Hamburger Beiträge
zur Angewandten Mathematik

Finite-Rank ADI Iteration for Operator Lyapunov Equations

Timo Reis and Winnifried Wollner

Nr. 2012-09
July 2012

FINITE-RANK ADI ITERATION FOR OPERATOR LYAPUNOV EQUATIONS

TIMO REIS*† AND WINNIFRIED WOLLNER*‡

Abstract. We give an algorithmic approach to the approximative solution of operator Lyapunov equations for controllability. Motivated by the successfully applied *alternating direction implicit (ADI)* iteration for matrix Lyapunov equations, we consider this method for the determination of Gramian operators of infinite-dimensional control systems. In the case where the input space is finite-dimensional, we show that this method provides approximative solutions of finite rank. Convergence in several norms is shown.
Particular emphasis is placed on systems governed by a heat equation with boundary control. We present that ADI iteration for the heat equation consists of solving a sequence of Helmholtz equations. A numerical example is presented.

Key words. Lyapunov equation, ADI iteration, numerical method in control theory, infinite-dimensional linear systems theory, heat equation

AMS subject classifications.

1. Introduction. A fundamental concept in linear finite-dimensional systems theory is the *Gramian matrix*, i.e., the solution $P \in \mathbb{R}^{n\times n}$ of the *Lyapunov equation*

$$AP + PA^T + BB^T = 0 \tag{1.1}$$

associated to a linear control system $\dot{x}(t) = Ax(t)+Bu(t)$ with $A \in \mathbb{R}^{n\times n}$, $B \in \mathbb{R}^{n\times m}$. These equations for instance arise in stability and controllability analysis [39, Sec. 3.8] and model reduction by balanced truncation [3, Chap. 7]. Due to their importance, a variety of numerical methods have been developed for Lyapunov equations, such as alternating direction implicit (ADI) iteration [22], Bartels-Stewart method [9], Smith's method [37], Krylov subspace method [21, 34], sign function method [30], and Hammarling's method [19] (see [3, Chap. 6] for an overview). Especially the ADI iteration, Smith method and matrix sign function method have in common that, in case of $m \ll n$, they typically provide so-called *low-rank approximative solutions*. That is, instead of the full Gramian matrix, a factor $S \in \mathbb{R}^{n\times k}$ with $k \ll n$ and $P \approx SS^T$ is computed iteratively. Besides memory savings, the advantage of low-rank approximative solutions is that they can be directly and efficiently used for balanced truncation model reduction without any evaluation of the Gramian matrix itself [38]. This feature makes these methods suitable for problems of large state space dimension $n \in \mathbb{N}$. An important class of large-scale systems are those emerging from fine spatial discretization of controlled systems which are governed by linear partial differential equations [10]. The latter however actually has (before discretization) infinite state space dimension; in the Lyapunov equation, the variables A, B and the to-be solved P are actually operators acting on infinite-dimensional spaces. It is hence natural to wonder about the following questions:

a) Can iterative algorithms be formulated for operator Lyapunov equations, and (when) do they converge?
b) What are the (computational) consequences for systems governed by PDEs?

*Fachbereich Mathematik, Universität Hamburg, Bundesstraße 55, 20146 Hamburg, Germany.
†timo.reis@math.uni-hamburg.de
‡winnifried.wollner@math.uni-hamburg.de

In this work we consider the ADI iteration for operator Lyapunov equations. Consequently, this iteration provides a factorization $P \approx SS^*$, where S is an operator acting on some space which is finite-dimensional, if the input space will be finite-dimensional. Convergence will be shown under only the assumption that the system $\dot{x}(t) = Ax(t)$ is exponentially stable and that the Gramian operator exists. We allow the input operator to be "unbounded", which is motivated by partial differential equations with boundary control. Particular emphasis is placed on systems governed by the boundary controlled heat equation, where it will turn out that the ADI method is consisting of the solution of a sequence of Helmholtz equations. Since the latter can be (approximatively) solved by using adaptive finite element methods, we will also discuss the impact of approximative solution in each step of the ADI iteration.

Let us mention that the general idea of transferring existing algorithms for matrix Lyapunov equations to the infinite-dimensional case is not new: The method of "proper orthogonal decomposition (POD)" was used [31, 44] to obtain low-rank approximative solutions of matrix Lyapunov equations. For $B = [b_1, \ldots, b_m] \in \mathbb{R}^{n,m}$, this method uses the representation $P = \int_0^\infty \sum_{k=1}^m x_k(t)x_k(t)^T$, where x_k solves the differential equation $\dot{x}_k(t) = Ax_k(t)$, $x_k(0) = b_k$. Approximative solutions of low rank are computed by determining a dominant subspace that is based on a singular value decomposition of a matrix that consists of several sampled values (so-called "snapshots") of the trajectories x_k. This method has been generalized to the infinite-dimensional case in [35] and has been applied to model reduction and linear-quadratic optimal control in [36]. In particular, convergence has been proven and confirmed by numerical examples. A slight drawback of this method that it is not directly generalizable to unbounded control operators B.

The paper is organized as follows. The subsequent Section 2 reviews the basic notational and functional analytic framework. Section 3 contains basic facts about semigroups and operator Lyapunov equations. In Section 4 we introduce the ADI iteration for the solution of operator Lyapunov equations and present results about convergence. In Section 5 we expand our analysis to an inexact ADI iteration that one would have to do in practical computations due to the necessity of discretization in the infinite-dimensional context. In Section 6 the developed theory is applied to a heat equation with Robin boundary control. A numerical example is presented.

2. Basic notation and functional analytic prerequisites. Throughout the paper $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$, $\mathbb{C}_+$, $\mathbb{C}_-$ and $\mathbb{C}^{n\times m}$ respectively denote the sets of positive real, nonnegative real, complex numbers with positive real part, complex numbers with negative real part, and the space of $n \times m$ complex matrices. $\mathbb{N}$ stands for the set of positive integers and by $\overline{z}$ we mean the complex conjugate of $z \in \mathbb{C}$.

For $p \geq 1$, ℓ_p stands for the p-summable complex sequences. We use the notation from [1] for Lebesgue and Sobolev spaces $L_p(\Omega)$ and $H^k(\Omega)$.

Throughout this work, integrals of functions with values in Hilbert space are understood in the sense of *Bochner*. For a brief overview on abstract integration theory we refer to [14, pp. 621] and the bibliography therein. For $p \in [1, \infty]$, some interval I and some separable Hilbert space X, $L_p(I, X)$ denotes the Lebesgue space of measurable functions $f : I \to X$ with the property that $\|f(\cdot)\|_X \in L_p(I)$.

Let Z and X be Hilbert spaces, such that $Z \subset X$ and the canonical injection $Z \to X$, $x \mapsto x$ is continuous and dense. By calling a Hilbert space X *pivot space*, we mean that X is identified its own topological dual X' (which is possible by the Riesz representation theorem [27, pp. 48]), and the dual of Z is defined in a way that the dual pairing $\langle \cdot, \cdot \rangle_{Z',Z}$ continuously (w.r.t. the norm in Z) extends the inner product in

X. It follows that $X \subset Z'$ and the canonical injection $X \to Z'$, $x \mapsto x$ is continuous and dense.

$\mathcal{B}(X, Y)$ and $\mathcal{K}(X, Y)$ are the spaces of bounded, resp. compact, linear operators $T : X \to Y$, and we abbreviate $\mathcal{B}(X) := \mathcal{B}(X, X)$, $\mathcal{K}(X) := \mathcal{K}(X, X)$. For a densely defined operator $T : D(T) \subset X \to X$, the symbols $\rho(T)$ and $\sigma(T)$ indicate its resolvent set and spectrum, respectively. The identity mapping on X is denoted by I_X and the zero operator from X to Y by $0_{X,Y}$. Given an operator $T : D(T) \subset X \to Y$, the *graph norm* is defined via $\|x\|^2_{D(T)} = \|x\|^2_X + \|Tx\|^2_Y$. If $D(T)$ associated with the graph norm $\|\cdot\|_{D(T)}$ is complete, then T is called *closed*.

A vector $v \in X$ is in a canonical way identified as an operator $v \in \mathcal{B}(\mathbb{C}, X)$ via $\lambda \mapsto \lambda v$. For a Hilbert space X and $m \in \mathbb{N}$, the product space X^m is equipped with the canonical inner product. For another Hilbert space Y and operators $T_1, \ldots, T_m \in \mathcal{B}(X, Y)$, the operator column matrix

$$T = \begin{bmatrix} T_1 & \cdots & T_m \end{bmatrix}$$

defines an operator $T \in \mathcal{B}(X^m, Y)$ in a straightforward manner.

The *adjoint* of $T \in \mathcal{B}(X, Y)$ is denoted by $T^* \in \mathcal{B}(Y, X)$ and the *dual* by $T' \in \mathcal{B}(Y', X')$. The adjoint of a densely defined operator $T : D(T) \subset X \to Y$, is defined on $T : D(T^*) \subset Y \to X$, where $D(T^*)$ consists of all $y \in Y$ with the property that that there exists some $z \in X$ with $\langle Tx, y\rangle_X = \langle x, z\rangle_X$ for all $x \in D(T)$ (in this case, we define $T^*y = z$). The dual of a densely defined operator $T : D(T) \subset X \to Y$, is defined on $T : D(T') \subset Y' \to X'$, where $D(T')$ consists of all $y \in Y$ with the property that that the mapping $D(T) \to \mathbb{C}$, $z \mapsto \langle Tz, x\rangle_X$ has an extension to an element in X'. For $y \in D(T')$, the element $T'y$ is defined via $\langle T'y, x\rangle_{X',X} = \langle y, Tx\rangle_{Y',Y}$ for all $x \in D(T)$. Note that T^* and T' coincide if both X and Y are considered to be pivot spaces. For further details concerning duals and adjoints, we refer to [5, pp. 49].

A densely defined operator $P : D(P) \subset X \to X$ is called *self-adjoint*, if $P = P^*$ (this also includes that $D(P) = D(P^*)$). A self-adjoint operator P is *nonnegative* if $\langle x, P_1 x\rangle_X \geq 0$ for all $x \in D(P)$. The notions of negativity, positivity and nonpositivity of an operator can be defined in straightforward manner. This induces a partial order on the set of self-adjoint operators: For two self-adjoint operators $P_1 : D(P_1) \subset X \to X$, $P_2 : D(P_2) \subset X \to X$ we say that $P_1 \geq P_2$, if $P_1 - P_2 \geq 0$. The *square root* of a nonnegative operator $P : D(P) \subset X \to X$ is denoted by $P^{1/2}$; its domain $D(P^{1/2})$ is the completion of $D(P)$ with the norm $\|x\|^2_{D(P^{1/2})} = \|x\|^2_X + \langle x, Px\rangle_X$ [14, p. 606].

Compact operators are known to admit a singular value decomposition

$$Tx = \sum_{i=1}^{\infty} \sigma_i \langle x, u_i\rangle_X \cdot v_i, \tag{2.1}$$

where the sequence of singular values $(\sigma_i)_i$ is monotonically decreasing and tends to zero, and $(u_i)_i$, $(v_i)_i$ are orthonormal systems in X and Y, respectively [27, pp. 203].

Subsequently, we introduce special classes and norms of operators which were originally introduced in [33].

Definition 2.1. *Let X, Y be separable Hilbert spaces and let $p \in [1, \infty[$. Then $T \in \mathcal{K}(X)$ is called a p-th* Schatten class operator*, if the sequence consisting of its singular values fulfill $(\sigma_i)_i \in \ell_p$. In this case we write $T \in \mathcal{S}_p(X, Y)$. Provided with the norm $\|T\|_{\mathcal{S}_p(X,Y)} = \|(\sigma_i)_i\|_{\ell_p}$ the space $\mathcal{S}_p(X, Y)$ becomes a Banach space. Operators of first Schatten class are called* nuclear *and those of second Schatten class are called* Hilbert-Schmidt*.*

We abbreviate $\mathcal{S}_p(X) := \mathcal{S}_p(X, X)$. For more details on the Schatten class, we refer to [24, pp. 126]. The *trace* of $T \in \mathcal{S}_1(X)$ is well-defined by the expression

$$\mathrm{tr}(T) = \sum_{i=1}^{\infty} \langle e_i, Te_i \rangle , \tag{2.2}$$

where (e_i) is an (arbitrary) orthonormal basis of X [27, pp. 206]. For self-adjoint and nonnegative $P \in \mathcal{S}_1(X)$, the spectral theorem implies that $\|P\|_{\mathcal{S}_1(X)} = \mathrm{tr}(P)$. Moreover, for $T \in \mathcal{S}_2(X, Y)$ it holds $T^*T \in \mathcal{S}_1(X)$, $TT^* \in \mathcal{S}_1(Y)$ with $\|T\|^2_{\mathcal{S}_2(X,Y)} = \|T^*\|^2_{\mathcal{S}_2(Y,X)} = \mathrm{tr}(T^*T) = \mathrm{tr}(TT^*)$.

3. Operator Lyapunov equations. We review basic facts about solvability of operator Lyapunov equations. Consider the following setup throughout this article: For Hilbert spaces U, X (which are assumed to be pivot spaces), let $A : D(A) \subset X \to X$ and $B \in \mathcal{B}(U, D(A^*)')$ be given. The operator Lyapunov equation is given by

$$2 \,\mathrm{Re} \, \langle Px, A^*x \rangle_X + \|B'x\|^2_U = 0 \quad \text{for all } x \in D(A^*) \tag{3.1}$$

and has to be solved for the self-adjoint operator $P \in \mathcal{B}(X)$. Indeed, (3.1) is equivalent to (1.1) in the case where A and B are real matrices. The property of B to possibly map to a larger space $D(A^*)' \supset X$ is motivated by partial differential equations with boundary control, see [12] and [42, Chap. 10].

To analyze solvability of operator Lyapunov equations we first need to introduce the concept of strongly continuous semigroups, stability and admissibility:

DEFINITION 3.1 (Strongly continuous semigroups, generators, exponential stability). *An operator-valued function* $T(\cdot) : \mathbb{R}_{\geq 0} \to \mathcal{B}(X)$ *is called* strongly continuous semigroup, *if* $T(0) = I_X$, $T(t+s) = T(t) \cdot T(s)$ *for all* $t, s \in \mathbb{R}_{\geq 0}$, *and*

$$\lim_{t \to 0, t>0} T(t)x = x \quad \text{for all } x \in X.$$

A strongly continuous semigroup is called exponentially stable, *if there exists some* $M \in \mathbb{R}_{\geq 0}$, $\omega \in \mathbb{R}_{>0}$ *such that*

$$\|T(t)\|_{\mathcal{B}(X)} \leq M \cdot e^{-\omega t} \quad \text{for all } t \in \mathbb{R}_{\geq 0}.$$

The operator $A : D(A) \subset X \to X$ *defined by*

$$\begin{aligned} Ax &= \lim_{t \to 0, t>0} \tfrac{1}{t}(T(t)x - x), \\ D(A) &= \left\{ x \in X \,\middle|\, \lim_{t \to 0, t>0} \tfrac{1}{t}(T(t)x - x) \in X \right\} \end{aligned}$$

is called generator *of the semigroup* $T(\cdot)$.

The domains of A and its adjoint are known to be dense in X [42, Cor. 2.1.8 & Prop. 2.8.1.]; in particular, every strongly continuous semigroup possesses a generator.

DEFINITION 3.2 (Admissible control operator, Gramian operator). *Let* U, X *be Hilbert spaces, let* $A : D(A) \subset X \to X$ *be the generator of a strongly continuous semigroup* $T(\cdot)$ *on* X, *and let* $B \in \mathcal{B}(U, D(A^*)')$. *Then we call* B *an* admissible control operator *for* $T(\cdot)$, *if for some (and then also any)* $t \in \mathbb{R}_{>0}$, *there holds*

$$\Phi_t u := \int_0^t T(\tau)Bu(\tau)d\tau \in X \quad \text{for all } u \in L_2(\mathbb{R}_{\geq 0}, U). \tag{3.2}$$

The control operator B is called infinite-time admissible*, if*

$$\Phi u := \int_0^\infty T(\tau)Bu(\tau)d\tau \in X \quad \textit{for all } u \in L_2(\mathbb{R}_{\geq 0}, U). \tag{3.3}$$

If B is infinite-time admissible for $T(\cdot)$, then

$$P = \Phi\Phi^* \in \mathcal{B}(X)$$

is called Gramian *of (A, B).*

Some comments on facts about the above definition are stated below.

Remark 3.3.

a) *Expression 3.2 has to be understood in the following way: As $T(\cdot)$ extends to a strongly continuous semigroup on $D(A^*)'$ [42, Prop. 2.10.4], the function $T(\cdot)Bu(\cdot)$ is $D(A^*)'$-valued and measurable. Admissibility means that the integral is even in the smaller space X.*

b) *Admissibility implies $\Phi_t \in \mathcal{B}(L_2(\mathbb{R}_{\geq 0}, U), X)$ [20, p. 6], infinite-time admissibility implies $\Phi \in \mathcal{B}(L_2(\mathbb{R}_{\geq 0}, U), X)$ [20, p. 5].*

c) *If B is admissible for an exponentially stable semigroup $T(\cdot)$, then B is infinite-time admissible [20, p. 6]. Indeed, we will assume exponential stability in most of our results.*

d) *Any $B \in \mathcal{B}(U, X)$ is admissible.*

Note that fully dual statements hold true for observability Gramians [42, pp. 134]; all results in this article can be formulated for that case in a straightforward manner.

Next, recall that the Gramian indeed solves the operator Lyapunov equation (3.1) see [20]. Note that in [20] a more general context is considered in which exponential stability is not presumed.

THEOREM 3.4. *[20, Thm. 3.1] Let U, X be Hilbert spaces and $A : D(A) \subset X \to X$ be the generator of an exponentially stable semigroup $T(\cdot)$ on X. Then there holds:*

a) *If $B \in \mathcal{B}(U, D(A^*)')$ is an admissible control operator for $T(\cdot)$, then the Gramian P of (A, B) is the unique self-adjoint and nonnegative solution of the operator Lyapunov equation* (3.1)*.*

b) *On the other hand, if $B \in \mathcal{B}(U, D(A^*)')$ and there exists some nonnegative self-adjoint $Q \in \mathcal{B}(X)$ that satisfies*

$$2\,\mathrm{Re}\,\langle Qx, A^*x\rangle_X + \|B'x\|_U^2 \leq 0 \quad \textit{for all } x \in D(A^*) \tag{3.4}$$

then B is an admissible control operator for $T(\cdot)$.

Remark 3.5. Being aware of $2\,\mathrm{Re}\,\langle Px, A^*x\rangle_X = \langle Px, A^*x\rangle_X + \langle A^*x, Px\rangle_X$, an application of $x + y$ and $x + iy$ to (3.1) *implies that the operator Lyapunov equation is equivalent to*

$$\langle Px, A^*y\rangle_X + \langle A^*x, Py\rangle_X + \langle B'x, B'y\rangle_U = 0 \quad \textit{for all } x, y \in D(A^*). \tag{3.5}$$

We briefly focus on the case where the generator A is self-adjoint and negative, and the control operator fulfills $B \in \mathcal{B}(U, D(((-A)^{\frac{1}{2}})'))$ (the heat equation considered in Sec. 6 is of this type). The latter property is equivalent to

$$(-A)^{-\frac{1}{2}}B \in \mathcal{B}(U, X). \tag{3.6}$$

PROPOSITION 3.6. *Assume that $A : D(A) \subset X \to X$ is self-adjoint, negative, and it has compact resolvent. Let $B : U \to D(A)'$ such that* (3.6) *holds true. Then the following holds true:*

a) A generates an exponentially stable semigroup $T(\cdot)$ on X.
b) B is admissible for $T(\cdot)$.
Moreover, the Gramian P of (A, B) is nuclear if and only if $(-A)^{-1/2}B : U \to X$ is Hilbert-Schmidt. In this case, there holds

$$\|P\|_{\mathcal{S}_1(X)} = \operatorname{tr}(P) = \tfrac{1}{2} \cdot \|(-A)^{-1/2}B\|^2_{\mathcal{S}_2(U,X)} = -\tfrac{1}{2} \cdot \operatorname{tr}(B'A^{-1}B). \tag{3.7}$$

Proof. Since A is negative and has compact resolvent, there exists some $\nu \in \mathbb{R}_{>0}$ with $\langle x, Ax\rangle \leq -\nu\|x\|^2_X$ for all $x \in D(A)$. Then assertion a) follows from the Lumer-Phillips theorem [23, pp. 76]. We can further infer from (3.6) that there exists some $\mu \in \mathbb{R}_{>0}$ such that for all $z \in X$ holds

$$\|B'(-A)^{-\frac{1}{2}}z\|^2_U \leq \mu\|z\|^2_X.$$

Since this inequality holds clearly for all $z \in D(A^{\frac{1}{2}})$ we can perform the substitution $x = (-A)^{-\frac{1}{2}}z$ to see that for all $x \in D(A) = D(A^*)$ holds

$$\|B'x\|^2_U \leq \mu \cdot \|(-A)^{\frac{1}{2}}x\|^2_X = -\mu \cdot \langle x, Ax\rangle_X .$$

That is, (3.4) holds true for $\Pi = \frac{\mu}{2}I_X$, and we may apply Theorem 3.4 c) to conclude admissibility of B.

Since A is negative compact resolvent, there exists some orthonormal basis $(e_i)_i$ of X and sequence of negative real numbers $(\lambda_i)_i$ with $\lim_{i\to\infty} \lambda_i = -\infty$ such that $e_i \in D(A)$ and $Ae_i = \lambda_i e_i$. Then

$$\begin{aligned} 0 =& 2\operatorname{Re}\langle Ae_i, Pe_i\rangle_X + \|B'e_i\|^2_U \\ =& 2\operatorname{Re}\langle \lambda_i e_i, Pe_i\rangle_X + \|B'e_i\|^2_U = 2\lambda_i \langle e_i, Pe_i\rangle_X + 2\|B'e_i\|^2_U. \end{aligned}$$

Solving this equation for $\langle e_i, Pe_i\rangle_X$ and using that $B'(-A)^{-1/2} = ((-A)^{-1/2}B)^*$, we obtain

$$\begin{aligned} \langle e_i, Pe_i\rangle_X =& -\tfrac{1}{2\lambda_i} \cdot \|B'e_i\|^2_U = \tfrac{1}{2} \cdot \|B'(-\lambda_i)^{-1/2}e_i\|^2_U \\ =& \tfrac{1}{2} \cdot \|B'(-A)^{1/2}e_i\|^2_U = \tfrac{1}{2} \cdot \left\langle e_i, (-A)^{-1/2}B\left((-A)^{-1/2}B\right)^* e_i\right\rangle_X . \end{aligned}$$

Since both P and $(-A)^{-1/2}B\left((-A)^{-1/2}B\right)^*$ are nonnegative, their nuclear norms equal to their respective traces. Moreover, since we have that $(-A)^{-1/2}B \in \mathcal{S}_2(U, X)$ if and only if $(-A)^{-1/2}B((-A)^{-1/2}B)^* \in \mathcal{S}_1(X)$ equivalence between $P \in \mathcal{S}_1(X)$ and $(-A)^{-1/2}B \in \mathcal{S}_2(U, X)$ follows immediately. In this case, we have

$$\begin{aligned} \operatorname{tr}(P) =& \tfrac{1}{2} \cdot \operatorname{tr}((-A)^{-1/2}BB'(-A)^{-1/2}) = \tfrac{1}{2} \cdot \operatorname{tr}((-A)^{-1/2}B((-A)^{-1/2}B)^*) \\ =& \tfrac{1}{2} \cdot \|(-A)^{-1/2}B\|^2_{\mathcal{S}_2(U,X)} = -\tfrac{1}{2} \cdot \operatorname{tr}(B'A^{-1}B). \end{aligned}$$

□

Remark 3.7. Note that, under the assumptions that $A : D(A) \subset X \to X$ is negative and has compact resolvent, the input operator B fulfills (3.6) *and, additionally, the input space is finite-dimensional (i.e., w.l.o.g., $U = \mathbb{C}^m$), we can immediately infer from Proposition 3.6 that the Gramian P of (A, B) is nuclear. Namely,the expression $\|P\|_{\mathcal{S}(X)} = \operatorname{tr}(P)$ coincides with the trace of the matrix $\frac{1}{2}B'A^{-1}B \in \mathbb{C}^{m,m}$.*

4. Alternating direction implicit (ADI) iteration for operator Lyapunov equations. We now present an algorithm, take a closer look to operator Lyapunov equations, and set up an iterative scheme for their solution; we consider the ADI iteration for the operator case and discuss convergence. Before presenting results about convergence, we first present the algorithm, which exactly reads as in the matrix case [32, p. 43]. Besides the pair (A, B) defining a control system, this algorithm involves so-called *shift parameters* $p_i \in \mathbb{C}$, which have to be chosen a priori. In the finite-dimensional case, they are known to determine the velocity of convergence [32, pp. 43]. Their choice in the case of operator Lyapunov equations is discussed at the end of this section.

Algorithm 1 ADI iteration for operator Lyapunov equations.

Input: The generator A of an exponentially stable semigroup $T(\cdot)$, an admissible control operator $B \in \mathcal{B}(U, D(A^*)')$, and shift parameters $p_1, \dots, p_{i_{\max}} \in \mathbb{C}_-$
Output: $S = S_{i_{\max}} \in \mathcal{B}(U^{i_{\max}}, X)$, such that $SS^* \approx P$, where P is the Gramian of (A, B).

```
V_1 = (A + p_1 I)^{-1} B
S_1 = sqrt(-2 Re(p_1)) · V_1
for i = 2, 3, ..., i_max do
    V_i = V_{i-1} - (p_i + conj(p_{i-1})) · (A + p_i I)^{-1} V_{i-1}
    S_i = [ S_{i-1} , sqrt(-2 Re(p_i)) · V_i ]
end for
```

Remark 4.1.

a) *In the case of finite-dimensional input space, i.e.,* $U = \mathbb{C}^m$, *we have* $S_i \in \mathcal{B}(\mathbb{C}^{m\times i}, X)$. *This means that,* $P_i = S_i^* S_i$ *has finite rank and, in the block operator notation,* S_i *is a* $m \cdot i$*-tuple of elements of the state space* X. *These elements are obtained by solving equations of type* $(p_i I + A)w = z$. *In practice,* A *is usually a differential operator, and each step of ADI iteration consists of a (numerical) solution of the corresponding differential equation (see Sec. 6).*

b) *We note that the choice of* $i_{\max}$ *has of course not to be done a priori. Rather one might use a suitable* stopping criterion. *Due to* $P_i - P_{i-1} = V_i V_i^*$, *we have for each* $\|\cdot\| \in \{\|\cdot\|_{\mathcal{B}(X)}, \|\cdot\|_{\mathcal{S}_p(X)}\}$ *that*

$$\|P_i - P_{i-1}\| = \|V_i V_i^*\| = \|V_i^* V_i\|.$$

A suitable criterion for termination of the ADI iteration is therefore to check whether the norm of the operator $V_i^* V_i \in \mathcal{B}(U)$ *(which is a matrix, if* $U = \mathbb{C}^m$*) goes below a given absolute or relative threshold. For an overview on stopping criteria for the ADI iteration to solve matrix Lyapunov equations, we refer to [32, Sec. 4.6]. Note that, the case treated in Proposition 3.6, we will derive an explicit expression for the approximation error* $P - P_i$ *in the nuclear norm (see Proposition 4.8).*

The main result of convergence is presented below. The remaining part of this section mainly consists of its proof.

Theorem 4.2. *Let* U, X *be Hilbert spaces and operators* $A : D(A) \subset X \to X$ *be the generator of an exponentially stable semigroup* $T(\cdot)$ *and* $B \in \mathcal{B}(U, D(A^*)')$ *be an admissible control operator for* $T(\cdot)$. *Let* $P \in \mathcal{B}(X)$ *be the Gramian of* (A, B) *and, for some* $J \in \mathbb{N}$, *let* $(p_i)_i$ *be a* J-cyclic *(that is,* $p_{J+i} = p_i$ *for all* $i \in \mathbb{N}$*) sequence in* $\mathbb{C}_-$. *Then Algorithm 1 is feasible and the operator sequence* $(P_i)_i = (S_i S_i^*)_i$ *is* *strongly convergent to P, i.e.,*

$$\lim_{i\to\infty} P_i x = Px \quad \text{for all } x \in X.$$

Moreover, the following holds true:
a) If the Gramian P is compact, then

$$\lim_{i\to\infty} \|P - P_i\|_{\mathcal{B}(X)} = 0.$$

b) If, for some $p \in [1, \infty)$, the Gramian P is of p-th Schatten class, then

$$\lim_{i\to\infty} \|P - P_i\|_{\mathcal{S}_p(X)} = 0.$$

The strategy for the proof is to exhibit that $(P_i)_i$ is monotone and bounded with respect to the ordering induced by nonnegativity. Convergence will then be inferred by using the results on monotone operator sequences in Appendix A.

Before starting to prove some auxiliary results needed for the proof of Theorem 4.2, we state some preliminary facts and remarks.

Remark 4.3.
a) It follows from exponential stability that $\mathbb{C}_- \subset \rho(-A)$ [42, Cor. 2.3.3]. The iteration in Algorithm 1 is therefore feasible.
b) For $q \in \mathbb{C}$ and $p \in \mathbb{C}_-$ the operator $(A - qI)(A + pI)^{-1}$ is in $\mathcal{B}(X)$ due to

$$(A - qI)(A + pI)^{-1}x = x - (p + q)(A + pI)^{-1}x \quad \text{for all } x \in X. \tag{4.1}$$

c) The operator $A\colon D(A) \subset X \to X$ uniquely extends to an operator $\tilde{A}\colon D(\tilde{A}) = X \subset D(A^*)' \to D(A^*)'$ [42, Cor. 2.10.3]. In the sequel we will denote both A and $\tilde{A}$ by A. Consequently, for $p \in \rho(A)$, the inverse of the shifted operator can be extended as a bounded operator $(pI - A)^{-1}\colon D(A^*)' \to X$.
d) Using the previous observations, the operator

$$L_i := (A + p_i I)^{-1} \prod_{j=1}^{i-1} (A - \overline{p_{i-j}} I)(A + p_{i-j} I)^{-1} \tag{4.2}$$

fulfills $L_i \in \mathcal{B}(D(A^*)', X)$, hence $L_i B \in \mathcal{B}(U, X)$. Furthermore, simple calculations lead to $(L_i B)^* = B' L_i^*$
e) By simple arithmetics we obtain that

$$S_i = \left[\sqrt{-2\operatorname{Re}(p_1)} L_1 B \ \cdots \ \sqrt{-2\operatorname{Re}(p_i)} L_i B\right] \in \mathcal{B}(U^i, X), \tag{4.3}$$

hence $(P_i)_i = (S_i S_i^*)_i$ fulfills

$$\begin{aligned} P_0 &= 0, \\ P_i &= P_{i-1} - 2\operatorname{Re}(p_i) \cdot (L_i B) \cdot (L_i B)^*, \qquad i \in \mathbb{N}. \end{aligned} \tag{4.4}$$

Monotonicity of $(P_i)_i$ thus immediately follows from $p_i \in \mathbb{C}_-$.
f) An important class of infinite-dimensional systems are the so-called boundary control systems *[13, 25]. They are, in an abstract setting, of the form*

$$\begin{aligned} \dot{x}(t) &= \mathfrak{A} x(t), \\ \mathfrak{p} x(t) &= u(t), \end{aligned} \tag{4.5}$$

where $\mathfrak{A} : D(\mathfrak{A}) \subset X \to X$ *is densely defined and closed, and* $\mathfrak{p} \in \mathcal{B}(D(\mathfrak{A}), U)$ *is onto. In case of well-posedness (that is, for all* $t \in \mathbb{R}_{>0}$*,* $x(t) \in X$ *continuously depends on* $x(0) \in X$ *and* $u \in L_2([0,t], U)$*), this system can be rewritten as* $\dot{x}(t) = Ax(t) + Bu(t)$*, where the operator* A *generates a strongly continuous semigroup* $T(\cdot)$ *on* X*, and* B *is admissible for* $T(\cdot)$*, see [25]. If* A *generates an exponentially stable semigroup, then we may use [13, Thm. 2.9] to infer that for all* $p \in \overline{\mathbb{C}_-}$*,* $z \in X$*, and* $u \in U$ *the vectors* $x_1 = (A + pI)^{-1}z$ *and* $x_2 = (A + pI)^{-1}Bu \in X$ *are the unique solutions of the so-called* abstract elliptic problems

$$\begin{aligned} px_1 + \mathfrak{A}x_1 &= z, & px_2 + \mathfrak{A}x_2 &= 0, \\ \mathfrak{p}x_1 &= 0, & \mathfrak{p}x_2 &= u. \end{aligned} \tag{4.6}$$

The Gramian P *can therefore be computed by the ADI algorithm without explicit use of* A *and* B*.*

The subsequent two lemmas result in the fact that $(P_i)_i$ is bounded from above by the Gramian P.

LEMMA 4.4. *Let* U, X *be Hilbert spaces,* A *be a generator of an exponentially stable semigroup* $T(\cdot)$ *on* X *and* $B \in \mathcal{B}(U, D(A^*)')$ *an admissible control operator for* $T(\cdot)$*. Let* $P \in \mathcal{B}(X)$ *be the Gramian of* (A, B)*, let* $p_1, \ldots, p_i \in \mathbb{C}_-$ *and define*

$$T_i = \prod_{j=0}^{i-1} (A - \overline{p_{i-j}}I)(A + p_{i-j}I)^{-1}. \tag{4.7}$$

Then

$$\begin{aligned} &(A + p_iI)^{-1}T_{i-1}PT_{i-1}^* + T_{i-1}PT_{i-1}^*(A^* + \overline{p_i}I)^{-1} \\ =& -(A + p_iI)^{-1}T_{i-1}BB'T_{i-1}^*(A^* + \overline{p_i}I)^{-1} \\ &+ 2\operatorname{Re}(p_i)(A + p_iI)^{-1}T_{i-1}PT_{i-1}^*(A^* + \overline{p_i}I)^{-1} \in \mathcal{B}(X). \end{aligned} \tag{4.8}$$

Proof. Using Remark 4.3 b) one sees that T_{i-1} and $(A+\overline{p_i}I)^{-1}$ (and thus also their respective adjoints) commute. The operator T_i is hence well-defined by the expression (4.7). Let $x \in X$. Then $(A^* + \overline{p_i}I)^{-1}x \in D(A^*)$ and $T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x \in D(A^*)$. Now using the Lyapunov equation (3.1) in the last reformulation, we obtain

$$\begin{aligned} &\left\langle x, \left((A + p_iI)^{-1}T_{i-1}PT_{i-1}^* + T_{i-1}PT_{i-1}^*(A^* + \overline{p_i}I)^{-1}\right)x\right\rangle_X \\ =&\left\langle x, (A + p_iI)^{-1}T_{i-1}PT_{i-1}^*x\right\rangle_X + \left\langle x, T_{i-1}PT_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X \\ =&\left\langle (A^* + \overline{p_i}I)^{-1}T_{i-1}^*x, PT_{i-1}^*x\right\rangle_X + \left\langle PT_{i-1}^*x, T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X \\ =&2\operatorname{Re}\left\langle PT_{i-1}^*x, T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X \\ =&2\operatorname{Re}\left\langle T_{i-1}^*x, PT_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X \\ =&2\operatorname{Re}\left\langle A^*T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x, PT_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X \\ &+ 2\operatorname{Re}\left\langle \overline{p_i}T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x, PT_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X \\ =&-\|B'T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\|_U^2 \\ &+ 2\operatorname{Re}\left\langle \overline{p_i}T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x, PT_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X. \end{aligned}$$

Then, due to

$$\|B'T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\|_U^2 = \left\langle x, (A + p_iI)^{-1}T_{i-1}BB'T_{i-1}^*(A^* + \overline{p_i}I)^{-1}x\right\rangle_X$$

and

$$\left\langle T_{i-1}^*(A^*+\overline{p_i}I)^{-1}x, PT_{i-1}^*(A^*+\overline{p_i}I)^{-1}x\right\rangle_X \in \mathbb{R},$$

we obtain

$$\begin{aligned}
&\left\langle x, \left((A+p_iI)^{-1}T_{i-1}PT_{i-1}^* + T_{i-1}PT_{i-1}^*(A^*+\overline{p_i}I)^{-1}\right)x\right\rangle_X \\
=& -\left\langle x, (A+p_iI)^{-1}T_{i-1}BB'T_{i-1}^*(A^*+\overline{p_i}I)^{-1}x\right\rangle_X \\
&\quad + 2\operatorname{Re}(p_i)\cdot\left\langle x, (A+p_iI)^{-1}T_{i-1}PT_{i-1}^*(A^*+\overline{p_i}I)^{-1}x\right\rangle_X \\
=& -\left\langle x, (A+p_iI)^{-1}T_{i-1}BB'T_{i-1}^*(A^*+\overline{p_i}I)^{-1}x\right\rangle_X \\
&\quad + \left\langle x, 2\operatorname{Re}(p_i)\cdot(A+p_iI)^{-1}T_{i-1}PT_{i-1}^*(A^*+\overline{p_i}I)^{-1}x\right\rangle_X.
\end{aligned}$$

Since this holds true for all $x \in X$, the desired operator equation follows immediately. □

LEMMA 4.5. *Let U, X be Hilbert spaces, A be a generator of an exponentially stable semigroup $T(\cdot)$ on X and $B \in \mathcal{B}(U, D(A^*)')$ an admissible control operator for $T(\cdot)$. Let $P \in \mathcal{B}(X)$ be the Gramian, let $p_1, \ldots, p_i \in \mathbb{C}_-$, $T_i \in \mathcal{B}(X)$ as in (4.7) and P_i recursively defined as in (4.4). Then*

$$P - P_i = T_iPT_i^*. \tag{4.9}$$

In particular, there holds $P_i \le P$.

Proof. The statement is shown by induction on i. For $i = 0$, the assertion is fulfilled due to $T_0 = I_X$ and $P_0 = 0$. Let $i \in \mathbb{N}$ and assume that (4.9) is fulfilled for some $i-1$. Making use of the induction assumption, Lemma 4.4 and

$$T_i = (A-\overline{p_i}I)(A+p_iI)^{-1}T_{i-1} = T_{i-1} - 2\operatorname{Re}(p_i)(A+p_iI)^{-1}T_{i-1},$$

we obtain

$$\begin{aligned}
T_iPT_i^* =& T_{i-1}PT_{i-1}^* - 2\operatorname{Re}(p_i)\left((A+p_iI)^{-1}T_{i-1}PT_{i-1}^* + T_{i-1}PT_{i-1}^*(A^*+\overline{p_i}I)^{-1}\right) \\
&+ 4\operatorname{Re}(p_i)^2(A+p_iI)^{-1}T_{i-1}PT_{i-1}^*(A^*+\overline{p_i}I)^{-1} \\
=& P - P_{i-1} + 2\operatorname{Re}(p_i)(A+p_iI)^{-1}T_{i-1}BB'T_{i-1}^*(A^*+\overline{p_i}I)^{-1} = P - P_i
\end{aligned}$$

□

We can already infer from Lemma 4.5 and the results from Appendix A that the sequence $(P_i)_i$ obtained from Algorithm 1 is strongly convergent towards a nonnegative operator that is below the Gramian. The following results will show that the strong limit is indeed given by the Gramian P.

LEMMA 4.6. *Let the assumptions of Lemma 4.5 hold true. Further, assume that the sequence $(p_i)_i$ in $\mathbb{C}_-$ is J-cyclic for some $J \in \mathbb{N}$ and let the operator sequence $(T_i)_i$ be defined as in (4.7). Then for all $x \in D(A^*)$ holds*

$$\lim_{i\to\infty} T_i^*x = 0 \quad (\textit{in } X). \tag{4.10}$$

Proof. Let $x \in X$ be given. J-cyclicity of the shift parameters implies $T_{nJ+k} = T_k \cdot T_J^n$ for all $n \in \mathbb{N}$, $k \in \{0, \ldots, J-1\}$. Thus it is clear that it suffices to show that $(T_J^*)^kx$ converges to zero. To see this, consider the recursively defined operator sequence $(\tilde{P})_i$ in $\mathcal{B}(X)$ with

$$\begin{aligned}
\tilde{P}_0 &= 0, \\
\tilde{P}_i &= \tilde{P}_{i-1} - 2\operatorname{Re}(p_i)L_iL_i^*, \qquad i \in \mathbb{N}.
\end{aligned} \tag{4.11}$$

with L_i as in (4.2), i.e., the recursion (4.4) for the case where the identity is the control operator, i.e., $B = I_X$. Hence, $\tilde{P}_{(k+1)J} \geq \tilde{P}_{kJ}$ for all $k \in \mathbb{N}$, and we may apply Lemma 4.5 to see that for all $i \in \mathbb{N}$ holds

$$\tilde{P}_i \leq \tilde{P},$$

where $\tilde{P}$ is the Gramian of (A, I_X). The J-cyclicity of the shift parameters gives rise to

$$\tilde{P}_{(k+1)J} = T_J \tilde{P}_{kJ} T_J^* + \tilde{P}_J.$$

By Lemma 4.5, we obtain $\tilde{P}_{kJ} \leq \tilde{P}$ for all $k \in \mathbb{N}$, i.e., the sequence $(\tilde{P}_{kJ})_k$ is monotonically increasing and bounded from above. Theorem A.1 now gives rise to strong convergence of $(\tilde{P}_i)_i$ towards some $\tilde{P} \in \mathcal{B}(X)$, which leads to the discrete-time operator Lyapunov equation

$$\tilde{P} = T_J \tilde{P} T_J^* + \tilde{P}_J.$$

Now using [26, Prop. 8(b)], we obtain that for all $z \in X$ holds

$$0 = \lim_{k\to\infty} \left\langle (T_J^*)^k z, \tilde{P}_J (T_J^*)^k z \right\rangle_X .$$

Using that

$$P_J = -\sum_{i=1}^{J} \operatorname{Re}(p_i) L_i L_i^*,$$

we particularly obtain

$$\begin{aligned} 0 &= \lim_{k\to\infty} \langle (T_J^*)^k z, L_1 L_1^* (T_J^*)^k z \rangle_X = \lim_{k\to\infty} \|L_1^* (T_J^*)^k z\|_X = \lim_{k\to\infty} \|L_1^* (T_J^*)^k z\|_X^2 \\ &= \lim_{k\to\infty} \|(A^* + \overline{p_1})^{-1} (T_J^*)^k z\|_X^2 = \lim_{k\to\infty} \|(T_J^*)^k (A^* + \overline{p_1})^{-1} z\|_X^2, \end{aligned}$$

where the latter equality holds true since $(T_J^*)^k$ and $(A^* + \overline{p_1})^{-1}$ commute (see Rem. 4.3 b)). Since this holds true for all $z \in X$ the fact that $\operatorname{im}(A^* + \overline{p_1})^{-1} = D(A^*)$ implies the desired result. □

We are now prepared to formulate the proof of the main result:

Proof of Theorem 4.2:

We can infer from $p_i \in \mathbb{C}_-$ and (4.4) that $P_{i-1} \leq P_i$ for all $i \in \mathbb{N}$. Lemma 4.5 further implies that $P_i \leq P$ and we can apply Theorem A.1 to obtain that there exists some self-adjoint and nonnegative $Q \in \mathcal{B}(X)$ with $Q \leq P$ and

$$\lim_{i\to\infty} P_i x = Qx \quad \text{for all } x \in X.$$

Assuming that $P \neq Q$, the density of $D(A^*)$ in X gives rise to the existence of some $z \in D(A^*)$ with $\langle z, Qz\rangle_X \neq \langle z, Pz\rangle_X$. By Lemma 4.5, we have

$$\langle z, P_i z\rangle_X - \langle z, Pz\rangle_X = \langle z, T_i P T_i^* z\rangle_X = \langle T_i^* z, P T_i^* z\rangle_X .$$

Now taking the limits on both sides of the equation and using Lemma 4.6, we obtain

$$\langle z, Qz\rangle_X - \langle z, Pz\rangle_X = \lim_{i\to\infty} \langle z, P_i z\rangle_X - \langle z, Pz\rangle_X = \lim_{i\to\infty} \langle T_i^* z, P T_i^* z\rangle_X = 0,$$

which is a contradiction to $\langle z, Qz\rangle_X \neq \langle z, Pz\rangle_X$. The sequence $(P_i)_i$ therefore strongly converges to the Gramian P.

If we now assume that P is moreover compact (p-th Schatten class), then Proposition A.2 (Proposition A.3) will imply that even convergence towards P in the operator (p-th Schatten) norm holds true. □

We finally present some remarks about the shift parameter choice:

*Remark 4.7 (*Shift parameters*).*

a) *Equation* (4.9) *gives rise to he fact that the approximation $P_i \approx P$ is the better, the "smaller is T_i". In practice, one is interested in fast convergence, since this gives rise to good approximants of low rank.*

b) *In the finite-dimensional case, the shift parameters are chosen in a way that the spectral radius of T_J is minimized. With $\sigma(A) = \{\lambda_1, \ldots, \lambda_n\}$, this leads to the optimization problem*

$$\min_{p_1,\ldots,p_J \in \mathbb{C}_-} \max_{\lambda \in \sigma(A)} \prod_{j=1}^{J} \left| \frac{\overline{p_j} - \lambda}{p_j + \lambda} \right|.$$

This optimization problem can be solved by using advanced techniques of complex analysis, in particular the theory of elliptic integrals [17]. Furthermore, several suboptimal choices of the shift parameters that do not require the full information of the spectrum have been proposed and successfully applied (see [32, pp. 43] for an overview).

c) *In infinite dimensions and, in particular, in the case where A is unbounded any choice of the shift parameters in $\mathbb{C}_-$ will lead to an iteration operator fulfilling $\rho(T_J) = 1$. In the numerical experiments, we will chose the shift parameters by applying the existing approaches to a sufficiently accurate discretization of A.*

d) *It is worthwhile to place particular emphasis on systems with diagonalizable A, the so-called Riesz-spectral operators [14, Sec. 2.3](which occur for the heat and wave equation as well as at time-delay systems): That is, there exists a Riesz basis of eigenvectors of A and, moreover $\overline{\sigma(A)}$ is totally disconnected. Roughly speaking, a Riesz basis is a family of vectors which is topologically equivalent to an orthonormal system. A Riesz-spectral operator A generates a strongly continuous $T(\cdot)$ if and only if $\sup \operatorname{Re}(\sigma(A)) < \infty$; exponential stability of $T(\cdot)$ is equivalent to $\sup(\operatorname{Re} \sigma(A)) < 0$. Denoting $\sigma(A) = \{\lambda_i \,|\, i \in \mathbb{N}\}$, let $(\phi_i)_i$ is a Riesz basis of the eigenvectors of A^* (which exists due to [14, Lem. 2.3.2]) with $A^*\phi_i = \overline{\lambda_i}\phi_i$ for all $i \in \mathbb{N}$. Then the Lyapunov equation* (3.5) *gives rise to*

$$\begin{aligned} 0 &= \langle P\phi_i, A^*\phi_j\rangle_X + \langle A^*\phi_i, P\phi_j\rangle_X + \langle B'\phi_i, B'\phi_j\rangle_U \\ &= (\overline{\lambda_j} + \lambda_i)\,\langle P\phi_i, \phi_j\rangle_X + \langle B'\phi_i, B'\phi_j\rangle_U \quad \text{for all } i, j \in \mathbb{N}. \end{aligned} \tag{4.12}$$

Then $\lambda_i, \lambda_j \in \mathbb{C}_-$ implies $\lambda_i + \overline{\lambda_j} \neq 0$, and thus

$$\langle P\phi_i, \phi_j\rangle_X = -\frac{1}{\lambda_i + \overline{\lambda_j}} \cdot \langle B'\phi_i, B'\phi_j\rangle_U \quad \text{for all } i, j \in \mathbb{N}.$$

Note that, by the property of $(\phi_i)_i$ being a Riesz basis, this relation uniquely determines P; it can be considered as a Cauchy matrix representation [4] of P.

Assuming that, for $s \in \mathbb{N}$, $\Pi_{[s]} \in \mathcal{B}(X)$ is a projector onto the s-dimensional space $X_{[s]} := \operatorname{span}\{\phi_1, \ldots, \phi_s\}$, we obtain $(A + p_j)^{-1}\Pi_{[s]}x = \Pi_{[s]}(A + p_j)^{-1}\Pi_{[s]}x$ for all

$x \in X$ and $j \in \mathbb{N}$, and thus also $T_i \Pi_{[s]} x = \Pi_{[s]} T_i \Pi_{[s]} x$ for all $x \in X$ and all $i \in \mathbb{N}$. Then (4.9) implies

$$\Pi_{[s]}(P - P_i)(\Pi_{[s]})^* = (\Pi_{[s]} T_i \Pi_{[s]})(\Pi_{[s]} P \Pi_{[s]}^*)(\Pi_{[s]} T_i \Pi_{[s]})^*. \tag{4.13}$$

On the other hand, the spectral radius of the projected iteration matrix is given by

$$\rho(\Pi_{[s]} T_i \Pi_{[s]}) = \max_{l=1,\dots,s} \prod_{j=1}^{i} \left| \frac{\overline{p_j} - \lambda_l}{p_j + \lambda_l} \right| < 1.$$

As a consequence, we have linear convergence of $(\Pi_{[s]} P_i \Pi_{[s]}^*)_i$ to $\Pi_{[s]} P \Pi_{[s]}^*$ in the operator norm for any $s \in \mathbb{N}$. If the Gramian P is compact (or even of Schatten class), then there exists some $s \in \mathbb{N}$ such that "P is almost vanishing outside $X_{[s]}$"; that is, the norms of $(I - \Pi_{[s]}) P \Pi_{[s]}^*$, $\Pi_{[s]} P (I - \Pi_{[s]})^*$ and $(I - \Pi_{[s]}) P (I - \Pi_{[s]})^*$ are small. Since $P_i \leq P$, the norms of $(I - \Pi_{[s]}) P_i \Pi_{[s]}^*$, $\Pi_{[s]} P_i (I - \Pi_{[s]})^*$ and $(I - \Pi_{[s]}) P_i (I - \Pi_{[s]})^*$ are small as well (for a more mathematical justification of this argumentation we refer to the proofs of Theorem A.2 and Theorem A.3). One has therefore to find shift parameters that guarantee fast convergence on the "dominant subspace of P". Anyway, various open questions in the (optimal) shift parameter selection are left; this is an interesting topic for further research.

We finally give an explicit representation of the ADI approximation error for the class of systems considered in Proposition 3.6. On the basis of this result, suitable stopping criteria, i.e., the determination of $i_{\max}$ in Algorithm 1, may be designed (see also Remark 4.1 b)).

PROPOSITION 4.8. *Let $A : D(A) \subset X \to X$ be self-adjoint with $A \leq 0$ and $0 \in \rho(A)$. Further assume that A has compact resolvent and let $B : U \to D((-A)^{\frac{1}{2}})'$ such that $(-A)^{-1/2} B : U \to X$ is Hilbert-Schmidt. Let P be the Gramian of (A, B) and let shift parameters $p_1, \dots, p_i \in \mathbb{C}_-$ be given. Then, in the notation of Algorithm 1, there holds*

$$\|P - P_i\|_{\mathcal{S}_1(X)} = -\frac{1}{2} \cdot \operatorname{tr}(B' A^{-1} B) + 2 \sum_{k=1}^{i} \operatorname{Re}(p_i) \cdot \operatorname{tr}(V_k^* V_k). \tag{4.14}$$

In particular, if $U = \mathbb{C}^n$, then $B' A^{-1} B, S_k^* S_k \in \mathbb{C}^{n \times n}$ are Hermitian matrices.

Proof. By Lemma 4.5, we have $P - P_i \geq 0$, whence

$$\|P - P_i\|_{\mathcal{S}_1(X)} = \operatorname{tr}(P) - \operatorname{tr}(P_i).$$

Then the desired result follows from (3.7) and

$$\begin{aligned} \operatorname{tr}(P_i) = & \operatorname{tr}\left(\sum_{k=1}^{i} -2 \operatorname{Re}(p_k) \cdot V_k V_k^* \right) = -2 \sum_{k=1}^{i} \operatorname{Re}(p_k) \cdot \operatorname{tr}\left(V_k V_k^*\right) \\ = & -2 \sum_{k=1}^{i} \operatorname{Re}(p_k) \cdot \operatorname{tr}\left(V_k^* V_k\right). \end{aligned}$$

∎

5. Inexact ADI iteration. Algorithm 1 can, in general, not be implemented for practical purposes, if the state space X is infinite-dimensional. Instead on will have to work with suitable approximations of the equations arising in each step of the ADI iteration. That is, we approximate $(A+p_iI)^{-1}$ by some operator acting on a finite-dimensional subspace. This is in the sequel referred to as *inexact ADI iteration.* For instance, if A is a differential operator, an approximation can be performed by using (adaptive) finite-element methods (see Section 6).

It is the goal of this section to present an error analysis for inexact ADI: we will derive estimates for the error in the obtained approximation of the Gramian operator.

In this part we assume finite-dimensionality of the input space, i.e., $U = \mathbb{C}^m$ for some $m \in \mathbb{N}$. Note that this assumption is justified by practice: only finitely many actuating variables are available to control a given system. As a consequence, we have a representation $B = [b_1, \dots, b_m] \in (D(A^*)')^m$; the operators in the ADI iteration may be written as $V_i = [v_{i1}, \dots, v_m] \in X^m$. In the first step of the ADI iteration we therefore have to solve m equations $(A + p_1 I)v_{1k} = b_k \in D(A^*)'$; the following steps consist of the solving equations $(A+p_iI)x_k = v_{i-1,k} \in X$. To suitably approximate the equations arising in the ADI iteration, let $(X^{(i)})$ be a sequence of (finite-dimensional) subspaces of X, let $\Pi^{(1)} \in \mathcal{B}(D(A^*)')$ be a projector onto $X^{(1)}$, and, for $i \geq 2$, let $\Pi^{(i)} \in \mathcal{B}(X)$ be a projector onto $X^{(i)}$. Further, assume that, for $i \in \mathbb{N}$, the operators $\widetilde{A}_{p_i}^{(i)} \in \mathcal{B}(X^{(i)})$ are approximations on $A + p_iI$ in the sense that the solution $x \in X$ of the equation $\widetilde{A}_{p_i}^{(i)} x = b$ is "close to $(\widetilde{A}_{p_i}^{(i)})^{-1}\Pi^{(i)} b$" for suitable right hand side $b \in X$ or $b \in D(A^*)'$.

With these preparations we can formulate our inexact ADI iteration as follows:

Algorithm 2 Inexact ADI iteration for operator Lyapunov equations.

Input: A Hilbert space X, a sequence $(X^{(i)})$ of subspaces of X, operators $A : D(A) \subset X \to X$ and $B \in \mathcal{B}(U, D(A^*)')$; projectors $\Pi^{(i)}$ onto $X^{(i)}$ with $\Pi^{(1)} \in \mathcal{B}(D(A^*)')$ and $\Pi^{(j)} \in \mathcal{B}(X)$ for $i \geq 2$; operators $\widetilde{A}_{p_i}^{(i)} \in \mathcal{B}(X^{(i)})$; shift parameters $p_i \in \mathbb{C}_-$.

Output: $\widetilde{S} = \widetilde{S}_{i_{\max}} \in \mathcal{B}(U^{i_{\max}}, X)$, such that $\widetilde{S}\widetilde{S}^* \approx P$, where P is the Gramian of (A, B).

$\widetilde{V}_1 = (A_{p_1}^{(i)})^{-1}\Pi^{(i)}B$

$\widetilde{S}_1 = \sqrt{-2\,\mathrm{Re}(p_1)} \cdot \widetilde{V}_1$

for $i = 2, 3, \dots, i_{\max}$ **do**

  $\widetilde{V}_i = \widetilde{V}_{i-1} - (p_i + \overline{p_{i-1}})(\widetilde{A}_{p_i}^{(i)})^{-1}\Pi^{(i)}\widetilde{V}_{i-1}$

  $\widetilde{S}_i = [\, \widetilde{S}_{i-1}\,,\, \sqrt{-2\,\mathrm{Re}(p_i)} \cdot \widetilde{V}_i \,]$

end for

We will now derive expressions and estimates for the error between exact and inexact ADI iteration. We will first derive estimates for $V_i - \widetilde{V}_i$. Thereafter, we will provide upper bounds for the difference between S_i and $\widetilde{S}_i$, and, respectively, $P_i = S_iS_i^*$ and $\widetilde{P}_i = \widetilde{S}_i\widetilde{S}_i^*$.

Denoting

$$
\begin{aligned}
E_i &= V_i - \widetilde{V}_i,\\
G_i^E &= (A + p_iI)^{-1} - (\widetilde{A}_{p_i}^{(i-1)})^{-1}\Pi^{(i)}
\end{aligned}
$$

the construction of V_i and $\widetilde{V}_i$ in Algorithm 1 and Algorithm 2 yields that the error recursively fulfills

$$E_i = (A + p_i I)^{-1}(A - \overline{p_i} I)E_{i-1} + (p_i + \overline{p_{i-1}}) \cdot G_i^E \widetilde{V}_{i-1}. \tag{5.1}$$

Using that the operators $(A + p_j)^{-1}$, $(A + p_k)^{-1}$, $(A - \overline{p_j} I)$ and $(A - \overline{p_k} I)$ commute for all $j, k \in \mathbb{N}$, we may inductively conclude from (5.1) that for all $i > 1$, there holds

$$\begin{aligned} E_i &= (A - \overline{p_1} I)(A + p_i I)^{-1} \left(\prod_{k=2}^{i-1} (A - \overline{p_k} I)(A + p_k I)^{-1} \right) \cdot G_1^E B \\ &+ \sum_{j=2}^{i-1} (p_j + \overline{p_{j-1}}) \cdot (A - \overline{p_j} I)(A + p_i I)^{-1} \left(\prod_{k=j+1}^{i-1} (A - \overline{p_k} I)(A + p_k I)^{-1} \right) G_j^E \widetilde{V}_{j-1} \\ &+ (p_i + \overline{p_{i-1}}) \cdot G_i^E \widetilde{V}_{i-1}. \end{aligned} \tag{5.2}$$

Proposition 5.1. *Let X be a Hilbert space, let $A : D(A) \subset X \to X$ be the generator of an exponentially stable semigroup $T(\cdot)$ and $B = [b_1, \ldots, b_m] \in \mathcal{B}(\mathbb{C}^m, D(A^*)')$ be an admissible control operator for $T(\cdot)$. Let $J \in \mathbb{N}$, let $(p_i)_i$ be a J-cyclic sequence in $\mathbb{C}_-$. Assume that for all $i \in \mathbb{N}$, the operator $V_i = [v_{i1}, \ldots, v_{im}] \in \mathcal{B}(\mathbb{C}^m, X)$ is obtained by Algorithm 1.*
Let $(X^{(i)})$ be a sequence of subspaces of X, let $\Pi^{(1)} \in \mathcal{B}(D(A^*)')$ be a projector onto $X^{(1)}$, and, for $i \geq 2$, let $\Pi^{(i)} \in \mathcal{B}(X)$ be a projector onto $X^{(i)}$. Further, let $\widetilde{A}_{p_i}^{(i)} \in \mathcal{B}(X^{(i)})$, and assume that the operators $\widetilde{V}_i = [\widetilde{v}_{i1}, \ldots, \widetilde{v}_{im}] \in \mathcal{B}(\mathbb{C}^m, X)$ are obtained by Algorithm 2.
Assume that

$$\begin{aligned} \|(A + p_1 I)^{-1} b_l - (\widetilde{A}_{p_1}^{(1)})^{-1} \Pi^{(1)} b_l\|_X &\leq c^{(1l)} \text{ for } l = 1, \ldots, m, \text{ and} \\ \|(A + p_i I)^{-1} v_{i-1,l} - (\widetilde{A}_{p_i}^{(i)})^{-1} \Pi^{(i)} v_{i-1,l}\|_X &\leq c^{(il)} \text{ for } l = 1, \ldots, m, \quad i > 1. \end{aligned} \tag{5.3}$$

Let either be $\|\cdot\| = \|\cdot\|_{\mathcal{B}(X)}$ or $\|\cdot\| = \|\cdot\|_{\mathcal{S}_p(X)}$ for some $p \in [1, \infty)$. Then the following assertions hold true:
a) There exists some $M > 0$ such that for all $i \in \mathbb{N}$ holds

$$\|E_i\| \leq M \cdot \sum_{k=1}^{i} \sum_{l=1}^{m} c^{(kl)}.$$

b) If A is a Riesz-spectral operator, then there exists some $M > 0$, an increasing sequence of finite-dimensional subspaces $X_{[s]}$ of X with $X = \overline{\bigcup_{j \in \mathbb{N}} X_{[s]}}$, a bounded sequence of projectors $\Pi_{[s]}$ with $\operatorname{im} \Pi_{[s]} = X_{[s]}$ and some $M > 0$, such that for all $s \in \mathbb{N}$ there exists some $\rho_{[s]} \in (0, 1)$, such that

$$\|\Pi_{[s]} E_i\| \leq M \cdot \left(\sum_{l=1}^{m} c^{(il)} + \sum_{k=1}^{i-1} \rho_{[s]}^{i-k-1} \sum_{l=1}^{m} c^{(kl)} \right).$$

Proof. The following argumentation make use of the fact that any $x \in X$ can be identified as an operator $x \in \mathcal{B}(\mathbb{C}, X)$ via scalar multiplication. It can be seen that this operator is also belonging to any Schatten space with

$$\|x\|_X = \|x\|_{\mathcal{B}(\mathbb{C}, X)} = \|x\|_{\mathcal{S}_p(\mathbb{C}, X)}.$$

By the triangular inequality, we obtain that for any $x_1, \dots, x_m \in X$, the operator $M = [x_1, \dots, x_m] \in \mathcal{B}(\mathbb{C}^m, X)$ fulfills

$$\|M\|_{\mathcal{B}(\mathbb{C},X)} \leq \|M\|_{\mathcal{S}_p(\mathbb{C},X)} \leq \sum_{l=1}^{m} \|x_l\|_X.$$

a) By Lemma 4.6, we know that for all $j \in \mathbb{N}$, the sequence

$$\left(\left(\prod_{k=j+1}^{i} (A - \overline{p_k} I)(A + p_k I)^{-1} \right)^* \right)_i$$

strongly converges to zero. Then we can conclude from the Banach-Steinhaus Theorem [27] that there exists some $m > 0$ such that for all $i, j \in \mathbb{N}$ holds

$$\left\| \prod_{k=j+1}^{i} (A - \overline{p_k} I)(A + p_k I)^{-1} \right\|_{\mathcal{B}(X)} \leq m.$$

Then, by setting

$$\begin{aligned} M =& m \cdot \max \big(\{ \|(A - \overline{p_1} I)(A + p_i I)^{-1}\|_{\mathcal{B}(X)}, |p_i + \overline{p_{i-1}}| \} \cup \\ & \big\{ |p_j + \overline{p_{j-1}}| \cdot \|(A - \overline{p_1} I)(A + p_j I)^{-1}\|_{\mathcal{B}(X)} \mid j \in \{1, \dots, i-2\} \big\} \big), \end{aligned} \tag{5.4}$$

the desired result follows by a combination of (5.2) and (5.3).

b) If A is a Riesz-spectral operator, then there exists some bounded and boundedly invertible similarity transformation that results into a diagonal operator on ℓ_2. That is, there exists some bijective $T \in \mathcal{B}(\ell_2, X)$ such that $T^{-1}AT = D_A$, where $D_A : D(D_A) \subset \ell_2 \to \ell_2$ fulfills $D_A(x_1, x_2, \dots) = (\lambda_1 x_1, \lambda_2 x_2, \dots)$ with

$$D(D_A) = T^{-1} D(A) = \{ (x_1, x_2, \dots) \in \ell_2 \,|\, (\lambda_1 x_1, \lambda_2 x_2, \dots) \in \ell_2 \}.$$

Define the projectors by $\Pi_{[s]} = T^{-1} \Pi_{[s]}^D T$, where $\Pi_{[s]}^D \in \mathcal{B}(\ell_2)$ truncates after the s-th position, i.e.,

$$\Pi_{[s]}^D (x_1, \dots, x_j, x_{j+1}, x_{j+2}, \dots) = (x_1, \dots, x_j, 0, 0, \dots),$$

and set $X_{[s]} = \operatorname{im} \Pi_{[s]}$. Then we have $\dim X_{[s]} = s$ and $X = \overline{\bigcup_{s \in \mathbb{N}} X_{[s]}}$. The sequence $(\Pi_{[s]})_s$ is bounded by to $\|\Pi_{[s]}\|_{\mathcal{B}(X)} \leq \|T\|_{\mathcal{B}(X)} \cdot \|T^{-1}\|_{\mathcal{B}(X)} =: C$. The construction of $\Pi_{[s]}$ leads to

$$\Pi_{[s]} (A + p_k I)^{-1} = \Pi_{[s]} (A + p_k I)^{-1} \Pi_{[s]} = (A + p_k I)^{-1} \Pi_{[s]},$$

we can make use of (5.2) to see that

$$\begin{aligned} \Pi_{[s]} E_i = & (A - \overline{p_1} I)(A + p_i I)^{-1} \left(\prod_{k=2}^{i-1} \Pi_{[s]} (A - \overline{p_k} I)(A + p_k I)^{-1} \Pi_{[s]} \right) \cdot G_1^E B \\ & + \sum_{j=1}^{i-2} (p_j + \overline{p_{j-1}}) \cdot (A - \overline{p_j} I)(A + p_i I)^{-1} \\ & \quad \cdot \left(\prod_{k=j+1}^{i-1} \Pi_{[s]} (A - \overline{p_k} I)(A + p_k I)^{-1} \Pi_{[s]} \right) G_j^E \widetilde{V}_{j-1} \\ & + (p_i + \overline{p_{i-1}}) \cdot G_i^E \widetilde{V}_{i-1}. \end{aligned} \tag{5.5}$$

As in Remark 4.7 d), we can infer that the spectral radius of the operator products in the above expression are given by

$$\rho\left(\prod_{k=j+1}^{i-1}\Pi_{[s]}(A-\overline{p_k}I)(A+p_kI)^{-1}\Pi_{[s]}\right)=\max_{l=1,\dots,s}\prod_{k=j+1}^{i-1}\left|\frac{\overline{p_j}-\lambda_l}{p_j+\lambda_l}\right|.$$

The latter expression is below one, since the function $z\mapsto\frac{\overline{p_j}-z}{p_j+z}$ maps $\mathbb{C}_-$ onto the open complex unit circle. By J-cyclicity of the shift parameters, we can now infer that there exists some $\rho_{[s]}\in(0,1)$, such that for all $i,j\in\mathbb{N}$, there holds

$$\max_{l=1,\dots,s}\prod_{k=j+1}^{i-1}\left|\frac{\overline{p_j}-\lambda_l}{p_j+\lambda_l}\right|\leq\rho_{[s]}^{i-j-1}.$$

This gives rise to

$$\left\|\prod_{k=j+1}^{i-1}\Pi_{[s]}(A-\overline{p_k}I)(A+p_kI)^{-1}\Pi_{[s]}\right\|_{\mathcal{B}(X)}\leq\underbrace{\|T\|_{\mathcal{B}(X,\ell_2)}\cdot\|T^{-1}\|_{\mathcal{B}(\ell_2,X)}}_{=:m}\cdot\rho_{[s]}^{i-j-1}.$$

Now defining M as in (5.4), we obtain the desired result. ∎

As an immediate consequence of Proposition 5.1 and the fact

$$\|S_i-\widetilde{S}_i\|\leq-\sum_{k=1}^{i}\mathrm{Re}(p_k)\cdot\|V_k-\widetilde{V}_k\|,\tag{5.6}$$

we may formulate the following estimates for $\|S_i-\widetilde{S}_i\|$:

Corollary 5.2. *Under the assumptions and notation of Proposition 5.1, Algorithm 1 and Algorithm 2, the following assertions hold true:*

a) There exists some $M>0$ such that for all $i\in\mathbb{N}$ holds

$$\|S_i-\widetilde{S}_i\|\leq M\cdot\sum_{k=1}^{i}\sum_{l=1}^{m}(i-k)\cdot c^{(kl)}.$$

b) If A is a Riesz-spectral operator, then there exists a sequence of finite-dimensional subspaces $X_{[s]}$ of X with $X=\overline{\bigcup_{j\in\mathbb{N}}X_{[s]}}$, a sequence of projectors $\Pi_{[s]}$ with $\operatorname{im}\Pi_{[s]}=X_{[s]}$, such that for all $s\in\mathbb{N}$ there exists some $M_{[s]}>0$, such that

$$\|\Pi_{[s]}(S_i-\widetilde{S}_i)\|\leq M_{[s]}\cdot\left(\sum_{k=1}^{i}\sum_{l=1}^{m}c^{(kl)}\right).$$

Proof. Statement a) follows from the triangular inequality in (5.6), and accordingly using the error bound from Proposition 5.1 a). To prove b), we construct $X_{[s]}$ and $\Pi_{[s]}$ as in the proof of Proposition 5.1 b). Thereafter, making use of

$$\|\Pi_{[s]}(S_i-\widetilde{S}_i)\|\leq-\sum_{k=1}^{i}\mathrm{Re}(p_k)\cdot\|\Pi_{[s]}(V_i-\widetilde{V}_i)\|,$$

the error bound in Proposition 5.1 b) and the formula for geometric sums give rise to the result. ∎

Now we present estimates for the difference between approximative Gramians provided by exact and inexact ADI iteration.

COROLLARY 5.3. *Under the assumptions and notation of Proposition 5.1, Algorithm 1 and Algorithm 2, the following assertions hold true:*

a) There exists some $M > 0$ such that for all $i \in \mathbb{N}$ and

$$L_i = \sum_{k=1}^{i} \sum_{l=1}^{m} (i-k) \cdot c^{(kl)},$$

there holds

$$\|P_i - \widetilde{P}_i\| \leq M \cdot \left(L_i + L_i^2\right).$$

b) If A is a Riesz-spectral operator, then there exists a sequence of finite-dimensional subspaces $X_{[s]}$ of X with $X = \overline{\bigcup_{j\in\mathbb{N}} X_{[s]}}$, a sequence of projectors $\Pi_{[s]}$ with $\operatorname{im}\Pi_{[s]} = X_{[s]}$, such that for all $s \in \mathbb{N}$ there exists some $M_{[s]} > 0$, such that for

$$K_i := \sum_{k=1}^{i} \sum_{l=1}^{m} c^{(kl)},$$

there holds

$$\|\Pi_{[s]}(P_i - \widetilde{P}_i)\Pi_{[s]}^*\| \leq M_{[s]} \cdot \left(K_i + K_i^2\right).$$

Proof. In the case of $\|\cdot\| = \|\cdot\|_{\mathcal{S}_p(X)}$, statement a) follows from Corollary 5.3 a), together with

$$\begin{aligned}
&\|P_i - \widetilde{P}_i\|_{\mathcal{S}_p(X)} = \|S_iS_i^* - \widetilde{S}_i\widetilde{S}_i^*\|_{\mathcal{S}_p(X)}\\
\leq&\|S_i(S_i^* - \widetilde{S}_i^*)\|_{\mathcal{S}_p(X)} + \|(S_i - \widetilde{S}_i)\widetilde{S}_i^*\|_{\mathcal{S}_p(X)}\\
\leq&\|S_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)} \cdot \|S_i^* - \widetilde{S}_i^*\|_{\mathcal{S}_{2p}(X,\mathbb{C}^{im})} + \|S_i - \widetilde{S}_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)} \cdot \|\widetilde{S}_i^*\|_{\mathcal{S}_{2p}(X,\mathbb{C}^{im})}\\
=&\Big(\|S_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)} + \|\widetilde{S}_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)}\Big) \cdot \|S_i - \widetilde{S}_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)}\\
=&\Big(2\|S_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)} + \|S_i - \widetilde{S}_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)}\Big) \cdot \|S_i - \widetilde{S}_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)}\\
\leq&\Big(2\|P_i\|_{\mathcal{S}_p(X)}^{1/2} + \|S_i - \widetilde{S}_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)}\Big) \cdot \|S_i - \widetilde{S}_i\|_{\mathcal{S}_{2p}(\mathbb{C}^{im},X)}.
\end{aligned} \tag{5.7}$$

If $\|\cdot\|$ is the standard operator norm, then we can analogously estimate

$$\|P_i - \widetilde{P}_i\|_{\mathcal{B}(X)} \leq \Big(2\|P_i\|_{\mathcal{B}(X)} + \|S_i - \widetilde{S}_i\|_{\mathcal{B}(X)}\Big) \cdot \|S_i - \widetilde{S}_i\|_{\mathcal{B}(X)},$$

and we can argument as for Schatten norms.

To prove b), we first construct $X_{[s]}$ and $\Pi_{[s]}$ as in the proof of Proposition 5.1 b). Then, by determining bounds for $\|\Pi_{[s]}(P_i - \widetilde{P}_i)\Pi_{[s]}^*\|$ analogous to (5.7), the desired result follows immediately from Corollary 5.2 b). ∎

Remark 5.4.

a) If A is unbounded, then the construction of $\rho_{[s]}$ in Proposition 5.1 leads to $\sup_{s\in\mathbb{N}} \rho_{[s]} = 1$. According to their construction in Corollary 5.2, the constants $M_{[s]}$ in Corollary 5.3 b) are therefore not uniformly bounded.

b) Under the assumptions of Corollary 5.3 b) and, additionally

$$K := \lim_{i\to\infty} K_i = \sum_{k=1}^{\infty}\sum_{l=1}^{m} c^{(kl)} < \infty,$$

the error bound in b) can be slightly reformulated to

$$\|\Pi_{[s]}(P_i - \widetilde{P}_i)\Pi_{[s]}^*\| \le N_{[s]} \cdot K_i,$$

for some $N_{[s]} > 0$.

Under the assumption that inexact ADI iteration converges, we now present estimates for the difference between the limit of inexact ADI iteration and the Gramian operator P.

Theorem 5.5. *Let X be a Hilbert space, let $A : D(A) \subset X \to X$ be a Riesz-spectral operator that generates an exponentially stable semigroup $T(\cdot)$, and let $B = [b_1, \dots, b_m] \in \mathcal{B}(\mathbb{C}^m, D(A^*)')$ be an admissible control operator for $T(\cdot)$. Let $J \in \mathbb{N}$, let $(p_i)_i$ be a J-cyclic sequence in $\mathbb{C}_-$. Assume that for all $i \in \mathbb{N}$, the operator $V_i = [v_{i1}, \dots, v_{im}] \in \mathcal{B}(\mathbb{C}^m, X)$ is obtained by Algorithm 1.*
Let $(X^{(i)})$ be a sequence of subspaces of X, let $\Pi^{(1)} \in \mathcal{B}(D(A^*)')$ be a projector onto $X^{(1)}$, and, for $i \ge 2$, let $\Pi^{(i)} \in \mathcal{B}(X)$ be a projector onto $X^{(i)}$. Further, let $\widetilde{A}_{p_i}^{(i)} \in \mathcal{B}(X^{(i)})$, and assume that the operators $\widetilde{V}_i = [\widetilde{v}_{i1}, \dots, \widetilde{v}_{im}] \in \mathcal{B}(\mathbb{C}^m, X)$ are obtained by Algorithm 2.

a) Assume that the Gramian P of (A, B) is compact and let $\varepsilon \in \mathbb{R}_{>0}$. Then for all inexact ADI iterations with the property that the error bounds (5.3) *are fulfilled in each step, and $\|\widetilde{P}_i\|_{\mathcal{B}(X)} \le C \in \mathbb{R}$ for all $i \in \mathbb{N}$, there exist $M \in \mathbb{R}_{>0}$, $k \in \mathbb{N}$ with the following property: For all $i \in \mathbb{N}$ with $i \ge k$, there holds*

$$\|P - \widetilde{P}_i\|_{\mathcal{B}(X)} \le \varepsilon + M \cdot \sum_{k=1}^{\infty}\sum_{l=1}^{m} c^{(kl)}.$$

b) Assume that the Gramian P of (A, B) is of p-th Schatten class and let $\varepsilon \in \mathbb{R}_{>0}$. Then for all inexact ADI iterations with the property that the error bounds (5.3) *are fulfilled in each step, and $\|\widetilde{P}_i\|_{\mathcal{S}_p(X)} \le C \in \mathbb{R}$ for all $i \in \mathbb{N}$, there exist $M \in \mathbb{R}_{>0}$, $k \in \mathbb{N}$ with the following property: For all $i \in \mathbb{N}$ with $i \ge k$, there holds*

$$\|P - \widetilde{P}_i\|_{\mathcal{S}_p(X)} \le \varepsilon + M \cdot \sum_{k=1}^{\infty}\sum_{l=1}^{m} c^{(kl)}.$$

Proof. We only prove a), since b) is analogous:
The sequence $(\widetilde{P}_i)$ is bounded by assumption. By construction we further have $\widetilde{P}_i \le \widetilde{P}_{i+1}$ for all $i \in \mathbb{N}$. Since each $\widetilde{P}_i$ has finite-dimensional range, we even have $\widetilde{P}_i \in \mathcal{K}(X)$ for all $i \in \mathbb{N}$. Proposition A.2 now gives rise to convergence in $\mathcal{B}(X)$. That is, there exists some $\widetilde{P} \in \mathcal{K}(X)$ such that

$$\lim_{i\to\infty} \|\widetilde{P} - \widetilde{P}_i\|_{\mathcal{B}(X)} = 0.$$

The monotonicity of the sequence further implies that for all $i \in \mathbb{N}$ holds

$$P_i \le P, \text{ and } \widetilde{P}_i \le \widetilde{P}.$$

Let $\varepsilon > 0$ and let $k \in \mathbb{N}$ such that for all $i \geq k$, there holds

$$\|P - P_i\|_{\mathcal{B}(X)} < \frac{\varepsilon}{5} \text{ and } \|\widetilde{P} - \widetilde{P}_i\|_{\mathcal{B}(X)} < \frac{\varepsilon}{5}.$$

Spectral decomposition implies that that there exists some orthogonal projector $\Pi \in \mathcal{B}(X)$, such that

$$\|P - \widetilde{P}\|_{\mathcal{B}(X)} < \|\Pi^*(P - \widetilde{P})\Pi\|_{\mathcal{B}(X)} + \frac{\varepsilon}{5}.$$

Let $(\Pi^{[s]})_s$ be constructed as in the proof of Proposition 5.1, i.e., for the sequence $(\lambda_i)_i$ of eigenvalues of A, $\Pi^{[s]}$ projects onto the eigenspace corresponding to the eigenvalues $\lambda_i, \ldots, \lambda_s$, and along the complementary eigenspace. By the property of A being a Riesz-spectral operator, the sequence $(\Pi^{[s]})_s$ converges in the strong operator topology towards the identity operator. In particular, the Banach-Steinhaus Theorem [27] implies the existence of some $C > 0$ with $\|\Pi^{[s]}\| < C$ for all $s \in \mathbb{N}$. Another consequence of strong convergence of $(\Pi^{[s]})_s$ to I is that the sequence of complementary projectors $(I - \Pi^{[s]})_s$ converges to zero in the strong operator topology. Since $\operatorname{im}\Pi$ is finite-dimensional, we have

$$\lim_{s\to\infty} \|(I - \Pi^{[s]})\Pi\|_{\mathcal{B}(X)} = 0.$$

Consequently, there exists some $s \in \mathbb{N}$ with

$$\|(I - \Pi^{[s]})\Pi\|_{\mathcal{B}(X)} \cdot \big(\|P\|_{\mathcal{B}(X)} + L\big) \cdot (3C + 1) < \frac{\varepsilon}{5}.$$

Further, by making use of the monotonicity of (inexact) ADI iteration, we have

$$\|P_i - \widetilde{P}_i\|_{\mathcal{B}(X)} \leq \|P\|_{\mathcal{B}(X)} + \|\widetilde{P}\|_{\mathcal{B}(X)} \leq \|P\|_{\mathcal{B}(X)} + L.$$

Incorporating the above findings, we find that for all $i \geq k$, there holds

$$\begin{aligned}
&\|P - \widetilde{P}_i\|_{\mathcal{B}(X)} \\
\leq &\|P - \widetilde{P}\|_{\mathcal{B}(X)} + \underbrace{\|\widetilde{P} - \widetilde{P}_i\|_{\mathcal{B}(X)}}_{<\frac{\varepsilon}{5}} \\
< &\|\Pi^*(P - \widetilde{P})\Pi\|_{\mathcal{B}(X)} + \frac{2\varepsilon}{5} \\
\leq &\underbrace{\|\Pi^*(P - P_i)\Pi\|_{\mathcal{B}(X)}}_{<\frac{\varepsilon}{5}} + \|\Pi^*(P_i - \widetilde{P}_i)\Pi\|_{\mathcal{B}(X)} + \underbrace{\|\Pi^*(\widetilde{P}_i - \widetilde{P})\Pi\|_{\mathcal{B}(X)}}_{<\frac{\varepsilon}{5}} + \frac{2\varepsilon}{5} \\
< &\frac{3\varepsilon}{5} + \|\Pi^*(P_i - \widetilde{P}_i)\Pi\|_{\mathcal{B}(X)} \\
< &\frac{4\varepsilon}{5} + \|\Pi^*(\Pi^{[s]})^*(P_i - \widetilde{P}_i)\Pi^{[s]}\Pi\|_{\mathcal{B}(X)} + 2\|\Pi^*(\Pi^{[s]})^*(P_i - \widetilde{P}_i)(I - \Pi^{[s]})\Pi\|_{\mathcal{B}(X)} \\
&\quad + \|((I - \Pi^{[s]})\Pi)^*(P_i - \widetilde{P}_i)((I - \Pi^{[s]})\Pi)\|_{\mathcal{B}(X)}
\end{aligned}$$

$$\leq \frac{4\varepsilon}{5} + \|(\Pi^{[s]})^*(P_i - \widetilde{P}_i)\Pi^{[s]}\|_{\mathcal{B}(X)} + 2\underbrace{\|\Pi^{[s]}\|_{\mathcal{B}(X)}}_{\leq C}\underbrace{\|P_i - \widetilde{P}_i\|_{\mathcal{B}(X)}}_{\leq \|P\|_{\mathcal{B}(X)}+L}\|(I - \Pi^{[s]})\Pi\|_{\mathcal{B}(X)}$$

$$+ \underbrace{\|\Pi(I - \Pi^{[s]})\|_{\mathcal{B}(X)}}_{\leq 1+C}\underbrace{\|P_i - \widetilde{P}_i\|_{\mathcal{B}(X)}}_{\leq \|P\|_{\mathcal{B}(X)}+L}\|(I - \Pi^{[s]})\Pi)\|_{\mathcal{B}(X)}$$

$$\leq \frac{4\varepsilon}{5} + \|(\Pi^{[s]})^*(P_i - \widetilde{P}_i)\Pi^{[s]}\|_{\mathcal{B}(X)} + \underbrace{(3C+1)\cdot\big(\|P\|_{\mathcal{B}(X)} + L\big)\cdot\|(I - \Pi^{[s]})\Pi)\|_{\mathcal{B}(X)}}_{<\frac{\varepsilon}{5}}$$

$$< \varepsilon + \|(\Pi^{[s]})^*(P_i - \widetilde{P}_i)\Pi^{[s]}\|_{\mathcal{B}(X)}.$$

Now using Corollary 5.3 (see also Remark 5.4 b)), there exists some $M > 0$ such that

$$\|\Pi_{[s]}(P_i - \widetilde{P}_i)\Pi_{[s]}^*\| \leq M \cdot \sum_{k=1}^{\infty}\sum_{l=1}^{m} c^{(kl)},$$

which implies a). ∎

Remark 5.6. The above assumption on inexact ADI iteration are for instance fulfilled, if the projectors $\Pi^{(i)} = \widetilde{\Pi}$ and spaces $X^{(i)} = \widetilde{X}$ are finite-dimensional and constant, with, moreover

$$(\widetilde{A}_{p_i}^{(i)})^{-1} = (\widetilde{A} + p_i\widetilde{M})^{-1} \cdot \widetilde{M},$$

where $\widetilde{M}, \widetilde{A} \in \mathcal{B}(\widetilde{X})$ are invertible operators with the additional property that $\widetilde{A} + p\widetilde{M}$ is invertible for all $p \in \overline{\mathbb{C}_-}$. This follows from the results in [6].

6. Systems governed by the heat equation. To demonstrate the applicability of the so far presented operator theoretic results, we consider an infinite-dimensional system that is governed by the heat equation with spatially (but not time) constant Robin boundary conditions; the latter is assumed to be the input variable $u(\cdot) : \mathbb{R}_{\geq 0} \to \mathbb{R}$ of the system.

More precisely, for a bounded domain $\Omega \subset \mathbb{R}^n$ with piecewise C^2 boundary $\partial\Omega$ [1], we consider the heat equation

$$\frac{\partial x}{\partial t}(\xi, t) = \Delta x(\xi, t), \quad (\xi, t) \in \Omega \times \mathbb{R}_{\geq 0} \tag{6.1a}$$

with boundary condition

$$\nu(\xi)^T \nabla x(\xi, t) + \alpha x(\xi, t) = u(t), \quad (\xi, t) \in \partial\Omega \times \mathbb{R}_{\geq 0}, \tag{6.1b}$$

where $\nu(\xi)$ denotes the outward normal to $\partial\Omega$ in $\xi \in \partial\Omega$, $\alpha \in \mathbb{R}_{>0}$, and $u \in L_2(\mathbb{R}_{\geq 0})$ is the input of the system.

In the first step, we rewrite this as a system $\dot{x}(t) = Ax(t) + Bu(t)$, where the state is given by the spatial temperature function at time t, that is, $x(t) := x(\cdot, t) \in L_2(\Omega) := X$; the input is one-dimensional, i.e., $U = \mathbb{C}$. The construction of A and B has been performed in [12] in the case where $\alpha = 0$; our case can be treated analogously (it can, for instance, be derived from the results in [12] by additionally employing the output feedback theory presented in [41]). The operators A and B are given by

$$\begin{aligned} D(A) &= \{x \in H^1(\Omega) \mid \Delta x \in L^2(\Omega),\ \nu^T \nabla x + \alpha x = 0 \text{ on } \partial\Omega\}, \\ Ax &= \Delta x \quad \text{for all } x \in D(A), \\ \langle Bu, z\rangle_{D(A^*)', D(A^*)} &= u \cdot \int_{\partial\Omega} z(\xi) d\sigma_\xi, \end{aligned} \tag{6.2}$$

where by $d\sigma_\xi$ we denote the surface measure on $\partial\Omega$. It follows by the Gauß Theorem that A is self-adjoint with $A \leq 0$. Furthermore, $0 \in \rho(A)$, since the elliptic problem

$$\begin{aligned} -\Delta x(\xi) &= z(\xi), && \xi \in \Omega, \\ \nu(\xi)^T \nabla x(\xi) + \alpha x(\xi) &= 0, && \xi \in \partial\Omega \end{aligned}$$

has a unique solution for all $z \in L_2(\Omega)$. By the Rellich-Kondrachov Theorem [1, Thm. 6.3], $H^1(\Omega)$ is compactly embedded in $L_2(\Omega)$. This gives rise to $A^{-1} \in \mathcal{K}(L_2(\Omega))$, whence, by the resolvent identity [42, Rem. 2.2.5], A has compact resolvent. The construction of B in (6.2) further leads to

$$B'z = \int_{\partial\Omega} z(\xi) d\sigma_\xi \quad \text{for all } z \in D(A^*) = D(A).$$

Since we can infer from the Gauß Theorem that

$$\begin{aligned} \|z\|^2_{D((-A)^{\frac{1}{2}})} &= \|z\|^2_{L_2(\Omega)} - \langle z, Az\rangle_{L_2(\Omega)} = \|z\|^2_{L_2(\Omega)} - \int_\Omega z(\xi)\Delta z(\xi) d\xi \\ &= \|z\|^2_{L_2(\Omega)} + \|\nabla z\|^2_{L_2(\Omega)} + \alpha \int_{\partial\Omega} z^2(\xi) d\sigma_\xi \quad \text{for all } z \in D(A), \end{aligned}$$

the control operator fulfills $B \in \mathcal{B}(\mathbb{C}, D((-A)^{\frac{1}{2}}))$, whence, due to the one-dimensionality of the input space, there further holds $B \in \mathcal{S}_2(\mathbb{C}, D((-A)^{\frac{1}{2}}))$. Altogether, we are now in the situation of Proposition 3.6 and are able formulate the following result:

Corollary 6.1. *Assume that $\Omega \subset \mathbb{R}^n$ is a bounded domain with piecewise C^2 boundary $\partial\Omega$. Then the operator A as defined in (6.2) generates an exponentially stable semigroup $T(\cdot)$. The control operator B as in (6.2) is moreover admissible for $T(\cdot)$. Furthermore, the Gramian P of (A, B) is nuclear with*

$$\|P\|_{\mathcal{S}_1(X)} = \frac{1}{2} \cdot \int_{\partial\Omega} x(\xi) d\sigma_\xi, \tag{6.3a}$$

where $x \in H^1(\Omega)$ solves

$$\begin{aligned} -\Delta x(\xi) &= 0, && \xi \in \Omega, \\ \nu(\xi)^T \nabla x(\xi) + \alpha x(\xi) &= 1, && \xi \in \partial\Omega \end{aligned} \tag{6.3b}$$

Proof. The considerations in front of this theorem imply that A is self-adjoint with compact resolvent, $A \leq 0$, $0 \in \rho(A)$ and $(-A)^{-\frac{1}{2}}B \in \mathcal{S}_2(\mathbb{C}, L_2(\Omega))$. Exponential stability of the semigroup $T(\cdot)$ generated by A, admissibility of B for $T(\cdot)$, and nuclearity of the Gramian P are then immediate consequences of Proposition 3.6.
Formula (6.3) follows by an application of the findings in Remark 4.2 f) to the expression in (3.7). □

We now consider the ADI iteration for the heat equation (6.1). Since A is self-adjoint in this case, it makes sense to only choose real shift parameters. A substitution $q_i = -p_i$ of the shift parameters leads, according to Remark 4.2 f), to the ADI algorithm in the following form:

Remark 6.2. Algorithm 3 requires the solution of a sequence of Helmholtz equations. These can be solved via a finite element method. Note that, if the grid is chosen to be the same in all equations, then Algorithm 3 will be arithmetically equivalent to

Algorithm 3 ADI iteration for heat equation with one-dimensional Robin boundary control (6.1).

Input: Bounded domain $\Omega \subset \mathbb{R}^n$ with piecewise C^2 boundary $\partial\Omega$, negatives of the shift parameters $q_1, \dots, q_{i_{\max}} \in \mathbb{R}_{>0}$
Output: $S = S_{i_{\max}} \in \mathcal{B}(\mathbb{R}^{i_{\max}}, X)$, such that $SS^* \approx P$, where P is the Gramian of (A, B) (with A, B as in (6.2)).

1: Solve

$$\begin{aligned} q_1 \cdot v_1(\xi) - \Delta v_1(\xi) &= 0, && \xi \in \Omega, \\ \nu(\xi)^T \nabla v_1(\xi) + \alpha v_1(\xi) &= 1, && \xi \in \partial\Omega \end{aligned}$$

for $v_1 \in L_2(\Omega)$.
2: Define $S_1 = \sqrt{2q_1} \cdot v_1 \in \mathcal{B}(\mathbb{C}, L_2(\Omega))$
3: **for** $i = 2, 3, \dots, i_{\max}$ **do**
4: Solve

$$\begin{aligned} q_i \cdot \hat{v}(\xi) - \Delta \hat{v}(\xi) &= v_{i-1}(\xi), && \xi \in \Omega, \\ \nu(\xi)^T \nabla \hat{v}(\xi) + \alpha \hat{v}(\xi) &= 0, && \xi \in \partial\Omega \end{aligned}$$

for $\hat{v} \in L_2(\Omega)$.
5: Set $v_i = v_{i-1} - (q_i + q_{i-1}) \cdot \hat{v}$
6: $S_i = [\, S_{i-1} \, , \, \sqrt{2q_i} \cdot v_i \,] \in \mathcal{B}(\mathbb{R}^i, L_2(\Omega))$
7: **end for**

the approach of semi-discretizing the heat equation with respect to space, and an accordant application of the matrix version of the ADI method to the semi-discretized system.

Applying Proposition 4.8 to the heat equation considered in this part, we can derive the following expression for the error of the ADI iteration:

COROLLARY 6.3. *Assume that $\Omega \subset \mathbb{R}^n$ is a bounded domain with piecewise C^2 boundary $\partial\Omega$. Then, in the notation of Algorithm 3, there holds*

$$\|P - P_i\|_{\mathcal{S}_1(X)} = \frac{1}{2} \cdot \int_{\partial\Omega} x(\xi) d\sigma_\xi - 2 \sum_{k=1}^{i} q_k \cdot \int_\Omega |v_k(\xi)|^2 d\xi, \tag{6.4}$$

where $x \in H^1(\Omega)$ solves (6.3b).

6.1. Numerical Results. Now, we illustrate our findings with a short numerical example. To this end we consider (6.1) on the L-shaped domain $(0,1)^2 \setminus (0.5,1)^2$. We fix $\alpha = 1$ and can evaluate (6.3) exactly as $\|P\|_{S_1(X)} = 2$ because the solution to (6.3b) is given by $x(\cdot) \equiv \frac{1}{\alpha} = 1$. Now, we can apply the inexact version of Algorithm 3, compare Algorithm 2 to calculate approximate values $\tilde{V}_i$ and $\tilde{S}_i$. To do so, we use a finite element discretization of the PDE's given in Algorithm 3. The discretization is done using a Cartesian mesh consisting of square elements with maximal diameter h. On this mesh we define a subspace $V_h \subset H^1(\Omega)$ using piecewise bilinear finite elements. The calculations are done using the toolkit `DOpElib` [16] based upon the C++-library `deal.II`, see [7,8]. In order to assert that the approximation error during the solution of the discrete PDE is below a given tolerance TOL > 0 we employ a standard residual based L^2-error estimator η, see, e.g., [2]. Thus we can allow for refinement of the discretization if the error is too large, i.e., $\eta > \text{TOL}$ and for optional coarsening of the discretization once the error is too small, i.e., $\eta < 0.1\,\text{TOL}$. Note, that this means that the different approximations are not obtained with the same discretization and thus the software needs to work with solutions given on different meshes which is done in the library `DOpElib`.

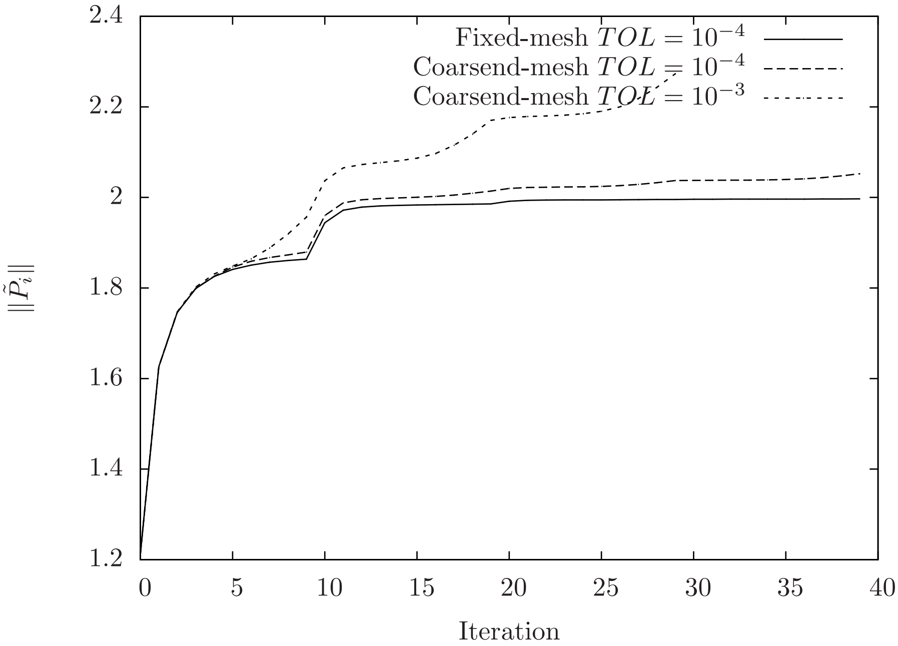

FIG. 6.1. *Behavior of the approximated Gramians for different algorithmic settings*

As a first test case we consider the behavior of $\|\tilde{P}_i\|_{S_1(X)}$ for various settings in the algorithm. The results are depicted in Figure 6.1. The shift parameters where chosen by applying the method of WACHSPRESS [40] on the basis of the lowest hundred eigenvalues of the Robin Laplacian, which are given by $\pi^2(i^2 + j^2)$ where $i, j = 1, \ldots, 10$.

Here, the solid black line corresponds to the case when one (fixed) uniform mesh of mesh size h is used to discretize the PDE. Then from a priori error estimates we can conclude that the approximation errors in (5.3) satisfy

$$c^{(il)} \leq Ch^2 \|\tilde{v}_{i-1}\|$$

Thus by the fact that the elements $\|\tilde{v}_i\|$ are summable we deduce that the error $\|P - \tilde{P}_i\|_{S_1(X)}$ will be bounded which can seen as well in the numerical result.

Further, to qualitatively test our error estimates we in addition employed refinement and coarsening during the ADI-iteration. To this end employed a standard residual based L^2-error estimator η. By the well known reliability and efficiency of η we can steer the meshes in such a way that the approximation errors $c^{(il)}$ $(i = 1, \ldots;$ $l = 1)$ given in (5.3) satisfy

$$c^{(il)} \approx \text{TOL}\,.$$

Since the $\sum_{i=1}^{\infty} c^{(il)} = \infty$ we expect the error to grow as $i \to \infty$.

Finally, we have a more fine grained look onto the iteration. As is shown in

Table 6.1
Convergence of the ADI-iterations with fixed mesh (left) and adjusted to tolerance 10^{-4} *(right)*

Iter. (i)	unknowns	$\|\tilde{P}_i\|_{S_1(X)}$	$\|\tilde{v}_i\|$	unknowns	$\|\tilde{P}_i\|_{S_1(X)}$	$\|\tilde{v}_i\|$
0	49665	1.21309	0.030728	49665	1.21309	0.030728
1	49665	1.62615	0.00910443	49665	1.62615	0.00910443
2	49665	1.74612	0.00121557	3201	1.74606	0.00121501
3	49665	1.79888	0.000267274	3201	1.79899	0.000268148
4	49665	1.82545	6.76124e-05	3201	1.82592	6.85014e-05
5	49665	1.841	2.18725e-05	225	1.8459	2.81147e-05
6	49665	1.85054	7.57344e-06	225	1.85887	1.02857e-05
7	49665	1.85676	2.61384e-06	225	1.86706	3.44507e-06
8	49665	1.86073	1.22589e-06	225	1.87312	1.87149e-06
9	49665	1.86353	7.09549e-07	225	1.87935	1.57927e-06
10	49665	1.94405	0.00203958	65	1.95994	0.00204132
11	49665	1.97195	0.000614905	833	1.98789	0.00061603
12	49665	1.97869	6.83044e-05	833	1.99473	6.93052e-05
13	49665	1.98125	1.29659e-05	833	1.99748	1.39308e-05
14	49665	1.98244	3.0429e-06	833	1.99903	3.94527e-06
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Table 6.1 including the possibility to coarsen the mesh allows an almost identical approximation of the Gramian with severely fewer unknowns needed in the calculation.

7. Conclusion and Outlook. In this work, the alternating direction implicit (ADI) iteration has been generalized to operator Lyapunov equations. It is shown that, for finite-dimensional input space, this method provides finite-rank approximations of the Gramian. Conditions for convergence in strong operator topology, operator and Schatten norms have been presented. Motivated by the fact that computations in infinite-dimensional spaces (such as, the solution of differential equations) can usually only be done approximatively, we have also presented an error analysis for inexact ADI iteration.

The presented theory and methods have been applied to a heat equation with boundary control. It turned out that, for this class, the ADI iteration for determination of Gramians requires the numerical solution of a sequence of Helmholtz equations. This has been done by employing an adaptive finite-element solver.

One of the most important application of Gramian operators is in model reduction of infinite-dimensional systems by balanced truncation [15, 18], i.e., the finite-dimensional approximation of infinite-dimensional input-output systems. The possible application of the presented ADI iteration to numerically solve operator Lyapunov equations is however not limited to only balanced truncation model reduction: For instance, the problem of linear-quadratic optimal control of infinite-dimensional systems requires the solution of operator Riccati equations [43]. A typical approach to numerical solution is the so-called *Kleinman iteration* [11], which consists of the (numerical) solution of a sequence of operator Lyapunov equations. A combination of the presented method with Kleinman iteration results in an algorithm for the solution of operator Riccati equations. A more detailed consideration of this class of problems is subject of further research.

Acknowledgement. The authors would like to thank Jens Saak for providing a routine for determining shift parameters.

Appendix A. Monotone operator sequences.

We present convergence results for sequences self-adjoint operators which are monotonic and bounded. These results are the basis for the proof that the ADI iteration converges for operator Lyapunov equations. First we repeat a classical result. Thereafter we focus on sequences of compact and p-th Schatten class operators.

THEOREM A.1. *[29, p. 263] Let X be a Hilbert space and $(P_i)_i$ be a sequence of self-adjoint operators in $\mathcal{B}(X)$ with $P_{i+1} \geq P_i$ for all $i \in \mathbb{N}$. Moreover, assume that there exists some $Q \in \mathcal{B}(X)$ such that $P_i \leq Q$ for all $i \in \mathbb{N}$. Then there exists some self-adjoint $P \in \mathcal{B}(X)$ such that $(P_i)_i$ converges to P in the strong operator topology, that is,*

$$\lim_{i\to\infty} P_i x = Px \qquad \text{for all } x \in X.$$

PROPOSITION A.2. *Let X be a Hilbert space and $(P_i)_i$ be a sequence of nonnegative bounded operators such that $P_i \leq P_{i+1}$ for all $i \in \mathbb{N}$. Moreover, assume that there exists some $Q \in \mathcal{K}(X)$ with $P_i \leq Q$ for all $i \in \mathbb{N}$. Then there exists some $P \in \mathcal{K}(X)$ such that $(P_i)_i$ converges to P in the operator norm, that is,*

$$\lim_{i\to\infty} \|P_i - P\|_{\mathcal{B}(X)} = 0.$$

Proof. We know from Theorem A.1 that $(P_i)_i$ converges to some $P \in \mathcal{B}(X)$ in the strong operator topology. Using $0 \leq P_i \leq P \leq Q \in \mathcal{K}(X)$, the min-max-Theorem of Courant-Fischer [28, Sec. 7.5] gives rise to $P_i, P \in \mathcal{K}(X)$. It remains to be shown that the sequence $(\|P - P_i\|_{\mathcal{B}(X)}$ converges to zero: The sequence $(P - P_i)_i$ is nonnegative and decreasing. Since P is compact and self-adjoint, spectral decomposition implies that there exists some orthogonal projector $\Pi_k \in \mathcal{B}(X)$ with k-dimensional range and $\|(I - \Pi_k)P(I - \Pi_k)\|_{\mathcal{B}(X)} < \frac{\varepsilon}{4}$. Since $(\Pi_k(P - P_i)\Pi_k)_i$ consists of operators with range contained in the finite-dimensional space $\operatorname{im}\Pi_k$ and, moreover, (P_i) converges to P in the strong operator topology, we obtain

$$\lim_{i\to\infty} \|\Pi_k(P - P_i)\Pi_k\|_{\mathcal{B}(X)} = 0.$$

Let $\varepsilon > 0$. Then there exists some $N \in \mathbb{N}$ such that for all $i > N$ holds

$$\|\Pi_k(P - P_i)\Pi_k\|_{\mathcal{B}(X)} < \frac{\varepsilon}{4}.$$

Let $x_1 \in \operatorname{im}\Pi$, $x_2 \in \ker\Pi$ with $\|x_1\|_X = \|x_2\|_X = 1$ and

$$\langle x_1, (I - \Pi_k)(P - P_i)\Pi_k x_2\rangle_X \leq -\|(I - \Pi_k)(P - P_i)\Pi_k\|_{\mathcal{L}(X)}.$$

Defining $x = x_1 + x_2$, the expansion

$$\begin{aligned}
0 \leq& \langle x, (P - P_i)x\rangle_X \\
=& \langle x, \Pi_k(P - P_i)\Pi_k x\rangle_X + 2\operatorname{Re}\langle x, (I - \Pi_k)(P - P_i)\Pi_k x\rangle_X \\
&+ \langle x, (I - \Pi_k)(P - P_i)(I - \Pi_k)x\rangle_X \\
\leq& \langle x_1, \Pi_k(P - P_i)\Pi_k x_1\rangle_X + 2\operatorname{Re}\langle x_2, (I - \Pi_k)(P - P_i)\Pi_k x_1\rangle_X \\
&+ \langle x_2, (I - \Pi_k)P(I - \Pi_k)x_1\rangle_X \\
\leq& \|\Pi_k(P - P_i)\Pi_k\|_{\mathcal{B}(X)} - 2\|(I - \Pi_k)(P - P_i)\Pi_k\|_{\mathcal{B}(X)} \\
&+ \|(I - \Pi_k)P(I - \Pi_k)\|_{\mathcal{B}(X)} \\
<& \frac{\varepsilon}{4} - 2\|(I - \Pi_k)(P - P_i)\Pi_k\|_{\mathcal{B}(X)} + \frac{\varepsilon}{4}
\end{aligned}$$

implies

$$\|(I - \Pi_k)(P - P_i)\Pi_k\|_{\mathcal{B}(X)} < \frac{\varepsilon}{4}.$$

By $\Pi_k(P - P_i)(I - \Pi_k) = ((I - \Pi_k)(P - P_i)\Pi_k)^*$ we further obtain

$$\|\Pi_k(P - P_i)(I - \Pi_k)\|_{\mathcal{B}(X)} < \frac{\varepsilon}{4},$$

which, altogether, leads to

$$\begin{aligned}
\|P - P_i\|_{\mathcal{B}(X)} \leq& \|\Pi_k(P - P_i)\Pi_k\|_{\mathcal{B}(X)} + \|(I - \Pi_k)(P - P_i)\Pi_k\|_{\mathcal{B}(X)} \\
&+ \|\Pi_k(P - P_i)(I - \Pi_k)\|_{\mathcal{B}(X)} + \|(I - \Pi_k)(P - P_i)(I - \Pi_k)\|_{\mathcal{B}(X)} \\
<& \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} + \frac{\varepsilon}{4} = \varepsilon.
\end{aligned}$$

□

PROPOSITION A.3. *Let X be a Hilbert space and $(P_i)_i$ be a sequence of nonnegative operators such that $P_i \leq P_{i+1}$ for all $i \in \mathbb{N}$. Moreover, assume that $p \in [1, \infty)$ and there exists some $Q \in \mathcal{S}_p(X)$ with $P_i \leq Q$ for all $n \in \mathbb{N}$. Then there exists some $P \in \mathcal{S}_p(X)$ such that $(P_i)_i$ converges to P in the nuclear norm, that is,*

$$\lim_{n\to\infty} \|P_i - P\|_{\mathcal{S}_p(X)} = 0.$$

Proof. Theorem A.1 implies that $(P_i)_i$ converges to some $P \in \mathcal{B}(X)$ in the strong operator topology. Since $0 \leq P_i \leq P \leq Q \in \mathcal{S}_p(X)$, we can infer from the "min-max-Theorem" that $P_i, P \in \mathcal{S}_p(X)$. Let $\varepsilon > 0$. Spectral decomposition implies that there exists some orthogonal projector $\Pi_k \in \mathcal{B}(X)$ with k-dimensional range and

$$\|(I - \Pi_k)P(I - \Pi_k)\|_{\mathcal{S}_p(X)} < \frac{\varepsilon}{4}.$$

Let σ_j and ς_j be the j-th singular value (of decreasing order) of $(I-\Pi_k)P(I-\Pi_k)$ and $(I-\Pi_k)(P-P_i)(I-\Pi_k)$, respectively. Since $0 \leq P - P_i \leq P$, the min-max-Theorem of Courant-Fischer implies $\sigma_j \leq \varsigma_j$ for all $j \in \mathbb{N}$, and thus

$$\|(I - \Pi_k)(P - P_i)(I - \Pi_k)\|_{\mathcal{S}_p(X)} \leq \|(I - \Pi_k)P(I - \Pi_k)\|_{\mathcal{S}_p(X)} \leq \frac{\varepsilon}{4}.$$

Operators of Schatten class are compact, whence we can apply Proposition A.2 to see that

$$\lim_{i\to\infty} \|\Pi_k(P-P_i)\Pi_k\|_{\mathcal{B}(X)} = \lim_{i\to\infty} \|\Pi_k(P-P_i)(I-\Pi_k)\|_{\mathcal{B}(X)} = 0. \tag{A.1}$$

The expressions $\|\Pi_k(P-P_i)\Pi_k\|_{\mathcal{B}(X)}$ and $\|\Pi_k(P-P_i)(I-\Pi_k)\|_{\mathcal{B}(X)}$ are given by the largest (out of k) singular values of $\Pi_k(P-P_i)\Pi_k$ and $\Pi_k(P-P_i)(I-\Pi_k)$, respectively. Hence we may estimate

$$\begin{aligned} \|\Pi_k(P-P_i)\Pi_k\|_{\mathcal{S}_p(X)} &\le \sqrt[p]{k}\cdot \|\Pi_k(P-P_i)\Pi_k\|_{\mathcal{B}(X)}, \\ \|\Pi_k(P-P_i)(I-\Pi_k)\|_{\mathcal{S}_p(X)} &\le \sqrt[p]{k}\cdot \|\Pi_k(P-P_i)(I-\Pi_k)\|_{\mathcal{B}(X)}. \end{aligned}$$

Relation (A.1) gives rise to the existence of some $N\in\mathbb{N}$ such that for all $i\ge N$, there holds

$$\|\Pi_k(P-P_i)\Pi_k\|_{\mathcal{B}(X)} < \frac{\varepsilon}{4\sqrt[p]{k}} \quad \text{and} \quad \|\Pi_k(P-P_i)(I-\Pi_k)\|_{\mathcal{B}(X)} < \frac{\varepsilon}{4\sqrt[p]{k}}.$$

Thus we have

$$\begin{aligned} \|P-P_i\|_{\mathcal{S}_p(X)} \le& \|\Pi_k(P-P_i)\Pi_k\|_{\mathcal{S}_p(X)} + 2\cdot\|\Pi_k(P-P_i)(I-\Pi_k)\|_{\mathcal{S}_p(X)} \\ &+ \|(I-\Pi_k)(P-P_i)(I-\Pi_k)\|_{\mathcal{S}_p(X)} \\ \le& \sqrt[p]{k}\cdot\|\Pi_k(P-P_i)\Pi_k\|_{\mathcal{B}(X)} + 2\sqrt[p]{k}\cdot\|\Pi_k(P-P_i)(I-\Pi_k)\|_{\mathcal{B}(X)} \\ &+ \|(I-\Pi_k)P(I-\Pi_k)\|_{\mathcal{S}_p(X)} \\ <& \sqrt[p]{k}\cdot\frac{\varepsilon}{4\sqrt[p]{k}} + 2\sqrt[p]{k}\cdot\frac{\varepsilon}{4\sqrt[p]{k}} + \frac{\varepsilon}{4} = \varepsilon. \end{aligned}$$

□

REFERENCES

[1] R.A. Adams, *Sobolev Spaces*, Academic Press, London, Sydney, Tokyo, Toronto, 1978.

[2] Mark Ainsworth and John Tinsley Oden, *A Posteriori Error Estimation in Finite Element Analysis*, Pure and Applied Mathematics (New York), Wiley-Interscience [John Wiley & Sons], New York, 2000.

[3] A.C. Antoulas, *Approximation of large-scale dynamical systems*, Advances in design and control, Society for Industrial and Applied Mathematics, Philadelphia/Pennsylvania, 2005.

[4] A.C. Antoulas, D.C. Sorensen, and Y. Zhou, *On the decay rate of Hankel singular values and related issues*, Systems Control Lett., 46 (2002), pp. 323–342.

[5] J.P. Aubin, *Applied Functional Analysis*, John Wiley & Sons, New York, Chichester,Brisbane, 1979.

[6] J.M. Badia, P. Benner, R. Mayo, E. Quintana-Ortí, G. Quintana-Ortí, and J. Saak, *Parallel order reduction via balanced truncation for optimal cooling of steel profiles;*, in Euro-Par 2005 Parallel Processing: 11th International Euro-Par Conference, J.C. Cunha and P.D. Mdeiros, eds., vol. 3548 of Lecture Notes in Computer Science, Springer-Verlag, Berlin, Heidelberg, New York, 2005, pp. 857–866.

[7] W. Bangerth, R. Hartmann, and G. Kanschat, `deal.II` *Differential Equations Analysis Library, Technical Reference*. http://www.dealii.org.

[8] ———, *deal.II — a general-purpose object-oriented finite element library*, ACM Trans. Math. Softw., 33 (2007), p. 24.

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

[10] P. Benner, S. Görner, and J. Saak, *Numerical solution of optimal control problems for parabolic systems*, in Parallel Algorithms and Cluster Computing: Implementations, Algorithms and Applications, K.-H. Hoffmann und A. Meyer, ed., vol. 52 of Lecture Notes in Computational Science and Engineering, Springer-Verlag, Berlin, Heidelberg, New York, 2006, pp. 151–169.

[11] J.A. Burns, E.W. Sachs, and L. Zietsman, *Mesh independence of Kleinman-Newton iterations for Riccati equations in Hilbert space*, SIAM J. Control Optim., 47 (2008), pp. 2663–2692.

[12] C.I. Byrnes, D.S. Gilliam, V.I. Shubov, and G. Weiss, *Regular linear systems governed by a boundary controlled heat equation*, J. Dyn. Control Syst., 8 (2002), pp. 341–370.

[13] A. Cheng and K. Morris, *Well-posedness of boundary control systems*, SIAM Journal of Control and Optimization, 42 (2003), pp. 1244–1265.

[14] R.F. Curtain and H.J. Zwart, *An Introduction to Infinite Dimensional Linear Systems Theory*, Springer, Berlin, Heidelberg, New York, 1995.

[15] K. Glover, R.F. Curtain, and J.R. Partington, *Realization and approximation of linear infinite-dimensional systems with error bounds*, SIAM J. Control Optim., 26 (1988), pp. 863–898.

[16] C. Goll, T. Wick, and W. Wollner, *The deal.II optimization environment: DOpElib*. www.dopelib.math.uni-hamburg.de.

[17] A.A. Gonchar, *Zolotarev problems connected with rational functions*, Math. USSR-Sb., 78 (1969), pp. 623–635.

[18] C. Guiver and M.R. Opmeer, *Model reduction by balanced truncation for systems with nuclear hankel operators*, tech. report, University of Bath, 2012. submitted for publication.

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

[20] S. Hansen and G. Weiss, *New results on the operator carleson measure criterion. distributed parameter systems: analysis, synthesis and applications, part 1*, IMA J. Math Control Inform., 14 (1997), pp. 3–32.

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

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

[23] A. Pazy, *Semigroups of linear operators and applications to partial differential equations*, vol. 44 of Applied Mathematical Sciences, Academic Press, New York, 1983.

[24] A. Pietsch, *Eigenvalues and s-Numbers*, no. 43 in Mathematik und ihre Anwendungen in Physik und Technik, Akademische Verlagsgesellschaft Greest & Portig K.-G., Leipzig, 1987.

[25] A.J. Pritchard and D. Salamon, *The linear quadratic control problem for infnite-dimensional systems with unbounded input and output operators*, SIAM J. Control Optim., 25 (1987), pp. 121–144.

[26] K.M. Przyluski, *The Lyapunov equation and the problem of stability for linear bounded discrete-time systems in Hilbert space*, Appl. Math. Optim., 6 (1980), pp. 97–112.

[27] M. Reed and B. Simon, *Methods of Modern Mathematical Physics I: Functional Analysis*, Academic Press, New York, London, 1972.

[28] ———, *Methods of Modern Mathematical Physics IV: Analysis of Operators*, Academic Press, New York, London, 1977.

[29] F. Riesz and B. Sz.-Nagy, *Functional Analysis*, Blackie & Son Limited, London,Glasgow, 1956.

[30] J.D. Roberts, *Linear model reduction and solution of the algebraic Riccati equation by use of the sign function*, Internat. J. Control, 32 (1980), pp. 677–687. (Reprint of Technical Report TR-13, CUED/B-Control, Engineering Deparment, Cambrige University, 1971).

[31] Y. Saad, *Numerical solution of large lyapunov equations*, in Signal Processing, Scattering and Operator Theory, and Numerical Methods, M.A. Kaarshoek, J.H. van Schuppen, and A.C.M. Ran, eds., vol. 5 of Progress in Systems and Control Theory, Birkhäuser, Boston, MA, 1990, pp. 503–511.

[32] J. Saak, *Efficient Numerical Solution of Large Scale Algebraic Matrix Equations in PDE Control and Model Order Reduction*, doctoral dissertation, Technische Universität Chemnitz, 2009.

[33] R. Schatten, *Norm ideals of completely continuous operators*, vol. 27 of Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer-Verlag, Berlin, New York, 1970.

[34] V. Simoncini, *A new iterative method for solving large-scale Lyapunov matrix equations*, SIAM J. Sci. Comput., 29 (2007), pp. 1286–1288.

[35] J.R. Singler, *Convergent snapshot algorithms algorithms for infinite-dimensional Lyapunov*

equations, IMA J. Numer. Anal., 31 (2011), pp. 1468–1496.

[36] ———, *Balanced POD for model reduction of linear PDE systems: Convergence theory*, Numer. Math., 121 (2012), pp. 127–164.

[37] R.A. Smith, *Matrix equation* $AX + XB = C$, SIAM J. Appl. Math., 16 (1968), pp. 198–201.

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

[39] H.L. Trentelman, A.A. Stoorvogel, and M. Hautus, *Control Theory for Linear Systems*, Communications and Control Engineering, Springer-Verlag, London, 2001.

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

[41] G. Weiss, *Regular Linear Systems with Feedback*, Math. Control Signals Systems, 7 (1993), pp. 23–57.

[42] G. Weiss and M. Tusznak, *Observation and Control for Operator Semigroups*, Birkhäuser Advanced Texts Basler Lehrbücher, Birkhäuser, Basel, 2009.

[43] M. Weiss, *Riccati equation theory for Pritchard-Salamon systems: a Popov function approach*, IMA J. Math Control Inform., 14 (1997), pp. 45–83.

[44] K. Willcox and J. Peraire, *Balanced model reduction via the proper orthogonal decomposition*, AIAA J., 40 (2002), pp. 2323–2330.