



Missouri University of Science and Technology
Scholars' Mine

Mathematics and Statistics Faculty Research &
Creative Works

Mathematics and Statistics


01 Jan 2011

Convergent Snapshot Algorithms for Infinite-Dimensional Lyapunov Equations

John R. Singler

Missouri University of Science and Technology, singlerj@mst.edu

Follow this and additional works at: https://scholarsmine.mst.edu/math_stat_facwork

 Part of the [Mathematics Commons](#), and the [Statistics and Probability Commons](#)

Recommended Citation

J. R. Singler, "Convergent Snapshot Algorithms for Infinite-Dimensional Lyapunov Equations," *IMA Journal of Numerical Analysis*, Oxford University Press, Jan 2011.

The definitive version is available at <https://doi.org/10.1093/imanum/drq028>

This Article - Journal is brought to you for free and open access by Scholars' Mine. It has been accepted for inclusion in Mathematics and Statistics Faculty Research & Creative Works by an authorized administrator of Scholars' Mine. This work is protected by U. S. Copyright Law. Unauthorized use including reproduction for redistribution requires the permission of the copyright holder. For more information, please contact scholarsmine@mst.edu.

Convergent Snapshot Algorithms for Infinite Dimensional Lyapunov Equations

John R. Singler

Department of Mathematics and Statistics
Missouri University of Science and Technology
Rolla, MO 65409-0020
Email: singlerj@mst.edu

April 16, 2013

Abstract

We consider two algorithms to approximate the solution Z of a class of stable operator Lyapunov equations of the form $AZ + ZA^* + BB^* = 0$. The algorithms utilize time snapshots of solutions of certain linear infinite dimensional differential equations to construct the approximations. Matrix approximations of the operators A and B are not required and the algorithms are applicable as long as the rank of B is relatively small. The first algorithm produces an optimal low rank approximate solution using proper orthogonal decomposition. The second algorithm approximates the product of the solution with a few vectors and can be implemented with a minimal amount of storage. Both algorithms are known for the matrix case, however the extension of the algorithms to infinite dimensions appears to be new. We establish easily verifiable convergence theory and a priori error bounds for both algorithms and present numerical results for two model problems.

1 Introduction

Approximating the solution of an operator Lyapunov equation of the form

$$AZ + ZA^* + BB^* = 0 \tag{1}$$

has many applications in model reduction and control problems for linear systems. For example, matrix Lyapunov equations are used in algorithms for balanced model reduction problems ([Antoulas, 2005](#); [Datta, 2004](#); [Zhou *et al.*, 1996](#)) and they arise in Newton iterations for Riccati equations ([Kleinman, 1968](#); [Banks & Ito, 1991](#); [Burns *et al.*, 2008](#); [Morris & Navasca, 2008](#)), which are used to compute optimal feedback control laws. Recent work on approximating large-scale Lyapunov equations include [Antoulas \(2005\)](#); [Baur & Benner \(2006\)](#); [Gavrilyuk *et al.* \(2004\)](#); [Grasedyck & Hackbusch \(2007\)](#); [Grasedyck *et al.* \(2003\)](#); [Gugercin *et al.* \(2003\)](#); [Li & White \(2002\)](#); [Penzl \(1990\)](#); [Rosen & Wang \(1995\)](#); [Simoncini \(2007\)](#). These problems often arise from the discretization of a partial differential equation (PDE).

In this work, we consider two snapshot algorithms to directly approximate solutions of a stable operator Lyapunov equation of the above form. Unlike many other large-scale algorithms, the algorithms presented here are not iterative; instead, the approximation is constructed by simulating m linear infinite dimensional differential equations, where m is the rank of B . Solving these

differential equations is the main computational cost of the algorithm; therefore, the proposed algorithm is applicable to large-scale systems when the rank of B is relatively small.

The first algorithm uses proper orthogonal decomposition to construct a low rank approximation to the solution Z of the Lyapunov equation. The second algorithm approximates Zx for a few vectors x with a minimal amount of storage; in certain applications (such as the Newton iterations for Riccati equations mentioned above) this is all that is needed. After an earlier version of this work was complete (Singler, 2008), the author became aware that both algorithms are known for the matrix case: the minimal storage algorithm was proposed by Saad (1990) and the low rank algorithm by Willcox and Peraire (2002). We extend both algorithms to an infinite dimensional case and prove convergence under minimal and easily verifiable assumptions. Furthermore, we establish a priori error bounds for both algorithms.

Let us briefly discuss other solution strategies for operator Lyapunov equations and the reasons for advocating the algorithms studied here. To clarify the ideas, we focus on a distributed parameter model problem: a one dimensional convection diffusion equation

$$w_t(t, x) = \mu w_{xx}(t, x) - \kappa w_x(t, x) + b(x)u(t), \quad (2)$$

$$w(t, 0) = 0, \quad w(t, 1) = 0, \quad (3)$$

where subscripts denote partial derivatives, μ and κ are constants, $b(x)$ is a given function, and $u(t)$ is an input. Roughly, the A operator corresponds to the convection diffusion operator, i.e., $Aw = \mu w_{xx} - \kappa w_x$, along with the boundary conditions (3). The B operator corresponds to multiplication by the function $b(x)$, i.e., $Bu = b(x)u$. (More details for this example can be found in Section 5.)

A standard approach to approximating the solution of the operator Lyapunov equation (1) is to first discretize the system and solve the corresponding finite dimensional Lyapunov equation. For example, discretizing the system in space (e.g., with finite differences or finite elements) leads to a matrix differential equation of the form

$$\dot{x}_N(t) = [A_N]x_N(t) + [B_N]u(t). \quad (4)$$

One can now use existing large-scale solvers for the matrix Lyapunov equation

$$[A_N][Z_N] + [Z_N][A_N]^* + [B_N][B_N]^* = 0. \quad (5)$$

This matrix approximation approach has been used extensively for model reduction and control computations for distributed parameter systems; see, e.g., Banks & Burns (1978); Banks & Ito (1991); Banks *et al.* (1996); Burns & Fabiano (1989); Burns *et al.* (1998); Camp & King (2002); Evans (2003); Gibson & Adamian (1991); King *et al.* (2006); Morris & Navasca (2008).

There may be difficulties with this matrix approximation approach that must be addressed in order to tackle complex problems.

Difficulties of matrix approximation approach:

- **Matrix approximations are required:** It may not be easy or even possible to extract matrix approximations ($[A_n], [B_N]$) of the operators (A, B) from existing black box or commercial simulation codes.
- **Difficulties verifying convergence:** For some problems and discretization schemes, the existing convergence theory (see, e.g., Corollary 4.11 in Curtain, 2003) can be difficult to verify. As the problems and discretization schemes become increasingly complex, it may be extremely difficult to theoretically and numerically verify convergence. For example:

- Some discretization methods for certain problems (such as linearized fluid flows) do not produce an approximate differential equation of the form (4); standard convergence theory does not apply and guarantees of convergence and accuracy may be difficult to obtain.
- A “natural” discretization scheme may fail to satisfy the requirements of the theory and produce an incorrect approximation (e.g., see Burns *et al.*, 1988; Borggaard *et al.*, 2004).
- **No known adaptive methods to increase accuracy:** There is no known method for estimating the error between the computed solution of the finite dimensional Lyapunov equation (5) and the solution of the corresponding infinite dimensional Lyapunov equation (1). Therefore, it may not be clear how to adaptive refine the discretization scheme to ensure accuracy.

The snapshot algorithms studied here have many advantages, including the potential to overcome the above difficulties.

Advantages of snapshot algorithms:

- **Computationally efficient:** The main computational cost of the algorithm is computing solutions of linear infinite dimensional differential equations. These computations can be performed very efficiently, using well developed computational methods and/or existing simulation code.
- **Matrix approximations are not required:** Of course discretization must be performed with the algorithms, however an approximating matrix differential equation of the form (4) is not required. This allows the use of existing simulation codes and specialized discretization schemes that produce approximations of different forms.
- **Straightforward to verify convergence:** The convergence theory is easy to verify and only requires convergence of the solutions of the linear differential equations.
- **Best possible approximation error:** For the low rank algorithm, the approximate error converges to the best possible error for approximations of a given rank.
- **Adaptive methods to increase accuracy:** Simple, computable error bounds indicate the quality of the approximation and can guide the rank and refinement of the approximation. In particular, the approximation error depends largely on the simulation error in solving the linear infinite dimensional differential equations; thus, it is possible to use adaptive solvers or error estimators to guide refinement and produce more accurate approximations.

Proper orthogonal decomposition (POD), which is described in detail in Section 3, has been used extensively for model reduction and control computations for partial differential equations; see, e.g., Atwell (2000); Atwell *et al.* (2001); Atwell & King (2001, 2004); Banks *et al.* (2002, 2000); Kepler *et al.* (2000); Lee & Tran (2005). The use of POD in this work is fundamentally different than in the above references. To be complete, we briefly outline the “standard” POD-based approach for the Lyapunov equation (1) and its difficulties:

1. Collect a dataset capturing various features of the partial differential equation (2)-(3). Data often comes from simulations using certain initial conditions and inputs $u(t)$.
2. Apply POD to the data and extract a low order POD basis (see Section 3 for details).

3. Project the partial differential equation onto the POD basis to create a low order matrix differential equation of the form (4). Solve the low order matrix Lyapunov equation (5).

Some researchers have had success on similar problems with this approach, however the method can produce unpredictably bad results. In general, there are *no guarantees* of accuracy or convergence for model reduction or control problems. (There are error bounds and error estimating procedures for the *simulation* problem; see [Kunisch & Volkwein, 2001, 2002](#); [Homescu et al., 2005](#)).

We emphasize that the use of POD in this work is *fundamentally different* from standard POD methods as described above. In particular, we prove convergence for the snapshot algorithms.

This work is primarily concerned with the introduction and analysis of the snapshot Lyapunov algorithms at the infinite dimensional level. The remainder of this work proceeds as follows. We begin with a presentation of the algorithms and an outline of the convergence theory and error bounds. Section 3 gives an overview of the continuous proper orthogonal decomposition, which is used in the low rank algorithm. This is followed by proofs of the theoretical results in Section 4. The following section presents numerical results for two model problems; the convergence analysis is confirmed, and we also find that accurate time stepping of the differential equations is important for obtaining accurate results.

2 The Snapshot Algorithms

We now describe the snapshot algorithms and give an overview of the approximation theory.

The algorithms are applicable to the matrix case and a class of infinite dimensional problems. As mentioned in the introduction, we believe these algorithms have great potential for infinite dimensional problems. Therefore, we concentrate on this case in the present work.

Throughout this work, let X be a separable Hilbert space with inner product (\cdot, \cdot) and corresponding norm $\|\cdot\| = (\cdot, \cdot)^{1/2}$. For simplicity, we assume the inner product is real valued. For the matrix Lyapunov equation, X is taken to be \mathbb{R}^n and the inner product can be taken as the standard dot product, $(a, b) = a^T b$, or a weighted dot product, $(a, b) = a^T M b$, where $M \in \mathbb{R}^{n \times n}$ is symmetric positive definite.

We suppose A and B have the following properties. In the matrix case, $A \in \mathbb{R}^{n \times n}$ is exponentially stable and $B \in \mathbb{R}^{n \times m}$. In the infinite dimensional case, $A : D(A) \subset X \rightarrow X$ generates an exponentially stable C_0 -semigroup¹ e^{At} over X and $B : \mathbb{R}^m \rightarrow X$ is finite rank and bounded. The latter assumption implies B must take the form

$$Bu = \sum_{j=1}^m b_j u_j,$$

where each $b_j \in X$ and $u = [u_1, \dots, u_m]^T \in \mathbb{R}^m$ (see Theorem 6.1 in [Weidmann, 1980](#)). This representation for B also holds for the matrix problem; in this case, b_j is the j th column of B .

2.1 Key to the Algorithms

We now show that the solution of the Lyapunov equation takes a special form. This is the key to both of the algorithms.

¹For an introduction to semigroup theory with applications to infinite dimensional control and systems theory, see [Curtain & Zwart \(1995\)](#).

In the infinite dimensional case, the operator Lyapunov equation (1) is understood as follows: the operator $Z : X \rightarrow X$ is a solution if Z maps $D(A^*)$ into $D(A)$ and

$$AZx + ZA^*x + BB^*x = 0$$

for all $x \in D(A^*)$; see, e.g., [Curtain & Zwart \(1995, pages 160–161\)](#).

Proposition 1. *Under the above assumptions, the unique solution $Z : X \rightarrow X$ of the Lyapunov equation (1) is given by*

$$Zx = \int_0^\infty \sum_{j=1}^m (x, w_j(t)) w_j(t) dt, \quad (6)$$

where $w_j(t) = e^{At}b_j$ is the unique solution of the linear evolution equation

$$\dot{w}_j(t) = Aw_j(t), \quad w_j(0) = b_j, \quad (7)$$

for $j = 1, \dots, m$.

Remark 1. If b_j is not in $D(A)$, then $w_j(t) = e^{At}b_j$ is not necessarily a classical solution of the differential equation (7). However, $w_j(t)$ is the unique solution of (7) in a generalized or weak sense; see, e.g., [Curtain & Zwart \(1995, Example A.5.29\)](#) or [Pazy \(1983, page 105\)](#). Throughout this work, a solution of an infinite dimensional differential equation is always understood in a generalized or weak sense.

Proof. Given the assumptions above, the exact solution $Z : X \rightarrow X$ of the Lyapunov equation is given by [Curtain & Zwart \(1995, Theorem 4.1.23\)](#)

$$Zx = \int_0^\infty e^{At} BB^* e^{A^*t} x dt.$$

As is well known, the solution may be factored as $Z = \mathcal{B}\mathcal{B}^*$, where $\mathcal{B} : L^2(0, \infty; \mathbb{R}^m) \rightarrow X$ is defined by

$$\mathcal{B}u = \int_0^\infty e^{At} Bu(t) dt$$

and $\mathcal{B}^* : X \rightarrow L^2(0, \infty; \mathbb{R}^m)$, the adjoint of \mathcal{B} , is given by $\mathcal{B}^*x = B^*e^{A^*t}x$. Again, given the assumptions above on B , the operator must have the form $Bu = \sum_{j=1}^m b_j u_j$, where $u = [u_1, \dots, u_m]^T \in \mathbb{R}^m$, and each b_j is in X . Then we have

$$\mathcal{B}u = \int_0^\infty e^{At} Bu(t) dt = \int_0^\infty \sum_{j=1}^m u_j(t) w_j(t) dt,$$

where $w_j(t) = e^{At}b_j$ is the solution of the linear evolution equation (7) for $j = 1, \dots, m$. Since e^{At} is exponentially stable, there are constants $M \geq 1$ and $\omega > 0$ so that $\|e^{At}x\| \leq Me^{-\omega t}\|x\|$ for any $x \in X$; therefore, each w_j is in $L^2(0, \infty; X)$. The adjoint operator $\mathcal{B}^* : X \rightarrow L^2(0, \infty; \mathbb{R}^m)$ is easily computed to be

$$[\mathcal{B}^*x](t) = [(x, w_1(t)), \dots, (x, w_m(t))]^T.$$

Again using $Z = \mathcal{B}\mathcal{B}^*$ gives the expression (6). □

The importance of this result is that the solution of the Lyapunov equation is exactly equal to the *continuous POD operator* for the set of functions $\{w_j(t)\}$. As we discuss in more detail below, the eigenvalues and normalized eigenvectors of the integral operator (6) are the POD eigenvalues and modes of the dataset $\{w_j(t)\}$. Since the Lyapunov solution Z equals the continuous POD operator for $\{w_j(t)\}$, the POD eigenvalues and modes equal, by definition, the eigenvalues and orthonormal eigenvectors of Z . For the low rank POD-based algorithm, we approximate the POD eigenvalues and modes and construct an approximate truncated eigenvalue expansion of the Lyapunov solution Z . The other snapshot based algorithm directly uses the integral representation of the solution Z to approximate the product of Z with a few vectors in the Hilbert space without ever forming or storing an approximation to Z .

2.2 The Algorithms

We now describe the convergent algorithms to approximate the solution $Z : X \rightarrow X$ of the Lyapunov equation (1). We comment on the dual Lyapunov equation $A^*Z + ZA + C^*C = 0$ below.

We first summarize the POD-based algorithm. Again, we assume $Bu = \sum_{j=1}^m b_j u_j$. Throughout this work we use a superscript N on a quantity to denote an approximation of that quantity.

Algorithm 1 (POD-Based Low Rank Approximate Solution):

1. For $j = 1, \dots, m$, compute an approximation $w_j^N(t)$ to the solution $w_j(t) = e^{At}b_j$ of the linear differential equation

$$\dot{w}_j(t) = Aw_j(t), \quad w_j(0) = b_j. \quad (8)$$

2. Compute $\{\lambda_k^N\}$ and $\{\varphi_k^N\}$, the POD eigenvalues and modes of the dataset $\{w_j^N\}_{j=1}^m$, e.g., by the method of snapshots or by quadrature (see Section 3.2).
3. Choose r and form the r th order approximate Lyapunov solution $Z_r^N : X \rightarrow X$ given by

$$Z_r^N x = \sum_{k=1}^r \lambda_k^N (x, \varphi_k^N) \varphi_k^N, \quad (9)$$

where (\cdot, \cdot) is the inner product over the Hilbert space.

The choice of r and the accuracy of the approximations $w_j^N(t)$ can be guided by error bounds, which are presented in Section 2.3. As mentioned above, this algorithm was proposed for the matrix case by Willcox & Peraire (2002).

Remark 2. If desired, the approximate solution can be factored as $Z_r^N = R^*R$, where $R : X \rightarrow \mathbb{R}^r$ and its adjoint $R^* : \mathbb{R}^r \rightarrow X$ are defined by

$$Rx = [(x, \psi_1^N), \dots, (x, \psi_r^N)]^T, \quad R^*a = \sum_{k=1}^r a_k \psi_k^N,$$

where $\psi_k^N = (\lambda_k^N)^{1/2} \varphi_k^N$ for $k = 1, \dots, N$ and $a = [a_1, \dots, a_r]^T$.

We now describe the snapshot based algorithm to approximate the product of the solution $Z : X \rightarrow X$ of the Lyapunov equation (1) with a vector $x \in X$. The algorithm is directly based on the integral representation of the solution given in Proposition 1 above.

Algorithm 2 (Snapshot-Based Approximate Solution/Vector Product):

1. For $j = 1, \dots, m$, compute an approximation $w_j^N(t)$ to the solution $w_j(t) = e^{At}b_j$ of the linear differential equation (8).
2. Compute the approximate product

$$Z^N x = \int_0^\infty \sum_{j=1}^m (x, w_j^N(t)) w_j^N(t) dt, \quad (10)$$

where (\cdot, \cdot) is the inner product over the Hilbert space.

As mentioned above, this algorithm with quadrature approximations for the integral was proposed for the matrix case by Saad (1990).

The integral can be approximated directly by the method of snapshots (see Section 3) or by quadrature; thus, one only needs “time snapshots” of the approximate solutions $w_j^N(t)$ to compute the approximate product. Again, the accuracy of the approximations $w_j^N(t)$ can be guided by error bounds, which are presented in Sections 2.3.

In the snapshot algorithm, the approximate operator Z^N does not need to be stored to compute the product $Z^N x$. Moreover, if storing the approximations $w_j^N(t)$ is a problem, one can use a time stepping algorithm for the differential equations (8) and at each time step proceed as follows: compute the solution of the differential equation at the next time step, update the integral approximation, and discard any solution data not required for the next time step. We give an example of this procedure using the trapezoid rule for the time stepping in Section 5.2. Products of Z^N with multiple vectors can be computed in a similar manner.

The low rank and snapshot algorithms can be applied to the dual Lyapunov equation

$$A^* Z + Z A + C^* C = 0, \quad (11)$$

by interchanging the roles of A and A^* and B and C^* above. Specifically, one must now approximate the solutions of the dual linear evolution equations

$$\dot{z}_j(t) = A^* z_j(t), \quad z_j(0) = c_j. \quad (12)$$

In the matrix case, c_j is the j th column of the matrix C^T . In the infinite dimensional case, we assume $C : X \rightarrow \mathbb{R}^p$ is bounded and finite rank so that C must have the form $Cx = [(x, c_1), \dots, (x, c_p)]$, where each $c_j \in X$ (again, see Theorem 6.1 in Weidmann, 1980). The remainder of the algorithms remains unchanged.

2.3 An Overview of the Approximation Theory

We now give a brief overview of results on convergence and accuracy of the Lyapunov approximations. Details of the results and the proofs are contained in Section 4 below.

Theorem 1 shows that the trace norm error between the Lyapunov solution Z and the operator Z^N in the snapshot based algorithm (Algorithm 2) is bounded in the trace norm (see Section 4.1) as follows:

$$\|Z - Z^N\|_{\text{tr}} \leq C^N \left(\sum_{j=1}^m \int_0^\infty \|w_j(t) - w_j^N(t)\|^2 dt \right)^{1/2},$$

where C^N is an approximately computable constant. As each w_j^N converges to w_j , the constant C^N converges and the trace norm error tends to zero. Note that the error is due to the error in approximating the solutions w_j of the linear differential equations (8). Therefore, adaptive

algorithms or error estimation techniques can be used to guide refined computations. Again, the author is not aware of any other algorithm that has the potential for such adaptive computations.

Corollary 1 then gives a bound on the trace norm error between the Lyapunov solution Z and the POD-based low rank approximation Z_r^N of Algorithm 1:

$$\|Z - Z_r^N\|_{\text{tr}} \leq \sum_{k>r} \lambda_k^N + C^N \left(\sum_{j=1}^m \int_0^\infty \|w_j(t) - w_j^N(t)\|^2 dt \right)^{1/2}.$$

We note that the sum of the neglected approximate POD eigenvalues λ_k^N is computable. Therefore, if one can estimate the approximation error between each w_j and w_j^N , then the approximation error between the solution Z and the low rank approximation Z_r^N can be estimated. Again, adaptive computations are possible. Furthermore, as each w_j^N converges to w_j , the error converges to the sum of the neglected POD eigenvalues:

$$\lim_{N \rightarrow \infty} \|Z - Z_r^N\|_{\text{tr}} = \sum_{k>r} \lambda_k,$$

and *this is the best possible error for a rank r approximation.*

Many researchers have noticed that the eigenvalues of solutions of Lyapunov equations (i.e., the POD eigenvalues) often decay rapidly when B has low rank; therefore, low rank approximations to a Lyapunov solution can be very accurate. This is the basis of most recent algorithms for solving large-scale Lyapunov equations. In the matrix case, the recent works by [Antoulas *et al.* \(2002\)](#); [Grasedyck \(2004\)](#); [Penzl \(2000\)](#) give theoretical reasons why the solutions of Lyapunov equations allow accurate low rank approximations. For examples and counterexamples, see the recent works on large-scale Lyapunov solvers referenced in Section 1.

In Theorem 3, we give a different bound for the operator norm error between the Lyapunov solution Z and the low rank approximation Z_r^N of Algorithm 1:

$$\|Z - Z_r^N\| \leq \lambda_{r+1} + \sum_{k=1}^r (|\lambda_k - \lambda_k^N| + 2\lambda_k^N \|\varphi_k - \varphi_k^N\|),$$

where λ_k and φ_k are the k th POD eigenvalue and mode of the set of functions $\{w_j\}_{j=1}^m$. If the first r POD eigenvalues are distinct, then the approximate POD eigenvalues and modes converge (when suitably normalized, see Theorem 2) and the second term in the error bound tends to zero. Thus, the speed of convergence of the approximate Lyapunov solution naturally depends on the speed of convergence of the approximate POD eigenvalues and modes. Very often in practice, the dominant POD eigenvalues and modes converge quickly, and therefore we expect fast convergence. Furthermore, the overall approximation error converges to the first neglected POD eigenvalue:

$$\lim_{N \rightarrow \infty} \|Z - Z_r^N\| = \lambda_{r+1},$$

and again this is the best possible error for a rank r approximation to Z .

Remark 3. Due to the approximation theory for the POD eigenvalues in Theorem 2 below, λ_{r+1} can be approximated by λ_{r+1}^N . Thus, if the first $r+1$ POD eigenvalues and the first r POD modes have converged, then λ_{r+1}^N is a good approximation of the operator norm error bound between Z and Z_r^N . This gives a simple way to assess the accuracy of the approximate solution.

3 The Continuous Proper Orthogonal Decomposition

As discussed above, the key to the proposed algorithms is that the Lyapunov solution is exactly the *continuous POD operator* for the set of functions $\{w_j(t)\}_{j=1}^m$. In this section, we summarize the continuous proper orthogonal decomposition from the recent works of Kunisch and Volkwein (2002) and Henri and Yvon (2002a; 2002b; 2005). These works focus on the continuous POD for a finite time interval, however the theory extends naturally to the case of an infinite time interval. For completeness, we present proofs of the theorems below in Section 4.2.

Section 3.1 reviews properties of the continuous proper orthogonal decomposition and Section 3.2 focuses on approximating the POD eigenvalues and modes.

3.1 Continuous POD and its Properties

Let $L^2(0, \infty; X)$ be the set of all functions w such that $w(t) \in X$ for all $t \geq 0$ and whose X norm is square integrable, i.e.,

$$\|w\|_{L^2(0, \infty; X)} = \left(\int_0^\infty \|w(t)\|^2 dt \right)^{1/2} < \infty.$$

A sequence of functions $\{w^N\} \subset L^2(0, \infty; X)$ converges to $w \in L^2(0, \infty; X)$ if $\|w^N - w\|_{L^2(0, \infty; X)} \rightarrow 0$ as $N \rightarrow \infty$.

We now define the continuous proper orthogonal decomposition and discuss its properties.

Definition 1. *The continuous POD operator $Z : X \rightarrow X$ for a dataset $\{w_j\}_{j=1}^m \subset L^2(0, \infty; X)$ is defined by*

$$Zx = \int_0^\infty \sum_{j=1}^m (x, w_j(t)) w_j(t) dt. \quad (13)$$

The continuous POD operator is self adjoint, compact, and nonnegative; thus, the eigenvalues of Z may be ordered $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ (with repetitions according to multiplicity) and the corresponding orthonormal eigenvectors $\{\varphi_k\} \subset X$ form a complete set.

Definition 2. *The eigenvalues $\{\lambda_k\}$ of the continuous POD operator Z are called the POD eigenvalues of $\{w_j\}$ and the orthonormal eigenvectors $\{\varphi_k\} \subset X$ of Z are called the POD modes of $\{w_j\}$.*

The POD eigenvalues are an indication of “energy content” and the POD modes are optimal for data reconstruction. First, the “total energy” in the dataset is contained in the POD eigenvalues:

Proposition 2. *Let $\{w_j\}_{j=1}^m$ be a collection of functions in $L^2(0, \infty; X)$ and let $\{\lambda_k\}$ be the eigenvalues of the POD operator corresponding to the dataset $\{w_j\}$. Then*

$$\int_0^\infty \sum_{j=1}^m \|w_j(t)\|^2 dt = \sum_{k \geq 1} \lambda_k < \infty.$$

If the dataset is projected onto a subset of the POD modes, the POD eigenvalues provide the exact data reconstruction error.

Proposition 3. Let $\{w_j\}_{j=1}^m$ be a collection of functions in $L^2(0, \infty; X)$ with POD eigenvalues $\{\lambda_k\}$ and modes $\{\varphi_k\}$. Let w_j^r be the r th order projection of w_j onto the POD basis, i.e.,

$$w_j^r(t) = \sum_{k=1}^r (w_j(t), \varphi_k) \varphi_k. \quad (14)$$

Then the data reconstruction error is given in terms of the sum of the neglected POD eigenvalues:

$$\sum_{j=1}^m \int_0^\infty \|w_j(t) - w_j^r(t)\|^2 dt = \sum_{k>r} \lambda_k.$$

In the case of a finite time interval, no other orthonormal basis yields a smaller reconstruction error. This optimal reconstruction property extends to the case of an infinite time interval, however we do not prove this here as it is not required for the current work.

We note in passing that approximating each $w_j(t)$ in the POD operator (13) by the r th order projection $w_j^r(t)$ produces the approximate POD operator Z_r in (9) computed in the low rank POD algorithm.

3.2 Approximating the Continuous POD

An important feature of proper orthogonal decomposition is that the POD eigenvalues and modes of a time varying dataset $\{w_j\}_{j=1}^m \subset L^2(0, \infty; X)$ can be approximated by a variety of algorithms. Two popular approaches are the method of snapshots and quadrature. In the method of snapshots, the main idea is to approximate each w_j with functions whose POD eigenvalues and modes are easily computable. In the quadrature approach, the POD integral operator is approximated using quadrature leading to easily computable approximate POD eigenvalues and modes. We describe both methods below. Furthermore, there are algorithms to compute the POD of very large datasets (Beattie *et al.*, 2006; Fahl, 2001).

The method of snapshots and the quadrature approach are related since both approximate a continuous POD operator with a *discrete* POD operator. The eigenvalues and orthonormal eigenvectors of the discrete POD operator are then used as approximations of the continuous POD eigenvalues and modes. Below, we describe the method of snapshots and the quadrature approach and then present the eigendecomposition of a discrete POD operator.

We begin with the method of snapshots introduced by Sirovich (1987). A popular approach to the method of snapshots is to use piecewise constant functions (in time) to approximate the functions w_j . It is possible to generalize this algorithm if more variation in time is desired.

Approach 1. The Method of Snapshots:

1. For a collection of functions $\{w_j\}_{j=1}^m \subset L^2(a, b; X)$, compute approximate snapshots $a_{j,k} \approx w_j(t_{j,k})$ of $w_j(t)$ at times $a = t_{j,0} < t_{j,1} < \dots < t_{j,N_j} = b$.
2. For each j , compute $v_{j,k} = (a_{j,k} + a_{j,k-1})/2$ to approximate the average value of $w_j(t)$ over the k th time interval for $k = 1, \dots, N_j$.
3. Set $\delta_{j,k} = t_{j,k} - t_{j,k-1}$, the k th time step for $k = 1, \dots, N_j$.
4. Define piecewise constant approximations $w_j^N(t)$ to $w_j(t)$ by $w_j^N(t) = v_{j,k}$ for $t_{j,k-1} \leq t \leq t_{j,k}$. Compute an approximate POD operator by substituting $w_j^N(t)$ for $w_j(t)$ in the continuous

POD operator:

$$Zx = \int_a^b \sum_{j=1}^m (x, w_j(t)) w_j(t) dt \approx Z^N x = \sum_{j=1}^m \sum_{k=1}^{N_j} \delta_{j,k} (x, v_{j,k}) v_{j,k}.$$

5. “Stack” the data scaled by the square roots of the time steps to arrive at a discrete POD operator as follows. Let $V_{j,k} = \delta_{j,k}^{1/2} v_{j,k}$ and define

$$W = [V_{1,1}, \dots, V_{1,N_1}, \dots, V_{m,1}, \dots, V_{m,N_m}].$$

Then $Z^N : X \rightarrow X$ is given by $Z^N x = \sum_{k=1}^N (x, W_k) W_k$, where $N = \sum_{j=1}^m N_j$.

6. Compute the eigenvalues $\{\lambda_k^N\}$ and orthonormal eigenvectors $\{\varphi_k^N\}$ of the discrete POD operator Z^N (as described in Proposition 4 below) to obtain approximations of the POD eigenvalues and modes for $\{w_j\}$.

We note that this algorithm is often implemented using an equally spaced time grid.

Next, we describe the related quadrature approach for approximating POD eigenvalues and modes of a time varying dataset.

Approach 2. Quadrature:

1. For a collection of functions $\{w_j\}_{j=1}^m \subset L^2(a, b; X)$, compute approximate snapshots $a_{j,k} \approx w_j(t_{j,k})$ of $w_j(t)$ at times $a = t_{j,0} < t_{j,1} < \dots < t_{j,N_j} = b$.
2. Use quadrature schemes to approximate the continuous POD operator:

$$Zx = \int_a^b \sum_{j=1}^m (x, w_j(t)) w_j(t) dt \approx Z^N x = \sum_{j=1}^m \sum_{k=1}^{N_j} \delta_{j,k} (x, a_{j,k}) a_{j,k}.$$

3. “Stack” the data scaled by the square roots of the quadrature weights to arrive at a discrete POD operator as follows. Let $V_{j,k} = \delta_{j,k}^{1/2} a_{j,k}$ and define

$$W = [V_{1,1}, \dots, V_{1,N_1}, \dots, V_{m,1}, \dots, V_{m,N_m}].$$

Then $Z^N : X \rightarrow X$ is given by $Z^N x = \sum_{k=1}^N (x, W_k) W_k$, where $N = \sum_{j=1}^m N_j$.

4. Compute the eigenvalues $\{\lambda_k^N\}$ and orthonormal eigenvectors $\{\varphi_k^N\}$ of the discrete POD operator Z^N (as described in Proposition 4 below) to obtain approximations of the POD eigenvalues and modes for $\{w_j\}$.

As is well known, the eigenvalues and eigenvectors of a discrete POD operator can be computed by solving a matrix eigenvalue problem. We provide a proof in Appendix A for completeness.

Proposition 4. *Let $\{W_k\}_{k=1}^N$ be a finite collection of elements in a Hilbert space X and let the $N \times N$ matrix Γ have ij entries (W_i, W_j) . Then the discrete POD operator $Z : X \rightarrow X$ defined by*

$$Zx = \sum_{k=1}^N (x, W_k) W_k,$$

is compact, self adjoint, and nonnegative. The nonzero eigenvalues of Z are equal to the nonzero eigenvalues of Γ and they may be ordered $\lambda_1 \geq \lambda_2 \geq \dots > 0$. If $\lambda_i \neq 0$, the corresponding orthonormal eigenvector φ_i of Z is given by

$$\varphi_i = \lambda_i^{-1/2} \sum_{j=1}^N (\gamma_i)_j W_j, \quad (15)$$

where $(\gamma_i)_j$ is the j th element of the i th orthonormal eigenvector of Γ .

3.2.1 Approximating the Continuous POD: Discretization

In the approximation procedures described above to compute the continuous POD of a time varying dataset, the data lies in an infinite dimensional Hilbert space. In order to perform the computations, a finite dimensional approximation must be performed somewhere in the algorithms. Below, we briefly describe one finite dimensional computational procedure for approximating the continuous POD of a dataset.

In many applications, one wishes to compute the POD of a dataset arising from a finite dimensional discretization of a partial differential equation or some other infinite dimensional system. Therefore, in the method of snapshots or the quadrature approach described above, the time snapshots of the data are often expressed as a Galerkin expansion. We describe the computation of the eigenvalues and eigenvectors of a discrete POD operator when the data takes the form

$$W_j = \sum_{k=1}^n d_{j,k} \Phi_k, \quad (16)$$

where each $d_{j,k}$ is a real number and each Φ_k is an element of the Hilbert space. This is not the most general expression possible, however this form leads to a simple algorithm.

Recall from above that in order to compute the eigenvalues and eigenvectors of a discrete POD operator, we need only compute the eigenvalues and eigenvectors of the matrix Γ with ij entries $\Gamma_{ij} = (W_i, W_j)$. Substituting the above Galerkin expansion in for each W_i gives that the matrix Γ is given by

$$\Gamma = D^T M D, \quad (17)$$

where the $n \times n$ ‘‘mass’’ matrix M and the $n \times N$ matrix D have ij entries

$$M_{ij} = (\Phi_i, \Phi_j), \quad D_{ij} = d_{j,i}.$$

As long as the above matrix product is not too expensive to compute, this leads to a simple algorithm to compute the discrete POD eigenvalues and modes.

Algorithm. Discrete POD with Data in Galerkin Expansions:

1. Given a collection of data $\{W_j\}_{j=1}^N \subset X$ of the form (16), compute the $n \times n$ matrix Γ defined above in (17).
2. Compute the eigenvalues $\{\lambda_k\}$ and orthonormal eigenvectors $\{\gamma_k\}$ of Γ .
3. Then the discrete POD eigenvalues are given by $\{\lambda_i\}_{i=1}^N$ and the orthonormal discrete POD modes $\{\varphi_i\} \subset X$ corresponding to nonzero eigenvalues are given by

$$\varphi_i = \lambda_i^{-1/2} \sum_{k=1}^n P_{ki} \Phi_k,$$

where the $n \times N$ matrix P is given by $P = D\Gamma_{\text{ev}}$, and the i th column of the $N \times N$ matrix Γ_{ev} is the i th orthonormal eigenvector of Γ .

In many applications, one only needs to compute a relatively small number of POD modes. In this case, the entire $n \times N$ matrix P does not need to be formed. Specifically, in order to compute the first r modes, first form the $N \times r$ matrix $\Gamma_{\text{ev},r}$, whose r columns are the first r orthonormal eigenvectors of Γ . Then to compute the first r POD modes (i.e., φ_i for $i = 1, \dots, r$), replace P with $P_r = D\Gamma_{\text{ev},r}$ in the above expression for each φ_i .

We also note that it may be beneficial to use special algorithms to compute the eigenvalues of $\Gamma = D^TMD$ without ever explicitly forming the product (see [Watkins, 2005](#)).

4 Approximation Theory and Error Bounds

We now prove the main results. We first review some background material and provide proofs for the continuous POD theory in Section 3.

4.1 Notation and Background

In order to discuss the properties of the approximate Lyapunov solution, we first introduce some notation and background material.

Let K be a compact linear operator from a Hilbert space X to a Hilbert space Y . The operator norm of K is given by $\|K\| = \sup \|Kx\|$, where $x \in X$ has unit norm. The Hilbert-Schmidt (HS) norm of K is given by

$$\|K\|_{HS} = \left(\sum_{j \geq 1} \|K\varphi_j\|^2 \right)^{1/2}$$

for any orthonormal basis $\{\varphi_j\} \subset X$. If $K : X \rightarrow X$ is self adjoint and nonnegative, then the stronger trace (or nuclear) norm is given by

$$\|K\|_{\text{tr}} = \sum_{j \geq 1} (K\varphi_j, \varphi_j),$$

for any orthonormal bases $\{\varphi_j\} \subset X$.

A compact operator K is called Hilbert-Schmidt if the HS norm of K is finite and K is called trace class (or nuclear) if the trace norm of K is finite. If two operators $K : X \rightarrow Y$ and $L : Y \rightarrow Z$ are HS, then the product $KL : X \rightarrow Z$ is trace class and $\|KL\|_{\text{tr}} \leq \|K\|_{HS}\|L\|_{HS}$. Also, $\|K\| = \|K^*\|$ for any of the above norms and $\|K\| \leq \|K\|_{HS} \leq \|K\|_{\text{tr}}$.

Now let $K : X \rightarrow X$ be compact, self-adjoint, and nonnegative. The eigenvalues of such an operator can be ordered $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ and the corresponding eigenvectors $\{\varphi_k\}$ form a complete orthonormal set. A best rank r approximation to K is given by a solution of the following problem: find the minimizer over all rank r operators F_r of the operator norm error $\|K - F_r\|$. A solution of this problem (which may not be unique) is given by the r th order truncated eigenvalue decomposition of K defined by

$$K_r x = \sum_{k=1}^r \lambda_k (x, \varphi_k) \varphi_k$$

The best value of the operator norm error $\|K - K_r\|$ is equal to λ_{r+1} , the first neglected eigenvalue. The truncated eigenvalue decomposition also gives a best rank r approximation of K in the trace norm. In this case, the best trace norm error is given by $\sum_{k>r} \lambda_k$, the sum of the neglected eigenvalues.

4.2 Continuous POD and its Properties

As mentioned above, we give proofs for the continuous POD results on an infinite time interval in Section 3 as they are crucial for the main results. Many of the proofs of the continuous POD results follow the theory for the finite time interval from the recent works of Kunisch and Volkwein (2002) and Henri and Yvon (2002b; 2002a; 2005).

Let $\{w_j\}_{j=1}^m$ be an arbitrary dataset in $L^2(0, \infty; X)$. The continuous POD operator $Z : X \rightarrow X$ for this dataset is given in Definition 1 in Section 3. The POD eigenvalues and modes are the eigenvalues and orthonormal eigenvectors of the continuous POD operator.

A fundamental property of the POD operator is that it can be factored as $Z = \mathcal{B}\mathcal{B}^*$ as follows.

Definition 3. For $\{w_j\}_{j=1}^m \subset L^2(0, \infty; X)$, define the bounded linear operator $\mathcal{B} : L^2(0, \infty; \mathbb{R}^m) \rightarrow X$ by

$$\mathcal{B}u = \int_0^\infty \sum_{j=1}^m u_j(t)w_j(t) dt.$$

It is now straightforward to check the factorization by direct computation.

Proposition 5. The adjoint operator $\mathcal{B}^* : X \rightarrow L^2(0, \infty; \mathbb{R}^m)$ is given by

$$[\mathcal{B}^*x](t) = [(x, w_1(t)), \dots, (x, w_m(t))]^T.$$

and therefore $Z = \mathcal{B}\mathcal{B}^*$.

The factorization allows us to obtain many properties of the POD operator. First, the factorization directly gives that the POD operator is self adjoint and nonnegative. Next, we show the POD operator is trace class.

Proposition 6. For $\{w_j\}_{j=1}^m \subset L^2(0, \infty; X)$, the operators \mathcal{B} and \mathcal{B}^* are Hilbert-Schmidt and therefore $Z = \mathcal{B}\mathcal{B}^*$ is trace class.

Proof. Our proof follows an argument in (Curtain & Sasane, 2001, Theorem 4).

For $i = 1, \dots, m$, define $L_i : X \rightarrow L^2(0, \infty)$ by $[L_i x](t) = (x, w_i(t))$. We have $|[L_i x](t)| \leq \|w_i(t)\| \|x\|$ and $\int_0^\infty \|w_i(t)\|^2 dt < \infty$ since $w_i \in L^2(0, \infty; X)$. Theorem 5 in Curtain & Sasane (2001) (which is a modification of Theorem 6.12, page 140, in Weidmann (1980)) shows each L_i is Hilbert-Schmidt.

The operator \mathcal{B}^* is given by

$$[\mathcal{B}^*x](t) = [L_1 x(t), \dots, L_m x(t)]^T.$$

Let $\{x_j\}$ be any orthonormal basis for X . Since each L_i is Hilbert-Schmidt, $\sum_{j \geq 1} \|L_i x_j\|_{L^2(0, \infty)}^2 < \infty$ for each i . Then

$$\sum_{j \geq 1} \|\mathcal{B}^* x_j\|_{L^2(0, \infty; \mathbb{R}^m)}^2 = \sum_{j \geq 1} \left(\sum_{i=1}^m \|L_i x_j\|_{L^2(0, \infty)}^2 \right) < \infty$$

by above. Therefore \mathcal{B}^* is Hilbert-Schmidt and the result follows. \square

The POD operator Z is compact since it is trace class. Since Z is also self adjoint, the eigenvalues of Z may be ordered $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ (with repetitions according to multiplicity) and the corresponding orthonormal eigenvectors $\{\varphi_k\} \subset X$ form a complete set.

We now prove that the “energy” in the dataset is given by the sum of the POD eigenvalues, i.e.,

$$\int_0^\infty \sum_{j=1}^m \|w_j(t)\|^2 dt = \sum_{k \geq 1} \lambda_k < \infty.$$

Proof of Proposition 2:

Proof. Let Z be the POD operator corresponding to $\{w_j\}$. We know the sum of the eigenvalues of Z is finite since Z is trace class. Also,

$$\sum_{k \geq 1} \lambda_k = \sum_{k \geq 1} (Z\varphi_k, \varphi_k) = \sum_{k \geq 1} \int_0^\infty \sum_{j=1}^m |(\varphi_k, w_j(t))|^2 dt.$$

Now expand each $w_j(t)$ in terms of the orthonormal basis $\{\varphi_k\}$ for X :

$$w_j(t) = \sum_{k \geq 1} (w_j(t), \varphi_k) \varphi_k.$$

Then

$$\int_0^\infty \sum_{j=1}^m \|w_j(t)\|^2 dt = \sum_{k \geq 1} \int_0^\infty \sum_{j=1}^m |(\varphi_k, w_j(t))|^2 dt$$

and the result follows from the computation above. \square

Next, we prove that if the dataset is projected onto a subset of the POD basis, then the reconstruction error is given by the sum of the neglected POD eigenvalues.

Proof of Proposition 3:

Proof. The proof is a direct computation using the definition (14) of w_j^r and the orthonormality of the POD basis.

$$\begin{aligned} \sum_{j=1}^m \int_0^\infty \|w_j(t) - w_j^r(t)\|^2 dt &= \sum_{j=1}^m \int_0^\infty \left\{ \|w_j\|^2 - 2(w_j, w_j^r) + (w_j^r, w_j^r) \right\} dt \\ &= \sum_{j=1}^m \int_0^\infty \left\{ \|w_j\|^2 - 2 \sum_{k=1}^r |(w_j, \varphi_k)|^2 + \sum_{k=1}^r |(w_j, \varphi_k)|^2 \right\} dt \\ &= \sum_{j=1}^m \left\{ \int_0^\infty \|w_j\|^2 dt - \int_0^\infty \sum_{k=1}^r |(w_j, \varphi_k)|^2 dt \right\} \\ &= \int_0^\infty \sum_{j=1}^m \|w_j\|^2 dt - \sum_{k=1}^r (Z\varphi_k, \varphi_k) \\ &= \sum_{k \geq 1} \lambda_k - \sum_{k=1}^r \lambda_k = \sum_{k > r} \lambda_k. \end{aligned}$$

Here we used the above proposition and the fact that $Z\varphi_k = \lambda_k\varphi_k$. \square

4.3 Main Results

We now give the main convergence results.

Throughout this section, w_j and w_j^N can be any functions in $L^2(0, \infty; X)$ for $j = 1, \dots, m$. We consider the POD operators $Z : X \rightarrow X$ for $\{w_j\}_{j=1}^m$ and $Z^N : X \rightarrow X$ for $\{w_j^N\}_{j=1}^m$. We let $\{\lambda_k, \varphi_k\}$ and $\{\lambda_k^N, \varphi_k^N\}$ be the eigenvalues and orthonormal eigenvectors of Z and Z^N , respectively. These are also the POD eigenvalues and modes of the datasets $\{w_j(t)\}$ and $\{w_j^N(t)\}$. Define $Z_r : X \rightarrow X$ and $Z_r^N : X \rightarrow X$ to be the r th order truncated eigenvalue expansions of Z and Z^N , namely

$$Z_r x = \sum_{k=1}^r \lambda_k(x, \varphi_k) \varphi_k, \quad Z_r^N x = \sum_{k=1}^r \lambda_k^N(x, \varphi_k^N) \varphi_k^N.$$

Also, for the functions $\{w_j^N\}_{j=1}^m$, define $\mathcal{B}^N : L^2(0, \infty; \mathbb{R}^m) \rightarrow X$ analogously to the operator \mathcal{B} defined above (Definition 3) with the functions $\{w_j\}_{j=1}^m$.

Below, we study how well Z^N and Z_r^N approximate Z .

The functions w_j and w_j^N are arbitrary in $L^2(0, \infty; X)$, however we want to think of them in terms of approximating the solution to the Lyapunov equation (1). Recall from Section 2 that the Lyapunov solution is precisely the POD operator Z defined above for $\{w_j\}_{j=1}^m$ the solutions of the linear infinite dimensional differential equations (8). In the algorithms to approximate the Lyapunov solution, the functions $\{w_j^N\}_{j=1}^m$ are approximations to the solutions $\{w_j\}_{j=1}^m$.

We begin with some preliminary lemmas.

Lemma 1. *For any finite collection $\{w_j\}_{j=1}^m$ of functions in $L^2(0, \infty; X)$, the Hilbert-Schmidt norm of \mathcal{B} is given by*

$$\|\mathcal{B}\|_{HS} = \left(\sum_{k \geq 1} \lambda_k \right)^{1/2} = \left(\int_0^\infty \sum_{j=1}^m \|w_j(t)\|^2 dt \right)^{1/2},$$

where $\lambda_1 \geq \lambda_2 \geq \dots \geq 0$ are the eigenvalues of $Z = \mathcal{B}\mathcal{B}^*$.

Proof. We proved that \mathcal{B} is Hilbert-Schmidt in Proposition 6. The Hilbert-Schmidt norm of \mathcal{B} can be computed as follows. Let $\{\lambda_k, \varphi_k\}$ be the eigenvalues and orthonormal eigenvectors of $Z = \mathcal{B}\mathcal{B}^*$. Then

$$\|\mathcal{B}\|_{HS}^2 = \|\mathcal{B}^*\|_{HS}^2 = \sum_{k \geq 1} (\mathcal{B}^* \varphi_k, \mathcal{B}^* \varphi_k) = \sum_{k \geq 1} (\varphi_k, Z \varphi_k) = \sum_{k \geq 1} \lambda_k.$$

Continuous POD theory (see Proposition 2) gives

$$\sum_{k \geq 1} \lambda_k = \int_0^\infty \sum_{j=1}^m \|w_j(t)\|^2 dt$$

and the result follows. \square

The following lemma is a known result in continuous POD theory; we present a proof for completeness.

Lemma 2. *Let $\{w_j\}_{j=1}^m$ and $\{w_j^N\}_{j=1}^m$ be two collections of functions in $L^2(0, \infty; X)$ with POD eigenvalues $\{\lambda_k\}$ and $\{\lambda_k^N\}$, respectively. If each $w_j^N \rightarrow w_j$ in $L^2(0, \infty; X)$ as $N \rightarrow \infty$, then the sum of the POD eigenvalues converge:*

$$\lim_{N \rightarrow \infty} \sum_{k \geq 1} \lambda_k^N = \sum_{k \geq 1} \lambda_k.$$

Proof. We know from continuous POD theory (Proposition 2) that

$$\sum_{k \geq 1} \lambda_k = \sum_{j=1}^m \|w_j\|_{L^2(0, \infty; X)}^2, \quad \sum_{k \geq 1} \lambda_k^N = \sum_{j=1}^m \|w_j^N\|_{L^2(0, \infty; X)}^2.$$

Since $w_j^N \rightarrow w_j$ in $L^2(0, \infty; X)$ for each j , we know $\|w_j^N\|_{L^2(0, \infty; X)}^2 \rightarrow \|w_j\|_{L^2(0, \infty; X)}^2$ as $N \rightarrow \infty$. This gives the result. \square

We now prove that Z^N converges to the Lyapunov solution Z in the trace norm as each w_j^N converges to w_j in $L^2(0, \infty; X)$. We also give a bound on the trace norm error involving the error between each w_j and w_j^N .

Theorem 1. *The trace norm error between Z and Z^N can be bounded as follows:*

$$\|Z - Z^N\|_{\text{tr}} \leq C^N \left(\sum_{j=1}^m \int_0^\infty \|w_j(t) - w_j^N(t)\|^2 dt \right)^{1/2}, \quad (18)$$

where the constant C^N is given by

$$C^N = \left(\sum_{k \geq 1} \lambda_k \right)^{1/2} + \left(\sum_{k \geq 1} \lambda_k^N \right)^{1/2} = \left(\sum_{j=1}^m \int_0^\infty \|w_j(t)\|^2 dt \right)^{1/2} + \left(\sum_{j=1}^m \int_0^\infty \|w_j^N(t)\|^2 dt \right)^{1/2}. \quad (19)$$

If $w_j^N \rightarrow w_j$ in $L^2(0, \infty; X)$ for each j , then C^N converges to the constant C given by

$$C = 2 \left(\sum_{k \geq 1} \lambda_k \right)^{1/2} = 2 \left(\sum_{j=1}^m \int_0^\infty \|w_j(t)\|^2 dt \right)^{1/2},$$

and therefore $\|Z - Z^N\|_{\text{tr}} \rightarrow 0$ as $N \rightarrow \infty$.

Proof. First, since $\{w_j - w_j^N\}_{j=1}^m$ is a finite collection of functions in $L^2(0, \infty; X)$ and

$$(\mathcal{B} - \mathcal{B}^N)u = \int_0^\infty \sum_{j=1}^m u_j(t) (w_j(t) - w_j^N(t)) dt,$$

Lemma 1 above gives that $\mathcal{B} - \mathcal{B}^N$ is Hilbert-Schmidt and

$$\|\mathcal{B} - \mathcal{B}^N\|_{HS} = \left(\int_0^\infty \sum_{j=1}^m \|w_j(t) - w_j^N(t)\|^2 dt \right)^{1/2}.$$

Next,

$$\|Z - Z^N\|_{\text{tr}} \leq \|Z - \mathcal{B}\mathcal{B}^{N*}\|_{\text{tr}} + \|\mathcal{B}\mathcal{B}^{N*} - Z^N\|_{\text{tr}}.$$

Factor $Z = \mathcal{B}\mathcal{B}^*$ and similarly for Z^N . Then

$$\begin{aligned} \|Z - Z^N\|_{\text{tr}} &\leq \|\mathcal{B}\|_{HS} \|\mathcal{B} - \mathcal{B}^N\|_{HS} + \|\mathcal{B} - \mathcal{B}^N\|_{HS} \|\mathcal{B}^N\|_{HS} \\ &= \left(\|\mathcal{B}\|_{HS} + \|\mathcal{B}^N\|_{HS} \right) \|\mathcal{B} - \mathcal{B}^N\|_{HS} \end{aligned}$$

and the error bound follows from the observation above and Lemma 1

If each w_j^N converges to w_j in $L^2(0, \infty; X)$, then the sequence C^N converges to the constant C defined above due Lemma 2 above. Since C^N converges as $w_j^N \rightarrow w_j$, the right hand side of the error bound tends to zero and therefore $\|Z - Z^N\|_{\text{tr}} \rightarrow 0$ as $N \rightarrow \infty$. \square

Corollary 1. *The trace norm error between Z and Z_r^N can be bounded as follows:*

$$\|Z - Z_r^N\|_{\text{tr}} \leq \sum_{k>r} \lambda_k^N + C^N \left(\sum_{j=1}^m \int_0^\infty \|w_j(t) - w_j^N(t)\|^2 dt \right)^{1/2}, \quad (20)$$

where the constant C^N is defined in (19). If $w_j^N \rightarrow w_j$ in $L^2(0, \infty; X)$ for each j , then

$$\lim_{N \rightarrow \infty} \|Z - Z_r^N\| = \sum_{k>r} \lambda_k,$$

which is the best possible error for a rank r approximation to Z .

Proof. We have

$$\|Z - Z_r^N\|_{\text{tr}} \leq \|Z - Z^N\|_{\text{tr}} + \|Z^N - Z_r^N\|_{\text{tr}}.$$

Since Z_r^N is the truncated eigenvalue expansion of Z^N , which is self adjoint and nonnegative, Z_r^N is the best rank r approximation to Z^N with trace norm error equal to the neglected eigenvalues:

$$\|Z^N - Z_r^N\|_{\text{tr}} = \sum_{k>r} \lambda_k^N.$$

The result now follows from the above theorem. \square

Next we give a different expression for the operator norm error between the Lyapunov solution Z and the low rank approximation Z_r^N . This error bound depends on the convergence of the approximate POD eigenvalues and modes, which we now describe. First, if the approximate data w_j^N converges to the true data w_j in $L^2(0, \infty; X)$ for each j , then the POD eigenvalues will converge and the POD modes corresponding to distinct POD eigenvalues will converge. If a POD eigenvalue is repeated, however, we are only guaranteed that a subsequence of the approximate POD modes will converge. Again, this is a known result in continuous POD theory and we provide a proof for completeness.

Theorem 2. *Let $\{w_j\}_{j=1}^m$ and $\{w_j^N\}_{j=1}^m$ be two collections of functions in $L^2(0, \infty; X)$ with POD eigenvalues and modes denoted by $\{\lambda_k, \varphi_k\}$ and $\{\lambda_k^N, \varphi_k^N\}$, respectively. If $w_j^N \rightarrow w_j$ in $L^2(0, \infty; X)$ as $N \rightarrow \infty$, then the following statements are true:*

1. *The individual POD eigenvalues converge as $N \rightarrow \infty$, i.e., for each k ,*

$$\lim_{N \rightarrow \infty} |\lambda_k^N - \lambda_k| = 0.$$

2. *If the k th POD eigenvalue is distinct, then the k th POD mode (suitably normalized) converges, i.e.,*

$$\lim_{N \rightarrow \infty} \|\varphi_k^N - \varphi_k\| = 0.$$

3. *If the k th POD eigenvalue is not distinct, then a subsequence of the k th POD mode (suitably normalized) converges, i.e., there is a subsequence $\{N_j\}$ so that*

$$\lim_{N_j \rightarrow \infty} \|\varphi_k^{N_j} - \varphi_k\| = 0.$$

Proof. Let Z and Z^N denote the POD operators for the datasets $\{w_j\}$ and $\{w_j^N\}$, respectively. By definition, the POD eigenvalues and modes for $\{w_j\}$ and $\{w_j^N\}$ are the eigenvalues and orthonormal eigenvectors of the POD operator corresponding to the data.

As each $w_j^N \rightarrow w_j$ in $L^2(0, \infty; X)$, Theorem 1 above gives that Z^N converges to Z in the trace norm and therefore also in the (weaker) operator norm. Since Z^N and Z are compact, the conclusions follow directly from the norm convergence and eigenvalue and eigenvector approximation theory; see, e.g., Chatelin (1983); Ahues *et al.* (2001). \square

We show the speed of convergence of the approximate Lyapunov solution is governed by the speed of convergence of the POD eigenvalues and modes.

Theorem 3. *The operator norm error between Z_r^N and Z is bounded as follows:*

$$\|Z - Z_r^N\| \leq \lambda_{r+1} + \sum_{k=1}^r (|\lambda_k - \lambda_k^N| + 2\lambda_k^N \|\varphi_k - \varphi_k^N\|).$$

If the first r eigenvalues of Z are distinct, then Z_r^N converges to Z_r , the r th order truncated eigenvalue expansion of Z , in the operator norm and therefore

$$\lim_{N \rightarrow \infty} \|Z - Z_r^N\| = \lambda_{r+1},$$

which is the best possible error for a rank r approximation to Z .

Proof. First,

$$\|Z - Z_r^N\| \leq \|Z - Z_r\| + \|Z_r - Z_r^N\|,$$

where $Z_r x = \sum_{k=1}^r \lambda_k(x, \varphi_k) \varphi_k$ is the truncated eigenvalue expansion of Z . The eigenvalue expansion truncation operator norm error is given by the first neglected eigenvalue, i.e.,

$$\|Z - Z_r\| = \lambda_{r+1}.$$

For the second term in the error bound, note

$$\|(Z_r - Z_r^N)x\| \leq \sum_{k=1}^r \|\lambda_k(x, \varphi_k) \varphi_k - \lambda_k^N(x, \varphi_k^N) \varphi_k^N\|.$$

Next, add and subtract both $\lambda_k^N(x, \varphi_k) \varphi_k$ and $\lambda_k^N(x, \varphi_k^N) \varphi_k$ inside of the norm. Using the triangle inequality, the Cauchy-Schwartz inequality, and the orthonormality of each POD basis gives the result.

If the first r eigenvalues of Z are distinct, then by the approximation theory for the POD eigenvalues and modes (Theorem 2) we have $\lambda_k^N \rightarrow \lambda_k$ and $\varphi_k^N \rightarrow \varphi_k$ for $k = 1, \dots, r$. Thus, the above error bound for $\|Z_r - Z_r^N\|$ goes to zero as $N \rightarrow \infty$. This implies Z_r^N converges to Z_r in the operator norm. \square

5 Numerical Results

In this section, we present numerical results for two infinite dimensional model problems. We begin with a simple problem derived from a one dimensional convection diffusion equation so that we may compare the matrix approximation approach with the low rank algorithm. The second model problem comes from a two dimensional convection diffusion equation. We leave experiments on other problem types and comparisons with existing large-scale matrix Lyapunov algorithms for another work.

5.1 Model Problem 1

We take the A and B operators from the one dimensional convection diffusion equation

$$\begin{aligned} w_t(t, x) &= \mu w_{xx}(t, x) - \kappa w_x(t, x) + b(x)u(t), \\ w(t, 0) &= 0, \quad w(t, 1) = 0, \quad w(0, x) = w_0(x), \end{aligned}$$

where subscript denote partial derivatives, μ is a positive constant, and κ is a real constant. The function $b(x)$ is square integrable.

Let the Hilbert space X equal $L^2(0, 1)$, the space of square integrable functions, with the standard inner product $(f, g) = \int_0^1 f(x)g(x) dx$. The A operator is defined by

$$Aw = \mu w_{xx} - \kappa w_x, \quad D(A) = H^2 \cap H_0^1,$$

and B is given by $[Bu](x) = b(x)u$. Here, H^m is the standard Sobolev space of functions with m derivatives all of which are square integrable; also, any function $w \in H_0^1$ must satisfy the Dirichlet boundary conditions $w(0) = 0$ and $w(1) = 0$.

The eigenvalues of the convection diffusion operator A are given by $\lambda_n = -\mu n^2 \pi^2 - \kappa^2/4\mu$. Since the eigenvalues are all negative and bounded away from the imaginary axis, the results in [Delattre et al. \(2003\)](#) and [Curtain & Zwart \(1995, Section 2.3\)](#) can be used to show that A generates an exponentially stable C_0 -semigroup.

5.1.1 Numerical Results

We now compare the numerical results of the POD-based algorithm with matrix Lyapunov computations using matrix approximations of the A and B operators.

For the computations, we chose $b(x) = 5(1-x)^2 \sin(\pi x)$, $\mu = 0.1$, and $\kappa = 1$. Standard piecewise linear finite elements were used for the spatial discretization of the partial differential equation (8). The discretized equations were integrated over $0 \leq t \leq 2$ using Matlab's `ode15s` solver with default error tolerances; at $t = 2$, the numerical solution is nearly zero. The time points returned from `ode15s` were used in the method of snapshots to approximate the POD eigenvalues and modes.

Standard piecewise linear finite elements were also used to provide the matrix approximations of the A and B operators for the matrix Lyapunov computations. Matlab's `lyap` function was used to solve the resulting matrix Lyapunov equations.

Figure 1 shows the POD eigenvalues computed by the method of snapshots for $N = 32, 64, 128$, and 256 equally spaced finite element nodes. Eigenvalue computations for the matrix Lyapunov solution using the standard matrix approximations produced similar results. The larger POD eigenvalues have converged at this level of refinement; the POD eigenvalues nearer to machine precision (10^{-16}) have not yet converged. Further refinement is unnecessary since only the larger POD eigenvalues are used to construct the approximate Lyapunov solution.

Figure 2 shows the first POD mode computed by the method of snapshots for $N = 32$ equally spaced finite element nodes. The mode has converged at this level of refinement. The other POD modes converged in a similar fashion, however the higher numbered modes were slower to converge under refinement. This behavior is likely due to the fact that the higher numbered modes tend to oscillate more than the lower numbered modes. Eigenvector computations for the matrix Lyapunov solution using the standard matrix approximations produced similar results.

Figure 3 shows approximate Lyapunov solutions acting on $f = \exp(x)$. POD-based approximations are shown with $N = 32$ equally spaced finite element nodes with orders $r = 1$ and $r = 2$. The matrix Lyapunov computations using the standard matrix approximations is shown with $N = 256$

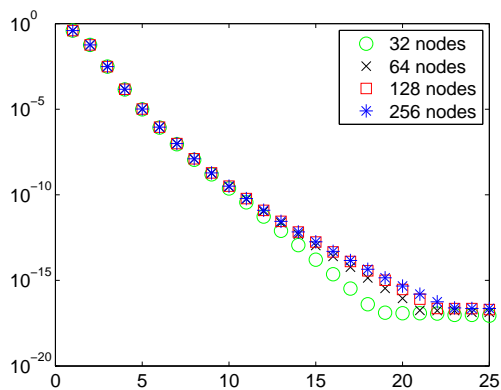


Figure 1: POD eigenvalues computed using $N = 32, 64, 128,$ and 256 equally spaced finite element nodes.

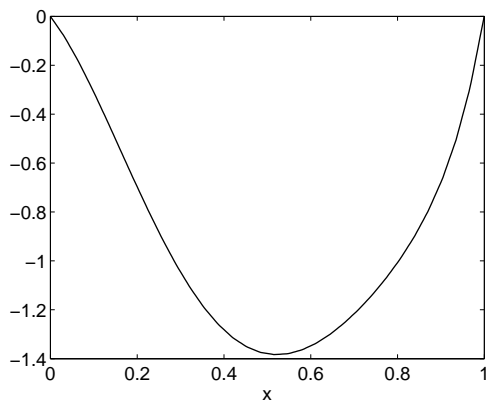


Figure 2: The first POD mode computed using $N = 32$ equally spaced finite element nodes.

equally spaced finite element nodes for comparison. The low order POD-based approximations give excellent agreement with the refined standard matrix approximation computations. In particular, for $r = 2$ the POD approximation is indistinguishable from the result of the standard computation.

The operator norm error bound in Theorem 3 gives a good indication of the accuracy of the POD-based approximation without comparison to other computations. Recall $\|(Z - Z_r^N)f\| \leq \|Z - Z_r^N\| \|f\|$. As discussed in Remark 3, we approximate $\|Z - Z_r^N\|$ by λ_{r+1}^N . For $f(x) = \exp(x)$, $\|f\| \approx 1.7873$. For $r = 1$, $\|Z - Z_r^N\| \approx 0.0569$; for $r = 2$, $\|Z - Z_r^N\| \approx 0.0031$. These values give approximate error bounds for $\|(Z - Z_r^N)f\|$ of 0.1016 for $r = 1$ and 0.0055 and $r = 2$. The above computations agree with these approximate error bounds.

We also look at the trace norm error bound in Corollary 1. For $r = 1$, the sum of the neglected eigenvalues is approximately 0.0601; for $r = 2$, this sum is approximately 0.0032. These values also give a good estimate of the approximation error. Of course, the full error bound involves the $L^2(0, \infty; X)$ error between the exact and approximate solution to the partial differential equation (8); we do not attempt to estimate this here.

We also note that we have seen similar performance of the algorithm when the rank of B is greater than one.

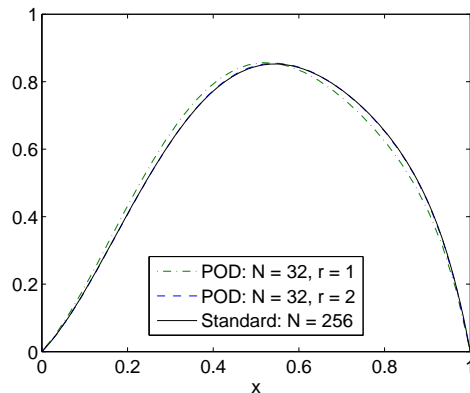


Figure 3: Approximate Lyapunov solutions acting on $f = \exp(x)$.

5.2 Model Problem 2

We take the A and B operators from the two dimensional convection diffusion equation with spatially varying convection

$$w_t(t, x, y) = \mu(w_{xx}(t, x, y) + w_{yy}(t, x, y)) - c_1 x w_x(t, x, y) - c_2 y w_y(t, x, y) + b(x, y)u(t),$$

over the unit square $[0, 1] \times [0, 1]$ with zero Dirichlet boundary conditions. Here, subscripts denote partial derivatives, μ , c_1 , and c_2 are constants, and $b(x, y)$ is a given square integrable function. The abstract formulation of the A and B operators are similar to the example above. We do not attempt to prove here that A generates an exponentially stable C_0 -semigroup; however, numerical results indicate that this is the case.

5.2.1 Numerical Results

For the computations, we chose $b(x, y) \equiv 1$, $\mu = 0.1$, $c_1 = 1$, and $c_2 = 1$. We use the low rank method and the snapshot method to approximate the solution of the operator Lyapunov equation Z acting on the function $f(x, y) = 5x^2 + y^2$.

For the low rank algorithm, standard piecewise bilinear finite elements were used for the spatial discretization of the partial differential equation (8). The discretized equations were integrated over $0 \leq t \leq 2$ using Matlab's `ode15s` solver with default error tolerances; at $t = 2$, the numerical solution is nearly zero. The time points returned from `ode15s` were used in the method of snapshots to approximate the POD eigenvalues and modes.

For the snapshot algorithm, we used the trapezoid rule for the time integration, again stopping at $t = 2$. The trapezoid rule can be derived by assuming the solution is piecewise linear in time. Substituting this approximation into the approximate POD operator (10) gives the following algorithm:

1. Approximate the solution of the PDE with the trapezoid rule:

$$(I - \Delta t A/2)w_{n+1} = (I + \Delta t A/2)w_n,$$

where I is the identity operator.

2. Update the approximation to $[Zf](x, y)$:

$$[Zf]_{n+1} = [Zf]_n + [(f, w_{n+1})\Delta t/3 + (f, w_n)\Delta t/6]w_{n+1} \\ + [(f, w_{n+1})\Delta t/6 + (f, w_n)\Delta t/3]w_n.$$

This algorithm is at the infinite dimensional level and the equations must be discretized in space to obtain an approximation to $[Zf](x, y)$. Again, we used piecewise bilinear finite elements for the discretization.

Remark 4. In practice the algorithm can be stopped whenever the norm of w_{n+1} becomes smaller than a certain tolerance. Also, the only solution data that must be stored is the solution at the previous time step (w_n). Furthermore, any time stepping method, spatial discretization, and approximation method for the POD integral operator (10) can be used.

Figure 4 shows snapshot approximations of $[Zf](x, y)$ computed using the above trapezoid rule algorithm with $\Delta t = 0.1$ and $\Delta t = 0.01$ and 64 equally spaced finite element nodes in each coordinate direction. Further refinement in space produced little change. The approximation using

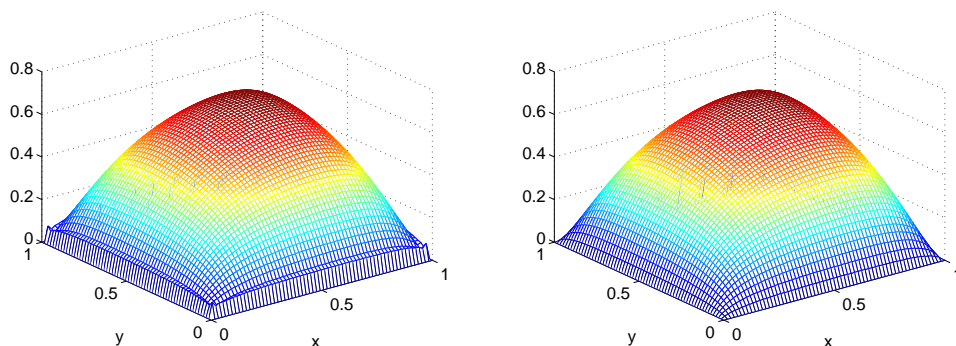


Figure 4: Snapshot approximations of $[Zf](x, y)$, where $f = 5x^2 + y^2$. Here, 64 equally spaced finite element nodes were used in each coordinate direction, and (left) $\Delta t = 0.1$ and (right) $\Delta t = 0.01$.

the larger time step $\Delta t = 0.1$ suffers from some loss of accuracy near the boundary. This is due to the incompatibility of the initial condition $w(0, x, y) = b(x, y) \equiv 1$ with the zero Dirichlet boundary conditions. Decreasing the time step to $\Delta t = 0.01$ increases the accuracy. The convergence is clear; further refinement in time produced little change. Also, the POD algorithm with $r = 5$ produces an approximation that is nearly identical to the $\Delta t = 0.01$ trapezoid rule approximation (not shown). For this computation, Matlab's adaptive solver `ode15s` takes very small time steps near $t = 0$ to obtain accurate results.

6 Conclusion

We presented two algorithms to compute approximate solutions of Lyapunov equations. The first algorithm is based on proper orthogonal decomposition and produces a low rank approximate solution; the second snapshot algorithm approximates the product of the Lyapunov solution with a few vectors and can be implemented with a minimal amount of storage. The algorithms are applicable to large-scale matrix problems as well as a class of infinite dimensional problems. Since the algorithms are based on approximating the solutions of linear evolution equations, the computations

can use existing simulation code as well as tools such as adaptive solvers and parallel algorithms. The quality of the approximate solutions can be ascertained by simple error bounds. Numerical results for parabolic model problems confirmed the convergence theory. The results indicate that accurate time stepping is important for these algorithms, especially if the initial condition to the PDE is not smooth (i.e., it is not in the domain of the A operator). Adaptive time stepping will be considered in future work.

Although we tested the algorithm on simple problems here, in another work (Dickinson *et al.*, 2009) we successfully applied the trapezoid snapshot algorithm to operator Lyapunov equations arising in feedback control computations for an incompressible fluid flow problem. The computations in that work used existing simulation code and no matrix approximations of the operators were ever extracted from the code.

In other future work, we intend to further test the algorithms on other types of infinite dimensional systems (such as delay equations and hyperbolic partial differential equations). We also plan to consider other classes of infinite dimensional systems, such as those with an unbounded B operator.

Furthermore, we plan to investigate the computational cost of the algorithms in the future. It is important to note that even if the snapshot algorithms are not as efficient as other matrix Lyapunov equation solvers, they will still be computationally tractable and therefore they may be preferable to use for certain operator Lyapunov equations due to the advantages discussed above.

Since the proposed algorithms depend on the solution of linear infinite dimensional differential equations, the algorithms may have difficulty for problems whose solutions decay slowly to zero or rapidly oscillate. We believe that most, if not all, of the recently developed matrix Lyapunov solvers also may have difficulty with such problems.

We also note that the solution of Lyapunov equations plays an important role in standard methods to compute truncated balanced reduced order models of linear systems (see, e.g., Antoulas, 2005; Datta, 2004; Zhou *et al.*, 1996). Although the POD-based algorithm presented here could be used for these Lyapunov computations, we propose that it is more natural to use Rowley's POD-based algorithm for approximate balanced truncation (Rowley, 2005). (In fact, Rowley's algorithm inspired the present work and also (Singler & Batten, 2009), which extends the algorithm in (Rowley, 2005) to an infinite dimensional case.) This method requires the solution of the linear differential equations (8) and (12) and bypasses the solution of Lyapunov equations (1) and (11).

A Appendix: Discrete POD Computation

We now prove Proposition 4, which shows that computing discrete POD eigenvalues and modes is equivalent to solving a matrix eigenvalue problem. To prove this, we use a general result on the eigenvalues of factored compact operators.

Proposition 7. *Let X and Y be two Hilbert spaces and let $L : X \rightarrow Y$ be a compact linear operator. Define $S : Y \rightarrow Y$ and $T : X \rightarrow X$ by $S = LL^*$ and $T = L^*L$. Then the nonzero eigenvalues of S and T are equal. For a nonzero eigenvalue λ , the corresponding orthonormal eigenvectors φ of S and ψ of T are related by*

$$\varphi = \lambda^{-1/2}L\psi, \quad \psi = \lambda^{-1/2}L^*\varphi.$$

Proof. Since L is compact, S and T are both compact. Also, S and T are self adjoint and non-negative due to their factored representation. Therefore, the eigenvalues of both S and T can be ordered and their corresponding orthonormal eigenvectors form complete sets.

Let λ_k be the k th (ordered) eigenvalue of $T = L^*L$ with corresponding orthonormal eigenvector ψ_k . Then

$$T\psi_k = \lambda_k\psi_k \Rightarrow L^*L\psi_k = \lambda_k\psi_k \Rightarrow (LL^*)(L\psi_k) = \lambda_k(L\psi_k) \Rightarrow S(L\psi_k) = \lambda_k(L\psi_k).$$

Therefore, λ_k is an eigenvalue of S with corresponding eigenvector $L\psi_k$. Let $\varphi_k = \lambda_k^{-1/2}L\psi_k$. Then

$$(\varphi_i, \varphi_j) = \lambda_i^{-1/2}\lambda_j^{-1/2}(L^*L\psi_i, \psi_j) = \lambda_i^{-1/2}\lambda_j^{-1/2}(T\psi_i, \psi_j) = \lambda_i^{1/2}\lambda_j^{-1/2}\delta_{ij} = \delta_{ij},$$

where δ_{ij} is the Kronecker delta. Thus, $\{\lambda_k, \varphi_k\}$ are eigenpairs for S and $\{\varphi_k\}$ is an orthonormal set.

Next, let λ_k be the k th (ordered) eigenvalue of $S = LL^*$ with corresponding orthonormal eigenvector φ_k . A similar argument shows λ_k is an eigenvalue of T with corresponding orthonormal eigenvector $\psi_k = \lambda_k^{-1/2}L^*\varphi_k$. \square

Now we proceed with the proof of Proposition 4, which gives a method for computing the POD of a discrete dataset.

Proof of Proposition 4:

Proof. The discrete POD operator can be factored as $Z = \mathcal{B}\mathcal{B}^*$ as follows. Define $\mathcal{B} : \mathbb{R}^N \rightarrow X$ by

$$\mathcal{B}u = \sum_{k=1}^N u_k W_k.$$

The adjoint operator $\mathcal{B}^* : X \rightarrow \mathbb{R}^N$ is computed to be $\mathcal{B}^*x = [(x, W_1), \dots, (x, W_N)]^T$ and therefore $Z = \mathcal{B}\mathcal{B}^*$.

Since the range of \mathcal{B}^* is finite dimensional, \mathcal{B}^* and \mathcal{B} are both compact. Due to the factorization $Z = \mathcal{B}\mathcal{B}^*$, Z is compact, self adjoint, and nonnegative. Proposition 7 above then gives that the nonzero eigenvalues of $Z = \mathcal{B}\mathcal{B}^*$ and $\hat{Z} := \mathcal{B}^*\mathcal{B}$ are equal and can be ordered $\lambda_1 \geq \lambda_2 \geq \dots > 0$. Furthermore, if λ_i is nonzero, the orthonormal eigenvector φ_i of Z is related to the orthonormal eigenvector $\hat{\varphi}_i$ of \hat{Z} by $\varphi_i = \lambda_i^{-1/2}\mathcal{B}\hat{\varphi}_i$.

A calculation shows that $\hat{Z}u = \Gamma u$, where the $N \times N$ matrix Γ has ij entries (W_i, W_j) . Thus, the nonzero eigenvalues of Z and Γ are equal. Using the definition of \mathcal{B} and the expression above for the orthonormal eigenvectors φ_i of Z shows that the eigenvectors are given by (15). \square

Acknowledgements

The author thanks the anonymous referees for pointing out the paper by Saad (1990) and for their comments which greatly improved the manuscript. Research supported in part by the Air Force Office of Scientific Research under grants FA9550-05-1-0041 and FA9550-07-1-0540.

References

- AHUES, M., LARGILLIER, A. & LIMAYE, B. V. (2001) *Spectral Computations for Bounded Operators*. Chapman & Hall/CRC, Boca Raton, FL. 4.3
- ANTOULAS, A. C., SORENSEN, D. C. & ZHOU, Y. (2002) On the decay rate of Hankel singular values and related issues. *Systems Control Lett.*, **46**, 323–342. 2.3

ANTOULAS, A. C. (2005) *Approximation of Large-Scale Dynamical Systems*. Philadelphia, PA: SIAM. [1](#), [6](#)

ATWELL, J. A. (2000) Proper Orthogonal Decomposition for Reduced Order Control of Partial Differential Equations. *Ph.D. thesis*, Department of Mathematics, Virginia Polytechnic Institute and State University. [1](#)

ATWELL, J. A., BORGGGAARD, J. T. & KING, B. B. (2001) Reduced order controllers for Burgers' equation with a nonlinear observer. *Int. J. Appl. Math. Comput. Sci.*, **11**, 1311–1330. [1](#)

ATWELL, J. A. & KING, B. B. (2001) Proper orthogonal decomposition for reduced basis feedback controllers for parabolic equations. *Math. Comput. Modelling*, **33**, 1–19. [1](#)

ATWELL, J. A. & KING, B. B. (2004) Reduced order controllers for spatially distributed systems via proper orthogonal decomposition. *SIAM J. Sci. Comput.*, **26**, 128–151. [1](#)

BANKS, H. T., SMITH, R. C. & WANG, Y. (1996) *Smart Material Structures*. Paris: Masson and Wiley. [1](#)

BANKS, H. T., DEL ROSARIO, R. C. H. & SMITH, R. C. (2000) Reduced-order model feedback control design: numerical implementation in a thin shell model. *IEEE Trans. Automat. Control*, **45**, 1312–1324. [1](#)

BANKS, H. T., BEELER, S. C., KEPLER, G. M. & TRAN, H. T. (2002) Reduced order modeling and control of thin film growth in an HPCVD reactor. *SIAM J. Appl. Math.*, **62**, 1251–1280. [1](#)

BANKS, H. T. & BURNS, J. A. (1978) Hereditary control problems: numerical methods based on averaging approximations. *SIAM J. Control Optimization*, **16**, 169–208. [1](#)

BANKS, H. T. & ITO, K. (1991) A numerical algorithm for optimal feedback gains in high-dimensional linear quadratic regulator problems. *SIAM J. Control Optim.*, **29**, 499–515. [1](#), [1](#)

BAUR, U. & BENNER, P. (2006) Factorized solution of Lyapunov equations based on hierarchical matrix arithmetic. *Computing*, **78**, 211–234. [1](#)

BEATTIE, C. A., BORGGGAARD, J., GUGERCIN, S. & ILIESCU, T. (2006) A domain decomposition approach to POD. *Proceedings of the 45th IEEE Conference on Decision and Control. Proceedings of the 45th IEEE Conference on Decision and Control.*, pp. 6750–6756. [3.2](#)

BORGGGAARD, J., BURNS, J. A., VUGRIN, E. & ZIETSMAN, L. (2004) On strong convergence of feedback operators for non-normal distributed parameter systems. *Proceedings of the IEEE Conference on Decision and Control*, vol. 2. *Proceedings of the IEEE Conference on Decision and Control*, vol. 2., pp. 1526 – 1531. [1](#)

BURNS, J., ITO, K. & PROPST, G. (1988) On nonconvergence of adjoint semigroups for control systems with delays. *SIAM J. Control Optim.*, **26**, 1442–1454. [1](#)

BURNS, J. A., KING, B. B. & RUBIO, D. (1998) Feedback control of a thermal fluid using state estimation. *Int. J. Comput. Fluid Dyn.*, **11**, 93–112. [1](#)

BURNS, J. A., SACHS, E. W. & ZIETSMAN, L. (2008) Mesh independence of Kleinman–Newton iterations for Riccati equations in Hilbert space. *SIAM Journal on Control and Optimization*, **47**, 2663–2692. [1](#)

- BURNS, J. A. & FABIANO, R. H. (1989) Feedback control of a hyperbolic partial-differential equation with viscoelastic damping. *Control Theory Adv. Tech.*, **5**, 157–188. [1](#)
- CAMP, K. A. E. & KING, B. B. (2002) A comparison of balanced truncation techniques for reduced order controllers. *Proceedings of 2002 MTNS. Proceedings of 2002 MTNS*. [1](#)
- CHATELIN, F. (1983) *Spectral Approximation of Linear Operators*. New York: Academic Press Inc. [4.3](#)
- CURTAIN, R. F. (2003) Model reduction for control design for distributed parameter systems. *Research Directions in Distributed Parameter Systems*. Philadelphia, PA: SIAM, pp. 95–121. [1](#)
- CURTAIN, R. F. & SASANE, A. J. (2001) Compactness and nuclearity of the Hankel operator and internal stability of infinite-dimensional state linear systems. *Internat. J. Control*, **74**, 1260–1270. [4.2](#)
- CURTAIN, R. F. & ZWART, H. J. (1995) *An Introduction to Infinite-Dimensional Linear System Theory*. New York: Springer-Verlag. [1](#), [2.1](#), [1](#), [2.1](#), [5.1](#)
- DATTA, B. N. (2004) *Numerical Methods for Linear Control Systems*. San Diego, CA: Elsevier Academic Press. [1](#), [6](#)
- DELATTRE, C., DOCHAIN, D. & WINKIN, J. (2003) Sturm-Liouville systems are Riesz-spectral systems. *Int. J. Appl. Math. Comput. Sci.*, **13**, 481–484. [5.1](#)
- DICKINSON, B. T., SINGLER, J. R. & BATTEN, B. A. (2009) A snapshot algorithm for linear feedback flow control design. *Proceedings of the AIAA Infotech@Aerospace Conference and AIAA Unmanned...Unlimited Conference. Proceedings of the AIAA Infotech@Aerospace Conference and AIAA Unmanned...Unlimited Conference*. AIAA paper number 2009-1961. [6](#)
- EVANS, K. A. (2003) Reduced order controllers for distributed parameter systems. *Ph.D. thesis*, Virginia Polytechnic Institute and State University. [1](#)
- FAHL, M. (2001) Computation of POD basis functions for fluid flows with Lanczos methods. *Math. Comput. Modelling*, **34**, 91–107. [3.2](#)
- GAVRILYUK, I. P., HACKBUSCH, W. & KHOROMSKIJ, B. N. (2004) Data-sparse approximation to the operator-valued functions of elliptic operator. *Math. Comp.*, **73**, 1297–1324. [1](#)
- GIBSON, J. S. & ADAMIAN, A. (1991) Approximation theory for linear-quadratic-Gaussian optimal control of flexible structures. *SIAM J. Control Optim.*, **29**, 1–37. [1](#)
- GRASEDYCK, L., HACKBUSCH, W. & KHOROMSKIJ, B. N. (2003) Solution of large scale algebraic matrix Riccati equations by use of hierarchical matrices. *Computing*, **70**, 121–165. [1](#)
- GRASEDYCK, L. (2004) Existence of a low rank or \mathcal{H} -matrix approximant to the solution of a Sylvester equation. *Numer. Linear Algebra Appl.*, **11**, 371–389. [2.3](#)
- GRASEDYCK, L. & HACKBUSCH, W. (2007) A multigrid method to solve large scale Sylvester equations. *SIAM J. Matrix Anal. Appl.*, **29**, 870–894. [1](#)
- GUGERCIN, S., SORENSEN, D. C. & ANTOULAS, A. C. (2003) A modified low-rank Smith method for large-scale Lyapunov equations. *Numer. Algorithms*, **32**, 27–55. [1](#)

- HENRI, T. & YVON, J.-P. (2002a). Institut de Recherche Mathématique de Rennes, Prépublication 02-48. [3](#), [4.2](#)
- HENRI, T. & YVON, J.-P. (2002b). Institut de Recherche Mathématique de Rennes, Prépublication 02-40. [3](#), [4.2](#)
- HENRI, T. & YVON, J.-P. (2005) Convergence estimates of POD-Galerkin methods for parabolic problems. *System Modeling and Optimization*. Boston, MA: Kluwer Acad. Publ., pp. 295–306. [3](#), [4.2](#)
- HOMESCU, C., PETZOLD, L. R. & SERBAN, R. (2005) Error estimation for reduced-order models of dynamical systems. *SIAM J. Numer. Anal.*, **43**, 1693–1714. [1](#)
- KEPLER, G. M., TRAN, H. T. & BANKS, H. T. (2000) Reduced order model compensator control of species transport in a CVD reactor. *Optimal Control Appl. Methods*, **21**, 143–160. [1](#)
- KING, B. B., HOVAKIMYAN, N., EVANS, K. A. & BUHL, M. (2006) Reduced order controllers for distributed parameter systems: LQG balanced truncation and an adaptive approach. *Math. Comput. Modelling*, **43**, 1136–1149. [1](#)
- KLEINMAN, D. L. (1968) On an iterative technique for Riccati equation computations. *IEEE Trans. Autom. Control*, **AC-13**, 114–115. [1](#)
- KUNISCH, K. & VOLKWEIN, S. (2001) Galerkin proper orthogonal decomposition methods for parabolic problems. *Numer. Math.*, **90**, 117–148. [1](#)
- KUNISCH, K. & VOLKWEIN, S. (2002) Galerkin proper orthogonal decomposition methods for a general equation in fluid dynamics. *SIAM J. Numer. Anal.*, **40**, 492–515. [1](#), [3](#), [4.2](#)
- LEE, C. H. & TRAN, H. T. (2005) Reduced-order-based feedback control of the Kuramoto-Sivashinsky equation. *J. Comput. Appl. Math.*, **173**, 1–19. [1](#)
- LI, J.-R. & WHITE, J. (2002) Low rank solution of Lyapunov equations. *SIAM J. Matrix Anal. Appl.*, **24**, 260–280. [1](#)
- MORRIS, K. & NAVASCA, C. (2008). To appear in Computational Optimization and Applications. [1](#), [1](#)
- PAZY, A. (1983) *Semigroups of Linear Operators and Applications to Partial Differential Equations*. New York: Springer-Verlag. [1](#)
- PENZL, T. (1999/00) A cyclic low-rank Smith method for large sparse Lyapunov equations. *SIAM J. Sci. Comput.*, **21**, 1401–1418. [1](#)
- PENZL, T. (2000) Eigenvalue decay bounds for solutions of Lyapunov equations: the symmetric case. *Systems Control Lett.*, **40**, 139–144. [2.3](#)
- ROSEN, I. G. & WANG, C. (1995) A multilevel technique for the approximate solution of operator Lyapunov and algebraic Riccati equations. *SIAM J. Numer. Anal.*, **32**, 514–541. [1](#)
- ROWLEY, C. W. (2005) Model reduction for fluids, using balanced proper orthogonal decomposition. *Internat. J. Bifur. Chaos Appl. Sci. Engrg.*, **15**, 997–1013. [6](#)

- SAAD, Y. (1990) Numerical solution of large Lyapunov equations. *Signal processing, scattering and operator theory, and numerical methods (Amsterdam, 1989)*. Progr. Systems Control Theory, vol. 5. Boston, MA: Birkhäuser Boston, pp. 503–511. [1](#), [2.2](#), [A](#)
- SIMONCINI, V. (2007) A new iterative method for solving large-scale Lyapunov matrix equations. *SIAM J. Sci. Comput.*, **29**, 1268–1288. [1](#)
- SINGLER, J. R. (2008) Approximate low rank solutions of Lyapunov equations via proper orthogonal decomposition. *Proceedings of the 2008 American Control Conference. Proceedings of the 2008 American Control Conference.*, pp. 267–272. [1](#)
- SINGLER, J. R. & BATTEN, B. A. (2009) A proper orthogonal decomposition approach to approximate balanced truncation of infinite dimensional linear systems. *International Journal of Computer Mathematics*, **86**, 355–371. [6](#)
- SIROVICH, L. (1987) Turbulence and the dynamics of coherent structures. I. Coherent structures. *Quart. Appl. Math.*, **45**, 561–571. [3.2](#)
- WATKINS, D. S. (2005) Product eigenvalue problems. *SIAM Rev.*, **47**, 3–40 (electronic). [3.2.1](#)
- WEIDMANN, J. (1980) *Linear operators in Hilbert spaces*. New York: Springer-Verlag. [2](#), [2.2](#), [4.2](#)
- WILLCOX, K. & PERAIRE, J. (2002) Balanced model reduction via the proper orthogonal decomposition. *AIAA Journal*, **40**, 2323–2330. [1](#), [2.2](#)
- ZHOU, K., DOYLE, J. C. & GLOVER, K. (1996) *Robust and Optimal Control*. Prentice-Hall. [1](#), [6](#)