A Variational Formulation for LTI-Systems and Model Reduction

Moritz Feuerle and Karsten Urban

Abstract We consider a variational formulation of Linear Time-Invariant (LTI)-systems and derive a model reduction in dimension and time inspired by space-time variational reduced basis (RB) methods for parabolic problems. A residual-type RB error estimator is derived whose effectivity is investigated numerically.

1 Introduction

Model order reduction (MOR) of (linear) systems is a huge field of research with an enormous amount of literature. On the other hand, the reduced basis (RB) method has become a widely spread technique for reducing parameterized partial differential equations. We refer e.g. to [2], where both model reduction techniques are reviewed. In this paper, we consider a variational formulation of Linear Time-Invariant (LTI) systems that allows us to introduce an RB-type residual error estimator inspired by space-time RB methods for parabolic problems, [6, 7]. This, in turn, yields a reduction not only of the dimension of the LTI system but also w.r.t. the temporal discretization, i.e., the number of time steps.

The paper is organized as follows: In §2, we introduce a variational formulation of LTI systems and show its well-posedness, §3 is devoted to Petrov-Galerkin discretizations which are used as a detailed solution for the Reduced Basis Method (RBM) in §4. We present some numerical results in §5 and end by conclusions as well as an outlook in §6.

Moritz Feuerle, Karsten Urban
Ulm University, Institute for Numerical Mathematics, Helmholtzstr. 20, 89081 Ulm (Germany), e-mail: {moritz.feuerle,karsten.urban}@uni-ulm.de

2 Variational formulation for LTI-systems

We consider LTI systems on some time interval $I := (0,T)$, $T > 0$. Given integers $m, n, p \in \mathbb{N}$, matrices $A \in \mathbb{R}^{n\times n}$ (which is assumed to be s.p.d. for simplicity), $B \in \mathbb{R}^{n\times p}$, $C \in \mathbb{R}^{m\times n}$, $D \in \mathbb{R}^{m\times p}$, a control $u : I \to \mathbb{R}^p$ and an initial state $x_0 \in \mathbb{R}^n$, determine the state $x : I \to \mathbb{R}^n$ and output $y : I \to \mathbb{R}^m$ s.t.

$$\dot{x}(t) + Ax(t) = Bu(t), \quad y(t) = Cx(t) + Du(t), \quad t \in I, \qquad x(0) = x_0. \tag{1}$$

W.l.o.g. we restrict ourselves to the homogeneous case, i.e., $x_0 = 0$, but note, that the inhomogeneous case can easily be incorporated.

A Variational formulation. We multiply the first equation in (1) with a test function $z : I \to \mathbb{R}^n$ and integrate over I, i.e.,

$$\int_0^T (\dot{x}(t), z(t))\, dt + \int_0^T (Ax(t), z(t))\, dt = \int_0^T (Bu(t), z(t))\, dt, \tag{2}$$

where $(\cdot,\cdot)$ denotes the Euclidean scalar product with induced norm $\|\cdot\|$ in $\mathbb{R}^d$, $d \in \{m, n, p\}$. Obviously, (2) makes sense for $z \in Z := L_2(I, \mathbb{R}^n) \equiv L_2(I)^n$, $\|z\|_Z := \|z\|_{L_2(I)^n}$. The desired state function $x : I \to \mathbb{R}^n$ is then sought in the Sobolev-Bochner Hilbert space $X := H^1_{(0)}(I)^n := \{x \in H^1(I)^n : x(0) = 0\}$. As in [6, 7] we consider a slightly stronger norm than the usual graph norm, namely

$$|||x|||^2_{X,\mathrm{Std}} := \|\dot{x}\|^2_{L_2(I)^n} + \|x\|^2_{L_2(I)^n} + \|x(T)\|^2, \tag{3}$$

with the corresponding inner product $(x, v)_{X,\mathrm{Std}} := (\dot{x}, \dot{v})_{L_2(I)^n} + (x, v)_{L_2(I)^n} + (x(T), v(T))$ for $x, v \in X$, which is well-defined recalling that $X \hookrightarrow C([0,T], \mathbb{R}^n)$. Then, setting $U := L_2(I)^p$ as parameter space, we obtain the following variational formulation of (1):

$$\text{for } u \in U \text{ find } x = x(u) \in X: \quad b(x, z) = f(z; u) := (Bu, z)_{L_2(I)^n} \quad \forall z \in Z, \tag{4}$$

where the parameter-independent bilinear form reads $b(x, z) := (\dot{x} + Ax, z)_{L_2(I)^n}$. We stress the fact that $f(\cdot; u)$ is *linear* in u (for $x_0 \neq 0$ affine-linear).

Well-posedness. In order to prove well-posedness of (4), we need to satisfy Nečas' conditions, namely boundedness, injectivity and inf-sup condition of $b(\cdot,\cdot)$. Since the verification is very similar to space-time variational formulation of parabolic initial value problems, we refer to it, [5, 6, 7]. In particular, the inf-sup constant can be detailed in similar way, see [6, Prop. 1] and [5, Thm. 5.1].

Proposition 2.1 *Let $A \in \mathbb{R}^{n\times n}$ be s.p.d. with constants $\alpha_A > 0$ and $\gamma_A < \infty$, such that $\alpha_A\|\phi\| \le \|A\phi\| \le \gamma_A\|\phi\|$ for all $\phi \in \mathbb{R}^n$. Then,*

$$\inf_{x\in X} \sup_{z\in Z} \frac{b(x, z)}{|||x|||_{X,\mathrm{Std}}\, \|z\|_Z} \ge \beta^{\mathrm{Std}} := \frac{\min\{1, \alpha_A \min\{1, \gamma_A^{-2}\}\}}{\sqrt{2}\max\{1, (\alpha_A)^{-1}\}} > 0. \tag{5}$$

In order to (quantitatively) improve the inf-sup-bound in (5), we consider an energy norm, namely $(\phi,\psi)_{\mathcal{A}} := (\phi,\mathcal{A}\psi)$, $\|\phi\|^2_{\mathcal{A}} := (\phi,\phi)_{\mathcal{A}}$ for an s.p.d. matrix $\mathcal{A} \in \mathbb{R}^{n\times n}$, $\phi,\psi \in \mathbb{R}^n$ and (with a slight double use of notation) $(z,w)_{\mathcal{A}} := \int_0^T (z(t),w(t))_{\mathcal{A}}\, dt$ as well as $\|z\|^2_{\mathcal{A}} := (z,z)_{\mathcal{A}}$ for $w,z \in L_2(I)^n$. Then, we set

$$|||x|||^2_X := \|\dot{x}\|^2_{A^{-1}} + \|x\|^2_A + \|x(T)\|^2, \qquad |||z|||_Z := \|z\|_A,$$

and following the reasoning in [7], we can easily show that

$$\inf_{x\in X}\sup_{z\in Z}\frac{b(x,z)}{|||z|||_Z\,|||x|||_X} = \sup_{x\in X}\sup_{z\in Z}\frac{b(x,z)}{|||z|||_Z\,|||x|||_X} = 1 \equiv \beta^{\mathrm{En}}. \tag{6}$$

3 Petrov-Galerkin (detailed) discretizations

In order to compute an approximation to the solution of (4), we use a standard Petrov-Galerkin approach. To this end, one constructs finite-dimensional trial and test spaces $X^{\mathcal{N}} \subset X$,$Z^{\mathcal{N}} \subset Z$ with $\dim(X^{\mathcal{N}}) = \dim(Z^{\mathcal{N}}) = \mathcal{N}$. For stability, these spaces need to satisfy a discrete inf-sup (LBB) condition, i.e.,

$$\beta^{\mathcal{N}} := \inf_{x^{\mathcal{N}}\in X^{\mathcal{N}}}\sup_{z^{\mathcal{N}}\in Z^{\mathcal{N}}}\frac{b(x^{\mathcal{N}},z^{\mathcal{N}})}{|||x^{\mathcal{N}}|||_X\,|||z^{\mathcal{N}}|||_Z} \geq \beta^{\mathrm{En}}_{\mathrm{LB}} > 0, \tag{7}$$

where the lower-bound $\beta^{\mathrm{En}}_{\mathrm{LB}}$ for the inf-sup-constant is independent of $\mathcal{N}$ as $\mathcal{N} \to \infty$. Then, the discrete version of (4) is a Petrov-Galerkin scheme of the form

$$\text{for } u \in U \text{ find } x^{\mathcal{N}}(u) \in X^{\mathcal{N}}:\ b(x^{\mathcal{N}}(u),z^{\mathcal{N}}) = (Bu,z^{\mathcal{N}})_{L_2(I)^n} \quad \forall z^{\mathcal{N}} \in Z^{\mathcal{N}}, \tag{8}$$

where $u \in U$ is possibly suitably discretized (see below). As usual, we define the *primal residual* $r^{\mathrm{pr}}(\cdot;u) \in Z'$ as

$$r^{\mathrm{pr}}(z;u) := f(z;u)_{L_2(I)^n} - b(x^{\mathcal{N}}(u),z) = b(x(u)-x^{\mathcal{N}}(u),z), \quad z \in Z, \tag{9}$$

and its norm by $R^{\mathrm{pr}}(u) := |||r^{\mathrm{pr}}(\cdot;u)|||_{Z'}$. Since $Z = L_2(I)^n$ is a Hilbert (pivot) space, we can identify $Z = Z'$, which significantly reduces the complexity in computing this dual norm (we do not need to determine Riesz representations). Then, the following error-residual relation is straightforward and well-known (recall $\beta^{\mathrm{En}} \equiv 1$)

$$|||x(u)-x^{\mathcal{N}}(u)|||_X \leq R^{\mathrm{pr}}(u) = \|Bu - \dot{x}^{\mathcal{N}}(u) - Ax^{\mathcal{N}}(u)\|_{A^{-1}} =: \Delta^{\mathrm{pr}}(u). \tag{10}$$

A Time-Marching Discretization. We start by introducing a Petrov-Galerkin discretization arising from a (finite element) discretization in time, which leads to a Crank-Nicolson (CN) time-marching scheme. To this end, we choose some integer $K > 1$ and set $\Delta t := T/K$ resulting in a temporal triangulation $\mathcal{T}^{\mathrm{time}}_{\Delta t} \equiv \{t^{k-1} \equiv (k-1)\Delta t < t \leq k\,\Delta t \equiv t^k, 1 \leq k \leq K\}$ in time. Denote by

$S_{\Delta t} = \text{span}\{\sigma^1,\dots,\sigma^K\}$ piecewise linear finite elements on I, where σ^k is the (interpolatory) hat-function with the nodes t^{k-1}, t^k and t^{k+1} (resp. truncated for $k \in \{0,K\}$) and $Q_{\Delta t} = \text{span}\{\tau^1,\dots,\tau^K\}$ piecewise constant finite elements, where $\tau^k := \chi_{I^k}$, the characteristic function on the temporal element $I^k := (t^{k-1},t^k)$. Then, we set $X^{\mathcal{N}}_{\text{CN}} := S_{\Delta t}\otimes\mathbb{R}^n$, $Z^{\mathcal{N}}_{\text{CN}} := Q_{\Delta t}\otimes\mathbb{R}^n$, i.e., the detailed dimension is $\mathcal{N} := Kn$. Within that framework, the detailed approximation amounts computing $x^{\mathcal{N}}_{\text{CN}} \in X^{\mathcal{N}}_{\text{CN}}$ represented as[1] $x^{\mathcal{N}}_{\text{CN}}(t;u) \equiv x^{\mathcal{N}}_{\text{CN}}(t;u)(t) = \sum_{k=1}^K \boldsymbol{x}^k_{\text{CN}}\sigma^k(t)$, $x(t^k) \approx \boldsymbol{x}^k_{\text{CN}} \in \mathbb{R}^n$, $k = 1,\dots,K$, $t \in I$, and $\boldsymbol{x}^{\mathcal{N}}_{\text{CN}} := (\boldsymbol{x}^k_{\text{CN}})_{k=1,\dots,K} \in \mathbb{R}^{K\times n} \cong \mathbb{R}^{Kn} = \mathbb{R}^{\mathcal{N}}$. Setting $\Pi_{\Delta t} := ([\Pi_{\Delta t}]_{k,\ell})_{k,\ell=0,\dots,K}$, $\check{\Pi}_{\Delta t} := ([\Pi_{\Delta t}]_{k,\ell})_{k=0,\dots,K-1,\ell=1,\dots,K}$, $\hat{\Pi}_{\Delta t} := ([\Pi_{\Delta t}]_{k,\ell})_{k=1,\dots,K,\ell=0,\dots,K}$ and $\bar{\Pi}_{\Delta t} := ([\Pi_{\Delta t}]_{k,\ell})_{k,\ell=1,\dots,K}$ for $\Pi \in \{K,L,M,N,O\}$,

$$[K_{\Delta t}]_{k,\ell} := (\dot\sigma^k,\dot\sigma^\ell)_{L_2(I)},\quad [L_{\Delta t}]_{k,\ell} := (\sigma^k,\sigma^\ell)_{L_2(I)},\quad [M_{\Delta t}]_{k,\ell} := (\sigma^k,\tau^\ell)_{L_2(I)}$$
$$[N_{\Delta t}]_{k,\ell} := (\dot\sigma^k,\tau^\ell)_{L_2(I)}\qquad [O_{\Delta t}]_{k,\ell} := (\dot\sigma^k,\sigma^\ell)_{L_2(I)}, \tag{11}$$

and recalling $[M_{\Delta t}]_{k,\ell} = \frac{\Delta t}{2}(\delta_{k,\ell}+\delta_{k+1,\ell})$ and $[N_{\Delta t}]_{k,\ell} = \delta_{k,\ell}-\delta_{k+1,\ell}$, we obtain $b(x^{\mathcal{N}}_{\text{CN}},\tau^\ell e_\mu)_{L_2(I)^n} = \big[[Id+\frac{\Delta t}{2}A]\boldsymbol{x}^\ell_{\text{CN}} - [Id-\frac{\Delta t}{2}A]\boldsymbol{x}^{\ell-1}_{\text{CN}}\big]_\mu$.

Discretization of the Control. Without any discretization, we can in general not evaluate the term $(Bu,z^{\mathcal{N}})_{L_2(I)^n}$ exactly. As a first attempt, it seems reasonable (as done in the literature of LTIs) to use the same temporal discretization, i.e., $U^{\mathcal{N}} := S_{\Delta t}\otimes\mathbb{R}^p$ and interpolate the control onto the temporal nodes $\mathcal{T}^{\text{time}}_{\Delta t}$, i.e., $u^{\mathcal{N}}(t) := \sum_{k=0}^K \boldsymbol{u}^k\,\sigma^k(t)$, $\boldsymbol{u}_{\Delta t} = (\boldsymbol{u}^k)_{k=0,\dots,K} \in \mathbb{R}^{(K+1)\times p}$, where we note that the initial value $u(0)$ does not need to vanish, which is the reason, why the above sum starts from $k=0$.

Crank-Nicolson Scheme. We finally obtain the following iteration: $\boldsymbol{x}^0 := x_0$ and

$$[Id-\tfrac{\Delta t}{2}A]\boldsymbol{x}^\ell_{\text{CN}} = [Id+\tfrac{\Delta t}{2}A]\boldsymbol{x}^{\ell-1}_{\text{CN}} + \tfrac{\Delta t}{2}B(\boldsymbol{u}^\ell+\boldsymbol{u}^{\ell-1}),\quad \ell = 1,2,\dots,K. \tag{12}$$

In particular, the reduction to homogeneous initial conditions has no effect to the temporal iteration. These considerations also show the well-posedness of the discrete problem (8). Note, that (12) yields an iteration so that one does not need to solve the potentially large linear system as for the second discretization in (15) below. Of course, (12) can also be written as a linear system $(\mathcal{B}^{\mathcal{N}}_{\text{CN}})^T\boldsymbol{x}^{\mathcal{N}}_{\text{CN}}(u^{\mathcal{N}}) = \boldsymbol{f}^{\mathcal{N}}_{\text{CN}}(u^{\mathcal{N}})$, where $[\mathcal{B}^{\mathcal{N}}_{\text{CN}}]_{(k,\nu),(\ell,\mu)} = [\check{N}_{\Delta t}]_{k,\ell}\,[Id]_{\nu,\mu} + [\check{M}_{\Delta t}]_{k,\ell}\,[A]_{\nu,\mu}$, which means that $\mathcal{B}^{\mathcal{N}}_{\text{CN}} = \check{N}_{\Delta t}\otimes Id + \check{M}_{\Delta t}\otimes A$, which is non-symmetric.

Standard Error Estimate. An error estimate is derived by using well-known techniques from studying iterations. Denoting by $x(t;u^{\mathcal{N}})$ the solution of (1), we have

$$\|x(t^\ell;u^{\mathcal{N}}) - x^{\mathcal{N}}_{\text{CN}}(t^\ell;u^{\mathcal{N}})\| \le 2\Delta t\sum_{k=0}^{\ell-1}\frac{\gamma^k_{\text{E}}}{\alpha^{k+1}_{\text{I}}}\|r^{\text{pr}}(t^{\ell-k};x^{\mathcal{N}}_{\text{CN}},u^{\mathcal{N}})\| =: \Delta^{\text{Std}}(u^{\mathcal{N}}), \tag{13}$$

[1] We often omit the dependency on the control for simplicity.

where $\alpha_I := 1 + \frac{\Delta t}{2}\alpha_A$, $\gamma_E := 1 + \frac{\Delta t}{2}\gamma_A$ with α_A, γ_A given in Proposition 2.1 and the residual $r^{\text{pr}}(t; x^{\mathcal{N}}, u^{\mathcal{N}}) := Bu^{\mathcal{N}}(t) - \dot{x}^{\mathcal{N}}(t) - Ax^{\mathcal{N}}(t)$.

Supremizers and a Linear System. Alternatively, given some choice for $X^{\mathcal{N}}$, we choose the test space in such a way that the inf-sup-constant $\beta^{\mathcal{N}}$ in (7) is maximized. This is typically done by using so called *supremizers*, [4], which reads here

$$z_{x^{\mathcal{N}}} = A^{-1}\dot{x}^{\mathcal{N}} + x^{\mathcal{N}}. \tag{14}$$

Let $\Xi^{\mathcal{N}} := \{\xi_1^{\mathcal{N}}, \dots, \xi_{\mathcal{N}}^{\mathcal{N}}\}$, $X^{\mathcal{N}} = \text{span}(\Xi^{\mathcal{N}})$, then we set $\Theta^{\mathcal{N}} := \{\theta_1^{\mathcal{N}}, \dots, \theta_{\mathcal{N}}^{\mathcal{N}}\}$, $\theta_i^{\mathcal{N}} := z_{\xi_i}$ and $Z_{\text{sup}}^{\mathcal{N}} := \text{span}(\Theta^{\mathcal{N}})$. We obtain a linear system for (8)

$$\boldsymbol{\mathcal{B}}_{\text{sup}}^{\mathcal{N}} \boldsymbol{x}_{\text{sup}}^{\mathcal{N}}(u) = \boldsymbol{f}_{\text{sup}}^{\mathcal{N}}(u), \tag{15}$$

where the (symmetric) stiffness matrix has the entries $[\boldsymbol{\mathcal{B}}_{\text{sup}}^{\mathcal{N}}]_{i,j} = b(\xi_i^{\mathcal{N}}, \theta_j^{\mathcal{N}}) = (\dot{\xi}_i^{\mathcal{N}} + A\xi_i^{\mathcal{N}}, A^{-1}\dot{\xi}_j^{\mathcal{N}} + \xi_j^{\mathcal{N}})_{L_2(I)^n} = (A\theta_i^{\mathcal{N}}, \theta_j^{\mathcal{N}})_{L_2(I)^n}$, $i, j = 1, \dots, \mathcal{N}$, and the right-hand side reads $(\boldsymbol{f}_{\text{sup}}^{\mathcal{N}}(u))_i := (Bu, A^{-1}\dot{\xi}_i^{\mathcal{N}} + \xi_i^{\mathcal{N}})_{L_2(I)^n}$, $i, j = 1, \dots, \mathcal{N}$. For the specific choice of the CN-trial functions $\xi_i^{\mathcal{N}} = \sigma^k \otimes e_\nu$, $i = (k, \nu)$, $k = 1, \dots, K$, $\nu = 1, \dots, n$, $\mathcal{N} = Kn$, we obtain $\boldsymbol{\mathcal{B}}_{\text{sup}}^{\mathcal{N}} = (\bar{K}_{\Delta t} \otimes A^{-1}) + (\bar{L}_{\Delta t} \otimes A) + ((\bar{O}_{\Delta t} + \bar{O}_{\Delta t}^T) \otimes Id)$.

RB-type Residual Error Estimate. This Pertov-Galerkin formualtion allows us to use a result in [7, Prop. 2.9] to derive an ‘RB-type residual’ error estimator to be described now. For the trial space $X^{\mathcal{N}}$ we will consider as in [7] a discrete norm $|||\cdot|||_{X,\Delta t}$. To define it, we set $\bar{\boldsymbol{x}}_k^{\mathcal{N}} := \frac{1}{\Delta t} \int_{I^k} x^{\mathcal{N}}(s)\, ds$ and $\bar{x}^{\mathcal{N}}(t) := \sum_{k=1}^{K} \bar{\boldsymbol{x}}_k^{\mathcal{N}} \tau^k(t)$, $t \in I$. Then, we set $|||x^{\mathcal{N}}|||_{X,\Delta t}^2 := \|\dot{x}^{\mathcal{N}}\|_{A^{-1}}^2 + \|\bar{x}^{\mathcal{N}}\|_A^2 + \|x^{\mathcal{N}}(T)\|^2$. With these settings, it was proven in [7, Prop. 2.9] that

$$\beta_{\text{sup}}^{\mathcal{N}} := \inf_{x^{\mathcal{N}} \in X^{\mathcal{N}}} \sup_{z^{\mathcal{N}} \in Z_{\text{sup}}^{\mathcal{N}}} \frac{b(x^{\mathcal{N}}, z^{\mathcal{N}})}{|||z^{\mathcal{N}}|||_Z \, |||x^{\mathcal{N}}|||_{X,\Delta t}} = 1.$$

Let us stress that $\beta_{\text{sup}}^{\mathcal{N}}$ is *independent of the control (parameter) u* and of T, Δt. Thus, for any approximation $x_N(u) \in X^{\mathcal{N}}$ (e.g., the RB approximation below), we get

$$|||x_{\text{sup}}^{\mathcal{N}}(u) - x_N(u)|||_{X,\Delta t} \le \Delta_N^{\text{pr}}(u) := \|Bu - \dot{x}_N(u) - Ax_N(u)\|_{A^{-1}}. \tag{16}$$

We may use a discretized control $u^{\mathcal{N}}$ or any u allowing to compute $\boldsymbol{f}_{\text{sup}}^{\mathcal{N}}(u)$, e.g. $\boldsymbol{f}_{\text{sup}}^{\mathcal{N}}(u^{\mathcal{N}}) = [(\hat{O}_{\Delta t} \otimes (A^{-1}B)) + (\hat{L}_{\Delta t} \otimes B)]\boldsymbol{u}_{\Delta t}$ for $\xi_i^{\mathcal{N}} = \sigma^k \otimes e_\nu$ as above.

4 Reduced Basis Method (RBM)

Now, we employ the RBM to the above introduced variational formulation of an LTI. As mentioned already earlier, we view the control u as a parameter, i.e., (1) is seen as a parametric linear system. Doing so, we can reduce both the dimension n of the LTI system and the number K of time steps by reducing $\mathcal{N} := Kn$ to some $N \ll \mathcal{N}$.

RBM for Petrov-Galerkin Problems. The starting point is the detailed discretization (8) of (4). Within a multi-query context, one would need to solve (8) for many different controls $u \in U$ and in a realtime scenario, a good approximation to $x^{\mathcal{N}}(u)$ would be needed extremely fast. This is precisely the situation one is facing within parameterized partial differential equations, where the RBM has proven to be a very useful tool for model reduction (at least in the elliptic and parabolic case).

We thus interpret (4) as a semi-discretized parabolic problem and follow [6, 7] to construct a RBM for the arising non-symmetric space-time-like problem. In order to do so, one looks for subspaces $X_N \subset X^{\mathcal{N}}$ and $Z_N \subset Z^{\mathcal{N}}$ of dimension $\dim(X_N) = \dim(Z_N) = N \ll \mathcal{N} = Kn$ and some $B_N \in \mathbb{R}^{N \times p}$ such that

$$\text{find } x_N \equiv x_N(u) \in X_N : \; b(x_N, z_N) = f_N(z_N; u) := (B_N u, z_N)_{L_2(I)^n} \quad \forall\, z_N \in Z_N \tag{17}$$

and in such a way that x_N can be computed *online efficient*, i.e., with a complexity independent of $\mathcal{N}$. Let us assume that we have (possibly orthonormal) bases $\{\xi^{(i)} : i = 1, \ldots, N\}$ and $\{z^{(j)} : j = 1, \ldots, N\}$ for X_N and Z_N, respectively, at hand. Then, (17) amounts solving a linear system $\boldsymbol{\mathcal{B}}_N^T \boldsymbol{x}_N(u) = \boldsymbol{f}_N(u)$ of dimension N, where $\boldsymbol{f}_N(u)$ (and hence the coefficient vector $\boldsymbol{x}_N(u)$) depend on the control u and we obtain a parameter-dependent solution $\boldsymbol{x}_N(u)$. Moreover, $[\boldsymbol{\mathcal{B}}_N]_{i,j} = b(\xi^{(i)}, z^{(j)})$ and $[\boldsymbol{f}_N(u)]_j = (B_N u, z^{(j)})_{L_2(I)^n}$. Of course, the reduced system depends on the choice of the detailed Petrov-Galerkin detailed discretization. Let $P_N : X^{\mathcal{N}} \to X_N$ and $Q_N : Z^{\mathcal{N}} \to Z_N$ denote projections onto the reduced spaces and let $\boldsymbol{P}_N$, $\boldsymbol{Q}_N : \mathbb{R}^{\mathcal{N}} \to \mathbb{R}^N$ denote the matrix representations w.r.t. the above bases, we get $\boldsymbol{\mathcal{B}}_{N,\text{disc}}^T = \boldsymbol{Q}_N (\boldsymbol{\mathcal{B}}_{\text{disc}}^{\mathcal{N}})^T \boldsymbol{P}_N^T$ and $\boldsymbol{f}_{N,\text{disc}}(u) = \boldsymbol{Q}_N \boldsymbol{f}_{\text{disc}}^{\mathcal{N}}(u)$ for disc $\in \{\text{sup}, \text{CN}\}$. Given some RB basis functions $\xi^{(1)}, \ldots, \xi^{(N)}$ in $X_{\text{CN}}^{\mathcal{N}}$ determined as $\xi^{(i)} := x_{\text{CN}}^{\mathcal{N}}(u^{(i)})$ by (12) (the selection of the 'snapshots' $u^{(i)}$ will be detailed below) and the supremizers $z^{(1)}, \ldots, z^{(N)}$ by (14), the system matrix of the reduced problem reads $\boldsymbol{\mathcal{B}}_{N,\text{sup}} = (\boldsymbol{\Xi}^{\mathcal{N}})^T \boldsymbol{\mathcal{B}}_{\text{sup}}^{\mathcal{N}} \boldsymbol{\Xi}^{\mathcal{N}}$ (recall (11)), where $\boldsymbol{\Xi}^{\mathcal{N}} := (\boldsymbol{\xi}_{\Delta t}^{(i)})_{i=1,\ldots,N}$. Note, that $\boldsymbol{\mathcal{B}}_{N,\text{sup}}$ is symmetric and independent of the parameter, i.e., the control. We can thus precompute and store a LU- or QR-decomposition, which reduces the online amount of work to solve the linear system to $\mathcal{O}(N^2)$. The right-hand side is parameter-dependent and reads $\boldsymbol{f}_{N,\text{sup}}(u^{\mathcal{N}}) = (\boldsymbol{\Xi}^{\mathcal{N}})^T \boldsymbol{f}_{\text{sup}}^{\mathcal{N}}(u^{\mathcal{N}})$ for some $u^{\mathcal{N}} \in U^{\mathcal{N}}$.

Reduced Basis Generation. We use a greedy procedure to compute a Reduced Basis, indicated in Algorithm 1 and which is based upon some error estimator Δ_N^{pr}. After execution of this scheme, we obtain a reduced space $X_N \equiv X_N^{\mathcal{N}} := \text{span}\{x_{\text{CN}}^{\mathcal{N}}(u^{(1)}), \ldots, x_{\text{CN}}^{\mathcal{N}}(u^{(N)})\}$ as well as a reduced test space $Z_N \equiv Z_{N,\text{sup}}^{\mathcal{N}} := \text{span}\{z^{\mathcal{N}}(u^{(1)}), \ldots, z^{\mathcal{N}}(u^{(N)})\}$ and also a reduced control space U_N. The general procedure is indicated by Algorithm 1, which is based upon the choice of a training parameter space $U_{\text{train}} \subset U^{\mathcal{N}}$. Note, that the state snapshots are computed by using the CN-time marching scheme (12) and the reduced system is then generated by the supremizers in (14), see line 2 in Algorithm 1.

Algorithm 1 (Primal) Greedy algorithm with CN-snapshots and RB-supremizers

Choose $U_{\text{train}} \subset U^{\mathcal{N}}$, $\texttt{tol}$, $\eta^{(1)} := u^{(1)}$; set $N := 1$
Compute $\xi^{(N)} := x^{\mathcal{N}}_{\text{CN}}(\eta^{(N)})$, $z^{(N)} := z^{\mathcal{N}}(\eta^{(N)})$ ▷ detailed solution (12) & supremizer (14)
set $X_N := \operatorname{span}\{\xi^{(1)}, \dots, \xi^{(N)}\}$, $Z_N := \operatorname{span}\{z^{(1)}, \dots, z^{(N)}\}$, orthonormalize bases
set $U_N := \operatorname{span}\{\eta^{(1)}, \dots, \eta^{(N)}\}$, orthonormalize
for $u \in U_{\text{train}}$ **do**
    Compute $x_N(u) \in X_N$ ▷ RB approximation with N d.o.f.
    Compute $\Delta^{\text{pr}}_N(u)$ ▷ primal error estimator, e.g., (16)
end for
Set $\eta^{(N+1)} := \arg\max_{u \in U_{\text{train}}} \Delta^{\text{pr}}_N(u)$ ▷ worst parameter
if $\Delta^{\text{pr}}_N(\eta^{(N+1)}) > \texttt{tol}$ set $N := N + 1$, goto 2 **else** break **end if**

Computation of the RB Error Bound. We can further detail the residual-based error estimate from (16) applied to the reduced problem, i.e.,

$$|||x^{\mathcal{N}}_{\text{sup}}(u) - x_{N,\text{sup}}(u)|||_{X,\Delta t} \leq \Delta^{\text{pr}}_N(u) := \|Bu - \dot{x}_{N,\text{sup}}(u) - Ax_{N,\text{sup}}(u)\|_{A^{-1}}. \quad (18)$$

First, we have $\Delta^{\text{pr}}(u)^2 = \|Bu\|^2_{A^{-1}} - 2\boldsymbol{f}_{N,\text{sup}}(u)^T \boldsymbol{x}_N(u) + \boldsymbol{x}_N(u)^T \boldsymbol{\mathcal{B}}_{N,\text{sup}} \boldsymbol{x}_N(u)$ for $\boldsymbol{x}_N \equiv \boldsymbol{x}_{N,\text{sup}}$. Obviously, the last two terms can easily and efficiently be evaluated. Hence, we consider the first part, namely $\|Bu\|_{A^{-1}} = \|A^{-1/2}Bu\|_{L_2(I)^n}$. At this point, it is now crucial how a reduced discretization of the control u is or can be chosen:

- If the control comes from temporal measurements, it will most likely be in form of a *detailed* control, i.e., $u^{\mathcal{N}}$. Then, $\|Bu\|^2_{A^{-1}} = \boldsymbol{u}^T_{\Delta t}(L_{\Delta t} \otimes B^T A^{-1} B)\boldsymbol{u}_{\Delta t}$, which is *not fully* online efficient since the computational amount depends on K.
- If the control can be reduced a priori, e.g., in a multi-query context (think of optimal control), then one would have some u_N with N degrees of freedom so that $\|Bu_N\|_{A^{-1}}$ can be computed in $\mathcal{O}(N^2)$ operations independent of $\mathcal{N} = nK$.

5 Numerical Experiments

We report on some results of our numerical experiments for a standard example, where A arises from a Finite Element discretization of a 1d heat equation with Neumann boundary conditions on the left end and homogeneous Dirichlet boundary conditions on the right end as well as homogeneous initial conditions. The control matrix is $B := n\kappa(-1, 0, \dots, 0)^T \in \mathbb{R}^{n \times 1}$, $m = 1$ and $\kappa > 0$ is the conductivity. On the left-hand side of Figure 1, we see the Greedy error sequence, i.e., the decay of Δ^{pr}_N over a training set of controls as $N \to \infty$. We observe a rate of about $10^{-0.1N}$. On the right-hand side, we increase the number K of time steps and observe that we can basically reach any desired accuracy. Moreover, we compare the exact error with the error estimator Δ^{pr}_N and obtain decreasing effectivities for increasing K. We stress that we measure the error in a quite strong norm $|||\cdot|||_{X,\Delta t}$, which is much stronger than what is usually used in model order reduction, namely $\|\cdot\|_{L_2(I)^n}$.

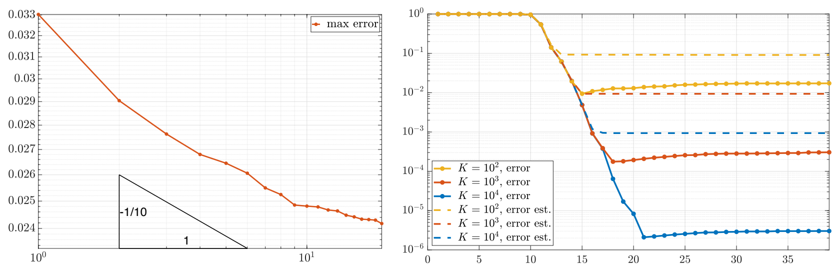

Fig. 1: Greedy error sequence (left), test error and error estimator for increasing K (right): relative error vs. N.

6 Summary and Outlook

We have introduced a space-time-type RB model reduction for LTI systems which allows to reduce both the state dimension n and the number of time steps K. We obtain exponential decay w.r.t. the reduced dimension N and reasonable effectivities, in a quite strong norm, however. The next step is to extend this framework to the output using adjoint techniques. At that stage, quantitative comparisons with well-established techniques like balanced truncation, will be performed. This should result in a clear picture together with other comparisons of model order reduction and POD-Greedy [1] as well as POD-Greedy versus space-time RBM, see [3].

References

1. U. Baur, P. Benner, B. Haasdonk, C. Himpe, I. Martini, and M. Ohlberger. Comparison of methods for parametric model order reduction of time-dependent problems. In *Model reduction and approximation*, volume 15 of *Comput. Sci. Eng.*, pages 377–407. SIAM, 2017.
2. P. Benner, M. Ohlberger, A. Patera, G. Rozza, and K. Urban. *Model Reduction of Parametrized Systems*. Springer International Publishing, Cham, 2017.
3. S. Glas, A. Mayerhofer, and K. Urban. Two ways to treat time in reduced basis methods. In *Model reduction of parametrized systems*, volume 17 of *MS&A. Model. Simul. Appl.*, pages 1–16. Springer, Cham, 2017.
4. G. Rozza. Reduced basis methods for stokes equations in domains with non-affine parameter dependence. *Computing and Visualization in Science*, 12(1):23–35, Jan 2009.
5. C. Schwab and R. Stevenson. Space-time adaptive wavelet methods for parabolic evolution problems. *Mathematics of Computation*, 78(267):1293–1318, 2009.
6. K. Urban and A. T. Patera. A new error bound for reduced basis approximation of parabolic partial differential equations. *C. R. Math. Acad. Sci. Paris*, 350(3-4):203–207, 2012.
7. K. Urban and A. T. Patera. An improved error bound for reduced basis approximation of linear parabolic problems. *Mathematics of Computation*, 83(288):1599–1615, 2014.