N}$ with $l \ll N$ also depends affinely on μ, i.e.,

$$C(\mu) = \sum_{j=1}^{n_C} \theta_j^C(\mu)C_j.$$

A main goal of model reduction is to approximate system (6.1) by a reduced-order model

$$\tilde{E}(\mu)\dot{\tilde{x}}(t,\mu) = \tilde{A}(\mu)\tilde{x}(t,\mu) + \tilde{B}(\mu)u(t,\mu),$$
$$\tilde{y}(t,\mu) = \tilde{C}(\mu)\tilde{x}(t,\mu), \tag{6.2}$$

where $\tilde{A}(\mu)$, $\tilde{E}(\mu) \in \mathbb{R}^{r \times r}$, $\tilde{B}(\mu) \in \mathbb{R}^{r \times m}$ and $C(\mu) \in \mathbb{R}^{l \times r}$ with $r \ll N$. This model can be computed by balanced truncation, a probably most effective model order reduction method for linear control systems. This method is based on the controllability and observability Gramians $X(\mu)$ and $Y(\mu)$ defined as the solutions of the PALE (1.1) and the dual PALE

$$A^T(\mu)Y(\mu)E(\mu) + E^T(\mu)Y(\mu)A(\mu) = -C^T(\mu)C(\mu), \tag{6.3}$$

respectively. For the brevity of this paper, we do not review the balanced truncation method here, but refer the reader who is not familiar with this method to [2, 29].

Based on the derived results, we develop a so-called parametric balanced truncation method as follows. One can observe that for the parametric system (6.1), the balanced truncation method admits the offline-online decomposition. In the offline phase, we determine the reduced basis matrices $V_X \in \mathbb{R}^{N \times n_X}$ and $V_Y \in \mathbb{R}^{N \times n_Y}$ by the Greedy algorithm applied to the PALEs (1.1) and (6.3), respectively. Then in the online phase, for any $\mu \in \mathcal{D}$, we first find the approximate Gramians

$$X(\mu) \approx V_X Z_X(\mu)Z_X^T(\mu)V_X^T, \qquad Y(\mu) \approx V_Y Z_Y(\mu)Z_Y^T(\mu)V_Y^T,$$

where $\hat{X}(\mu) = Z_X(\mu)Z_X^T(\mu)$ and $\breve{Y}(\mu) = Z_Y(\mu)Z_Y^T(\mu)$ solve, respectively, the reduced Lyapunov equations (4.3) with $\hat{A}(\mu) = V_X^T A(\mu)V_X$, $\hat{E}(\mu) = V_X^T E(\mu)V_X$ and $\hat{B}(\mu) = V_X^T B(\mu)$ and

$$\breve{A}^T(\mu)\breve{Y}(\mu)\breve{E}(\mu) + \breve{E}^T(\mu)\breve{Y}(\mu)\breve{A}(\mu) = -\breve{C}^T(\mu)\breve{C}(\mu) \tag{6.4}$$

with $\breve{A}(\mu) = V_Y^T A(\mu)V_Y$, $\breve{E}(\mu) = V_Y^T E(\mu)V_Y$ and $\breve{C}(\mu) = C(\mu)V_Y$. Computing the SVD

$$Z_Y^T(\mu)V_Y^T E(\mu)V_X Z_X(\mu) = \sum_{j=1}^{n_E} \theta_j^E(\mu)Z_Y^T(\mu)V_Y^T E_j V_X Z_X(\mu) \tag{6.5}$$

$$= [U_1(\mu), \ U_2(\mu)] \begin{bmatrix} \Sigma_1(\mu) & 0 \\ 0 & \Sigma_2(\mu) \end{bmatrix} [V_1(\mu), \ V_2(\mu)]^T,$$

with $\Sigma_1(\mu) \in \mathbb{R}^{r \times r}$, we obtain the projection matrices

$$W(\mu) = V_Y Z_Y(\mu) U_1(\mu) \Sigma_1^{-1/2}(\mu), \quad T(\mu) = V_X Z_X(\mu) V_1(\mu) \Sigma_1^{-1/2}(\mu).$$

Then the reduced-order system matrices in (6.2) take the form

$$\begin{aligned}
\tilde{E}(\mu) &= W^T(\mu) E(\mu) T(\mu), & \tilde{A}(\mu) &= W^T(\mu) A(\mu) T(\mu), \\
\tilde{B}(\mu) &= W^T(\mu) B(\mu), & \tilde{C}(\mu) &= C(\mu) T(\mu).
\end{aligned}$$

Exploiting the affine structure and the parametric formulations of the Gramians, we get

$$\begin{aligned}
\tilde{E}(\mu) &= \sum_{j=1}^{n_E} \theta_j^E(\mu) \tilde{W}^T(\mu) V_Y^T E_j V_X \tilde{T}(\mu), & \tilde{B}(\mu) &= \sum_{j=1}^{n_B} \theta_j^B(\mu) \tilde{W}^T(\mu) V_Y^T B_j, \\
\tilde{A}(\mu) &= \sum_{j=1}^{n_A} \theta_j^A(\mu) \tilde{W}^T(\mu) V_Y^T A_j V_X \tilde{T}(\mu), & \tilde{C}(\mu) &= \sum_{j=1}^{n_C} \theta_j^C(\mu) C_j V_X \tilde{T}(\mu).
\end{aligned} \tag{6.6}$$

with $\tilde{W}(\mu) = Z_Y(\mu) U_1(\mu) \Sigma_1^{-1/2}(\mu)$ and $\tilde{T}(\mu) = Z_X(\mu) V_1(\mu) \Sigma_1^{-1/2}(\mu)$. All parameter-independent terms should be computed and stored before running the online stage. We summarize the parametric balanced truncation model reduction method as follows.

Offline: Given the parametric system (6.1),
- Compute the reduced basis matrices V_X and V_Y.
- Compute and store all parameter-independent matrices
  $V_X^T E_j V_X$, $V_Y^T E_j V_Y$ and $V_Y^T E_j V_X$ for $j = 1, \ldots, n_E$;
  $V_X^T A_j V_X$, $V_Y^T A_j V_Y$ and $V_Y^T A_j V_X$ for $j = 1, \ldots, n_A$;
  $V_X^T B_j$ and $V_Y^T B_j$ for $j = 1, \ldots, n_B$; $C_j V_X$ and $C_j V_Y$ for $j = 1, \ldots, n_C$.

Online: Given $\mu \in \mathcal{D}$,
- Compute $\hat{A}(\mu) = V_X^T A(\mu) V_X$, $\hat{E}(\mu) = V_X^T E(\mu) V_X$, $\hat{B}(\mu) = V_X^T B(\mu)$ and $\breve{A}(\mu) = V_Y^T A(\mu) V_Y$, $\breve{E}(\mu) = V_Y^T E(\mu) V_Y$, $\breve{C}(\mu) = C(\mu) V_Y$ using precomputed parameter-independent terms and the affine structure.
- Solve (4.3) and (6.4) for the Cholesky factors $Z_X(\mu)$ and $Z_Y(\mu)$, respectively.
- Compute the SVD (6.5).
- Compute the reduced-order model (6.2), (6.6).

Ignoring the small numbers n_A, n_E, n_B, and n_C, one can verify that the computational complexity of the online stage does not exceed $\mathcal{O}(n_X^3 + n_Y^3)$.

REMARK 6.1. *Note that the parametric reduced-order system (6.2) can also be determined by applying the reduced basis method directly to (6.1), see [11]. However, the model reduction approach presented here has several advantages over that method. It has the control-theoretic background and gains all benefits of balanced truncation. Furthermore, in contrast to the reduced basis method applied to the dynamical system, our approach does not rely on the state snapshots and the projection subspaces are independent of the training input.*

7. Numerical examples. In this section, we present some results of numerical experiments to demonstrate the properties of the reduced basis method. For solving the Lyapunov equations for fixed parameter values, we use the LR-ADI method as described in [5]. All computations are performed with MATLAB on a laptop using Windows 8.1, equipped with 2.40 GHz 8 GB Intel(R) Core(TM) i7-4500U CPU.

7.1. A heat equation. The first model is taken from [17]. Consider the heat equation

$$\frac{\partial \vartheta}{\partial t} - \nabla(\sigma(\xi)\nabla\vartheta) = f \quad \text{in} \quad \Omega \times (0,T),$$
$$\vartheta = 0 \quad \text{on} \quad \partial\Omega \times (0,T),$$

(7.1)

with the heat conductivity coefficient

$$\sigma(\xi) = \begin{cases} 1 + \mu_i & \text{for} \quad \xi \in D_i, \ i = 1, \ldots, 4, \\ 1 & \text{for} \quad \xi \in \Omega \backslash (\cup_{i=1}^4 D_i), \end{cases}$$

(7.2)

where $D_i \subset \Omega = (0,4)^2$, $i = 1, \ldots, 4$, are four discs of radius 0.5 centered at $(1,1)$, $(3,1)$, $(1,3)$ and $(3,3)$, respectively, and the parameter $\mu = [\mu_1, \mu_2, \mu_3, \mu_4]^T$ varies in $\mathcal{D} = [0.1, 10]^4$. Equation (7.1) with the source term $f \equiv 1$ is discretized using the finite element method with piecewise linear basis functions resulting in a system (6.1) of dimension $N = 1580$ with the symmetric positive definite mass matrix $E(\mu) \equiv E$ and the stiffness matrix

$$A(\mu) = \mu_1 A_1 + \mu_2 A_2 + \mu_3 A_3 + \mu_4 A_4 + A_5,$$

(7.3)

where A_i, $i = 1, \ldots, 4$, are symmetric negative semidefinite, and A_5 is symmetric negative definite. The input matrix $B(\mu) \equiv B \in \mathbb{R}^N$ originates from the source function f, and the output matrix $C(\mu) \equiv C \in \mathbb{R}^{1 \times N}$ is chosen at random.

For solving the controllability and observability PALEs (1.1) and (6.3), we employ the reduced basis method with the same setting for both equations: the training set $\mathcal{D}_{\text{train}}$ consisting of 10000 random points, the tolerance $tol_{\text{rb}} = 10^{-4}$ for stopping the Greedy iteration, the maximal number of iterations $k_{\max} = 40$, the test set $\mathcal{D}_{\text{test}}$ containing 50 random points (different from that in the training set) on which the online phase is run, and the tolerance $tol_{\text{cc}} = 10^{-10}$ for column compression in the online phase. In Figure 7.1(a), we present the convergence history of the largest error estimate in the Greedy algorithms for both PALEs. One can observe that the convergence of the Greedy algorithm is not very satisfactory: the iteration indeed stops before the tolerance is reached. The true errors in the solutions are, however, acceptable, and the computed reduced bases provide reasonably good approximations to the solutions for the test parameter samples. Figure 7.1(b) shows the absolute errors $\|X(\mu) - X_{\text{RB}}(\mu)\|_F$ and $\|X(\mu) - \hat{X}_{\text{RB}}(\mu)\|_F$, the error estimates $\Delta_k(\mu)$ and $\hat{\Delta}_k(\mu)$ as in (3.9) and (4.5), respectively, and the error efficiencies

$$\frac{\gamma_{\text{UB}}(\mu)}{\alpha_{\text{LB}}(\mu)} \|X(\mu) - X_{\text{RB}}(\mu)\|_F, \qquad \frac{\gamma_{\text{UB}}(\mu)}{\alpha_{\text{LB}}(\mu)} \|X(\mu) - \hat{X}_{\text{RB}}(\mu)\|_F$$

for the controllability PALE (1.1) with $\mu \in \mathcal{D}_{\text{test}}$. One can see that the corresponding errors and error estimates are different from each other by a multiplicative factor of about 10^3. Moreover, $\|X(\mu) - \hat{X}_{RB}(\mu)\|_F$ is smaller than $\|X(\mu) - X_{RB}(\mu)\|_F$ which supports our observation in Section 4.2 and encourages the use of the method by the fact that the error in the online phase is even smaller than that in the offline phase.

Finally, by our setting of inputs, we reduce the system order from 1580 down to about 20 (the reduced orders may vary for different parameters). Figure 7.2(a) shows the relative errors for the solution of the PALEs (1.1) and (6.3) in the online phase. In Figure 7.2(b), we present the approximate $\mathcal{H}_\infty$-norm of the absolute error in the
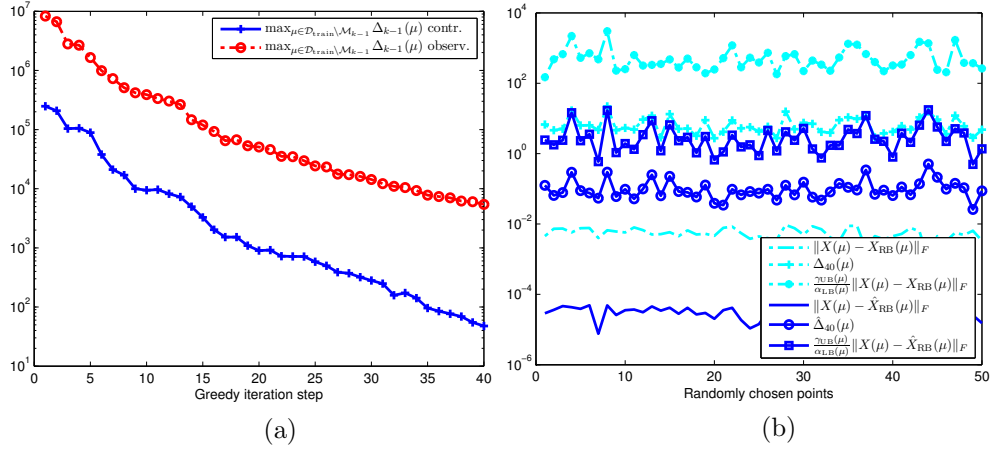
FIG. 7.1. *Heat equation: (a) the largest error estimate during the Greedy iteration; (b) errors, error estimates and error efficiencies measured in two ways for the controllability PALE (1.1).*
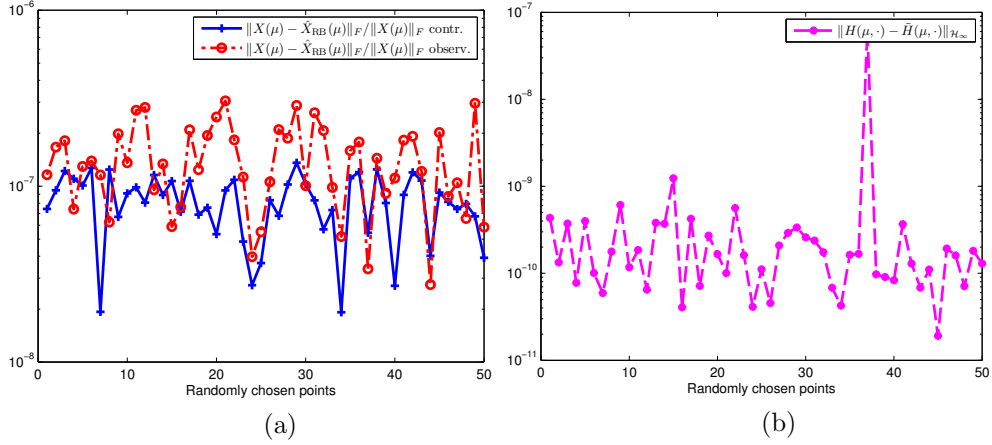

FIG. 7.2. *Heat equation: (a) the relative errors of approximate solutions of the controllability and observability PALEs; (b) the absolute error of parametric model order reduction.*

frequency response defined as

$$
\begin{aligned}
\|H(\mu,\cdot) - \tilde{H}(\mu,\cdot)\|_{\mathcal{H}_\infty} &= \sup_{\omega \in \mathbb{R}} \|H(\mu,i\omega) - \tilde{H}(\mu,i\omega)\|_2 \\
&\approx \sup_{\omega_j \in [\omega_{\min},\omega_{\max}]} \|H(\mu,i\omega_j) - \tilde{H}(\mu,i\omega_j)\|_2,
\end{aligned}
$$

where $H(\mu,s) = C\big(sE - A(\mu)\big)^{-1}B$ and $\tilde{H}(\mu,s) = \tilde{C}\big(s\tilde{E} - \tilde{A}(\mu)\big)^{-1}\tilde{B}$ are the transfer functions of the original and the reduced-order systems and $\mu \in \mathcal{D}_{\text{test}}$. The relative errors, which are not depicted here, are even smaller as the full response of the original system vary between 1 and 10.

7.2. An anemometer model. We consider now an anemometer model describing a thermal based flow sensor, see [21] and references therein. Simulation of this

device requires solving a convection-diffusion partial differential equation of the form

$$\rho c\frac{\partial \vartheta}{\partial t} = \nabla(\kappa \nabla \vartheta) - \rho c v \nabla \vartheta + \dot{q}, \tag{7.4}$$

where ρ denotes the mass density, $c \in [0,1]$ is the specific heat, $\kappa \in [1,2]$ is the thermal conductivity, $v \in [0.1,2]$ is the fluid velocity, ϑ is the temperature, and $\dot{q}$ is the heat flow into the system caused by the heater. The considered model is restricted to the case $\rho = 1$. The finite element discretization of (7.4) leads to system (6.1) of order $N = 29008$ with the mass matrix $E(\mu) = E_1 + cE_2$, where E_1 and E_2 are symmetric positive definite and the stiffness matrix $A(\mu) = A_1 + kA_2 + cvA_3$, where A_1 is symmetric negative definite, A_2 is nonsymmetric but negative semidefinite, A_3 is symmetric negative semidefinite, and $\mu = [c,k,v]^T$. The input matrix $B \in \mathbb{R}^N$ and the output matrix $C \in \mathbb{R}^{1\times N}$ are parameter-independent. The data can be downloaded from [1].

In this example, we want to test the error for nonsymmetric system and the reliability of the method when applied to really large system. We run the Greedy algorithm for 20 steps on the training set $\mathcal{D}_{\text{train}}$ with 10000 points and, in the online phase, we examine the error at 50 test points. In both cases, the points are chosen randomly. We compute the errors, the nonsymmetric error estimates and the efficiencies as in Theorem 5.2 and Theorem 5.3, see Figure 7.3(b) and Figure 7.4(a). The situation is quite similar to the previous example except for the fact that the error of the approximate solutions $\hat{X}_{RB}(\mu)$ of the observability PALE (6.3) shown in Figure 7.4(a) is rather large comparing to that of the controllability PALE (1.1). The reason is most probably that the corresponding Greedy iteration stagnates after first five steps, see Figure 7.3(a). Together with the numerical results in Section 7.1, we believe that the reduced basis approximation is better when the maximal error in the Greedy search (almost) monotonously decreases. In Figure 7.3(b), we present the approximate $\mathcal{H}_\infty$-norm of the absolute error for the reduced models at the test parameters. One can see that the errors remain small for all these parameters.

Now we turn our attention to computation time shown in Figure 7.5. One can see that most of the time (63%) used in offline phase is spent for solving the Lyapunov equation at twenty different points seek by the Greedy algorithm. Thanks to suitable arrangement of parameter-dependent and parameter-independent terms, the computation time of seemingly expensive search on the training set can be almost ignored (4%). The time for computing the parameter independent terms is quite remarkable (27%). Note, however, that these terms will be stored and used in the online phase which helps to reduce the cost. To wit, in online phase, for the true error, we have to solve the original Lyapunov equation and the reduced Lyapunov equation at 50 points. The first task takes 1663 sec, while the second one takes only 31 sec, which accelerates the computation by a factor of 53.

8. Conclusion. In this paper, we presented a reduced basis method for solving large-scale PALEs. For deriving the error estimates for approximate solutions to PALEs with symmetric and nonsymmetric matrix coefficients depending affinely on parameters, we used min-Θ approach. The reduced basis method was then used to extend the standard balanced truncation model reduction approach to the parametric systems which does not required interpolation. Numerical examples shown that, on the one hand, the proposed method worked well for large problems, but on the other hand, the error estimate can be poor especially in the nonsymmetric case. Tightening the error estimates by using other matrix norms and/or other techniques such as
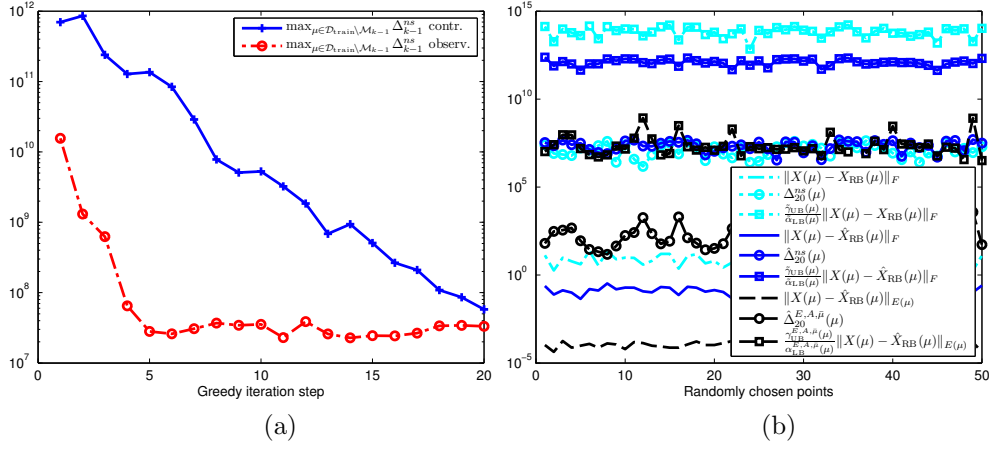
(a)   (b)

FIG. 7.3. *Anemometer model:* (*a*) *the largest error estimate during the Greedy iteration;* (*b*) *errors, error estimates and error efficiencies measured in two ways for the controllability PALE* (1.1).
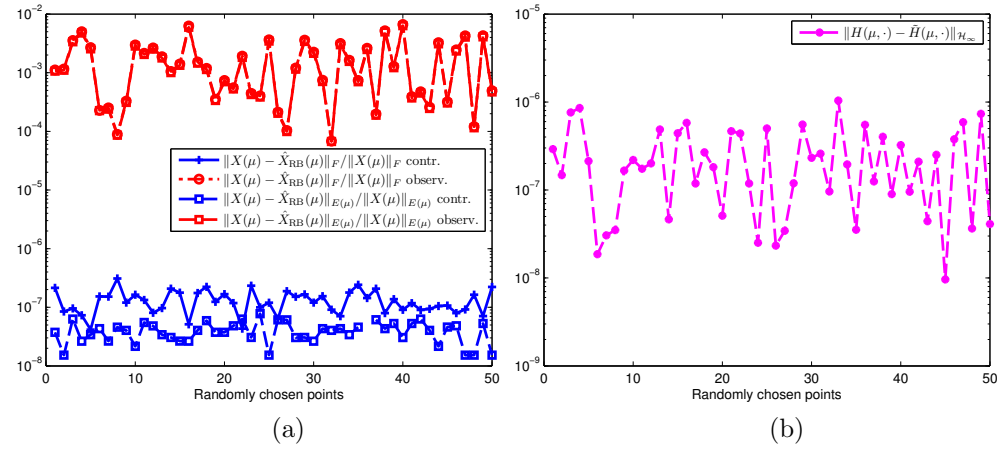

(a)   (b)

FIG. 7.4. *Anemometer model:* (*a*) *the relative errors of approximate solutions of the controllability and observability PALEs; (b) the absolute error of parametric model order reduction.*

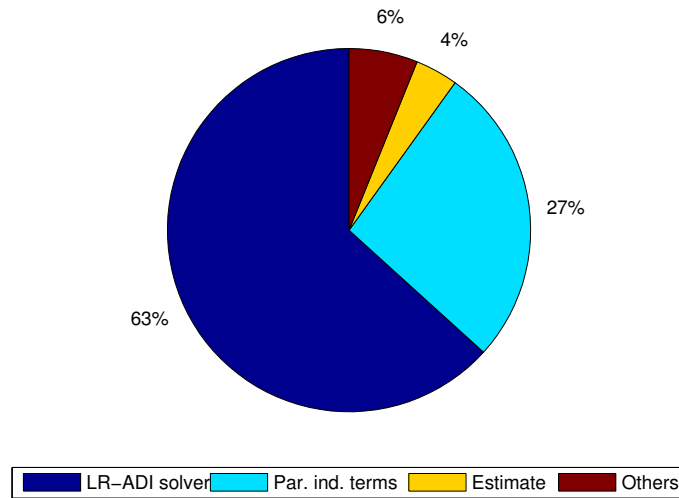natural norm approach, successive constraint method or their combination remains for the future work.

Fig. 7.5. *Anemometer model: time spent for different tasks.*

REFERENCES

[1] *MOR Wiki - Model Order Reduction Wiki.* Url: `http://morwiki.mpi-magdeburg.mpg.de/morwiki/`.

[2] A.C. Antoulas, *Approximation of Large-Scale Dynamical Systems*, SIAM, Philadelphia, PA, 2005.

[3] R.H. Bartels and G.W. Stewart, *Solution of the equation $AX + XB = C$*, Comm. ACM, 15 (1972), pp. 820–826.

[4] U. Baur and P. Benner, *Modellreduktion für parametrisierte Systeme durch balanciertes Abschneiden und Interpolation*, at-Automatisierungstechnik, 57 (2009), pp. 411–419.

[5] P. Benner and J. Saak, *Numerical solution of large and sparse continuous time algebraic matrix riccati and lyapunov equations: a state of the art survey*, GAMM-Mitteilungen, 36 (2013), pp. 32–52.

[6] R. Byers, *Solving the algebraic Riccati equation with the matrix sign function*, Linear Algebra Appl., 85 (1987), pp. 267–279.

[7] G. Dahlquist, *Stability and Error Bounds in the Numerical Integration of Ordinary Differential Equations*, vol. 130 of Transactions of Royal Institute of Technology, Stockholm, 1959.

[8] S. Deperis, *Reduced basis error bound computation of parameter-dependent Navier-Stokes equations by the natural norm approach*, SIAM J. Numer. Anal., 46 (2008), pp. 2039–2067.

[9] V. Druskin, L. Knizhnerman, and V. Simoncini, *Analysis of the rational Krylov subspace and ADI methods for solving the Lyapunov equation*, SIAM J. Numer. Anal., 49 (2011), pp. 1875–1898.

[10] Z. Gajić and M.T.J. Qureshi, *Lyapunov Matrix Equation in System Stability and Control*, Academic Press, San Diego, CA, 1995.

[11] B. Haasdonk and M. Ohlberger, *Efficient reduced models and a posteriori error estimation for parametrized dynamical systems by offline/online decomposition*, Math. Comput. Model. Dyn. Syst., 17 (2011), pp. 145–161.

[12] B. Haasdonk and A. Schmidt, *Reduced basis approximation of large scale parametric algebraic Riccati equations*, SimTech Preprint, Universität Stuttgart, Juni 2015.

[13] S.J. Hammarling, *Numerical solution of the stable non-negative definite Lyapunov equation*, IMA J. Numer. Anal., 2 (1982), pp. 303–323.

[14] R.A. Horn and C.R. Johnson, *Matrix Analysis*, Cambridge University Press, Cambridge, 1985.

[15] ———, *Topics in Matrix Analysis*, Cambridge University Press, Cambridge, 1991.

[16] I.M. Jaimoukha and E.M. Kasenally, *Krylov subspace methods for solving large Lyapunov equations*, SIAM J. Numer. Anal., 31 (1994), pp. 227–251.

[17] D. Kressner, M. Plešinger, and C. Tobler, *A preconditioned low-rank CG method for*

parameter-dependent Lyapunov matrix equations, Numer. Linear Algebra Appl., 21 (2014), pp. 666–684.

[18] P. Lancaster and M. Tismenetsky, *The Theory of Matrices*, Academic Press, Orlando, FL, 2nd ed., 1985.

[19] J.-R. Li and J. White, *Low rank solution of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 24 (2002), pp. 260–280.

[20] A. Lu and E. Wachspress, *Solution of Lyapunov equations by alternating direction implicit iteration*, Comput. Math. Appl., 21 (1991), pp. 43–58.

[21] C. Moosmann, E.B. Rudnyi, A. Greiner, J.G. Korvink, and M. Hornung, *Parameter preserving model order reduction of a flow meter*, in Technical Proceedings of the 2005 Nanotechnology Conference and Trade Show (Nanotech 2005, May 8-12, 2005, Anaheim, California, USA), vol. 3, NSTINanotech, 2005, pp. 684–687.

[22] H. Panzer, B. Kleinherne, and B. Lohmann, *Analysis, interpretation and generalization of strictly dissipative state space formulation of second order systems*, in Methoden und Anwendungen der Regelungstechnik, G. Roppenecker and B. Lohmann, eds., Shaker-Verlag, 2013.

[23] A. T. Patera and G. Rozza, *Reduced Basis Approximation and A Posteriori Error Estimation for Parametrized Partial Differential Equations*, MIT Pappalardo Graduate Monographs in Mechanical Engineering, MIT, MA, 2007.

[24] T. Penzl, *A cyclic low-rank Smith method for large sparse Lyapunov equations*, SIAM J. Sci. Comp., 21 (1999/2000), pp. 1401–1418.

[25] Y. Saad, *Numerical solution of large Lyapunov equations*, in Signal Processing, Scattering, Operator Theory, and Numerical Methods (Amsterdam, 1989), M.A. Kaashoek, J.H. Van Schuppen, and A.C.M. Ran, eds., Birkhäuser, Boston, MA, 1990, pp. 503–511.

[26] V. Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput., 29 (2007), pp. 1268–1288.

[27] ———, *Computational methods for linear matrix equations*, Technical Report, 2013.

[28] G. Söderlind, *The logarithmic norm. History and modern theory*, BIT, 46 (2006), pp. 631–652.

[29] M.S. Tombs and I. Postlethweite, *Truncated balanced realization of a stable non-minimal state-space system*, Internat. J. Control, 46 (1987), pp. 1319–1330.

[30] B. Vandereycken and S. Vandewalle, *A Riemannian optimization approach for computing low-rank solutions of Lyapunov equations*, SIAM J. Matrix Anal. Appl., 31 (2010), pp. 2553–2579.

[31] K. Veroy, C. Prud'home, and A.T. Patera, *Reduced-basis approximation of the viscous Burgers equation: rigorous a posteriori error bounds*, C. R. Acad. Sci. Paris, Ser. I, 337 (2003), pp. 619–624.

[32] K. Veroy, C. Prud'home, D. Rovas, and A.T. Patera, *A posteriori error bounds for reduced-basis approximation of parametrized noncoercive and nonlinear elliptic partial differential equations*, in Proceedings of the 16th AIAA Computational Fluid Dynamics Conference, vol. AIAA, 2003, pp. Paper 2003–3847.

[33] E.L. Wachspress, *Iterative solution of the Lyapunov matrix equation*, Appl. Math. Lett., 1 (1988), pp. 87–90.

[34] B. Zhou, G. Duan, and Z. Lin, *A parametric Lyapunov equation approach to the design of low gain feedback*, IEEE Trans. Automat. Control, 53 (2008), pp. 1548–1554.