52nd IEEE Conference on Decision and Control
December 10-13, 2013. Florence, Italy

Minimax projection method for linear evolution equations

Sergiy Zhuk

***Abstract*— In this paper we present a minimax projection method for linear evolution equations in Hilbert space. The method extends classical Galerkin approach: it builds a differential-algebraic equation with uncertain parameters that models dynamics of exact projection coefficients representing the projection of the evolution equation's solution onto a finite-dimensional subspace. The a priori ellipsoidal bounding set for uncertain parameters is also constructed. The output of the method is an ellipsoid enclosing exact projection coefficients. The ellipsoid can be constructed numerically: we illustrate this applying the method to 1D heat equation.**

I. Introduction

Semigroups of linear operators in Hilbert space H and associated linear evolution equations are commonly used for analysis of Partial Differential Equations (PDEs). From the theoretical perspective, semigroups allow one to treat a linear PDE as a linear dynamical system with infinite dimensional state space H:

$$\frac{dV}{dt} + AV = f, \quad V(0) = v\,. \tag{1}$$

This, in turn, reveals a deep relation between PDEs and ODEs: given a finite dimensional subspace H_N of H one can construct an ODE representing the restriction of (1) onto H_N. This relation is used in applications to get an approximation of V by means of linear systems with finite dimensional state space H_N. In fact, given a linear PDE in the form (1) and a finite dimensional subspace $H_N := \mathrm{Lin}\{\varphi_1 \dots \varphi_N\}$ one can construct an ODE with state space H_N that approximates dynamics of projection coefficients $\mathbf{a}_N^{\mathrm{true}} = (a_1^{\mathrm{true}} \dots a_N^{\mathrm{true}})^T$ representing the projection V_N of V onto H_N. Assuming that $\{H_N\}_{N\in\mathbb{N}}$ forms a stable approximation for H (see [1] for details) one can get $\|V - V_N\|_H \to 0$ for $N \to \infty$. This represents a basic idea behind the so-called projection methods.

Motivation. One of the most popular projection methods is Petrov-Galerkin method. It is built upon the following requirement (see [7, p.43]):

$$\frac{dV_N}{dt} + AV_N - f \perp \mathrm{Lin}\{\varphi_1 \dots \varphi_N\}\,. \tag{2}$$

This condition yields an ODE for determining the projection coefficients $\mathbf{a}_N = (a_1 \dots a_N)^T$:

$$\langle \frac{dV_N}{dt} + AV_N - f, \varphi_n\rangle = 0\,, \quad n = 1 \dots N\,. \tag{3}$$

We stress that the vector of exact projection coefficients $\mathbf{a}_N^{\mathrm{true}}$ solves (3) in very special cases (see discussion in subsection III-A). In the general case, the solution of ODE (3) deviates from $\mathbf{a}_N^{\mathrm{true}}$ for any finite N and in this sense the system (3) is not closed, that is it does not retain all information which is necessary to describe the evolution of $\mathbf{a}_N^{\mathrm{true}}$. However, the solution of (3) converges to $\mathbf{a}_N^{\mathrm{true}}$ in the limit $N \to \infty$, provided the approximation is consistent (that is, the norm of the projection of $A(V - V_N)$ onto H_N vanishes if $N \to \infty$) and stable (see [7, p.251] for details). Convergence results in functional spaces (when one is interested in convergence of $V - V_N$ to 0 as $N \to \infty$ in the norm of H) were discussed, for instance, in [12] for linear hyperbolic PDEs and in [17] for linear parabolic PDEs. Basic theory of projection methods for elliptic PDEs may be found, for example, in [1].

We stress that in practice it is often desirable to have a closed finite dimensional system for projection coefficients where the impact of coefficients $(a_{N+1} \dots)$ onto $\mathbf{a}_N^{\mathrm{true}}$ is modeled by uncertain inputs with a given uncertainty description. In this case it would be necessary to construct a priori estimates of the approximation error for a fixed N that are suitable for finite or infinite time interval. It turned out (see [10], [9]) that for finite-dimensional systems these points could be addressed combining the framework of Differential-Algebraic Equations (DAEs) with classical results in state estimation for uncertain linear ODEs and a posteriori basis functions $\{\varphi_1 \dots \varphi_N\}$ generating H_N and obtained by means of principal component analysis. This served us as a motivation to investigate the subject in more general framework: for a priori basis functions and linear infinite dimensional systems.

Contribution and related work. In this paper we propose to model $\mathbf{a}_N^{\mathrm{true}}$ as a solution of a DAE with uncertain parameters representing the projection error (which contains projection of $A(V - V_N)$ onto H_N). We use semigroup properties in order to construct a priori bounding set for the projection error. This, in turn, allows us to claim that $\mathbf{a}_N^{\mathrm{true}}$ solves the proposed DAE for a particular choice of uncertain parameters within the bounding set. In fact, we close the system (3) by plugging in there an uncertain input modelling the projection error. We allow this input to vary in the a priori bounding set and introduce an additional algebraic constraint allowing to filter out un-admissible inputs. Assembling the differential and algebraic equation one gets an uncertain DAE which models $\mathbf{a}_N^{\mathrm{true}}$.

Since DAE has uncertain but bounded input we apply the minimax state estimation approach proposed for DAE in [18]. It allows to construct the minimax projection coefficients – a robust estimate for $\mathbf{a}_N^{\mathrm{true}}$ – and describe the worst-case estimation error. Namely, we construct an ellipsoid containing $\mathbf{a}_N^{\mathrm{true}}$ and centered around the minimax coefficents.

IBM Research, Dublin, Ireland, sergiy.zhuk@ie.ibm.com

The shape of the ellipsoid is defined by a symmetric positive definite matrix solving Differential Riccati Equation (DRE). The largest eigen-value of this matrix defines the worst-case estimation error. We refer the reader to [11], [3], [14] and [8] for the basic information on the minimax framework. We note that the minimax estimate does not require that projection of $A(V-V_N)$ onto H_N vanishes for a fixed N (or for $N \to \infty$ by consistency condition above). We only need to have it bounded and this bound is reflected in the a priori bounding set for uncertain input. To construct this bound, we use a priori estimates available for strongly continuous semigroups. We stress that the asymptotic behaviour of DRE's solutions is well understood [2] so the analysis of the approximation error for infinite horizon problems reduces to application of well known facts from control theory. On the other hand, there exists a number of numerical methods suitable for solving DRE [4]: we discuss in details how to compute the numerical minimax projection coefficients for 1D heat equation.

Although the literature on projection methods and related control techniques is very rich, to our best knowledge, modelling of projection coefficients for the linear evolution equations by means of uncertain DAEs and subsequent application of minimax state estimation approach to get robust estimates for DAE's states has not been discussed neither in control nor in approximation literature yet. In contrast to projection methods discussed in [1], [7], [12], [17] we focus on "closing the system" by adding uncertain parameters and giving a robust estimate of the state for a fixed dimension of the space H_N. As a byproduct, our DAE-oriented projection method brings an additional algebraic equation that controls the projection of $A(V-V_N)$ onto the completion of H_N and, hence, allows to narrow down the set of admissible uncertain inputs. The latter helps to improve the estimation quality: in the considered example, minimax projection coefficients converge to $\mathbf{a}_N^{\mathrm{true}}$ faster compared to classical Galerkin approximation for a fixed N and $t \to \infty$. To sum up, we can state the contribution of this paper as follows. For a given subspace $H_N := \mathrm{Lin}\{\varphi_1 \dots \varphi_N\}$ we

1) extend Galerkin model (3) to a DAE with uncertain parameters modelling exact projection coefficients $\mathbf{a}_N^{\mathrm{true}}$;
2) construct an a priori uncertainty description for uncertain parameters in the form of ellipsoidal set;
3) derive the minimax estimate and error for DAE's state vector that, in turn, constitute the minimax projection coefficients for V solving (1).

Outline. This paper is organised as follows. Subsection I-A contains notations, section II describes the problem statement, section III comprises main results: DAE model for exact projection coefficients with comparisons to the classical Galerkin approach (subsection III-A), and derivation of minimax projection coefficients (subsection III-B). Subsection IV presents case-study: 1D heat equation. Conclusions are in section V.

A. Notation

$\mathbb{R}^n$ denotes the n-dimensional Euclidean space; $(\mathbf{x},\mathbf{y})$ denotes the canonical inner product for $\mathbf{x},\mathbf{y} \in \mathbb{R}^n$; $\|\mathbf{x}\|^2_{\mathbb{R}^n} := (\mathbf{x},\mathbf{x})$; $\langle u, v\rangle$ denotes an inner product of two elements $u, v \in H$; $\|f\|^2_H := \langle f, f\rangle$; $L^2(0,t_1;H) := \{f : f(t) \in H$ and $\int_0^T \|f(t)\|^2_H dt < +\infty\}$; $C^1(0,T;H)$ is a set of all H-valued functions f, defined on $[0,T]$ such that a Fréchet derivative $f'(t) = \lim_{h\to 0} \frac{\|f(t+h)-f(t)\|_H}{h}$ exists for every $0 < t < T$ and is a continuous function on $(0,T)$; I stands for an identity operator or matrix; the prime $'$ denotes the operation of taking the adjoint: A' denotes adjoint operator, $\mathbf{A}'$ denotes the transposed matrix; $\|\mathbf{A}\|$ is the largest singular value of $\mathbf{A}$; for symmetric matrix $\mathbf{A}^{\frac{1}{2}}$ denotes a square root of $\mathbf{A}$; $D(A)$ denotes the domain of linear operator A; $\delta_{i,j} = 1$ if $i = j$ and 0 otherwise.

II. Problem statement

Assume that $V \in L^2(0,T;H)$ is a solution of the following evolution equation:

$$\frac{dV}{dt} + AV = f, \quad V(0) = v\,, \tag{4}$$

where $A : D(A) \subset H \to H$ is a closed linear operator with domain $D(A)$ which is dense in a Hilbert space H, $v \in H$ and $f \in L^2(0,T;H)$. In what follows we focus on a case of separable Hilbert spaces only.

Definition 1 (strong solution): $V \in L^2(0,T;H)$ is said to be a solution of (4) if $V \in C^1(0,T;H)$, $V(t,\cdot) \in D(A)$ for all $t \in (0,T)$ and (4) holds true on $[0,T)$.

Definition 2 (strongly continuous semigroups): $\{G(t)\}_{t\geq 0}$ is a strongly continuous semigroup of bounded linear operators if

$$\begin{cases} G(t+s) = G(t)G(s), & \forall t,s \geq 0 \\ G(0) = I \end{cases}$$

and for any $\delta > 0$, $t > 0$ and $x \in H$ there exists $\varepsilon(t,x) > 0$ such that $\|G(t)x - G(s)x\|_H \leq \delta$ if $|t-s| < \varepsilon(t,x)$.

Definition 3: A linear operator $A : D(A) \subset H \to H$ is said to be a generator of a strongly continuos semigroup $\{G(t)\}_{t\geq 0}$ on a Hilbert space H if $Av = \lim_{h\downarrow 0} \frac{1}{h}(G(h)v - v)$ and $D(A) = \{v \in H : \lim_{h\downarrow 0} \frac{1}{h}(G(h)v - v)$ exists$\}$.

If $-A$ generates a strongly continuous semigroup $\{G(t)\}_{t\geq 0}$ and $v \in D(A)$ and $f \in C^1(0,T;H)$ then (see for instance [15, p.105]) (4) has a unique solution V in the form:

$$V(t) = G(t)v + \int_0^t G(t-s)f(s)ds\,. \tag{5}$$

Although (5) is valid for the case $v \in H$ and $f \in L^2(0,T;H)$ and the corresponding V becomes a mild solution in the latter case, we will be working with strong solutions in order to simplify the presentation.

If $H = \mathbb{R}^n$ then (5) gives a well-known representation for the solution of LTI system. If H is a functional space[1], the semigroup generated by A might be associated with a linear

[1]For instance, $H = H^2(\Omega)$ for a bounded open set $\Omega \subset \mathbb{R}^n$ with smooth boundary

parabolic or hyperbolic equation (see [6, p.421]). We refer the reader to [5] for an extensive list of specific semigroups related to linear PDEs.

Let us briefly state few results from the general semigroup theory serving a basis for our presentation:

A) for a strongly continuous semigroup $\{G(t)\}_{t\geq 0}$ there exists $M \geq 1$ and $\omega \in R$ such that $\|G(t)v\|_H \leq Me^{\omega t}\|v\|_H$ (see [5, p.41]);
B) if A generates a strongly continuous semigroup $\{G(t)\}_{t\geq 0}$ then the equality $AG(t)v = G(t)Av$ holds true for $v \in D(A)$ (see [5, p.50]);
C) if (4) has a solution then $\int_0^t G(t-s)f(s)ds \in D(A)$ for $t \in (0,T)$ (see [15, p.107,T.2.4]).

In order to construct the solution of (4) one can choose a basis $\Phi := \{\varphi_k\}_{k\in\mathbb{N}} \subset D(A)$ in H and derive an Ordinary Differential Equation (ODE) describing the dynamics of coefficients representing V in Φ. Namely, assume that $\sum_{k=1}^N a_k(t)\varphi_k \to V$ and $\sum_{k=1}^N a_k(t)A\varphi_k \to AV(t)$ in H for all $t \in [0,T)$. Then $AV(t) = \sum_{k=1}^{+\infty} a_k(t)A\varphi_k$. Now, plugging these expansions into (4) and taking the inner product of both sides with φ_s, we get:

$$\sum_{k=1}^{+\infty} M_{s,k}\frac{da_k}{dt} + S_{s,k}a_k(t) = \langle f, \varphi_s\rangle\,, s \in \mathbb{N} \qquad (6)$$

where $M_{k,s} = \langle\varphi_k, \varphi_s\rangle$, $S_{k,s} = \langle A\varphi_k, \varphi_s\rangle$. In fact, equation (6) is a realisation of the abstract problem (4) in a coordinate Hilbert space ℓ_2: infinite matrix $\mathbf{A} := \{S_{s,k}\}_{s,k\in\mathbb{N}}$ and vectors $\mathbf{a} := (a_1, \dots, a_n, \dots)$, $\mathbf{v} := (\langle v, \varphi_1\rangle \dots \langle v, \varphi_n\rangle \dots)$ and $\mathbf{f} := (\langle f, \varphi_1\rangle \dots \langle f, \varphi_n\rangle \dots)^T$ represent the operator A, solution V, v and input f in the basis Φ. We stress that the procedure described so far is used to define a weak solution for (4). In this paper we assume that the (strong) solution, defined as in Definition 1, exists and is represented by a semigroup G. This guarantees existence and uniqueness of solution for (6). We refer the reader to [16] where a connection between the semigroup approach adopted here and the weak formulation[2] for (4) is discussed.

The system (6) allows one to construct $\mathbf{a}$ explicitly, provided Ψ is composed of eigen-vectors of A. It is a non-trivial problem, though, to get eigen-vectors for a given differential operator. In practice, one usually approximates infinite matrix $\mathbf{A}$ by finite-dimensional sub-matrices and solves the resulting ODE to estimate dynamics of the vector of exact projection coefficients $\mathbf{a}_N^{\text{true}} := (a_1 \dots a_N)^T$. This constitutes the well known Galerkin projection method. Now we are ready to formulate the problem statement:

1) construct a finite-dimensional linear system describing dynamics of the vector of exact projection coefficients $\mathbf{a}_N^{\text{true}}$ over time;
2) derive a robust estimate $\hat{\mathbf{a}}_N$ for $\mathbf{a}_N^{\text{true}}$ and compute the estimation error in H_N.

[2]In the week formulation the unbounded operator A is a restriction of a linear bounded operator $\mathcal{A} : D(A) \to D(A)'$ associated with a bilinear continuous form $a : D(A) \times D(A)' \to R$ onto a set $\{v \in D(A) : \mathcal{A} \in H\}$. As a result the semigroup approach yields more regular solutions (see [16, p.21]).

III. Main result

Define the following linear mappings: projection $\mathbf{P}_{\text{N}} : H \to \mathbb{R}^N$ and reconstruction $\mathbf{P}_{\text{N}}^+ : \mathbb{R}^N \to H$:

$$\mathbf{P}_{\text{N}}\,\psi := \begin{bmatrix}\langle\psi,\varphi_1\rangle \\ \dots \\ \langle\psi,\varphi_N\rangle\end{bmatrix}, \quad \mathbf{P}_{\text{N}}^+\,\mathbf{a} := \sum_{i=1}^N a_i\varphi_i\,, \mathbf{a} \in \mathbb{R}^N\,. \quad (7)$$

We set by definition:

$$\mathbf{M}_{\text{N}} := \mathbf{P}_{\text{N}}\,\mathbf{P}_{\text{N}}^+, \quad \mathbf{A}_{\text{N}} := \mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,A\,\mathbf{P}_{\text{N}}^+,$$
$$\mathbf{S} = \{\langle A\varphi_i, A\varphi_j\rangle\}_{i,j=1}^N, \mathbf{H}_{\text{N}} := (\mathbf{S} - \mathbf{A}_{\text{N}}{}'\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{A}_{\text{N}})^{\frac{1}{2}}\,.$$

In the next proposition we derive the DAE with uncertain parameters $\mathbf{e}^m, \mathbf{e}^o$ describing dynamics of $\mathbf{a}_N^{\text{true}}$ and construct the bounding set for $\mathbf{e}^m, \mathbf{e}^o$.

Proposition 1: Assume $v, f(s) \in D(A)$ and define

$$\mu_1(t) := Me^{\omega t}(\|v\|_H + \int_0^t e^{-s\omega}\|f(s)\|_H ds)\,,$$
$$\mu_2(t) := Me^{\omega t}(\|Av\|_H + \int_0^t e^{-s\omega}\|Af(s)\|_H ds)\,,$$
$$\beta(t) := \int_0^t 2(\mu_2^2(s) + \|\mathbf{S}\|\mu_1^2(s))ds\,.$$

There exists $\mathbf{e}^m$ and $\mathbf{e}^o$ such that:

$$\int_0^T (\mathbf{M}_{\text{N}}\,\mathbf{e}^m, \mathbf{e}^m) + (\mathbf{e}^o, \mathbf{e}^o)dt \leq \beta(T)\,, \qquad (8)$$

and the vector of the exact projection coefficients $\mathbf{a}_N^{\text{true}} \in \mathbb{R}^N$ solves the following DAE:

$$\begin{aligned}\frac{d\mathbf{a}}{dt} &= -\,\mathbf{A}_{\text{N}}\,\mathbf{a} + \mathbf{e}^m + \mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,f\,,\\ 0 &= \mathbf{H}_{\text{N}}\,\mathbf{a} + \mathbf{e}^o\,, \mathbf{a}(0) = \mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,v\,.\end{aligned} \qquad (9)$$

Proof: Let us prove that $\mathbf{a}_N^{\text{true}} \in \mathbb{R}^N$ solves (9) for a particular choice of $\mathbf{e}^m$ and $\mathbf{e}^o$. We stress that $\mathbf{a}_N^{\text{true}}$ is a unique minimum point of the projection problem: $\min_{a_1\dots a_N} \|V(t) - \sum_{i=1}^N a_i\varphi_i\|_H$ so it is straightforward to check that $\mathbf{a}_N^{\text{true}} = \mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,V$. Define

$$e(t) := A\,\mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,V(t) - \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,AV(t)\,. \quad (10)$$

Since $V(t) = \mathbf{P}_{\text{N}}^+\,\mathbf{a}_N^{\text{true}} + (I - \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}})V(t)$ it follows that $\mathbf{a}_N^{\text{true}}$ solves

$$\begin{aligned}\frac{d\,\mathbf{P}_{\text{N}}^+\,\mathbf{a}}{dt} &= \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,\frac{dV}{dt} = \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}(-AV + f)\\ &= -A\,\mathbf{P}_{\text{N}}^+\,\mathbf{a}(t) + e(t) + \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,f\,.\end{aligned} \qquad (11)$$

As $\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,\mathbf{P}_{\text{N}}^+ = I$ and $\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,\mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}} = \mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}$, we get, multiplying (11) by $\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}$, that $\mathbf{a}_N^{\text{true}}$ solves the first equation in (9) for $\mathbf{e}^m = \mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}}\,e(t)$. On the other hand, (11) has a solution if and only if $-A\,\mathbf{P}_{\text{N}}^+\,\mathbf{a} + e(t) \in R(\mathbf{P}_{\text{N}}^+)$. This holds true, in turn, if

$$(I - \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}})A\,\mathbf{P}_{\text{N}}^+\,\mathbf{a} = (I - \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}})e(t)\,. \quad (12)$$

After simple algebra one gets:

$$\|(I - \mathbf{P}_{\text{N}}^+\,\mathbf{M}_{\text{N}}^{-1}\,\mathbf{P}_{\text{N}})A\,\mathbf{P}_{\text{N}}^+\,\mathbf{a}_N^{\text{true}}\|_H = \|\,\mathbf{H}_{\text{N}}\,\mathbf{a}_N^{\text{true}}\|_{\mathbb{R}^N}^2\,. \quad (13)$$

Since $\| \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e(t)\|_H^2 = (\mathbf{M}_{\mathrm{N}} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e, \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e)$ and

$$\begin{aligned} \|e(t)\|_H^2 = &\| \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e(t)\|_H^2 \\ &+ \|(I - \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}}) e(t)\|_H^2 . \end{aligned} \tag{14}$$

(12)-(14) imply that the following equalities hold true for $\mathbf{a}_N^{\mathrm{true}}$ and $\mathbf{e}^m = \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e(t)$ ($\mu := \int_0^T \|e(t)\|_H^2 dt$):

$$\begin{aligned} &\frac{d\mathbf{a}}{dt} = -\mathbf{A}_{\mathrm{N}} \mathbf{a} + \mathbf{e}^m + \mathbf{P}_{\mathrm{N}} f , \\ &0 = \mathbf{H}_{\mathrm{N}} \mathbf{a} + \mathbf{e}^o , \\ &\int_0^T (M\mathbf{e}^m, \mathbf{e}^m) + (\mathbf{e}^o, \mathbf{e}^o) dt = \mu . \end{aligned} \tag{15}$$

To conclude the proof it is sufficient to show that $\mu \le \int_0^T 2(\mu_1^2(s) + \|S\| \mu_2^2(s)) ds$. We stress that:

$$\begin{aligned} &\| \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e(t)\|_H^2 = (\mathbf{M}_{\mathrm{N}} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e, \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e) \\ &= \langle \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} AV(t), AV(t) \rangle + (\mathbf{A}_{\mathrm{N}} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{A}_{\mathrm{N}} \mathbf{a}_N^{\mathrm{true}}, \mathbf{a}_N^{\mathrm{true}}) \\ &- 2\langle AV, \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{A}_{\mathrm{N}} \mathbf{a}_N^{\mathrm{true}} \rangle \end{aligned}$$

and so, using (14) and (13) one gets:

$$\begin{aligned} &\|e(t)\|_H^2 = \langle \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} AV(t), AV(t) \rangle + (\mathbf{S}\mathbf{a}_N^{\mathrm{true}}, \mathbf{a}_N^{\mathrm{true}}) \\ &- 2\langle AV, \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{A}_{\mathrm{N}} \mathbf{a}_N^{\mathrm{true}} \rangle \le 2\|AV(t)\|_H^2 + 2\|\mathbf{S}\| \|V(t)\|_H^2 . \end{aligned}$$

Now, using (5) and A) we get: $\|V(t)\|_H \le \mu_1(t)$. On the other hand, B) and C) imply $A \int_0^t G(t-s) f(s) ds = \int_0^t G(t-s) Af(s) ds$, so that $\|AV(t)\|_H \le \mu_2(t)$. Therefore, $\|e(t)\|_H^2 \le 2(\mu_2^2(t) + \|\mathbf{S}\| \mu_1^2(t))$. ■

A. Comparison to the classical Galerkin approach

To simplify the presentation we assume for a moment that $f = 0$. The classical Galerkin projection approach is built upon the following requirement [7, p.43]:

$$\frac{dV_N}{dt} + AV_N \perp \mathrm{Lin}\{\varphi_1 \dots \varphi_N\} . \tag{16}$$

where $V_N(t) = \sum_{i=1}^N a_i(t) \varphi_i$ is an approximation of the projection of V onto $\mathrm{Lin}\{\varphi_1 \dots \varphi_N\}$. This condition yields the following ODE for determining the coefficients $\mathbf{a} = (a_1 \dots a_N)^T$:

$$\frac{d\mathbf{a}}{dt} = -\mathbf{A}_{\mathrm{N}} \mathbf{a}, \quad \mathbf{a}(t_0) = \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} v . \tag{17}$$

Let us investigate connections between (9) and (17). We note that the basic assumption (16) of Galerkin method holds true for $\mathbf{a}_N^{\mathrm{true}}$ if and only if

$$\frac{d\, \mathbf{P}_{\mathrm{N}}^{+} \mathbf{a}_N^{\mathrm{true}}}{dt} + A\, \mathbf{P}_{\mathrm{N}}^{+} \mathbf{a}_N^{\mathrm{true}} \perp \mathrm{Lin}\{\varphi_1 \dots \varphi_N\} .$$

Now, by (11), the latter is true if and only if $\mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e(t) = 0$. Recalling (10), we rewrite e as follows:

$$\begin{aligned} e = &(I - \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}}) A\, \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} V \\ &+ \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} A(\mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} - I) V . \end{aligned} \tag{18}$$

Intuitively, term $\|(I - \mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}}) V\|_H$ measures the energy of the projection error for the basis $\{\varphi_1 \dots \varphi_N\}$ and, therefore, $\|e\|_H$ is a sum of the energy of the projection error of A-image of the projected solution $\mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} V$, and the energy of the projection of A-image of the projection error $(\mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} - I)V$. Now, we compute $\mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e(t) = \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} A(\mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} - I)V$ and so $\mathbf{a}(t) = \mathbf{a}_N^{\mathrm{true}}$ if and only if

$$\mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} A(\mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} - I) V = 0 . \tag{19}$$

In other words, $\mathbf{a}(t) = \mathbf{a}_N^{\mathrm{true}}$ if the A-image of the projection error $(\mathbf{P}_{\mathrm{N}}^{+} \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} - I)V$ is orthogonal to the span of $\{\varphi_k\}_{k=1}^N$ and, therefore, has no impact onto the dynamics of $\mathbf{a}_N^{\mathrm{true}}$. We stress that $\mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e(t) \ne 0$ in the general case but there is a number of important special cases when this holds true. We note that $e = 0$ provided $A\varphi_k = \alpha_k \varphi_k$. This suggests the following interpretation for (18): norm of e quantifies "how far off is the subspace generated by $\{\varphi_k\}_{k=1}^N$ from an eigen-subspace of A". More generally, $\mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e = 0$ if $\mathbf{P}_{\mathrm{N}}^{+} \mathbf{P}_{\mathrm{N}}$ commutes with A. In this case, (17) gives a closed system for $\mathbf{a}_N^{\mathrm{true}}$: it contains all required information to describe how the exact projection coefficients evolve in time. We stress that, in practice, the assumption $\mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e = 0$ is not easy to check (for a given set of basis functions) as e depends on the solution V which is unknown. On the other hand, it is not easy, given A, to choose basis functions such that $\mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e = 0$. Therefore, in practice, Galerkin system (17) is usually non-closed.

The solution proposed by Proposition 1 is to consider $\mathbf{e}^m = \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e$ as an uncertain input for (17) and construct an a priori estimate for $\mathbf{e}^m = \mathbf{M}_{\mathrm{N}}^{-1} \mathbf{P}_{\mathrm{N}} e$ in the form (8) using the information about A and data v, f. As a result, the true coefficients $\mathbf{a}_N^{\mathrm{true}}$ belong to the set of solutions of (9). The information, provided by the second equation in (9), allows to filter out un-admissible $\mathbf{e}^m$. Indeed, (15) shows that $\mathbf{a}_N^{\mathrm{true}}$ solves the algebraic equation in (9) for some $\mathbf{e}^o$ belonging to the ellipsoid defined by (8). This allows one, in turn, to narrow down the set of all admissible $\mathbf{a}$ solving (9). As a result, DAE (9) together with ellipsoid (8) represents a closed system for projection coefficients $\mathbf{a}_N^{\mathrm{true}}$ in the general case.

The system (9) contains uncertain parameters and is not suitable for computations in the form it has been stated so far. In contrast, Galerkin system does not require any additional considerations to approximate $\mathbf{a}_N^{\mathrm{true}}$. Fortunately, the uncertainty description (8) and DAE equation (9) fall into the framework of minimax state estimation approach [18]. In the next subsection we will show how one can use (9) in order to construct the minimax estimate for $\mathbf{a}_N^{\mathrm{true}}$ and describe the estimation error.

B. Minimax projection coefficients

Definition 4: Assume $\mathbf{a}_0$ solves the linear ODE

$$\frac{d\mathbf{a}_0}{dt} = -\mathbf{A}_{\mathrm{N}} \mathbf{a}_0 + \mathbf{e}^m, \quad \mathbf{a}_0(0) = 0 ,$$

and satisfies an algebraic equation $\mathbf{y}(s) = \mathbf{H}_{\mathrm{N}} \mathbf{a}_0(s) + \mathbf{e}^o(s)$, $0 \le s \le t$ for some $\mathbf{y}$, $\mathbf{e}^m$ and $\mathbf{e}^o$ such that for the given $\mathbf{Q} = \mathbf{Q}' > 0$ and $\mathbf{R} = \mathbf{R}' > 0$ the following inequality holds true:

$$\int_0^t (\mathbf{Q}\mathbf{e}^m, \mathbf{e}^m) + (\mathbf{R}\mathbf{e}^o, \mathbf{e}^o) dt \le 1 .$$

The minimax estimate $\hat{\mathbf{a}}(t)$ of $\mathbf{a}_0$ is an output of a LTV system $\frac{d\mathbf{a}}{dt} = -\mathbf{A}_\mathrm{N}\,\mathbf{a} + \hat{\mathbf{u}}$ such that:

$$\hat{\sigma}(t) := \max_{\mathbf{e}^m,\mathbf{e}^o} \|\hat{\mathbf{a}}(t) - \mathbf{a}_0(t)\|_{\mathbb{R}^n} \le \max_{\mathbf{e}^m,\mathbf{e}^o} \|\mathbf{a}(t) - \mathbf{a}_0(t)\|_{\mathbb{R}^n}$$

for any LTV in the form $\frac{d\mathbf{a}}{dt} = -\mathbf{A}_\mathrm{N}\,\mathbf{a} + \mathbf{u}$. The number $\hat{\sigma}(t)$ represents the minimax error.

As one can see from the above definition, the minimax estimate is robust to realisations of uncertain $\mathbf{e}^m$ and $\mathbf{e}^o$ and the worst-case estimation error is represented by $\hat{\sigma}(t)$.

Define $\mathbf{Q}(t) := \beta^{-1}(t)\,\mathbf{M}_\mathrm{N}$ and $\mathbf{R}(t) := \beta^{-1}(t)I$. Note that (see [7, p.251]) the convergence of solution $\mathbf{a}$ of (3) to $\mathbf{a}_N^{\mathrm{true}}$ depends on whether the term $\|\mathbf{e}^m\|_H = \|\,\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,e\|_{\mathbb{R}^n} = \|\,\mathbf{P}_\mathrm{N}^+\,\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,A(\mathbf{P}_\mathrm{N}^+\,\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N} - I)V\|_H$ tends to 0 or not for $N \to \infty$. In terms of Proposition 1 this term regulates the size of the ellipsoid (8) through μ (see (15)) which is bounded by $\beta(T)$ from above. We note that the minimax estimate depends on the ratio between $\mathbf{Q}$ and $\mathbf{R}$ only. Now, since both $\mathbf{Q}$ and $\mathbf{R}$ are uniform in β it follows that the actual minimax estimate does not depend on the particular norm of e. In other words, multiplying e by a positive number C yields the same minimax estimate but amplifies the minimax error by C.

The following theorem presents the minimax estimate $\hat{\mathbf{a}}_N$ for the vector of exact projection coefficients $\mathbf{a}_N^{\mathrm{true}}$ and the minimax estimation error $\hat{\sigma}(t)$.

Theorem 1: Assume that $\mathbf{K}(0) = 0$ and $\mathbf{K}$ solves the following Riccati equation:

$$\frac{d\mathbf{K}}{dt} = -\mathbf{A}_\mathrm{N}{}'\,\mathbf{K} - \mathbf{K}\,\mathbf{A}_\mathrm{N} + \mathbf{Q}^{-1} - \mathbf{K}\,\mathbf{H}_\mathrm{N}{}'\,\mathbf{R}\,\mathbf{H}_\mathrm{N}\,\mathbf{K}\,, \quad (20)$$

Set $\hat{\mathbf{a}}_N(0) = \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,v$ and let $\hat{\mathbf{a}}_N$ solve the following ODE:

$$\frac{d\hat{\mathbf{a}}_N}{dt} = -\mathbf{A}_\mathrm{N}\,\hat{\mathbf{a}}_N - \mathbf{K}\,\mathbf{H}_\mathrm{N}{}'\,\mathbf{R}\,\mathbf{H}_\mathrm{N}\,\hat{\mathbf{a}}_N + \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,f\,. \quad (21)$$

Then the vector of exact projection coefficients $\mathbf{a}_N^{\mathrm{true}}$ belongs to the ellipsoid centered around $\hat{\mathbf{a}}_N$:

$$(\mathbf{K}^{-1}(\mathbf{a}_N^{\mathrm{true}}(t) - \hat{\mathbf{a}}_N(t)), \mathbf{a}_N^{\mathrm{true}}(t) - \hat{\mathbf{a}}_N(t)) \le 1\,, t > 0\,. \quad (22)$$

The minimax error is given by:

$$\hat{\sigma}(t) = \|\mathbf{a}_N^{\mathrm{true}}(t) - \hat{\mathbf{a}}_N(t)\|_{\mathbb{R}^N} \le \|\mathbf{K}^{\frac{1}{2}}(t)\|\,.$$

Proof: The proof is omitted due to the lack of space. ∎

IV. Case study: 1D heat equation

In this subsection we apply the minimax projection method to 1D heat equation. We set $H = L^2(-1,1)$, $A = -c^2\partial_{xx}^2$ with $D(A) = \{v \in H : v(-1) = v(1) = 0, v_{x,x}, v_x \in H\}$. Then (4) reads as (assuming $f = 0$):

$$V_t - c^2 V_{x,x} = 0, V(-1,t) = V(1,t) = 0, V(x,0) = v(x)\,.$$

It is easy to check integrating by parts that $A = A^*$ and $D(A) = D(A^*)$. On the other hand, $-A$ generates (see [16, p.19,Example 4.C]) a strongly continuous semigroup $\{G(t)\}_{t\ge 0}$ which is a contraction and $G(t)v(x) = V(x,t) = \sum_{n=1}^{\infty} a_n \exp\{-n^2c^2\pi^2 t\}\sin(n\pi x)$ provided $v(x) = \sum_{n=1}^{\infty} a_n \sin(n\pi x)$. For the numerical experiment we take $v(x) = \sin(n\pi x)$ so that the exact solution reads as $V(x,t) = \exp\{-n^2c^2\pi^2 t\}\sin(n\pi x)$.

Following [7, p.121] we take basis functions φ_k in the form of Legendre polynomials $\varphi_k = P_{k+1} - P_{k-1}$. Then $\varphi_k(-1) = \varphi_k(1) = 0$ and $\{\varphi_k\}_{k\in\mathbb{N}}$ spans H. This system is not orthogonal, though: the corresponding mass matrix $\mathbf{M}_\mathrm{N} = \mathbf{P}_\mathrm{N}\,\mathbf{P}_\mathrm{N}^+$ is tridiagonal:

$$M_{n,s} = \delta_{n+2,s}\frac{-2}{2n+1} + \delta_{n-2,s}\frac{-2}{2n-1} + \delta_{s,n}\frac{4(2n+1)}{(2n-1)(2n+3)}\,.$$

In this case Galerkin system (17) is defined by: $\frac{d\mathbf{a}}{dt} = -\mathbf{A}_\mathrm{N}\,\mathbf{a}$, $\mathbf{a}(0) = \mathbf{a}_N^{\mathrm{true}}(0) = \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\sin(n\pi x)$. The matrix $\mathbf{A}_\mathrm{N}$ can be easily computed exactly noting that:

$$\langle AP_n, P_k\rangle = \begin{cases} 0, n+2 > k \text{ or } n+2 \le k, & k+n \text{ is odd,} \\ k(k+1) - n(n+1), & \text{otherwise}\,. \end{cases}$$

Now we have to verify whether Galerkin system is closed, that is $\mathbf{a}(t) = \mathbf{a}_N^{\mathrm{true}}$. To this end we set $c := n^{-1}$ and evaluate (see (19)): $\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,e(t) = \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,AV(t) - \mathbf{A}_\mathrm{N}\,\mathbf{a}_N^{\mathrm{true}} = e^{-\pi^2 t}(-\pi^2 I - \mathbf{A}_\mathrm{N})\mathbf{a}_N^{\mathrm{true}}(0)$. Since $\mathbf{A}_\mathrm{N} = \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,A\,\mathbf{P}_\mathrm{N}^+ \ne -\pi^2 I$, it follows that Galerkin system is not closed. In fact, the norm of $\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,e(t)$ defines the error of the Galerkin approximation. Indeed, setting $\mathrm{err}(t) := \mathbf{a}(t) - \mathbf{a}_N^{\mathrm{true}}$ we obtain:

$$\frac{d\,\mathrm{err}(t)}{dt} = \mathbf{A}_\mathrm{N}\,\mathrm{err}(t) + \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,e(t)\,.$$

We get: $\|\,\mathbf{P}_\mathrm{N}^+\,\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,e(t)\|_H^2 = e^{-2\pi^2 t}\|(\pi^2 I + \mathbf{A}_\mathrm{N})\mathbf{a}_N^{\mathrm{true}}(0)\|_{\mathbb{R}^n}$. Now, since $\mathbf{A}_\mathrm{N} = \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,A\,\mathbf{P}_\mathrm{N}^+$ is multiplied by $c^2 = n^{-2}$, it follows that the spectrum of $\mathbf{A}_\mathrm{N}$ decreases proportionally to n^{-2}. Thus, $err(t)$ decays slow for large n, provided the norm of $\mathbf{a}_N^{\mathrm{true}}(0) = \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\sin(n\pi x)$ was high. Recalling the definition of $M_{i,j}$ given above we see that the norm of $\mathbf{a}_N^{\mathrm{true}}(0)$ is high indeed: in fact, $\|(\pi^2 I + \mathbf{A}_\mathrm{N})\mathbf{a}_N^{\mathrm{true}}(0)\|_{\mathbb{R}^n} \approx 426$ for $n = 30$ and $N = 10$ whereas the corresponding spectrum of $\mathbf{A}_\mathrm{N}$ is in $[-0.515, -.01]$. Therefore, we expect a slow convergence for Galerkin method and this is justified by Figure 1. On the other hand, $V(x,t) = \exp\{-\pi^2 t\}\sin(n\pi x)$ and so $\mathbf{a}_N^{\mathrm{true}} = \exp\{-\pi^2 t\}\mathbf{a}_N^{\mathrm{true}}(0)$ that converges to 0 with rate $\exp\{-2\pi^2 t\}$.

Let us compute the minimax estimate (21). To identify the ellipsoid (8) we will make use of our knowledge of $V(x,t)$ and evaluate $\mu = \int_0^T \|e(t)\|_H^2 dt$ exactly (see (15)). Combining (13),(12) and (14) we get:

$$\begin{aligned}\mu &= \int_0^T (\mathbf{M}_\mathrm{N}\,\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,e, \mathbf{M}_\mathrm{N}^{-1}\,\mathbf{P}_\mathrm{N}\,e) + \|\,\mathbf{H}_\mathrm{N}\,\mathbf{a}_N^{\mathrm{true}}\|_{\mathbb{R}^N}^2 dt \\ &= \alpha_n(\mathbf{M}_\mathrm{N}(-\pi^2 I - \mathbf{A}_\mathrm{N})\mathbf{a}_N^{\mathrm{true}}(0), (-\pi^2 I - \mathbf{A}_\mathrm{N})\mathbf{a}_N^{\mathrm{true}}(0)) \\ &+ \alpha_n\|\,\mathbf{H}_\mathrm{N}\,\mathbf{a}_N^{\mathrm{true}}(0)\|_{\mathbb{R}^N}^2 := \alpha_n(\mu_1 + \mu_2)\,,\end{aligned}$$

where $\alpha_n := \int_0^T e^{-2\pi^2 t}dt = (1 - e^{-2\pi^2 T})(2\pi^2)^{-1} \approx 0.0507$ and $\mu_1 \approx 69$. We stress that $\mathbf{H}_\mathrm{N} = (\mathbf{S} - \mathbf{A}_\mathrm{N}{}'\,\mathbf{M}_\mathrm{N}^{-1}\,\mathbf{A}_\mathrm{N})^{\frac{1}{2}}$ may be computed exactly. As a result we get $\mu_2 \approx 0.32$ so

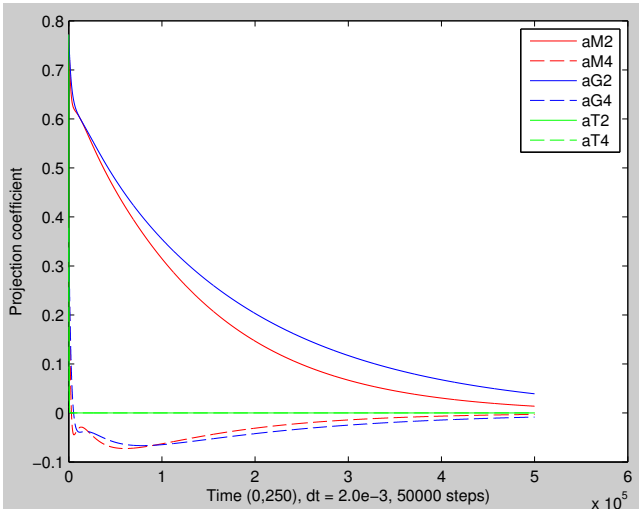

Fig. 1. Comparisons for a_2^{true} and a_4^{true} projection coefficients: aM2, aM4 represent minimax estimates (solid and dashed red), aG2, aG4 – Galerking approximations (solid and dashed blue), aT2, aT4 are exact projection coefficients (solid and dashed green) respectively.

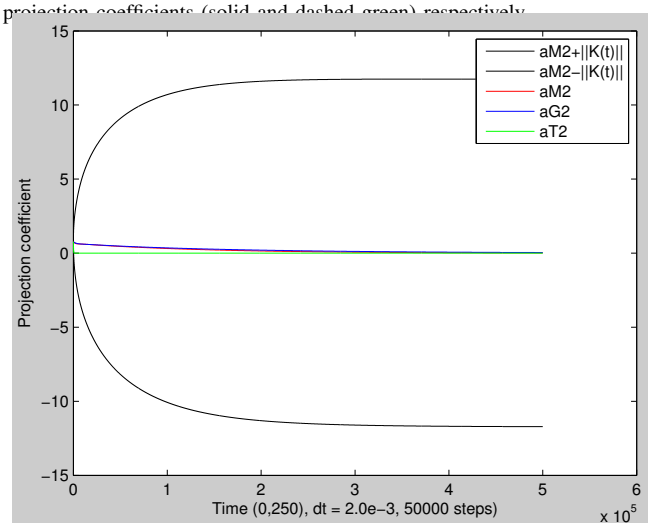

Fig. 2. Ellipsoid (in black) is centered around the minimax estimate: $[\|K(t)\| - \hat{a}_2(t), \|K(t)\| + \hat{a}_2(t)]$. It contains the truth and Galerkin approximation.

the contribution of the term $\mathbf{H}_{\mathrm{N}}\,\mathbf{a}_N^{\text{true}}(0)$ into the total error is negligible compared to $\mathbf{M}_{\mathrm{N}}^{-1}\,\mathbf{P}_{\mathrm{N}}\,e$. In total, we get $\mu = \alpha_n(\mu_1+\mu_2) \approx 3.5$. Now, to construct a numerical solution of DRE (20) we applied implicit backward differentiation multistep method of 4th order which requires to solve an ARE on each time step. The spectrum of $-\mathbf{A}_{\mathrm{N}}\,\hat{\mathbf{a}}_N - \mathbf{K}\,\mathbf{H}_{\mathrm{N}}{}'\,\mathbf{R}\,\mathbf{H}_{\mathrm{N}}$ is in $[-1.36, -0.01]$ which improves the convergence rate for the minimax estimate: although the minimax error $\hat{\sigma}(t)$ stabilizes around 11 (see Fig.2), that shows a high level of uncertainty in the system, the minimax estimate $\hat{\mathbf{a}}_N$ converges to $\mathbf{a}_N^{\text{true}}$ faster ($\approx$ 10% faster) than Galerkin approximation (see Fig.1).

V. Conclusion

We presented the minimax projection method for linear evolution equations. The main idea behind it is to model a vector of exact projection coefficients using a DAE with uncertain inputs representing the projection error. We allow the uncertain inputs to vary within an a priori ellipsoid which is constructed using a priori estimates available for the corresponding strongly continuous semigroup. The minimax state estimation approach, applied to uncertain DAE, allows to construct the minimax projection coefficients and estimate the worst-case error. The resulting estimate is governed by an ODE which differs from the classical Galerkin model in that it has a new term. In terms of control theory this term represents a stabilizing feedback. Minimax error is described by the largest eigenvalue of the feed-back gain which solves differential Riccati equation.

For the future it would be desirable mainly from the practical point of view to study infinite horizon case, that is to relate the choice of basis functions to detectability for $\mathbf{A}_{\mathrm{N}}, \mathbf{H}_{\mathrm{N}}$ and corresponding convergence of $\mathbf{K}$ to the solution of algebraic Riccati equation. In this case, the projection coefficients of the non-stationary problem would tend to the coefficients of the equation $LV = f$. Another promising direction is to minimize $\|\mathbf{K}(T)\|$ by the choice of basis functions for a fixed N.

References

[1] J. Aubin. *Approximation of elliptic boundary-value problems*. Wiley, 1972.

[2] F. Callier and J. Willems. Criterion for the convergence of the solution of the riccati differential equation. *IEEE TAC*, AC-26(6), 1981.

[3] F. L. Chernousko. *State Estimation for Dynamic Systems*. Boca Raton, FL: CRC, 1994.

[4] L. Dieci. Numerical integration of differential riccati equation and some related issues. *SIAM Journal on numerical analysis*, 29(3), 1992.

[5] K. Engel and R. Nagel. *One-parameter semigroups for linear evolution equations*. Springer, 2000.

[6] L. Evans. *Partial Differential Equations*, volume 19 of *Graduate studies in mathematics*. AMS, 2nd edition, 2010.

[7] J. Hesthaven, S. Gottlieb, and D. Gottlieb. *Spectral Methods for Time-Dependent Problems*. Cambridge University Press, 2007.

[8] Alexander Kurzhanski and István Vályi. *Ellipsoidal calculus for estimation and control*. Systems & Control: Foundations & Applications. Birkhäuser Boston Inc., Boston, MA, 1997.

[9] V. Mallet and S. Zhuk. Reduced minimax state estimation. Technical Report RR-7500, INRIA, Paris-Rocquencourt, 2010. `http://hal.archives-ouvertes.fr/inria-00550729/en/`.

[10] V. Mallet and S. Zhuk. Reduced minimax filtering by means of differential-algebraic equations. In *Proc. of 5th Int. Conf. on Physics and Control*, 2011. Available at: `lib.physcon.ru`.

[11] Mario Milanese and Roberto Tempo. Optimal algorithms theory for robust estimation and prediction. *IEEE Trans. Automat. Control*, 30(8):730–738, 1985.

[12] K. Morton and E. Süli. Evolution-galerkin methods and their supra-convergence. *Numer. Math.*, 71:331–355, 1995.

[13] P. Müller. Descriptor systems: pros and cons of system modelling by differential-algebraic equations. *Mathematics and Computers in Simulation*, 53(4-6):273–279, 2000.

[14] A. Nakonechny. A minimax estimate for functionals of the solutions of operator equations. *Arch. Math. (Brno)*, 14(1):55–59, 1978.

[15] A. Pazy. *Semigroups of Linear Operators and Applications to Partial Differential Equations*. Springer, 1992.

[16] R.E. Showalter. *Monotone Operators in Banach Space and Nonlinear Partial Differential Equations*. AMS, 1997.

[17] V. Thomée. *Galerkin finite element methods for parabolic problems*. Computational mathematics. Springer, 1997.

[18] S. Zhuk. Minimax state estimation for linear stationary differential-algebraic equations. In *Proc. of 16th IFAC Symposium on System Identification*, 2012. Available at: `www.ifac-papersonline.net`.