2013 European Control Conference (ECC)
July 17-19, 2013, Zürich, Switzerland.

Iterative Solution of Operator Lyapunov Equations arising in Heat Transfer

Timo Reis and Winnifried Wollner

Abstract— **We consider an iterative method for the numerical solution of Lyapunov equations of infinite-dimensional control systems governed by an heat equation. Inspired by the 'alternating direction implicit (ADI)' iteration, which has been successfully applied to the solution of matrix Lyapunov equations, we present a method to determine approximations of the Gramian operator corresponding to the heat equation system. This method provides approximations of finite rank and is shown to be convergent in the nuclear norm.**

I. INTRODUCTION

The *Lyapunov equation* is fundamental in the theory of linear time-invariant input-output systems. Besides testing for stability and controllability of a given system, the solutions of Lyapunov equations give rise to the optimal input energy that is required to steer the system to a prescribed state [1, Chap. 3]. A further field where Lyapunov equations arise is in *model reduction by balanced truncation* [2, Chap. 7], i.e., the approximation of linear time-invariant input-output systems by systems of considerably lower order. Lyapunov equations also occur in linearizations of *algebraic Riccati equations*, which arise in linear-quadratic optimal control; such linearizations occur when Newton's method is applied to solve algebraic Riccati equation numerically [3].

Its wide applicability has led to a variety of efficient numerical methods for the solution of Lyapunov equations for finite-dimensional linear time-invariant systems (in this case, the variable to be found is the so-called *Gramian matrix* $P \in \mathbb{R}^{n,n}$), which are also applicable to systems of considerable large state space dimension. To mention only a few methods, Lyapunov equations for finite-dimensional linear systems can be solved via Bartels-Stewart method [4], Hammarling's method [5], alternating direction implicit (ADI) iteration [6], Smith's method [7], Krylov subspace method [8], [9], and sign function method [10] (see [2, Chap. 6] for an overview). Especially the latter four mentioned methods have in common that, in the case where the input dimension is much smaller than the state space dimension, 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. An essential advantage of approximative solutions of low rank is that they are memory-saving; only $k \cdot n$ instead of n^2 entries have to be stored.

In this article, we present that the ADI iteration can as well be used to determine Gramian operators of infinite-dimensional linear time-invariant control systems. The latter class typically occurs when partial differential equations with distributed or boundary control are considered. Convergence of the ADI method in several norms will be discussed. If the input space is moreover finite-dimensional, then the ADI method will provide a sequence of approximants all having finite rank. We will place particular emphasis on the heat equation with boundary control. ADI iteration for this class will consist of the solution of a sequence of Helmholtz equations. These can been solved by using adaptive finite-element methods. For further details on the proofs we refer to the forthcoming publication [11].

II. NOTATION

Throughout the paper $\mathbb{R}_{>0}$, $\mathbb{R}_{\geq 0}$, $\mathbb{C}_+$, $\mathbb{C}_-$ and $\mathbb{C}^{n,m}$ 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 $z \in \mathbb{C}$.

We use the notation from [12] for Lebesgue $L_p(\Omega)$ and Sobolev spaces $H^k(\Omega)$.

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 [13, 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 $\|f(\cdot)\|_X \in L_p(I)$.

$\mathcal{B}(X,Y)$ is the space of bounded and linear operators $T : X \to Y$, and we abbreviate $\mathcal{B}(X) := \mathcal{B}(X,X)$. The identity mapping on X is denoted by I_X and the zero operator from X to Y by $0_{X,Y}$.

The *graph norm* of a linear operator $T : D(T) \subset X \to Y$ 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.

This was supported by the Lothar Collatz Center for Computing in Science

T. Reis and W. Wollner are at the Department of Mathematics, University of Hamburg, Bundesstr. 55, 20146 Hamburg, Germany `{timo.reis|winnifried.wollner}@uni-hamburg.de`

The *adjoint* of $T \in \mathcal{B}(X,Y)$ is denoted by $T^* \in \mathcal{B}(Y,X)$ and the *dual* by $T' \in \mathcal{B}(Y',X')$ where Y' and X' denote the dual spaces to Y and 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 there exists some $z \in X$ with $\langle Tx, y\rangle_Y = \langle x, z\rangle_X$ 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 [14, 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* ($P \geq 0$) if $\langle x, Px\rangle_X \geq 0$ 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 $P_1 - P_2 \geq 0$.

Compact operators are known to admit a singular value decomposition

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

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 [15, pp. 203].

A compact operator $T \in \mathcal{K}(X)$ is called of *p-th Schatten class operator*, if the sequence consisting of its singular values fulfill $\sum_{i=1}^{\infty} \sigma_i^p < \infty$. In this case we write $T \in \mathcal{S}_p(X,Y)$. Elements of $\mathcal{S}_p(X,Y)$ are provided with the $\|T\|_{\mathcal{S}_p(X,Y)} = (\sum_{k=1}^{\infty} \sigma_i^p)^{1/p}$. 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 [16, pp. 126]. Nuclear operators have a well-defined *trace*. That is

$$\operatorname{tr}(T) = \sum_{i=1}^{\infty} \langle e_i, Te_i\rangle, \tag{2}$$

where (e_i) is an (arbitrary) orthonormal basis of X [15, pp. 206].

III. INFINITE-DIMENSIONAL SYSTEMS AND OPERATOR LYAPUNOV EQUATIONS

For Hilbert spaces X and U, a linear time-invariant control system

$$\dot{x}(t) = Ax(t) + Bu(t), \tag{3}$$

where $A : D(A) \subset X \to X$ is closed and $B \in \mathcal{B}(U, D(A^*)')$. In the following we collect some special properties which are needed throughout this article.

Definition 1 (Semigroups, generators, stability) Let X be a Hilbert space. 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 $T(\cdot)$ is called weakly stable, *if for all $x_1, x_2 \in X$ holds*

$$\lim_{t\to\infty} \langle x_1, T(t)x_2\rangle_X = 0,$$

strongly stable, *if for all $x \in X$ holds*

$$\lim_{t\to\infty} \|T(t)x\|_X = 0,$$

and 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 $T(\cdot)$.*

Remark 2 If $T(\cdot)$ is a strongly continuous semigroup generated by A, then the adjoint semigroup *$T^*(\cdot)$ is strongly continuous, as well. Its generator is given by A^*. Exponential stability of $T(\cdot)$ implies exponential stability of $T^*(\cdot)$; weak stability of $T(\cdot)$ implies weak stability of $T^*(\cdot)$. However, strong stability of $T^*(\cdot)$ can, in general, not be inferred from strong stability of $T(\cdot)$.*

For various physically motivated control systems (in particular partial differential equations with boundary control), it is reasonable that the input operator maps to the larger space $D(A^*)' \supset X$. Though this naming is a bit misleading, the input operator is called *bounded*, if its range is contained in X, and unbounded, otherwise. Subsequently we introduce a property of the input operator which, roughly speaking, means that (3) has a well-defined and continuous state trajectory for any square integrable input.

A strongly continuous semigroup $T(\cdot)$ on X extends to a strongly continuous semigroup on $D(A^*)'$ [17, Prop. 2.10.4]. Therefore, the operator product $T(t)B$ can be understood as an element in $\mathcal{B}(U, D(A^*)')$, and we can formulate the following definition.

Definition 3 (Admissible control operator) Let U, X be Hilbert spaces, let $T(\cdot)$ be a strongly continuous semigroup 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 for all $u \in L_2(\mathbb{R}_{\geq 0}, U)$*

$$\Phi_t u := \int_0^t T(t-\tau)Bu(\tau)d\tau \in X. \tag{4}$$

The control operator B is called infinite-time admissible*, if for all $u \in L_2(\mathbb{R}_{\geq 0}, U)$*

$$\Phi_\infty u := \int_0^\infty T(\tau)Bu(\tau)d\tau \in X. \tag{5}$$

The pair (A, B) consisting of a closed operator $A : D(A) \subset X \to X$ and $B \in \mathcal{B}(U, D(A^*)')$ is called (infinite-time) admissible on X, U, *if A generates a strongly continuous semigroup $T(\cdot)$ on X, and B is (infinite-time) admissible for $T(\cdot)$.*

The operator Φ_t applied to $u \in L_2(\mathbb{R}_{\geq 0}, U)$ has to be understood as the state at time t of the system (3) with trivial initial condition $x(0) = x_0$. Further, for $u \in L_2(\mathbb{R}_{\geq 0}, U)$, the expression $\Phi_\infty u$ corresponds to the state $x(0)$ of the system $\dot{x}(t) = Ax(t) + Bu(-t)$. The operator Φ_∞ (Φ_t) is therefore called *controllability map (at time t)*. An application of the closed-graph theorem consequences that (infinite-time) admissibility implies that these mappings are bounded, i.e., $\Phi_t \in \mathcal{B}(L_2(\mathbb{R}_{\geq 0}, U), X)$ ($\Phi_\infty \in \mathcal{B}(L_2(\mathbb{R}_{\geq 0}, U), X)$).

Definition 4 Let (A, B) be admissible on X, U. Then the Gramian of (A, B) at time $t > 0$ *is given by*

$$P_t = \Phi_t \Phi_t^* \in \mathcal{B}(X).$$

If (A, B) is moreover infinite-time admissible, then the Gramian of (A, B) *is defined to be*

$$P = \Phi_\infty \Phi_\infty^* \in \mathcal{B}(X).$$

The operators P, P_t are obviously self-adjoint and nonnegative. In the following we present their relation to the operator Lyapunov equation.

Theorem 5 [18, Thm. 3.1], see also [17, Thm. 5.1.1], [19, Thm. 10.4.3]. Let U, X be Hilbert spaces and let $A : D(A) \subset X \to X$ be the generator of the strongly continuous semigroup $T(\cdot)$ on X. Let $B \in \mathcal{B}(U, D(A^*)')$. Then the following four statements are equivalent:

a) (A, B) is infinite-time admissible on X, U.

b) There exists some operator $P \in \mathcal{B}(X)$, such that for all $x \in D(A)$ holds

$$\lim_{t\to\infty} \left\| Px - \int_0^t T(t)BB'T^*(t)x \right\|_X = 0. \tag{6}$$

c) There exists some nonnegative and self-adjoint $Q \in \mathcal{B}(X)$, such that

$$2\,\mathrm{Re}\,\langle Qx, A^*x\rangle_X + \|B'x\|_U^2 \leq 0 \quad \textit{for all } x \in D(A^*). \tag{7}$$

d) There exists some nonnegative and self-adjoint $\Pi \in \mathcal{B}(X)$, such that

$$2\,\mathrm{Re}\,\langle \Pi x, A^*x\rangle_X + \|B'x\|_U^2 = 0 \quad \textit{for all } x \in D(A^*). \tag{8}$$

Moreover, if (A, B) is infinite-time admissible on X, U, then the following statements are true:

(I) P defined in (6) is the Gramian of (A, B).

(II) P is the smallest nonnegative and self-adjoint solution of (7) (and hence also of (8)). In other words, $P \geq 0$, P fulfills (7), and if, for $0 \geq \Pi = \Pi^* \in \mathcal{B}(X)$ holds (7), then $P \leq \Pi$.

(III) For all $x \in X$ holds

$$\lim_{t\to\infty} \langle T^*(t)x, PT^*(t)x\rangle_X = 0. \tag{9}$$

(IV) If $0 \leq \Pi = \Pi^* \in \mathcal{B}(X)$ fulfills (8) and

$$\lim_{t\to\infty} \langle T^*(t)x, \Pi T^*(t)x\rangle_X = 0, \tag{10}$$

then $P = Q$. In particular, if $T^*(\cdot)$ is strongly stable (or even exponentially stable), then $\Pi = P$ is the only nonnegative and self-adjoint solution of (8).

(V) If $\ker P = \{0\}$, then $T(\cdot)$ is weakly stable.

Remark 6 a) If B is admissible for an exponentially stable semigroup $T(\cdot)$, then B is infinite-time admissible [18, p. 6]. Indeed, we will assume exponential stability in most of our results.

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

In the case where A is negative and self-adjoint and $B'A^{-1}B$ exists with

$$B'A^{-1}B \in \mathcal{S}_1(U, X), \tag{11}$$

it can be shown that the Gramian is nuclear. We can furthermore give an explicit expression for the nuclear norm of P;

Theorem 7 [11, Prop. 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 (11) holds true. Then the following holds true:

a) A generates an exponentially stable semigroup $T(\cdot)$ on X.

b) B is admissible for $T(\cdot)$.

c) The Gramian P of (A, B) is nuclear with, in particular,

$$\|P\|_{\mathcal{S}_1(X)} = -\tfrac{1}{2} \cdot \mathrm{tr}(B'A^{-1}B). \tag{12}$$

Remark 8 Note that, under the assumptions of the above theorem and the additional assumption that the input space U is finite-dimensional (i.e., w.l.o.g., $U = \mathbb{C}^m$), the latter expression is moreover the nuclear norm of the Gramian. In this case, nuclearity of the Gramian P is therefore a consequences of only the existence of $B'A^{-1}B \in \mathbb{C}^{m,m}$.

To go from abstraction to concretion, consider the heat equation

$$\frac{\partial x}{\partial t}(\xi, t) = \Delta x(\xi, t), \quad (\xi, t) \in \Omega \times \mathbb{R}_{\geq 0} \tag{13a}$$

evolving on a bounded domain $\Omega \subset \mathbb{R}^d$ with piecewise twice differentiable boundary $\partial\Omega$, compare [12], apply the Robin boundary condition

$$\nu(\xi)^T \nabla x(\xi, t) + \alpha x(\xi, t) = u(t), \tag{13b}$$

for any $(\xi, t) \in \partial\Omega \times \mathbb{R}_{\geq 0}$, where $\nu(\xi)$ denotes the outward normal to $\partial\Omega$ in $\xi \in \partial\Omega$, $\alpha \in \mathbb{R}_{>0}$. The scalar function $u \in L_2(\mathbb{R}_{\geq 0})$ is supposed to be the input of the system.

To rewrite this as a system (3), set $x(t) := x(\cdot, t) \in L_2(\Omega) := X$ and $U = \mathbb{C}$. Using the results from [20], the operators A and B are given by

$$\begin{aligned} D(A) = \{&x \in H^1(\Omega) \,|\, \Delta x \in L^2(\Omega), \\ &\qquad\qquad \nu^T \nabla x + \alpha x = 0 \text{ on } \partial\Omega\}, \\ Ax = \Delta x& \quad \text{for all } x \in D(A), \\ \langle Bu, z\rangle_{D(A^*)', D(A^*)} &= u \cdot \int_{\partial\Omega} z(\xi) d\sigma_\xi, \end{aligned} \tag{14}$$

where by $d\sigma_\xi$ we denote the surface measure on $\partial\Omega$. A is self-adjoint and negative, since, the Gauss's theorem leads to

$$\langle x, Ax\rangle_X = -\|\nabla x\|^2_{L_2(\Omega, \mathbb{R}^d)} \text{ for all } x \in D(A).$$

By the Rellich-Kondrachov Theorem [12, Thm. 6.3], $H^1(\Omega)$ is compactly embedded in $L_2(\Omega)$. Therefore, A has compact resolvent. Since, furthermore, the boundary integral continuously depends on $\|\nabla x\|_{L_2(\Omega, \mathbb{R}^d)}$, we may infer the existence of $c > 0$, such that

$$\|\nabla x\|_{L_2(\Omega, \mathbb{R}^d)} \leq c \cdot |u| \int_{\partial\Omega} z(\xi) d\sigma_\xi \text{ for all } x \in D(A).$$

This gives rise to the existence of $B' A^{-1} B$, whence, by Proposition 7, the system is infinite-time admissible and has nuclear Gramian with

$$\|P\|_{\mathcal{S}_2(X)} = -\frac{B' A^{-1} B}{2}.$$

Using [21, Thm. 2.9], the expression $x_h = A^{-1} B \in H^1(\Omega)$ is the solution of the Laplace equation with Robin boundary condition $\nu(\xi)^T \nabla x_h(\xi) + \alpha x_h(\xi) = -1$. The nuclear norm of the Gramian therefore reads

$$\|P\|_{\mathcal{S}_1(X)} = -\frac{1}{2} \cdot \int_{\partial\Omega} x_h(\xi) d\sigma_\xi. \tag{15}$$

IV. ALTERNATING DIRECTION IMPLICIT (ADI) ITERATION FOR OPERATOR LYAPUNOV EQUATIONS

In Algorithm 1 we formulate the ADI iteration for the determination of Gramian operators. This algorithm exactly reads as in the matrix case [22, p. 43]; it 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 [22, pp. 43]. Their choice in the case of operator Lyapunov equations will be briefly discussed.

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 $p_1, \ldots, p_{i_{\max}} \in \mathbb{C}_-$

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).

```
$V_1 = (A + p_1 I)^{-1} B$
$S_1 = \sqrt{-2\,\mathrm{Re}(p_1)} \cdot V_1$
for $i = 2, 3, \ldots, i_{\max}$ do
    $V_i = V_{i-1} - (p_i + \overline{p_{i-1}}) \cdot (A + p_i I)^{-1} V_{i-1}$
    $S_i = [\, S_{i-1} \,,\, \sqrt{-2\,\mathrm{Re}(p_i)} \cdot V_i \,]$
end for
```

Remark 9 a) In the case of finite-dimensional input space $U = \mathbb{C}^m$, we have $S_i \in \mathcal{B}(\mathbb{C}^{m \cdot i}, X)$. This means that, $P_i = S_i^* S_i$ has finite rank and S_i has a representation by means of an $m \cdot i$-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.

b) The number $i_{\max}$ has to be determined during the algorithm by using 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.

The convergence result is presented below.

*Theorem 10 [11, Thm. 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.,*

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

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

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

Remark 11 Strong stability implies that the spectrum of A is contained in $\mathbb{C}_-$ [17, Cor. 2.3.3]. The iteration in Algorithm 1 is therefore feasible, if the shift parameters all have negative real part.

Remark 12 (Shift parameters) In the finite-dimensional case, the shift parameters are chosen by means of the spectrum of A, such that fast convergence may be achieved (see [22, pp. 43] for an overview). For systems governed by *partial differential equation, we propose a shift parameter choice by applying these existing approaches to a sufficiently accurate spatial discretization of the system.*

Having a system being of the class considered in Proposition 7, we can give an explicit expression for the nuclear norm of the difference between the Gramian and the operators obtained in the ADI iteration.

Proposition 13 Let the assumptions of Proposition 7 be valid. 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

$$\|P-P_i\|_{\mathcal{S}_1(X)} = 2\sum_{k=1}^{i}\mathrm{Re}(p_i)\cdot\mathrm{tr}(V_k^*V_k) - \frac{1}{2}\cdot\mathrm{tr}(B'A^{-1}B). \tag{16}$$

In particular, if $U = \mathbb{C}^m$, then $B'A^{-1}B, S_k^*S_k \in \mathbb{C}^{m,m}$ are Hermitian matrices; for systems with single input, the expressions $B'A^{-1}B$, $S_k^*S_k$ are real numbers.

Now we consider again the system (13) governed by the heat equation, i.e., the operators A and B are of the form (14). Using [21, Thm. 2.9], we can infer that for $f \in X = L_2(\Omega)$ and $p \in \mathbb{C}_-$, the expression $x = (A+pI)^{-1}f$ is the solution of $p\cdot x(\xi) + \Delta x(\xi) = f(\xi)$ with boundary condition $\nu(\xi)^T\nabla x(\xi) + \alpha x(\xi) = 0$. Further, the expression $y = (A+pI)^{-1}B$ is given by the solution of $p\cdot y(\xi) + \Delta y(\xi) = 0$ with boundary condition $\nu(\xi)^T\nabla v_1(\xi) + \alpha v_1(\xi) = -1$. Therefore, ADI iteration for the heat equation consists of the solution of a sequence of Helmholtz equations. By a substitution $q_i = -p_i$ of the shift parameters, we are led to the ADI algorithm in the following form:

Algorithm 2 ADI iteration for heat equation with one-dimensional Robin boundary control (13).

Input: Bounded domain $\Omega \subset \mathbb{R}^d$ 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 (14)).

```
Solve
    $q_1 \cdot v_1(\xi) - \Delta v_1(\xi) = 0, \qquad \xi \in \Omega,$
    $\nu(\xi)^T\nabla v_1(\xi) + \alpha v_1(\xi) = 1, \qquad \xi \in \partial\Omega$
for $v_1 \in L_2(\Omega)$.
Define $S_1 = \sqrt{2q_1} \cdot v_1 \in \mathcal{B}(\mathbb{C}, L_2(\Omega))$
for $i = 2, 3, \dots, i_{\max}$ do
    Solve
        $q_i \cdot \hat{v}(\xi) - \Delta\hat{v}(\xi) = v_{i-1}(\xi), \qquad \xi \in \Omega,$
        $\nu(\xi)^T\nabla\hat{v}(\xi) + \alpha\hat{v}(\xi) = 0, \qquad \xi \in \partial\Omega$
    for $\hat{v} \in L_2(\Omega)$.
    Set $v_i = v_{i-1} - (q_i + q_{i-1}) \cdot \hat{v}$
    $S_i = [\, S_{i-1}, \sqrt{2q_i} \cdot v_i \,] \in \mathcal{B}(\mathbb{R}^i, L_2(\Omega))$
end for
```

Remark 14 a) If the shift parameters are chosen to be real (which is reasonable due to $\sigma(A) \subset \mathbb{R}$ and the findings in Remark 12), then all equations that have to be solved in Algorithm 2 are real as well.

b) The Helmholtz equations occurring in Algorithm 2 can be solved by using (adaptive) finite elements. Note that, if the grid is not changed during iteration, then Algorithm 2 will be arithmetically equivalent to the approach of discretizing the heat equation with respect to space, and an accordant application of the matrix version of the ADI method to the spatially discretized finite-dimensional system. For an error analysis of inexact solutions of the equations in the ADI iteration, we refer to [11, Sec. 5]

Applying the error expression (16) to the system (13), we obtain that, in the notation of Algorithm 2 and $P_i = S_iS_i^*$, there holds

$$\|P-P_i\|_{\mathcal{S}_1(X)} = \frac{1}{2}\cdot\int_{\partial\Omega} x_h(\xi)d\sigma_\xi - 2\sum_{k=1}^{i} q_k\cdot\int_{\Omega}|v_k(\xi)|^2 d\xi, \tag{17}$$

where $x \in H^1(\Omega)$ solves the Laplace equation with Robin boundary condition $\nu(\xi)^T\nabla x_h(\xi) + \alpha x_h(\xi) = 1$.

A. Numerical Results

Consider the heat equation with single Robin boundary control (13) on the L-shaped domain $\Omega = (0,1)^2 \setminus (0.5,1)^2$, and $\alpha = 1$. Using (15) and the fact that the function $x_h \equiv -1$ solves the Laplace equation with $\nu(\xi)^T\nabla x_h(\xi) + x_h(\xi) = -1$, the nuclear norm of the Gramians is given by $\|P\|_{S_1(X)} = 2$. Now we apply Algorithm 2, where, in each iteration, we perform a finite-element discretization. This 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` [23] based upon the C++-library `deal.II`, see [24], [25]. 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., [26]. Thus we can allow for refinement of the discretization if the error is too large, i.e., $\eta >$ TOL and for optional coarsening of the discretization once the error is too small, i.e., $\eta < 0.1\,$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`.

As a test case we consider the behavior of $\|P_i\|_{S_1(X)}$ for a fixed mesh and an adaptive coarsening and refinement adjusted to tolerance 10^{-4}. The results are depicted in Figure I. The shift parameters where chosen by applying the method of WACHSPRESS [27] on the basis of the lowest hundred eigenvalues of the Robin Laplacian on the unit square, which are given by $\pi^2(i^2 + j^2)$ where $i, j = 1, \dots, 10$.

As is shown in Table I including the possibility to coarsen the mesh allows an almost identical approximation of the

TABLE I

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
⋮	⋮	⋮	⋮	⋮	⋮	⋮

Gramian with severely fewer unknowns needed in the calculation.

REFERENCES

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

[2] A. Antoulas, *Approximation of Large-Scale Dynamical Systems*, ser. Advances in design and control. Philadelphia/Pennsylvania: Society for Industrial and Applied Mathematics, 2005.

[3] D. Kleinman, "On an iterative technique for Riccati equation computations," *IEEE Trans. Automat. Control*, vol. AC-13, pp. 114–115, 1968.

[4] R. Bartels and G. Stewart, "Solution of the matrix equation $AX + XB = C$," *Communications of the ACM*, vol. 15, no. 9, pp. 820–826, 1972.

[5] S. Hammarling, "Numerical solution of the stable, nonnegative definite Lyapunov equation," *IMA J. Numer. Anal.*, vol. 2, no. 3, pp. 303–323, 1982.

[6] A. Lu and E. Wachspress, "Solution of Lyapunov equations by alternating direction implicit iteration," *Comput. Math. Appl.*, vol. 21, no. 9, pp. 43–58, 1991.

[7] R. Smith, "Matrix equation $AX + XB = C$," *SIAM J. Appl. Math.*, vol. 16, no. 1, pp. 198–201, 1968.

[8] I. Jaimoukha and E. Kasenally, "Krylov subspace methods for solving large Lyapunov equations," *SIAM J. Numer. Anal.*, vol. 31, no. 1, pp. 227–251, 1994.

[9] V. Simoncini, "A new iterative method for solving large-scale Lyapunov matrix equations," *SIAM J. Sci. Comput.*, vol. 29, no. 3, pp. 1286–1288, 2007.

[10] J. Roberts, "Linear model reduction and solution of the algebraic Riccati equation by use of the sign function," *Internat. J. Control*, vol. 32, no. 4, pp. 677–687, 1980, (Reprint of Technical Report TR-13, CUED/B-Control, Engineering Deparment, Cambrige University, 1971).

[11] T. Reis, M. Opmeer, and W. Wollner, "Finite-rank ADI iteration for operator Lyapunov equations," Universität Hamburg, Tech. Rep. 2012-09, 2012.

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

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

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

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

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

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

[18] S. Hansen and G. Weiss, "New results on the operator carleson measure criterion," *IMA J. Math Control Inform.*, vol. 14, no. 1, pp. 3–32, 1997.

[19] O. Staffans, *Well-posed Linear systems*, ser. Encyclopedia of Mathematics and its Applications. Cambridge: Cambridge University Press, 2005, vol. 103.

[20] D. Salamon, "Infinite-dimensional linear systems with unbounded control and observation: a functional analytic approach," *Trans. Amer. Math. Soc.*, vol. 300, no. 2, pp. 383–431, 1987.

[21] A. Cheng and K. Morris, "Well-posedness of boundary control systems," *SIAM Journal of Control and Optimization*, vol. 42, no. 4, pp. 1244–1265, 2003.

[22] J. Saak, "Efficient numerical solution of large scale algebraic matrix equations in pde control and model order reduction," Doctoral dissertation, Technische Universität Chemnitz, 2009.

[23] C. Goll, T. Wick, and W. Wollner, "The differential equations and optimization environment: DOpElib." [Online]. Available: www.dopelib.math.uni-hamburg.de

[24] W. Bangerth, R. Hartmann, and G. Kanschat, "deal.II — a general-purpose object-oriented finite element library," *ACM Trans. Math. Softw.*, vol. 33, no. 4, p. 24, 2007.

[25] W. Bangerth, T. Heister, and G. Kanschat, `deal.II` *Differential Equations Analysis Library, Technical Reference*. [Online]. Available: http://www.dealii.org

[26] M. Ainsworth and J. T. Oden, *A Posteriori Error Estimation in Finite Element Analysis*, ser. Pure and Applied Mathematics (New York). New York: Wiley-Interscience [John Wiley & Sons], 2000.

[27] E. Wachspress, "Iterative solution of the Lyapunov matrix equation," *Appl. Math. Lett.*, vol. 1, pp. 87–90, 1988.