2019

Approximation of the Generalized Singular Value Expansion

Matthew Jacob Roberts
Michigan Technological University, majrober@mtu.edu

APPROXIMATION OF THE GENERALIZED SINGULAR VALUE EXPANSION

By
Matthew J. Roberts

A DISSERTATION
Submitted in partial fulfillment of the requirements for the degree of
DOCTOR OF PHILOSOPHY
In Mathematical Sciences

MICHIGAN TECHNOLOGICAL UNIVERSITY
2019

This dissertation has been approved in partial fulfillment of the requirements for the Degree of DOCTOR OF PHILOSOPHY in Mathematical Sciences.

Department of Mathematical Sciences

Dissertation Advisor:	*Dr. Mark S. Gockenbach*
Committee Member:	*Dr. Benjamin W. Ong*
Committee Member:	*Dr. Jiguang Sun*
Committee Member:	*Dr. Hassan Masoud*
Department Chair:	*Dr. Mark S. Gockenbach*

Contents

List of Figures

List of Tables

Abstract

Let X, Y, and Z be real separable Hilbert spaces, let $T : X \to Y$ be a compact operator, and let $L : D(L) \to Z$ be a closed and densely defined linear operator. Then the generalized singular value expansion (GSVE) is an expansion that expresses T and L in terms of a common orthonormal basis. Under certain hypotheses on discretization, the GSVE of an approximate operator pair (T_j, L_j), where $T_j : X_j \to Y_j$ and $L_j : X_j \to Z_j$, converges to the GSVE of (T, L). Error estimates establish a rate of convergence that is consistent with numerical experiments in the case of discretization using piecewise linear finite elements. Further numerical testing suggests that a higher rate of convergence is attained by using higher order elements. However, the theory does not cover this case.

Chapter 1

Introduction

A linear inverse problem is a problem of the form $Tx = y$, where $T : X \to Y$ is a linear operator and we wish to estimate the exact solution x^* from a noisy measurement y of the exact data y^* (see [1], [2], [3], and [4]). We consider only linear inverse problems defined on real separable Hilbert space. A particular class of inverse problems that is of interest consists of those in which the operator T is an integral operator defined by

$$(Tf)(s) = \int_{\Omega_1} k(s,t) f(t) d\Omega_1(t), \tag{1.1}$$

where Ω_1 and Ω_2 are either bounded intervals in $\mathbb{R}$, or are bounded two-dimensional domains. The corresponding separable Hilbert spaces are $X = L^2(\Omega_1)$ and $Y = L^2(\Omega_2)$ (see [5]).

We call $Tx = y$ an inverse problem only when the problem is unstable (that is, x does not depend continuously on y); for this reason, it is necessary to use regularization of some sort to produce an acceptable approximation of the true solution x.

Let us consider the following model inverse problem. Let $X = L^2(0,1)$ and define $T : X \to X$ by the integral operator

$$(Tx)(s) = \int_0^1 k(s,t) x(t) \; dt,$$

where $k(s,t) = \dfrac{1}{2}(s + t - |s - t|) - st$. It is a quick exercise to verify that T is the solution operator to the following two-point boundary value problem.

$$\begin{aligned} -y''(t) &= x(t) \text{ in } (0,1) \\ y(0) &= 0 \\ y(1) &= 0. \end{aligned} \tag{1.2}$$

More precisely, given the right hand side $x(t)$ to the above two-point boundary value problem (1.2), the operator T gives back the solution $y(t)$. Therefore,

$$(Tx)(s) = \int_0^1 k(s,t)x(t)\ dt = y(s).$$

In this example, the two-point boundary value problem (1.2) is the forward problem and the equation

$$Tx = y \tag{1.3}$$

defines the corresponding inverse problem.

To see that (1.3) actually defines an inverse problem, consider the case where $x(t) = t$ is the exact solution. This produces the exact data $y(s) = -\dfrac{s^3}{6} + \dfrac{s}{6}$. Let $\beta \in \mathbb{R}^n$ be a measurement of y at n equally spaces points on $[0, 1]$, subject to uniformly distributed random noise, scaled to 1% in the Euclidean norm.

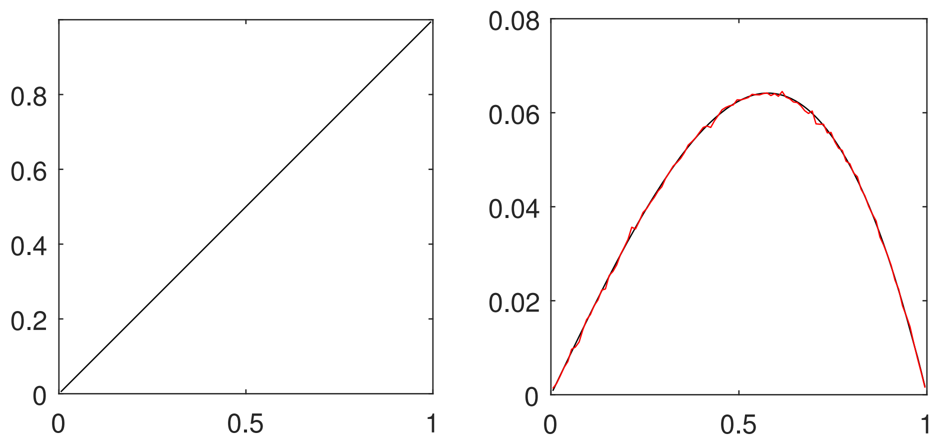

Figure 1.1: Exact solution (left) and exact and noisy data (right) for the model inverse problem.

We will now discretize the interval $[0, 1]$ in the following way. Let $h = \dfrac{1}{n}$ and consider the mesh

$$M = \{[0, h], [h, 2h], \cdots, [(n-1)h, 1]\}.$$

We can discretize the space X by the finite dimensional subspace $X_n = \text{span}\{x_1, x_2, \cdots, x_n\}$, where $\{x_1, x_2, \cdots, x_n\}$ is the standard nodal basis for the space of continuous piecewise constant functions relative to the mesh M. We discretize the operator T by the matrix $A \in \mathbb{R}^{n\times n}$ using the Galerkin method (see [1], Section 3.2).

By replacing the equation $Tx = y$ with the discretized equation $A\alpha = \beta$, we then have the vector $\hat{y} = \sum_{k=0}^{n-1} \beta_k x_k$ as our approximation (measurement) of y, and the resulting vector $\hat{x} = \sum_{k=0}^{n-1} \alpha_k x_k$ (where α is the solution of $A\alpha = \beta$) as our approximation of x. Figure 1.2 gives the plot of the approximated solution $\hat{x}$.

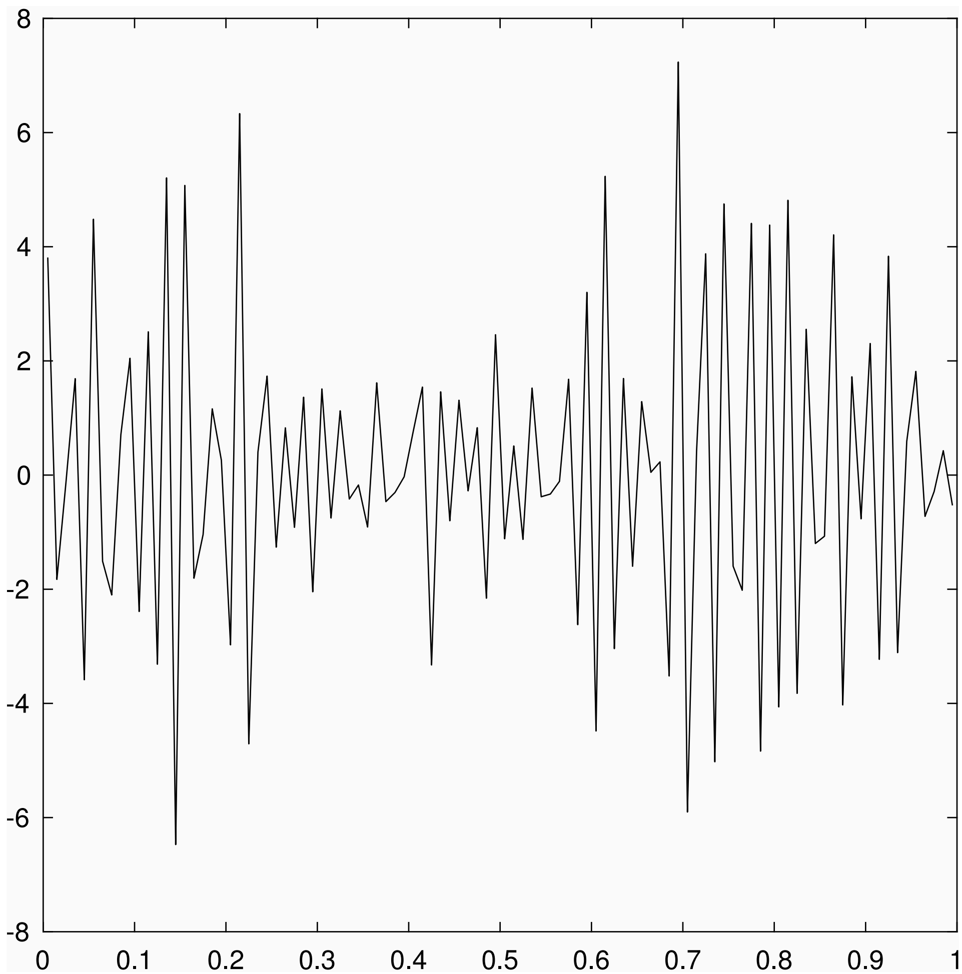

Figure 1.2: Naive solution of the model inverse problem.

As we can see, the approximate solution $\hat{x}$ is not close to the actual solution x. This illustrates the fact that the solution x to equation (1.3) does not continuously depend on the data y. Hence, the equation $Tx = y$ defines a linear inverse problem, and we will need to use regularization of some sort in order to solve it. We begin with considering classical *Tikhonov regularization* (see [6] and [5]).

In Tikhonov regularization, we consider the solution to the regularized problem

$$\min_{x\in X} \|Tx - y\|^2 + \lambda\|x\|^2, \tag{1.4}$$

where $\lambda > 0$ is a constant. In this case, $k \in L^2((0,1)\times(0,1))$ and T is a compact operator, so the singular value expansion (SVE) of T is invaluable for analyzing the solution to this regularized problem. The SVE of T can be given as follows:

$$T = \sum_{k=1}^{\infty} \sigma_k \psi_k \otimes \phi_k. \tag{1.5}$$

Here, $\{\phi_k\}$ is a complete orthonormal sequence for the space $\mathcal{N}(T)^\perp$, $\{\psi_k\}$ is an orthonormal sequence in X, and $\{\sigma_k\}$ is the sequence of singular values of T, a sequence of positive numbers monotonically decreasing to 0 (see [7] Section 2.8 or see [4]). It is easy to show that the unique solution of (1.4) lies in $\mathcal{R}(T^*)$ (see [4]).

Using the SVE of T, we have that for any $x = \sum_{k=1}^{\infty} \alpha_k \phi_k \in \mathcal{N}(T)^\perp$,

$$\begin{aligned}
\|Tx - y\|^2 + \lambda\|x\|^2 &= \left[\sum_{k=0}^{\infty} \left(\sigma_k \langle x, \phi_k\rangle_X - \langle y, \psi_k\rangle_Y\right)^2 + \lambda \langle x, \phi_k\rangle_X^2\right] + \|\hat{y}\|^2 \\
&= \left[\sum_{k=0}^{\infty} (\sigma_k^2 + \lambda)\left(\langle x, \phi_k\rangle_X - \frac{\sigma_k}{\sigma_k^2 + \lambda}\langle y, \psi_k\rangle_Y\right)^2 + \left(1 - \frac{\sigma_k^2}{\sigma_k^2 + \lambda}\right)\langle y, \psi_k\rangle_Y\right] \\
&\quad + \|\hat{y}\|^2,
\end{aligned}$$

where $\hat{y}$ is the orthogonal projection of y onto $R(T)^\perp$. Therefore, the unique solution to problem (1.4) is given by

$$x_{\lambda,y} = \sum_{k=0}^{\infty} \frac{\sigma_k}{\sigma_k^2 + \lambda} \langle y, \psi_k\rangle_Y \phi_k. \tag{1.6}$$

Using an appropriate choice for λ, the plots for the exact solution x and the regularized solution $x_{\lambda,y}$ can be seen in Figure 1.3.

As we can see from figure 1.3, the regularized solution $x_{\lambda,y}$ inherits the Dirichlet boundary conditions of the forward problem (1.2). To understand why this is so, it is easy to show that $x_{\lambda,y} \in \mathcal{R}(T^*)$. In this particular example, the operator T is self-adjoint. This follows from the fact the that kernel $k(s,t)$ is symmetric about the line $s = t$ (i.e. $k(s,t) = k(t,s)$. Since T is the solution operator to the forward problem (1.2) with Dirichlet boundary conditions, every element in the range of T has Dirichlet boundary conditions. Thus, the regularized solution $x_{\lambda,y} \in \mathcal{R}(T^*) = \mathcal{R}(T)$ inherits

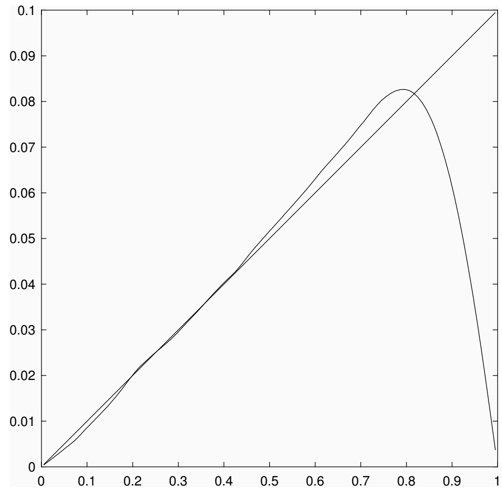

Figure 1.3: Solution of model inverse problem produced by Tikhonov regularization.

Dirichlet boundary conditions. For this reason, classical Tikhonov regularization does not work well for this problem.

We can generalize classical Tikhonov regularization by considering the unique solution to the problem

$$\min_{x \in D(L)} \|Tx - y\|^2 + \lambda \|Lx\|^2, \tag{1.7}$$

where again $\lambda > 0$ is a constant. Here, L is a closed operator with domain $D(L)$ and is densely defined in X. It is chosen to have the property that $\|Lx\|$ is small for reasonable solutions x and large for those x with some undesirable feature. Such regularization is called *Tikhonov regularization with seminorms* since $\|x\|_L = \|Lx\|$ defines a seminorm on $D(L)$ (see [6] or [4]). In the case of classical Tikhonov regularization, L is the identity operator, and the undesirable feature of x is being large in magnitude.

In many cases, the true solution x has a certain level of smoothness or regularity. In other words, the undesirable approximations for x in the problem $Tx = y$ are those approximations in which the derivative is undefined or large in magnitude. In the case that $X = L^2(a, b)$, the derivative operator $L : D(L) \to Z$ is linear, closed, and densely defined in X. Thus, a method of regularization appropriate for problems of this kind will be Tikhonov regularization with seminorms defined by the derivative operator. Similarly, if $\Omega \subseteq \mathbb{R}^d$ ($d = 2$ or $d \geq 3$), the gradient operator is a suitable regularization operator for many problems.

Let problem (1.7) be discretized to produce the following linear algebra problem:

$$\min_{\vec{x} \in \mathbb{R}^n} \|A\vec{x} - \vec{y}\|^2 + \lambda \|B\vec{x}\|^2 \tag{1.8}$$

Here, $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times n}$, $\vec{x} \in \mathbb{R}^n$, and $\vec{y} \in \mathbb{R}^m$, and we assume that $\mathcal{N}(A)\cap\mathcal{N}(B) = \{0\}$, where $\mathcal{N}(A)$ represents the null space of the matrix A. In order to solve this problem, it is beneficial to simultaneously diagonalize the matrices A and B. This is done by computing the generalized singular value decomposition (GSVD) of the matrix pair (A, B). We present one version of the GSVD that is relevant to our discussion (see [8], Theorem 22.2) .

Theorem 1. *Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times n}$ be matrices such that $m \geq n$ and $\mathcal{N}(A) \cap \mathcal{N}(B) = \{0\}$. Then there exist a nonsingular matrix $W \in \mathbb{R}^{n\times n}$, matrices $U \in \mathbb{R}^{m\times n}$ and $V \in \mathbb{R}^{p\times p}$ with orthonormal columns, and diagonal matrices $S \in \mathbb{R}^{n\times n}$ and $M \in \mathbb{R}^{p\times n}$ such that*

$$A = USW^{-1}\ ,\ \ B = VMW^{-1}.$$

Moreover, the diagonal entries s_i of S and m_1 of M are nonnegative, and satisfy

$$\begin{aligned} &s_i^2 + m_i^2 = 1 \ \textit{for}\ i = 1, 2, \cdots, p, \\ &s_i = 1 \ \textit{for}\ i = p+1, p+2, \cdots, n \end{aligned}$$

(assuming for convenience that $n \geq p$). In matrix form,

$$S^T S + M^T M = I.$$

Let $A = USW^{-1}$ and $B = VMW^{-1}$ be the GSVD of (A, B) as given in the theorem. Then

$$\begin{aligned} \|A\vec{x} - \vec{y}\|^2 + \lambda\|B\vec{x}\|^2 &= \|USW^{-1}\vec{x} - \vec{y}\|^2 + \lambda\|VMW^{-1}\vec{x}\|^2 \\ &= \|US\vec{w} - \vec{y}\|^2 + \lambda\|VM\vec{w}\|^2 \\ &= \sum_{k=1}^{n}(s_k w_k - \langle u_k, \vec{y}\rangle)^2 + \sum_{k=1}^{p}\lambda(m_k w_k)^2 \\ &= \sum_{k=1}^{p}\left[(s_k w_k - \langle u_k, \vec{y}\rangle)^2 + \lambda m_k^2 w_k^2\right] + \sum_{k=p+1}^{n}(w_k - \langle u_k, \vec{y}\rangle)^2. \end{aligned}$$

By regrouping the terms of these sums, we have

$$\begin{aligned} &\sum_{k=1}^{p}\left[(s_k w_k - \langle u_k, \vec{y}\rangle)^2 + \lambda m_k^2 w_k^2\right] + \sum_{k=p+1}^{n}(w_k - \langle u_k, \vec{y}\rangle)^2 \\ &= \sum_{k=1}^{p}\left[(s_k^2 + \lambda m_k^2)\left(w_k - \frac{s_k\langle u_k, \vec{y}\rangle}{s_k^2 + \lambda m_k^2}\right)^2 + \left(1 - \frac{s_k^2}{(s_k^2 + \lambda m_k^2)^2}\right)\langle u_k, \vec{y}\rangle\right] \end{aligned}$$

$$+ \sum_{k=p+1}^{n} (w_k - \langle u_k, \vec{y} \rangle)^2$$

Therefore, the solution to the discretized problem is $\vec{x} = W\vec{w}$ where

$$w_k = \frac{s_k}{s_k^2 + \lambda m_k^2}(u_k \cdot \vec{y}).$$

Figure 1.4 shows the plots of the true solution x with the regularized solution $x_{\lambda,y}$ coming from seminorm regularization. (A good value of λ was chosen by trial and error.)

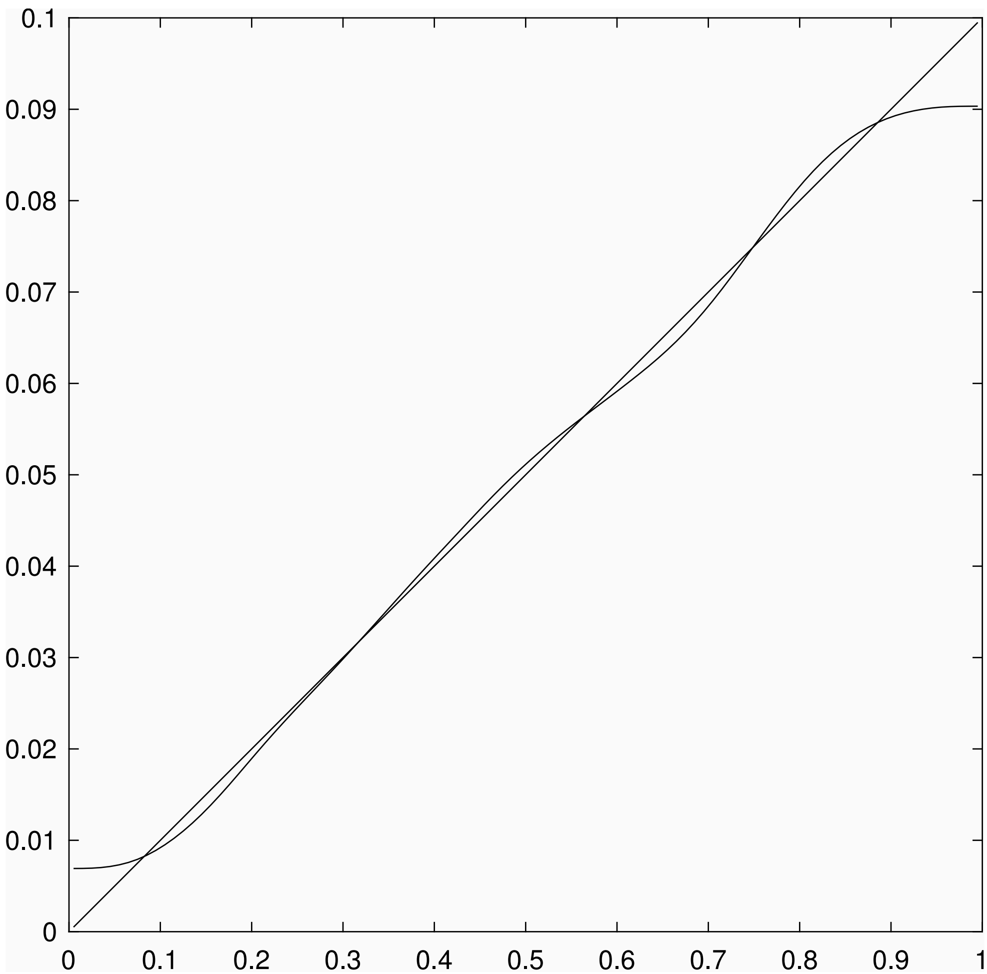

Figure 1.4: The exact solution of the model inverse problem, together with a solution produced by seminorm regularization.

The generalized singular value expansion (GSVE) of an operator pair (T, L), introduced in [4], allows the two operators T and L to be simultaneously diagonalized in the same way that the GSVD of a matrix pair (A, B) simultaneously diagonalizes the matrices A and B. Therefore, the GSVE of the operator pair (T, L) makes the analysis of problem (1.7) relatively transparent in the same way the GSVD of the matrix pair (A, B) does for the discretized problem (1.8). To describe the GSVE, we establish the following conditions on the operators T and L. Let X, Y, and Z be separable Hilbert spaces, let $T : X \to Y$ be a compact linear operator, and let $L : D(L) \to Z$ be a closed linear operator, where $D(L)$ is a dense subspace of X. We assume that there exists $\gamma > 0$ such that

$$\|Tx\|_Y^2 + \|Lx\|_Z^2 \geq \gamma \|x\|_X^2 \ , \text{ for all } x \in D(L). \tag{1.9}$$

We define the inner product $\langle \cdot, \cdot \rangle_*$ on $D(L)$ by

$$\langle u, v \rangle_* = \langle Tu, Tv \rangle_Y + \langle Lu, Lv \rangle_Z. \tag{1.10}$$

and write $\| \cdot \|_*$ for the corresponding norm. It is well known that $D(L)$ is a Hilbert space under the inner product $\langle \cdot, \cdot \rangle_*$ given that condition (1.9) is met (see Section 5.2 of [4]). The following theorem asserts the existence of the generalized singular value expansion of the operator pair (T, L) (see [9], Theorem 4.2).

Theorem 2. *Let X, Y, and Z be real separable Hilbert spaces. Assume that $T : X \to Y$ is a compact linear operator and $L : D(L) \to Z$ is a closed densely defined linear operator. Assume that there exists $\gamma > 0$ such that (1.9) holds. Then there exists a complete orthonormal set $\{\phi_k : k \in I\}$ for $D(L)$ where I is a countable index set, a partition $M_0 \cup M_a \cup M_b$ of I, orthonormal sets $\{\psi_k : k \in M_0 \cup M_b\}$ in Y, $\{\theta_k : k \in M_0 \cup M_a\}$ in Z, and subsets $\{a_k : k \in I\}$, $\{b_k : k \in I\}$ of the nonnegative real numbers such that*

$$T = \sum_{k \in M_0 \bigcup M_b} a_k \psi_k \otimes_* \phi_k \quad , \quad L = \sum_{k \in M_0 \bigcup M_a} b_k \theta_k \otimes_* \phi_k, \tag{1.11}$$

and $0 \leq a_k, b_k \leq 1$, $a_k^2 + b_k^2 = 1$ for every $k \in I$. Here, $\otimes_*$ refers to the outer product with respect to the $*$-norm. (i.e. $(\psi_k \otimes_* \phi_k)x = \langle \phi_k, x \rangle_* \psi_k$ for any $x \in D(L)$).

Using the GSVE of the operator pair (T, L) given by (1.11), we have, for any $x \in D(L)$,

$$\begin{aligned}
&\|Tx - y\|_Y^2 + \lambda \|Lx\|_Z^2 = \\
&= \|\hat{y}\|^2 + \sum_{k \in M_0 \cup M_b} \left(a_k \langle x, \phi_k \rangle_* - \langle y, \psi_k \rangle_Y\right)^2 + \sum_{k \in M_0 \cup M_a} \lambda b_k^2 \langle x, \phi_k \rangle_*^2 \\
&= \|\hat{y}\|^2 + \sum_{k \in M_b} \left(a_k \langle x, \phi_k \rangle_* - \langle y, \psi_k \rangle_Y\right)^2 + \sum_{k \in M_0} \left[\left(a_k \langle x, \phi_k \rangle_* - \langle y, \psi_k \rangle_Y\right)^2 + \lambda b_k^2 \langle x, \phi_k \rangle_*^2\right] \\
&\qquad + \sum_{k \in M_a} \lambda b_k^2 \langle x, \phi_k \rangle_*^2
\end{aligned}$$

$$= \sum_{k\in M_0} \left[(a_k^2 + \lambda b_k^2) \left(\langle x, \phi_k \rangle_* - \frac{a_k}{a_k^2 + \lambda b_k^2} \langle y, \psi_k \rangle_Y \right)^2 + \left(1 - \frac{a_k^2}{a_k^2 + \lambda b_k^2} \right) \langle y, \psi_k \rangle_Y \right]$$
$$+ \sum_{k\in M_a} \lambda b_k^2 \langle x, \phi_k \rangle_*^2 + \sum_{k\in M_b} (a_k \langle x, \phi_k \rangle_* - \langle y, \psi_k \rangle_Y)^2 + \sum_{k\in M_a} \lambda b_k^2 \langle x, \phi_k \rangle_*^2 + \|\hat{y}\|^2.$$

Here, $\hat{y}$ is the orthogonal projection of y onto $R(T)^\perp$. Therefore, the solution of problem (1.7) is given by

$$x_{\lambda,y} = \sum_{k\in M_0} \frac{a_k}{a_k^2 + \lambda b_k^2} \langle y, \psi_k \rangle_Y \phi_k + \sum_{k\in M_b} \langle y, \psi_k \rangle_Y \phi_k. \tag{1.12}$$

In practice, the GSVE of the operator pair (T, L) is a useful tool for analyzing methods such as Tikhonov regularization with seminorms. The GSVE of (T, L) can also be used directly to make computations, as can be seen in the above derivations for the regularized problem (1.7). In the next section, we provide an algorithm for computing the GSVE of (T_j, L_j), where the operator pair (T_j, L_j) is a finite dimensional discretization of the operator pair (T, L). In order to compute the GSVE of the discretized operator pair (T_j, L_j), we compute the GSVD of an associated pair of matrices.

Chapter 2

The approximate GSVE of (T, L)

The purpose of this thesis is to propose and analyze a general approach to estimating the GSVE of an operator pair (T, L). Two approaches were presented in [9]. The first is based on recognizing that the pairs (a_k^2, ϕ_k) for $k \in I$ (with $a_k = 0$ for $k \in M_a$) are the eigenpairs of the compact self-adjoint operator $T^{\#}T$. These eigenpairs can be estimated using the general theory for symmetric, variationally posed eigenvalue problems, as presented in [10]. However, this approach has two shortcomings. We must choose a finite-dimensional subspace X_j of $D(L)$ with basis $\{x_1, x_2, \cdots, x_{n_j}\}$ and solve the generalized (matrix) eigenvalue problem

$$G\alpha = \lambda M\alpha,$$

where $G \in \mathbb{R}^{n\times n}$ and $M \in \mathbb{R}^{n\times n}$ are defined by

$$G_{ij} = \langle x_j, x_i\rangle_* \ , \ M_{ij} = \langle Tx_i, Tx_j\rangle_Y.$$

The first issue with this approach is the need to compute the matrix M; generally, this matrix is expensive to compute. (For example, if T is a Fredholm integral operator, then each M_{ij} is defined by a triple integral.) The second difficulty is that, in the typical application ($R(T)$ infinite-dimensional and not closed), M_0 has infinite cardinality and $a_k \to 0$ as $k \to \infty$. It follows that by using an algorithm that computes a_k^2 (instead of computing a_k directly), we artificially restrict the ability to compute small singular values; roughly speaking, at best we can compute values of a_k down to $\sqrt{u}$ (where u is the unit round), rather than down to u itself.

It should be noted that the approach described in the previous paragraph, which is described fully in [9], does have the advantage that its convergence follows directly from the theory of symmetric, variationally posed eigenvalue problems.

The second approach, as described in this chapter, is based on reducing the computation to that of a (matrix) generalized singular value decomposition. The GSVE of a pair of operators (T_j, L_j) is related to the GSVD of a pair of related matrices, where T_j and L_j are finite dimensional operators that approximate T and L respectively in some sense. We now elaborate on this.

Let $\{X_j\}$, $\{Y_j\}$, and $\{Z_j\}$ be sequences of finite dimensional spaces contained in $D(L)$, Y, and Z, respectively, such that for each $j \in \mathbb{Z}^+$,

$$\begin{aligned} X_j &= \operatorname{span}\{x_1^{(j)}, x_2^{(j)}, \cdots, x_{n_j}^{(j)}\}, \\ Y_j &= \operatorname{span}\{y_1^{(j)}, y_2^{(j)}, \cdots, y_{m_j}^{(j)}\}, \\ Z_j &= \operatorname{span}\{z_1^{(j)}, z_2^{(j)}, \cdots, z_{p_j}^{(j)}\}. \end{aligned}$$

Suppose that the sequence of spaces $\{X_j\}$ approximate the space $D(L)$ in that for any $x \in D(L)$,

$$\|\Pi_{X_j} x - x\|_* \to 0 \text{ as } j \to \infty.$$

Here, $\Pi_{X_j} : D(L) \to X_j$ is the orthogonal projection of $D(L)$ onto X_j with respect to the $*$-norm as defined in equation (1.10). Similarly, suppose that $\{Y_j\}$ and $\{Z_j\}$ approximate the spaces Y and Z, respectively, such that for any $y \in Y$ and for any $z \in Z$,

$$\begin{aligned} \|P_{Y_j} y - y\|_Y &\to 0 \text{ as } j \to \infty, \\ \|P_{Z_j} z - z\|_Z &\to 0 \text{ as } j \to \infty. \end{aligned}$$

Here, $P_{Y_j} : Y \to Y_j$ and $P_{Z_j} : Z \to Z_j$ are the orthogonal projections of Y onto Y_j and Z onto Z_j respectively. For each $j \in \mathbb{Z}^+$, let $T_j : X_j \to Y_j$ and $L_j : X_j \to Z_j$ be linear operators that approximate T and L in some sense. The conditions under which T_j and L_j should approximate T and L are made clear in the next chapter. For each $j \in \mathbb{Z}^+$, we define the $*_j$-inner product on the space X_j by

$$\langle x, y \rangle_{*_j} = \langle T_j x, T_j y \rangle_Y + \langle L_j x, L_j y \rangle_Z \text{ for all } x, y \in X_j \tag{2.1}$$

In general, $\langle \cdot, \cdot \rangle_{*_j}$ need not be positive definite on X_j . To ensure that $\langle \cdot, \cdot \rangle_{*_j}$ defines an inner product, we will assume that

$$\mathcal{N}(T_j) \cap \mathcal{N}(L_j) = 0 \text{ for every } j \in \mathbb{Z}^+.$$

Under the assumptions placed on (T_j, L_j) in the next chapter, this must hold for all $j \in \mathbb{Z}^+$ sufficiently large. Define the matrices $A_j \in \mathbb{R}^{m_j \times n_j}$ and $B_j \in \mathbb{R}^{p_j \times n_j}$ by

$$\begin{aligned} (A_j)_{k\ell} &= \langle y_k^{(j)}, T_j x_\ell^{(j)} \rangle_Y, \\ (B_j)_{k\ell} &= \langle z_k^{(j)}, L_j x_\ell^{(j)} \rangle_Z. \end{aligned}$$

Let $H_j \in \mathbb{R}^{m_j \times m_j}$ and $J_j \in \mathbb{R}^{p_j \times p_j}$ be the Gram matrices for $\text{span}\{y_1^{(j)}, y_2^{(j)}, \cdots, y_{m_j}^{(j)}\}$ and $\text{span}\{z_1^{(j)}, z_2^{(j)}, \cdots, z_{p_j}^{(j)}\}$, respectively, which are defined by

$$
\begin{aligned}
(H_j)_{k\ell} &= \langle y_k^{(j)}, y_\ell^{(j)} \rangle_Y, \\
(J_j)_{k\ell} &= \langle z_k^{(j)}, z_\ell^{(j)} \rangle_Z.
\end{aligned}
$$

The next theorem shows how to compute the GSVE of (T_j, L_j) using the GSVD of the matrix pair $(H_j^{-1/2}A, J_j^{-1/2}B)$.

Theorem 3. *Let A_j, B_j, H_j, and J_j be as defined above, and let*

$$
H_j^{-1/2} A_j = U S_j W_j^{-1} \ , \ J_j^{-1/2} B_j = V M_j W_j^{-1/2}
$$

be the GSVD of the matrix pair $(H_j^{-1/2} A_j, J_j^{-1/2} B_j)$. Define the matrices $U_j = H_j^{-1/2} U$ and $V_j = J_j^{-1/2} V$. Then the GSVE of the operator pair (T_j, L_j) is given by

$$
\begin{aligned}
T_j &= \sum_{k=1}^{\min\{m_j, n_j\}} a_k^{(j)} \psi_k^{(j)} \otimes_{*_j} \phi_k^{(j)}, \\
L_j &= \sum_{k=1}^{\min\{p_j, n_j\}} b_k^{(j)} \theta_k^{(j)} \otimes_{*_j} \phi_k^{(j)}.
\end{aligned}
$$

The values $a_1^{(j)}, a_2^{(j)}, \cdots, a_{\min\{m_j, n_j\}}^{(j)}$ are the diagonal entries of S_j, the values $b_1^{(j)}, b_2^{(j)}, \cdots, b_{\min\{p_j, n_j\}}^{(j)}$ are the diagonal entries of M_j, and

$$
\begin{aligned}
\phi_k^{(j)} &= \sum_{i=1}^{n_j} (W_j)_{ik} x_k, \\
\psi_k^{(j)} &= \sum_{i=1}^{m_j} (U_j)_{ik} y_k, \\
\theta_k^{(j)} &= \sum_{i=1}^{p_j} (V_j)_{ik} z_k.
\end{aligned}
$$

The sets $\{\phi_1^{(j)}, \phi_2^{(j)}, \cdots, \phi_{n_j}^{(j)}\}$, $\{\psi_1^{(j)}, \psi_2^{(j)}, \cdots, \psi_{m_j}^{(j)}\}$, and $\{\theta_1^{(j)}, \theta_2^{(j)}, \cdots, \theta_{p_j}^{(j)}\}$ are orthonormal in X_j, Y, and Z, respectively, where the $*_j$-inner product is used on X_j.

The proof the Theorem 3 is similar to that of Theorem 4.4 of [9]. In that paper, the special case of $T_j = P_{Y_j} T|_{X_j}$ and $L_j = P_{Z_j} L|X_j$ is considered. However, the derivation of the GSVE of an arbitrary pair of discretized operators (T_j, L_j) is similar to the special case covered in [9]. For this reason, the proof of Theorem 3 is omitted.

The GSVE of the operator pair (T_j, L_j) can be seen to be directly related to the GSVD of the pair of matrices from the last theorem. One advantage for computing this GSVE is that the approximate generalized singular vectors computed are orthogonal with respect to the spaces X_j, Y, and Z_j with respect to the $*_j$-norm, Y-norm, and Z-norm respectively. It is this orthogonality that makes analysis and computations transparent in much of applied mathematics.

As noted at the beginning of the chapter, our main goal is to analyze the convergence of the GSVE of (T_j, L_j) to the GSVE of (T, L). The next example demonstrates that a seemingly natural discretization need not lead to convergence of the GSVE.

Example 1. *Let $X = D(L) = H^1(0,1)$ and $Y = Z = L^2(0,1)$. Define operators $T : X \to Y$ and $L : D(L) \to Z$ by $Tx = x$ and $Lx = x'$, respectively. By Rellich's lemma, T (the identity operator) is compact. In this example, the $*$-norm is precisely the $H^1(0,1)$-norm.*

We can easily derive the GSVE of (T, L) using Fourier analysis; the result is

$$T = \sum_{k=0}^{\infty} a_k \psi_k \otimes_* \phi_k,$$
$$L = \sum_{k=1}^{\infty} b_k \theta_k \otimes_* \phi_k,$$

where, for $k \geq 1$,

$$\phi_k(t) = \sqrt{\frac{2}{k^2\pi^2 + 1}} \cos(k\pi t),\ \ \psi_k = \sqrt{2}\cos(k\pi t),\ \ \theta_k(t) = -\sqrt{2}\sin(k\pi t),$$
$$a_k = \frac{1}{\sqrt{k^2\pi^2+1}},\ \ b_k = \frac{k\pi}{\sqrt{k^2\pi^2+1}},$$

and $\phi_0(t) = 1$, $\psi_0(t) = 1$, $a_0 = 1$, and $b_0 = 0$. It can be verified that $\{\phi_k\}_{k=1}^{\infty}$, $\{\psi_k\}_{k=1}^{\infty}$, and $\{\theta_k\}_{k=1}^{\infty}$ are orthonormal in the $*$, Y, and Z inner products, respectively. Also

$$a_k^2 + b_k^2 = 1,\ T\phi_k = a_k\psi_k,\ \text{and } L\phi_k = b_k\theta_k \text{ for all } k \in \mathbb{Z}^+.$$

In the notation of Theorem 1.11, we have $M_0 = \mathbb{Z}^+$, $M_a = \varnothing$, and $M_b = \{0\}$.

We discretize (T, L) by defining $X_j = Y_j = Z_j$ to be the space of continuous piecewise linear functions on a uniform mesh of $[0,1]$ with j elements. Let $\{x_0, x_1, \cdots, x_j\}$ be the standard nodal basis. Define $T_j = P_{Y_j}T|_{X_j}$ and $L_j = P_{Z_j}L|X_j$. We compute the GSVE of (T_{100}, L_{100}) as described in Theorem 3 and graph $\phi_k^{(100)}$, $\psi_k^{(100)}$, and $\theta_k^{(100)}$ for $k = 1, 2, 3$ (see Figures 1-3). We see that $\phi_1^{(100)}$, $\psi_1^{(100)}$, $\theta_1^{(100)}$ and $\phi_2^{(100)}$, $\psi_2^{(100)}$, $\theta_2^{(100)}$

are accurate approximations of the corresponding exact functions, but $\phi_3^{(100)}$, $\psi_3^{(100)}$,

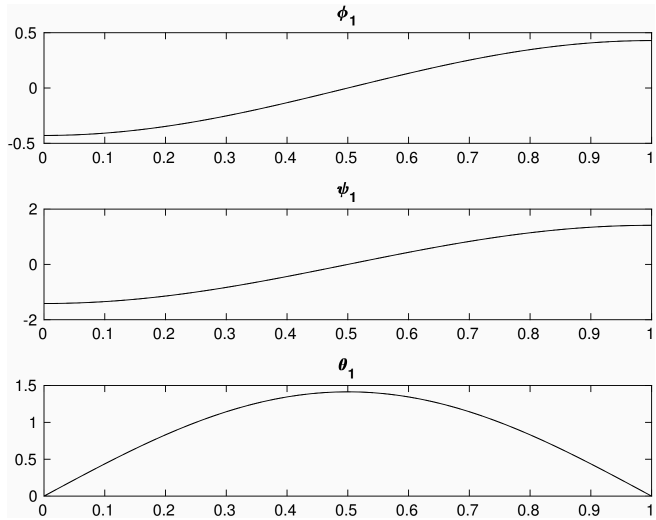

Figure 2.1: The computed functions $\phi_1^{(100)}$ (top), $\psi_1^{(100)}$ (middle), and $\theta_1^{(100)}$ (bottom) for Example 1, together with the corresponding exact functions ϕ_1, ψ_1, and θ_1. In each graph, the approximate function is the solid curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.

and $\theta_3^{(100)}$ *are completely wrong. The behavior seen in Figure 3 is consistent with the type of "spurious modes" observed in the numerical solution of variationally posed eigenvalue problems (see [10]). The spurious mode persists as the mesh is refined.*

Although we do not show any more results here, in fact every triple $(\phi_k^{(100)}, \psi_k^{(100)}, \theta_k^{(100)})$ *for* $k > 3$ *is far from the exact generalized singular functions* $(\phi_k, \psi_k, \theta_k)$*. Moreover, this behavior is not eliminated by refining the mesh. Every fourth generalized singular mode is spurious.*

In the next chapter, we analyze the convergence of the GSVE of (T_j, L_j) to that of (T, L), presenting a condition on the convergence of (T_j, L_j) to (T, L) that guarantees that the corresponding GSVEs converge. We will see that the condition fails for the discretization in Example 1 and also see how to modify the discretization to obtain convergence.

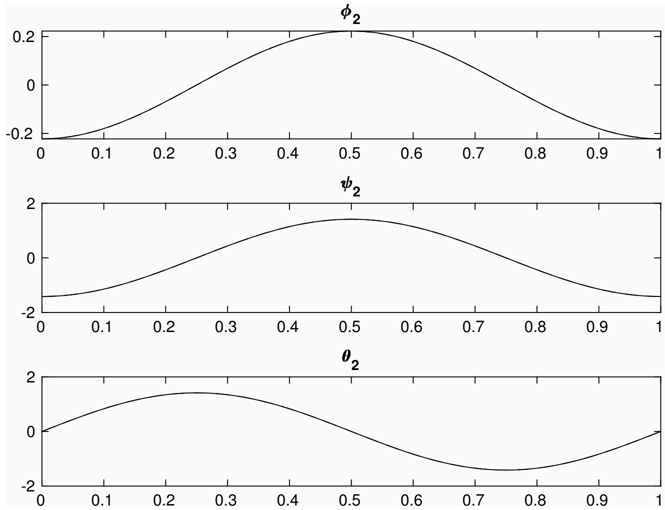

Figure 2.2: The computed functions $\phi_2^{(100)}$ (top), $\psi_2^{(100)}$ (middle), and $\theta_2^{(100)}$ (bottom) for Example 1, together with the corresponding exact functions ϕ_2, ψ_2, and θ_2. In each graph, the approximate function is the solid curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.

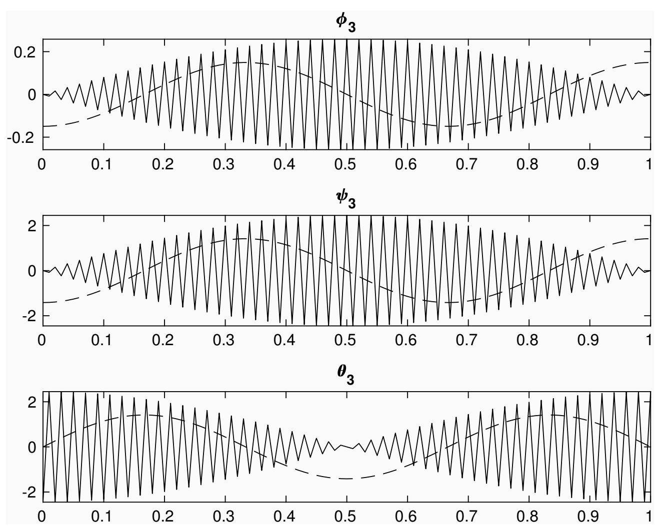

Figure 2.3: The computed functions $\phi_3^{(100)}$ (top), $\psi_3^{(100)}$ (middle), and $\theta_3^{(100)}$ (bottom) for Example 1, together with the corresponding exact functions ϕ_3, ψ_3, and θ_3. In each graph, the approximate function is the solid curve and the exact function is the dashed curve.

Chapter 3

Convergence

Let X, Y, and Z be separable Hilbert spaces, let $T : X \to Y$ be a compact linear operator, and let $L : D(L) \to Z$ be a closed linear operator, where $D(L)$ is a dense subspace of X. We assume that there exists $\gamma > 0$ such that

$$\langle Tx, Tx \rangle_Y + \langle Lx, Lx \rangle_Z \geq \gamma \|x\|_X^2 \quad \forall x \in D(L).$$

We define the bilinear form $\langle \cdot, \cdot \rangle_* : D(L) \times D(L) \to \mathbb{R}$ by

$$\langle x, y \rangle_* = \langle Tx, Ty \rangle_Y + \langle Lx, Ly \rangle_Z.$$

Condition (1.9) from Chapter 1, which has been restated above, guarantees $\langle \cdot, \cdot \rangle_*$ defines an inner product on $D(L)$. Then, by Theorem 2, the GSVE of (T, L) is given by

$$\begin{aligned} T &= \sum_{k \in M_0 \bigcup M_b} a_k \psi_k \otimes_* \phi_k, \\ L &= \sum_{k \in M_0 \bigcup M_a} b_k \theta_k \otimes_* \phi_k. \end{aligned}$$

Here, $\{(a_k, b_k)\}$ are the generalized singular values of (T, L), and the sets $\{\phi_k\}$, $\{\psi_k\}$, and $\{\theta_k\}$ are the generalized singular vectors of (T, L).

In the last section, we provided an algorithm for computing the approximate GSVE of (T_j, L_j), which is given by

$$\phi_k^{(j)} = \sum_{i=1}^{n_j} (W_j)_{ik} x_k, \quad \psi_k^{(j)} = \sum_{i=1}^{m_j} (U_j)_{ik} y_k, \quad \theta_k^{(j)} = \sum_{i=1}^{p_j} (V_j)_{ik} z_k.$$

We next provide sufficient conditions under which the GSVE of (T_j, L_j) is guaranteed to converge to the GSVE of (T, L). Informally, this means that the approximate generalized singular values $(a_k^{(j)}, b_k^{(j)})$ converge to the exact generalized singular values (a_k, b_k), and the approximate generalized singular vectors $\{\phi_k^{(j)}\}$, $\{\psi_k^{(j)}\}$, and $\{\theta_k^{(j)}\}$ converge to the exact generalized singular vectors $\{\phi_k\}$, $\{\psi_k\}$, and $\{\theta_k\}$. Convergence of the generalized singular vectors poses a complicated matter since the sets of generalized singular vectors correspond to subspaces of the Hilbert spaces $D(L)$, Y, and Z. The issues are comparable to those faced in approximating the eigenvalues and eigenvectors of a linear operator $A : X \to X$ by the eigenvalues and eigenvectors of an approximation A_j of A. We refer the reader to Boffi's survey article [10] for a detailed discussion. In the case of eigenvalues and eigenvectors, we can expect that the eigenvalues of A_j to converge to the corresponding eigenvalues of A in the expected manner. However, since a given eigenspace does not have a unique basis, there is no reason that the computed basis of the corresponding eigenspace of A_j to converge directly to a given basis of an eigenspace of A. Therefore, we have to refer to convergence of a sequence of subspaces to a given subspace, not the convergence of individual eigenvectors. Moreover, if λ is an eigenvalue of A of multiplicity k, then there are probably k simple eigenvalues of A_j that converge to λ as $j \to \infty$.

When discussing the convergence of the GSVE of (T_j, L_j) to the GSVE of (T, L), we have an additional complication, namely that both T and L can have a nontrivial null space. It is straightforward to show that $\mathcal{N}(L)$ must be finite-dimensional (otherwise, the inequality (1.9) is incompatible with the compactness of T). However, $\mathcal{N}(T)$ could be infinite-dimensional. We will assume throughout our discussion that $\mathcal{R}(T)$ is infinite-dimensional, since this is the interesting case in applications.

In terms of the GSVE of the operator pair (T, L),

$$\begin{aligned} T &= \sum_{k \in M_0 \bigcup M_b} a_k \psi_k \otimes_* \phi_k, \\ L &= \sum_{k \in M_0 \bigcup M_a} b_k \theta_k \otimes_* \phi_k, \end{aligned}$$

the generalized singular values of (T, L) have the following properties:

$$\begin{aligned} k \in M_b &\implies a_k = 1 \text{ and } b_k = 0, \\ k \in M_0 &\implies 0 < a_k, b_k < 1, \\ k \in M_a &\implies a_k = 0 \text{ and } b_k = 1. \end{aligned}$$

To compare the singular values of (T_j, L_j) with those of (T, L), we have to order the generalized singular values of (T, L) consistently. Since $\mathcal{N}(L)$ is finite-dimensional, we will assume that $\dim(\mathcal{N}(L)) = \ell$ and that $M_b = \{1, 2, \cdots, \ell\}$. Since $\mathcal{R}(T)$ is infinite-dimensional by assumption, we will define the index set M_0 by $M_0 = \{\ell+1, \ell+2, \ldots\}$ and assume that $a_{\ell+1} \geq a_{\ell+2} \geq \ldots$. Since $a_k^2 + b_k^2 = 1$, this implies that $b_{\ell+1} \leq b_{\ell+2} \leq$

. . ..

With these definitions for M_b and M_a, we see that $\{a_k : k \in \mathbb{Z}^+\}$ is a nonincreasing sequence of positive real numbers, and $\{b_k : k \in \mathbb{Z}^+\}$ is a nondecreasing sequence of nonnegative real numbers. However, if M_a is nonempty (that is, if T has a nontrivial null space), then there is no natural definition for M_a that maintains the monotonicity of the sequences $\{a_k\}$ and $\{b_k\}$. Therefore, we will continue to denote M_a as a (countable) abstract index set. We can now write the GSVE of (T, L) as follows:

$$T = \sum_{k=1}^{\infty} a_k \psi_k \otimes_* \phi_k, \tag{3.1}$$

$$L = \sum_{k=1}^{\infty} b_k \theta_k \otimes_* \phi_k \ + \sum_{k \in M_a} b_k \theta_k \otimes_* \phi_k. \tag{3.2}$$

Recall that the GSVE of (T_j, L_j) is given by

$$T_j = \sum_{k=1}^{\min\{m_j,n_j\}} a_k^{(j)} \psi_k^{(j)} \otimes_{*_j} \phi_k^{(j)},$$

$$L_j = \sum_{k=1}^{\min\{p_j,n_j\}} b_k^{(j)} \theta_k^{(j)} \otimes_{*_j} \phi_k^{(j)}.$$

Here, we assume that $a_1^{(j)} \geq a_2^{(j)} \geq \cdots \geq a_{n_j}^{(j)}$ and $b_1^{(j)} \leq b_2^{(j)} \leq \cdots \leq b_{n_j}^{(j)}$.

To describe the convergence of the singular vectors of (T_j, L_j) to those of (T, L), we will use the concept of the gap between two subspaces (see [10]).

Definition 4. *Let H be a Hilbert space, and let U and V be closed subspaces of H. The gap between U and V is defined to be $\hat{\delta}(U, V)$, where*

$$\delta(U, V) = \sup_{\substack{u \in U \\ \|u\|=1}} \inf_{v \in V} \|u - v\|,$$

$$\hat{\delta}(U, V) = \max\{\delta(U, V), \delta(V, U)\}.$$

We now introduce the concept of angle between subspaces U and V of a Hilbert space H. We define the asymmetric angle $\theta(U, V)$, denoted more simply by θ, by

$$\cos(\theta(U, V)) = \cos(\theta) = \inf_{\substack{u \in U \\ \|u\|=1}} \sup_{\substack{v \in V \\ \|v\|=1}} \langle u, v \rangle.$$

Notice that the above quantity is bounded between 0 and 1, so $\theta \in [0, \pi/2]$ is well defined. We next derive the following properties about the asymmetric angle and asymmetric gap.

Theorem 5. *Let U and V be closed subspaces of a Hilbert space H. Then the asymmetric angle $\theta = \theta(U,V)$ has the following property:*

$$\cos(\theta(U,V)) = \cos(\theta) = \inf_{\substack{u\in U\\ \|u\|=1}} \|P_V u\|.$$

Proof. Let $u \in U$ such that $\|u\| = 1$. If $\|P_V u\| = 0$, then $u \in V^\perp$ and hence

$$\langle u, v\rangle = \langle P_V u, v\rangle = 0.$$

Thus, $\sup_{\substack{v\in V\\ \|v\|_H=1}} \langle u, v\rangle_H = \|P_V u\|_H$ in this case. Suppose that $\|P_V u\| \neq 0$. Then,

$$\begin{aligned}\|P_V u\| = \langle P_V u, \frac{P_V u}{\|P_V u\|}\rangle &\leq \sup_{\substack{v\in V\\ \|v\|=1}} \langle P_V u, v\rangle \\ &\leq \sup_{\substack{v\in V\\ \|v\|=1}} \|P_V u\|\|v\| = \|P_V u\|.\end{aligned}$$

Since $\langle u, v\rangle = \langle P_V u, v\rangle$ for every $v \in V$, we have

$$\sup_{\substack{v\in V\\ \|v\|=1}} \langle u, v\rangle = \sup_{\substack{v\in V\\ \|v\|=1}} \langle P_V u, v\rangle = \|P_V u\|.$$

Thus we have shown that $\sup_{\substack{v\in V\\ \|v\|_H=1}} \langle u, v\rangle_H = \|P_V u\|_H$ in every case. We then have

$$\inf_{\substack{u\in U\\ \|u\|=1}} \|P_V u\| = \inf_{\substack{u\in U\\ \|u\|=1}} \sup_{\substack{v\in V\\ \|v\|=1}} \langle u, v\rangle = \cos(\theta).$$

□

Theorem 6. *Let U, V be closed subspaces of a Hilbert space H. Then*

$$\delta(U,V) = \sin(\theta(U,V)).$$

Proof. This follows directly from the definition of asymmetric gap. Writing $\theta = \theta(U,V)$,

$$\begin{aligned}\delta(U,V)^2 &= \left(\sup_{\substack{u\in U\\ \|u\|=1}} \inf_{v\in V} \|u - v\|\right)^2 \\ &= \sup_{\substack{u\in U\\ \|u\|=1}} \|u - P_V u\|^2\end{aligned}$$

$$
\begin{aligned}
&= \sup_{\substack{u\in U\\ \|u\|=1}} \left(1-\|P_V u\|^2\right)\\
&= 1-\inf_{\substack{u\in U\\ \|u\|=1}} \|P_V u\|^2\\
&= 1-\cos^2(\theta)=\sin^2(\theta).
\end{aligned}
$$

Therefore, $\delta(U,V)=\sin(\theta)$. □

Theorem 7. *If the asymmetric angles $\theta(U,V)$ and $\theta(V,U)$ are strictly less than $\pi/2$, then $\theta(U,V)=\theta(V,U)$ and*

$$\delta(U,V)=\delta(V,U).$$

Proof. Let $\theta=\theta(U,V)$ and let $\omega=\theta(V,U)$. Suppose that $\theta,\omega<\pi/2$. Then $\cos(\theta),\cos(\omega)>0$, and

$$\inf_{\substack{u\in U\\ \|u\|=1}} \|P_V u\|>0\ ,\ \inf_{\substack{v\in V\\ \|v\|=1}} \|P_U v\|>0.$$

Therefore, the projections P_U and P_V are bounded below when restricted to V and U respectively. Hence, $P_U:V\to U$ and $P_V:U\to V$ each have closed range in U and V, respectively, and are both injective. Also, we have for any $u\in U$ and for any $v\in V$,

$$\langle P_V u,v\rangle=\langle u,v\rangle=\langle u,P_U v\rangle.$$

Thus, as operators between the spaces U and V ,$P_V^*=P_U$ and $P_U^*=P_V$. Then

$$
\begin{aligned}
\mathcal{R}(P_U)=\mathcal{N}(P_V)^\perp=U,\\
\mathcal{R}(P_V)=\mathcal{N}(P_U)^\perp=V,
\end{aligned}
$$

which shows that P_U and P_V are bijections between U and V. Therefore, for any unit vector $u\in U$, there exists a unit vector $v\in V$ such that $u=\frac{P_U v}{\|P_U v\|}$. It follows that

$$
\begin{aligned}
\|P_V u\| &= \sup_{\substack{x\in V\\ \|x\|=1}} \langle P_V u,x\rangle\\
&= \sup_{\substack{x\in V\\ \|x\|=1}} \langle u,P_U x\rangle\\
&= \sup_{\substack{x\in V\\ \|x\|=1}} \langle \frac{P_U v}{\|P_U v\|},P_U x\rangle \geq \langle \frac{P_U v}{\|P_U v\|},P_U v\rangle=\|P_U v\|.
\end{aligned}
$$

Therefore we have

$$\inf_{\substack{v \in V \\ \|v\|=1}} \|P_U v\| \leq \|P_V u\| \quad \forall u \in U, \; \|u\| = 1$$
$$\implies \inf_{\substack{v \in V \\ \|v\|=1}} \|P_U v\| \leq \inf_{\substack{u \in U \\ \|u\|=1}} \|P_V u\|.$$

By symmetry of U and V in the above formulation, we then have

$$\inf_{\substack{v \in V \\ \|v\|=1}} \|P_U v\| \geq \inf_{\substack{u \in U \\ \|u\|=1}} \|P_V u\|.$$

Thus, $\cos(\theta(U, V)) = \cos(\theta(V, U))$, and we have

$$\delta(U, V) = \sin(\theta(U, V)) = \sin(\theta(V, U)) = \delta(V, U).$$

□

Corollary 8. *If $\delta(U, V)$ and $\delta(V, U)$ are strictly less than 1, then*

$$\delta(U, V) = \delta(V, U).$$

Proof. Suppose $\delta(U, V), \delta(V, U) < 1$. Then $\sin(\theta(U, V)), \sin(\theta(V, U)) < 1$ and therefore the angles $\theta(U, V)$ and $\theta(V, U)$ are strictly greater than 0, and the previous theorem then follows. □

Given the sequences $\{a_k\}$ and $\{b_k\}$ of singular values and the sequences $\{\phi_k\}$, $\{\psi_k\}$, and $\{\theta_k\}$ of singular vectors for (T, L), we define the corresponding singular spaces by

$$S_k(\phi) = \operatorname{span}\{\phi_i : a_i = a_k\}$$
$$S_k(\psi) = \operatorname{span}\{\psi_i : a_i = a_k\}$$
$$S_k(\theta) = \operatorname{span}\{\theta_i : a_i = a_k\}.$$

Typically, if a_k is a multiple singular value (that is, $\dim(S_k(\phi)) > 1$), then each approximate singular value $a_k^{(j)}$ converging to a_k will be a simple singular value of (T_j, L_j), meaning that

$$\dim\{\operatorname{span}\{\phi_i^{(j)} : a_i^{(j)} = a_k^{(j)}\} = 1.$$

For this reason, we define the approximate singular spaces of (T_j, L_j) by

$$S_k^{(j)}(\phi) = \operatorname{span}\{\phi_i^{(j)} : a_i^{(\ell)} \to a_k \text{ as } \ell \to \infty\}$$

$$S_k^{(j)}(\psi) = \text{span}\{\psi_i^{(j)} : a_i^{(\ell)} \to a_k \text{ as } \ell \to \infty\}$$
$$S_k^{(j)}(\theta) = \text{span}\{\phi_i^{(j)} : a_i^{(\ell)} \to a_k \text{ as } \ell \to \infty\}.$$

Note that because $a_k^2 + b_k^2 = 1$ for every $k \in \mathbb{Z}^+$, it follows that

$$\{i \in \mathbb{Z}^+ : a_i = a_k\} = \{i \in \mathbb{Z}^+ : b_i = b_k\}$$

Therefore, we could have defined the above subspaces with reference to $\{b_k\}$ instead of $\{a_k\}$.

We can now define what it means for the GSVE of (T_j, L_j) to converge to the GSVE of (T, L) (see [11], Definition 5).

Definition 9. *We say that the GSVE of (T_j, L_j), $j \in \mathbb{Z}^+$, converges to the GSVE of (T, L) if, for all $N \in \mathbb{Z}^+$ and all $\varepsilon > 0$, there exists an integer j_0 such that for all integers $j \geq j_0$,*

$$\left|a_k^{(j)} - a_k\right| < \varepsilon \text{ for every } k = 1, 2, \cdots, N,$$
$$\left|b_k^{(j)} - b_k\right| < \varepsilon \text{ for every } k = 1, 2, \cdots, N,$$
$$\hat{\delta}\left(S_k^{(j)}(\phi), S_k(\phi)\right) < \varepsilon \text{ for every } k = 1, 2, \cdots, N,$$
$$\hat{\delta}\left(S_k^{(j)}(\psi), S_k(\psi)\right) < \varepsilon \text{ for every } k = 1, 2, \cdots, N,$$
$$\hat{\delta}\left(S_k^{(j)}(\theta), S_k(\theta)\right) < \varepsilon \text{ for every } k = 1, 2, \cdots, N.$$

In computing the gaps, we use the $*$, Y, and Z norms for

$$\hat{\delta}\left(S_k^{(j)}(\phi), S_k(\phi)\right) \ , \ \hat{\delta}\left(S_k^{(j)}(\psi), S_k(\psi)\right) \ , \ \hat{\delta}\left(S_k^{(j)}(\theta), S_k(\theta)\right)$$

respectively.

Notice that Definition (9) does not refer to $\{\phi_k : k \in M_a\}$ or $\{\theta_k : k \in M_a\}$. Our theory with show that, in the representation

$$T = \sum_{k=1}^{\infty} a_k \psi_k \otimes_* \phi_k,$$
$$L = \sum_{k=1}^{\infty} b_k \theta_k \otimes_* \phi_k \ + \ \sum_{k \in M_a} b_k \theta_k \otimes_* \phi_k,$$

the series for T and the first series in the representation of L are approximated. It is not guaranteed that we can approximate the second series in the representation of L.

For each $j \in \mathbb{Z}^+$, we refer to three different inner products on the space X_j, namely the $*$-inner product, the $*_j$-inner product, and the X-inner product. Therefore, there are three different adjoint operators for the operator T_j . The adjoint of T with respect to the $*$-inner product is denoted by $T^{\#}$, the adjoint of T_j with respect to the $*_j$-inner product by $T_j^{\#_j}$, and the adjoint of T with respect to the X-inner product by T^*. To study the convergence of the GSVE of (T_j, L_j) to that of (T, L), we consider the operators $T_j^{\#_j} T_j$ and $T^{\#} T$. Using the expansion for T as in equation (3.1), we see that the eigenpairs of $T^{\#} T$ are $(a_k^2, \phi_k), k = 1, 2, \ldots$. Similarly, the eigenpairs for the operator $T_j^{\#_j} T_j$ are $\left((a_k^{(j)})^2, \phi_k^{(j)}\right), k = 1, 2, \ldots, n_j$. Our goal is to show that the eigensystem of $T_j^{\#_j} T_j$ converges to that of $T^{\#} T$; we can then show that the GSVE of (T_j, L_j) converges to the GSVE of (T, L).

We note that the operators $T_j^{\#_j} T_j : X_j \to X_j$ and $T_j^{\#_j} T_j P_{X_j} : X \to X$, where P_{X_j} is the orthogonal projection onto X_j with respect to the X-inner product, have the same eigenpairs. Indeed, since $T_j^{\#_j} T_j$ is just the restriction of $T_j^{\#_j} T_j P_{X_j}$ to X_j, it is immediate that an eigenpair of $T_j^{\#_j} T_j$ is an eigenpair of $T_j^{\#_j} T_j P_{X_j}$. Conversely, if $T_j^{\#_j} T_j P_{X_j} x = \lambda x$, then, since $T_j^{\#_j} T_j P_{X_j}$ maps X into X_j , it follows that $x \in X_j$, and hence (λ, x) is also an eigenpair of $T_j^{\#_j} T_j$.

The theory of Babuska and Osborn ([12]; see also [10], Sections 6 and 9) shows that if a sequence $\{A_j\}$ of compact operators $A_j : X \to X$ converges in the operator norm to the compact operator $A : X \to X$, then eigensystems of A_j converge to the eigensystem of A as $j \to \infty$, provided we exclude the zero eigenvalues of A from consideration. Specifically, we have the following theorem ([10], Theorem 9.1) (in which $\rho(A)$ denotes the resolvent set of A).

Theorem 10. *Let $A : X \to X$ be a compact linear operator, and let $\{A_j\}$ be a sequence of compact linear operators from X to X such that*

$$\|A - A_j\|_{\mathcal{L}(X,X} \to 0 \text{ as } j \to \infty.$$

Then for any compact set $K \subseteq \rho(A)$, there exists $j_0 \in \mathbb{Z}^+$ such that for every $j \geq j_0$, we have $K \subseteq \rho(A_j)$. If λ is a non-zero eigenvalue of A with multiplicity m, then there are m eigenvalues $\lambda_1^{(j)}, \lambda_2^{(j)}, \cdots, \lambda_m^{(j)}$ of A_j, repeated according to their algebraic multiplicities, such that each $\lambda_i^{(j)}$ converges to λ as $j \to \infty$. Moreover, if we define $E_j(\lambda)$ to be the direct sum of the eigenspaces corresponding to the eigenvalues of $\lambda_1^{(j)}, \lambda_2^{(j)}, \cdots, \lambda_m^{(j)}$, then the gap between $E_j(\lambda)$ and the eigenspace $E(\lambda)$ corresponding to the eigenvalue λ tends to 0 as $j \to \infty$.

By the above discussion, if we show that $A_j = T_j^{\#_j} T_j P_{X_j}$ converges to $A = T^{\#} T$ in the operator norm, then it will follow that the eigensystem of $T_j^{\#_j} T_j P_{X_j}$ converges to the eigensystem of $T^{\#} T$. We will use the following fundamental result (See [13]).

Theorem 11. *Let U, V, and W be Hilbert spaces. Let $M : V \to W$ be a bounded linear operator, let $T : U \to V$ be a compact linear operator, and let $M_j : V \to W$ be a bounded linear operator for each $j \in \mathbb{Z}^+$. Suppose that $M_j \to M$ pointwise on V. Then*

$$\|(M_j - M)T\|_{\mathcal{L}(U,W)} \to 0 \text{ as } j \to \infty.$$

Example 1 shows that the GSVE of (T_j, L_j) need not converge to the GSVE of (T, L). We now describe the fundamental assumption on the sequences $\{T_j\}$ and $\{L_j\}$ that will allow us to prove convergence of the GSVE. For each $j \in \mathbb{Z}^+$, we define

$$t_{j,1} = \max_{\substack{x \in X_j \\ x \neq 0}} \frac{\|(T - T_j)x\|_Y}{\|x\|_X}, \tag{3.3}$$

$$t_{j,2} = \max_{\substack{x \in X_j \\ x \neq 0}} \frac{\|(T - T_j)x\|_Y}{\|x\|_*}, \tag{3.4}$$

$$t_j = \max\{t_{j,1}, t_{j,2}\}, \tag{3.5}$$

$$\ell_j = \max_{\substack{x \in X_j \\ x \neq 0}} \frac{\|(L - L - j)x\|_Z}{\|x\|_*}, \tag{3.6}$$

$$c_j = \sqrt{t_j^2 + \ell_j^2}. \tag{3.7}$$

Henceforth, we will assume that $c_j \to 0$ as $j \to \infty$. We will see that this is enough to imply that the GSVE of (T_j, L_j) converges to the GSVE of (T, L).

By (3.5), we have

$$\|(T_j - T)x\|_Y \le t_j \|x\|_X \text{ and } \|(T_j - T)x\|_Y \le t_j \|x\|_* \text{ for all } x \in X_j,$$

and, by 3.6,

$$\|(L_j - L)x\|_Z \le \ell_j \|x\|_* \text{ for all } x \in X_j.$$

Therefore, for all $x \in X_j$,

$$\begin{aligned} \|T_j x\|_Y^2 = \langle T_j x, T_j x\rangle_Y &= \langle (T_j - T)x, T_j x\rangle_Y + \langle Tx, T_j x\rangle_Y \\ &\le \|(T_j - T)x\|_Y \|T_j x\|_Y + \|Tx\|_Y \|T_j x\|_Y \\ &\le t_j \|x\|_* \|T_j x\|_Y + \|Tx\|_Y \|T_j x\|_Y. \end{aligned}$$

Hence,

$$\|T_j x\|_y \le t_j \|x\|_* + \|Tx\|_Y \le (1 + t_j)\|x\|_* \text{ for all } x \in X_j \tag{3.8}$$

(since obviously $\|Tx\|_Y \le \|x\|_*$ for all $x \in D(L)$). Similarly,

$$\|L_j x\|_Z \le (1+\ell_j)\|x\|_* \text{ for all } x \in X_j. \tag{3.9}$$

In our analysis, it will useful to define the quantity η_j by

$$\eta_j = c_j^2 + 2(t_j + \ell_j). \tag{3.10}$$

We will need the following bound.

Lemma 12. *For every $j \in \mathbb{Z}^+$, and for every $x, y \in X_j$,*

$$|\langle x, y\rangle_* - \langle x, y\rangle_{*_j}| \le \eta_j \|x\|_* \|y\|_*.$$

Proof. Let $j \in \mathbb{Z}^+$, and let $x, y \in X_j$. Then

$$\begin{aligned}
|\langle x, y\rangle_* - \langle x, y\rangle_{*_j}| &= |\langle Tx, Ty\rangle_Y + \langle Lx, Ly\rangle_Z - \langle T_j x, T_j y\rangle_Y - \langle L_j x, L_j y\rangle_Z| \\
&= |\langle Tx, (T-T_j)y\rangle_Y + \langle Lx, (L-L_j)y\rangle_Z + \langle (T-T_j)x, T_j y\rangle_Y \\
&\qquad + \langle (L-L_j)x, L_j y\rangle_Z| \\
&\le \|Tx\|_Y \|(T-T_j)y\|_Y + \|Ly\|_Z \|(L-L_j)y\|_Z \\
&\qquad + \|(T-T_j)x\|_Y \|T_j x\|_Y + \|(L-L_j)x\|_Z \|L_j y\|_Z \\
&\le t_j \|Tx\|_Y \|y\|_* + \ell_j \|Lx\|_Z \|y\|_* + t_j \|x\|_* \|T_j y\|_Y + \ell_j \|x\|_* \|L_j y\|_Z \\
&\le (t_j^2 + \ell_j^2)\|x\|_* \|y\|_* + t_j(1+t_j)\|x\|_* \|y\|_* + \ell_j(1+\ell_j)\|x\|_* \|y\|_* \\
&= \left(t_j^2 + \ell_j^2 + 2(t_j + \ell_j)\right) \|x\|_* \|y\|_* \\
&= \eta_j \|x\|_* \|y\|_*.
\end{aligned}$$

□

From this, we have the following Corollary.

Corollary 13. *For any $j \in \mathbb{Z}^+$ and for any $x \in X_j$,*

$$(1-\eta_j)\|x\|_*^2 \le \|x\|_{*_j}^2 \le (1+\eta_j)\|x\|_*^2, \tag{3.11}$$

$$\frac{1}{1+\eta_j}\|x\|_{*_j} \le \|x\|_*^2 \le \frac{1}{1-\eta_j}\|x\|_{*_j}^2. \tag{3.12}$$

Next, we define $M_j : X_j \to X_j$ and $M : D(L^*L) \to X$ by

$$\begin{aligned}
M &= T^*T + L^*L \\
M_j &= T_j^* T_j + L_j^* L_j.
\end{aligned}$$

Let $x \in D(L)$ and $y \in D(L^*L)$. Notice that

$$\langle x, y\rangle_* = \langle Tx, Ty\rangle_Y + \langle Lx, Ly\rangle_Z = \langle x, (T^*T + L^*L)y\rangle_X$$

$$= \langle x, My \rangle_X.$$

Therefore, M has the following property:

$$\langle x, y \rangle_* = \langle x, My \rangle_X \quad \forall x \in D(L) \quad \forall y \in D(L^* L). \tag{3.13}$$

Similarly, the operator M_j has the following property:

$$\langle x, y \rangle_{*_j} = \langle x, M_j y \rangle_X \quad \forall x, y \in X_j. \tag{3.14}$$

These operators will be central to our analysis; the following three results come from these two properties of M and M_j.

Theorem 14. *The operator M is a bijection with bounded inverse, and*

$$\|M^{-1}\|_{\mathcal{L}(X, D(L))} < \frac{1}{\sqrt{\gamma}},$$

that is,

$$\|M^{-1}x\|_* \leq \frac{\|x\|_X}{\sqrt{\gamma}} \quad \forall x \in X.$$

Proof. See [4], Theorem 5.25. □

Theorem 15. *For each $j \in \mathbb{Z}^+$, the operator M_j is invertible and*

$$\|M_j\|_{\mathcal{L}(X, D(L))} \leq \frac{1}{(1 - \eta_j)\sqrt{\gamma}}.$$

That is,

$$\|M_j^{-1}x\|_* \leq \frac{1}{(1 - \eta_j)\sqrt{\gamma}} \|x\|_X \quad \forall x \in X_j.$$

Proof. Let $j \in \mathbb{Z}^+$ and let $x \in X_j$. By Corollary 13, we have

$$\begin{aligned}
\|M_j^{-1}x\|_*^2 \leq \frac{1}{1 - \eta_j} \|M_j^{-1}x\|_{*_j}^2 &= \frac{1}{1 - \eta_j} \langle M_j^{-1}x, M_j^{-1}x \rangle_{*_j} \\
&= \frac{1}{1 - \eta_j} \langle M_j^{-1}x, x \rangle_X \\
&\leq \frac{1}{1 - \eta_j} \|M_j^{-1}x\|_X \|x\|_X \\
&\leq \frac{1}{(1 - \eta_j)\sqrt{\gamma}} \|M_j^{-1}x\|_* \|x\|_X.
\end{aligned}$$

The desired result follows from dividing both sides of the inequality by $\|M_j^{-1}x\|_*$. $\square$

Next, recall that $\Pi_{X_j} : D(L) \to X_j$ denotes the orthogonal projection with respect to the $*$-inner product onto the subspace X_j . The following result allows us to compare M^{-1} and M_j^{-1}.

Theorem 16. *For every $x \in X$,*

$$\|\Pi_{X_j} M^{-1}x - M_j^{-1} P_{X_j} x\|_* \leq \frac{\eta_j}{(1-\eta_j)\sqrt{\gamma}} \|x\|_X.$$

Proof. Let $x \in X$. Then

$$\begin{aligned}
&\|\Pi_{X_j} M^{-1}x - M_j^{-1} P_{X_j} x\|_*^2 \\
&= \langle (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_* \\
&= \langle M^{-1}x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_* - \langle M_j^{-1} P_{X_j} x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_*
\end{aligned}$$

Notice that

$$\begin{aligned}
\langle M^{-1}x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_* &= \langle x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_X \\
&= \langle P_{X_j} x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_X \\
&= \langle M_j^{-1} P_{X_j} x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_{*_j}.
\end{aligned}$$

Therefore we have

$$\begin{aligned}
&\|\Pi_{X_j} M^{-1}x - M_j^{-1} P_{X_j} x\|_*^2 \\
&= \langle M^{-1}x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_* - \langle M_j^{-1} P_{X_j} x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_* \\
&= \langle M_j^{-1} P_{X_j} x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_{*_j} - \langle M_j^{-1} P_{X_j} x, (\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x \rangle_* \\
&\leq \eta_j \|M_j^{-1} P_{X_j} x\|_* \|(\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x\|_*,
\end{aligned}$$

where we have applied Lemma 12 for the last inequality. Hence,

$$\|(\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x\|_* \leq \eta_j \|M_j^{-1} P_{X_j} x\|_*.$$

Applying Theorem 15 (and the fact that $\|P_{X_j} x\|_X \leq \|x\|_X$), we obtain

$$\|(\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x\|_* \leq \frac{\eta_j}{(1-\eta_j)\sqrt{\gamma}} \|x\|_X,$$

as desired. $\square$

We can now prove that $M_j P_{X_j}$ converges pointwise to M^{-1} on X.

Theorem 17. *For every $x \in X$,*

$$\|M^{-1}x - M_j^{-1}P_{X_j}x\|_* \to 0 \text{ as } j \to \infty.$$

Proof. Let $x \in X$. Then

$$\begin{aligned}\|(M^{-1} - M_j^{-1}P_{X_j})x\|_* &\le \|M^{-1}x - \Pi_{X_j}M^{-1}x\|_* + \|(\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\|_* \\ &\le \|(I - \Pi_{X_j})M^{-1}x\|_* + \frac{\eta_j}{1-\eta_j}\|x\|_X.\end{aligned}$$

By assumption, Π_{X_j} converges pointwise to the identity operator on $D(L)$, and $\eta_j \to 0$ as $j \to \infty$. The desired result then follows. □

For every $y \in Y_j$ and for every $x \in X_j$, we have

$$\begin{aligned}\langle T_jx, y\rangle_Y = \langle x, T_j^{\#_j}y\rangle_{*_j} &= \langle T_jx, T_jT_j^{\#_j}y\rangle_Y + \langle L_jx, L_jT_j^{\#_j}y\rangle_Z \\ &= \langle x, (T_j^*T_j + L_j^*L_j)T_j^{\#_j}y\rangle_X.\end{aligned}$$

Also,

$$\langle T_jx, y\rangle_Y = \langle x, T_j^*y\rangle_X.$$

Because this is true for every $x \in X_j$ and for every $y \in Y_j$, we see that

$$T_j^* = (T_j^*T_j + L_j^*L_j)T_j^{\#_j} = M_jT_j^{\#_j}.$$

Similarly,

$$T^* = (T^*T + L^*L)T^{\#} = MT^{\#}.$$

We define $S_j : X_j \to Y$ by $S_j = T_j - T|_{X_j}$. By definition, we have that $t_j = \|S_j\|_{\mathcal{L}(X_j,Y)}$ and hence, by assumption, $\|S_j\|_{\mathcal{L}(X_j,Y)} \to 0$ as $j \to \infty$. We now compute the adjoint of S_j. To do this, let $x \in X_j$ and let $y \in Y$. Then

$$\begin{aligned}\langle S_jx, y\rangle_Y = \langle (T_j - T)x, y\rangle_Y &= \langle T_jx, y\rangle_Y - \langle Tx, y\rangle_Y \\ &= \langle T_jx, P_{Y_j}y\rangle_Y - \langle x, T^*y\rangle_X \\ &= \langle x, T_j^*P_{Y_j}y\rangle_X - \langle x, P_{X_j}T^*x\rangle_X \\ &= \langle x, (T_j^*P_{Y_j} - P_{X_j}T^*)y\rangle_X.\end{aligned}$$

Therefore, $S_j^* = T_j^*P_{Y_j} - P_{X_j}T^*$, and since $\|S_j^*\|_{\mathcal{L}(Y,X_j)} = \|S_j\|_{\mathcal{L}(X_j,Y)}$, we see that

$$\|P_{X_j}T^* - T_j^*P_{Y_j}\|_{\mathcal{L}(Y,X_j)} \to 0 \text{ as } j \to \infty. \tag{3.15}$$

The following theorem will be used to show that $T_j^{\#_j} T_j P_{X_j} \to T^{\#} T$ uniformly.

Theorem 18. *$T_j^{\#_j} P_{Y_j} T \to T^{\#} T$ in the $\mathcal{L}(D(L), D(L))$ norm.*

Proof. We have shown that $T^{\#} = M^{-1} T^*$ and $T_j^{\#_j} = M_j^{-1} T_j^*$. From this, it follows that

$$T_j^{\#_j} P_{Y_j} T - T^{\#} T = (M_j^{-1} T_j^* P_{Y_j} - M^{-1} T^*) T.$$

By Theorem 11, it suffices to prove that $M_j^{-1} T_j^* P_{Y_j} \to M^{-1} T^*$ pointwise on Y as $j \to \infty$. Let $y \in Y$. Then

$$\begin{aligned}
&\|M^{-1} T^* y - M_j^{-1} T_j^* P_{Y_j} y\|_* \\
&\leq \|(M^{-1} - M_j^{-1} P_{X_j}) T^* y\|_* + \|M_j^{-1} P_{X_j} T^* y - M_j^{-1} T_j^* P_{Y_j} y\|_* \\
&= \|(M^{-1} - M_j^{-1} P_{X_j}) T^* y\|_* + \|M_j^{-1} (P_{X_j} T^* y - T_j^* P_{Y_j} y)\|_* \\
&\leq \|(M^{-1} - M_j^{-1} P_{X_j}) T^* y\|_* + \frac{1}{(1-\eta_j)\sqrt{\gamma}} \|(P_{X_j} T^* - T_j^* P_{Y_j}) y\|_X.
\end{aligned}$$

It now follows from Theorem 17 and (3.15) that $\|M^{-1} T^* y - M_j^{-1} T_j^* P_{Y_j} y\|_* \to 0$ as $j \to \infty$. $\square$

We need two more results.

Lemma 19. *If $\{v_j\} \subseteq X$ and $v_j \to v$ weakly as $j \to \infty$, then $P_{X_j} v_j \to v$ weakly.*

Proof. For any $x \in X$, we have

$$\langle P_{X_j} v_j, x \rangle_X = \langle v_j, x \rangle_X + \langle v_j, (P_{X_j} - I) x \rangle_X \to \langle v, x \rangle_X$$

(notice that $\{v_j\}$ is a bounded sequence in X, and that $(P_{X_j} - I)x \to 0$ in norm). This shows that $P_{X_j} v_j \to v$ weakly as $j \to \infty$. $\square$

Theorem 20. *$T_j P_{X_j} \to T$ in the $\mathcal{L}(X, Y)$ norm.*

Proof. We argue by contradiction and assume that there exist $\varepsilon_0 > 0$ and a subsequence $\{j_k\}$ of $\mathbb{Z}^+$ such that for every $k \in \mathbb{Z}^+$, there exists $v_{j_k} \in X$ satisfying

$$\|v_{j_k}\|_X = 1 \text{ and } \|T_{j_k} P_{X_{j_k}} v_{j_k} - T v_{j_k}\|_Y \geq \varepsilon_0. \tag{3.16}$$

Since T is compact and X is separable, without loss of generality, we can assume that there exists $v \in X$ and $y \in Y$ such that $v_{j_k} \to v$ weakly in X and $T v_{j_k} \to y$ in Y. We then have

$$T_{j_k} P_{X_{j_k}} v_{j_k} = T P_{X_{j_k}} v_{j_k} + (T_{j_k} - T) P_{X_{j_k}} v_{j_k} \to Tv + 0 = y$$

($\|(T_{j_k} - T)P_{X_{j_k}} v_{j_k}\|_Y \le t_j \|P_{X_{j_k}} v_{j_k}\|_X \to 0$, and $TP_{X_{j_k}} \to Tv$ because $P_{X_{j_k}} v_{j_k} \to v$ weakly from Lemma 19 and T is compact). But then we have

$$T_{j_k} P_{X_{j_k}} v_{j_k} - Tv_{j_k} \to y - y = 0,$$

contradicting (3.16). The contradiction completes the proof. □

We have been working towards the following result.

Theorem 21. *$T_j^{\#_j} T_j P_{X_j} \to T^{\#}T$ in the $\mathcal{L}(D(L), D(L))$ norm.*

Proof. We have

$$\begin{aligned}
&\|T_j^{\#_j} T_j P_{X_j} - T^{\#}T\|_{\mathcal{L}(D(L),D(L))} \\
&\le \|T_j^{\#_j} T_j P_{X_j} - T_j^{\#_j} P_{Y_j} T\|_{\mathcal{L}(D(L),D(L))} + \|T_j^{\#_j} P_{Y_j} T - T^{\#}T\|_{\mathcal{L}(D(L),D(L))} \\
&= \|M_j^{-1} T_j^* P_{Y_j}(T_j P_{X_j} - T)\|_{\mathcal{L}(D(L),D(L))} + \|T_j^{\#_j} P_{Y_j} T - T^{\#}T\|_{\mathcal{L}(D(L),D(L))}.
\end{aligned}$$

The second term to the right of the equals sign goes to 0 by Theorem 18. Therefore, it suffices to show that the first term goes to 0. Applying Theorem 15, we have

$$\begin{aligned}
&\|M_j^{-1} T_j^* P_{Y_j}(T_j P_{X_j} - T)\|_{\mathcal{L}(D(L),D(L))} \\
&\le \frac{1}{(1-\eta_j)\sqrt{\gamma}} \|T_j^* P_{Y_j}(T_j P_{X_j} - T)\|_{\mathcal{L}(D(L),X)} \\
&\le \frac{\|T_j^*\|_{\mathcal{L}(Y_j,X_j)}}{(1-\eta_j)\sqrt{\gamma}} \|T_j P_{X_j} - T\|_{\mathcal{L}(D(L),Y)} \\
&= \frac{\|T_j\|_{\mathcal{L}(X_j,Y_j)}}{(1-\eta_j)\sqrt{\gamma}} \|T_j P_{X_j} - T\|_{\mathcal{L}(D(L),Y)} \\
&\le \frac{t_j + \|T\|_{\mathcal{L}(X,Y)}}{(1-\eta_j)\sqrt{\gamma}} \|T_j P_{X_j} - T\|_{\mathcal{L}(D(L),Y)} \\
&\le \frac{t_j + \|T\|_{\mathcal{L}(X,Y)}}{(1-\eta_j)\gamma} \|T_j P_{X_j} - T\|_{\mathcal{L}(X,Y)}.
\end{aligned}$$

Since $t_j \to 0$ as $j \to \infty$ and T is a bounded operator from X to Y, it follows from Theorem 20 that $\|M_j^{-1} T_j^* P_{Y_j}(T_j P_{X_j} - T)\|_{\mathcal{L}(D(L),D(L))} \to 0$ as $j \to \infty$. This completes the proof. □

Theorem 3 and Theorem 10 show that the eigensystem of $T_j^{\#_j} T_j$, which is the same as the eigensystem of $T_j^{\#_j} T_j P_{X_j}$, converges to the eigensystem of $T^{\#}T$. We can now prove the following theorem.

Theorem 22. *Assuming that $c_j \to 0$ as $j \to \infty$ (where c_j is defined by (3.7)), the GSVE of (T_j, L_j) converges to the GSVE of (T, L) in the sense of Definition 9.*

Proof. Since $((a_k^{(j)})^2, \phi_k^{(j)})$, $k = 1, 2, \cdots, n_j$, are the eigenpairs of $T_j^{\#_j} T_j P_{X_j}$, (a_k^2, ϕ_k) are the eigenpairs of $T^{\#}T$, and $T_j^{\#_j} T_j P_{X_j} \to T^{\#}T$ in the operator norm, it follows from Theorem 10 that the set of approximate generalized singular values $\{a_k^{(j)}\}$ converges to the set of true generalized singular values $\{a_k\}$, and the set of approximate generalized singular functions $\{\phi_k^{(j)}\}$ converges to the set of true generalized singular functions $\{\phi_k\}$ in the manner described by Definition 9. Moreover, since $\left(a_k^{(j)}\right)^2 + \left(b_k^{(j)}\right)^2 = 1$ for every $k = 1, 2, \cdots, n_j$, and $a_k^2 + b_k^2 = 1$ for every $k \in \mathbb{Z}^+$, it follows that the set of approximate generalized singular values $\{b_k^{(j)}\}$ also converges to the true set of generalized singular values $\{b_k\}$ in the manner described in Definition 9.

It now remains only to show that $\{\psi_k^{(j)}\}$ converges to $\{\psi_k\}$ and $\{\theta_k^{(j)}\}$ converges to $\{\theta_k\}$ as $j \to \infty$ in the sense of Definition 9. To show this, let k be an arbitrary positive integer and let $\varepsilon > 0$ be given. We must show that there exists $j_0 \in \mathbb{Z}^+$ such that

$$j \geq j_0 \implies \max\left\{\delta\left(S_k(\psi), S_k^{(j)}(\psi)\right), \delta\left(S_k^{(j)}(\psi), S_k(\psi)\right)\right\} < \varepsilon.$$

First, we show that $j_0 \in \mathbb{Z}^+$ can be chosen such that $\delta(S_k(\phi), S_k^{(j)}(\phi)) < \varepsilon$ for every $j \geq j_0$. That is, we show that j_0 can be chosen so that

$$j \geq j_0 \implies \sup_{\substack{y \in S_k(\psi) \\ \|y\|_Y = 1}} \inf_{v \in S_k^{(j)}} \|y - v\|_Y < \varepsilon. \tag{3.17}$$

We know that there exists $j_0 \in \mathbb{Z}^+$ such that

$$j \geq j_0 \implies \sup_{\substack{x \in S_k(\phi) \\ \|x\|_* = 1}} \inf_{v \in S_k^{(j)}(\phi)} \|x - v\|_* < \frac{a_k \varepsilon}{4} \text{ and } t_j < \min\left\{\frac{a_k \varepsilon}{2}, 1\right\}.$$

We will show that this value of j_0 satisfies (3.17). It suffices to show that for any $j \geq j_0$ and for any $y \in S_k(\psi)$ satisfying $\|y\|_Y = 1$, there exists $v \in S_k^{(j)}(\psi)$ such that $\|y - v\|_Y < \varepsilon$. Suppose

$$S_k(\psi) = \operatorname{span}\{\psi_{k_1}, \psi_{k_2}, \cdots, \psi_{k_q}\}.$$

Then there exists real numbers $\alpha_1, \alpha_2, \cdots, \alpha_q$ such that

$$y = a_k^{-1} \sum_{i=1}^{q} \alpha_i a_k \psi_{k_i} = a_k^{-1} \sum_{i=1}^{q} \alpha_i T \phi_{k_i} = a_k^{-1} T x,$$

where $x = \sum_{i=1}^{q} \alpha_i \phi_{k_i}$. Moreover, since $\{\phi_{k_1}, \phi_{k_2}, \cdots, \phi_{k_q}\}$ is orthonormal in $D(L)$ with respect to the $*$-inner product, we see that $\|x\|_* = \|y\|_Y = 1$. Hence, there exists

$u \in S_k^{(j)}(\phi)$ such that

$$\|x - u\|_* < \frac{a_k \varepsilon}{4}.$$

By construction, $Tx = a_k y$, and the vector v defined by $v = a_k^{-1} T_j u$ lies in $S_k^{(j)}(\psi)$. Moreover,

$$\begin{aligned}
\|y - v\|_Y = a_k^{-1} \|Tx - T_j u\|_Y &\leq \frac{\|(T - T_j)x\|_Y + \|T_j x - T_j v\|_Y}{a_k} \\
&\leq \frac{t_j + \|T_j\|_{\mathcal{L}(D(L),Y)} \|x - u\|_*}{a_k} \\
&\leq \frac{t_j + (1 + t_j)\|x - u\|_*}{a_k} \\
&\leq \frac{1}{a_k} \left(\frac{a_k \varepsilon}{2} + \frac{2 a_k \varepsilon}{4} \right) = \varepsilon
\end{aligned}$$

Thus, $\delta(S_k(\psi), S_k^{(j)}(\psi)) < \varepsilon$ for every $j \geq j_0$. The proof that j_0 can be chosen such that $\delta(S_k^{(j)}(\theta), S_k(\theta)) < \varepsilon$ for every $j \geq j_0$ is similar. Thus, we have shown that $\{\psi_k^{(j)}\}$ converges to $\{\psi_k\}$ in the sense of Definition 9.

The proof that $\{\theta_k^{(j)}\}$ converges to $\{\theta_k\}$ in the sense of Definition 9 is exactly the same, and the proof is complete. □

In Example 1, it appeared that the GSVE of (T_j, L_j) did not converge to the GSVE of (T, L). Thus, the sequence of discretized operator pairs (T_j, L_j) must fail to satisfy the hypotheses of Theorem 22.

Example 2. *In this example, we analyze the discretization of Example 1. In Example 1, $T_j = P_{Y_j} T|_{X_j}$, and since T is the identity operator, it follows that $T_j = T|_{X_j}$ ($X_j = Y_j$ for each $j \in \mathbb{Z}^+$). Therefore, $t_j = 0$ for every $j \in \mathbb{Z}^+$. However (recalling that x_i is the ith standard nodal basis function), a direct calculation shows that*

$$\ell_j = \sup_{\substack{x \in X_j \\ x \neq 0}} \frac{\|(L_j - L)x\|_{L^2(0,1)}}{\|x\|_{H^1(0,1)}} \geq \frac{\|(L_j - L)x_j\|_{L^2(0,1)}}{\|x_j\|_{H^1(0,1)}} \geq \frac{1}{2\sqrt{2}\sqrt{1 + h^2/6}}$$

(where $h = 1/j$) and hence ℓ_j is bounded away from 0. Therefore, Theorem 22 does not apply to this example.

We now present a discretization of the operators of Example 1 that satisfies the hypotheses of Theorem 22 and hence leads to convergence of the GSVE.

Example 3. *Let T, L, X_j, and Y_j be defined as in Example 1, but now define Z_j to be the space of piecewise constant functions on the uniform mesh with j elements.*

As before, T_j is to be $P_{Y_j}T|_{X_j} = T|_{X_j}$, and we define $L_j = P_{Z_j}L$. Since L maps X_j into Z_j, it follows that $L_j = L|_{X_j}$. Therefore, for this discretization, we have that $t_j = \ell_j = 0$ for all $j \in \mathbb{Z}^+$, and hence Theorem 22 guarantees that the GSVE of (T_j, L_j) converges to the GSVE of (T, L) in the sense of Definition 9 as $j \to \infty$.

Figures 3.1-3.3 show the approximate and exact singular functions for $k = 1, 2, 3$ (analogous to Figures 2.1-2.3 from Example 1). As in Example 1, we use $j = 100$ to obtain these numerical results. In contrast to Example 1, now all three of the examined singular modes are well approximated.

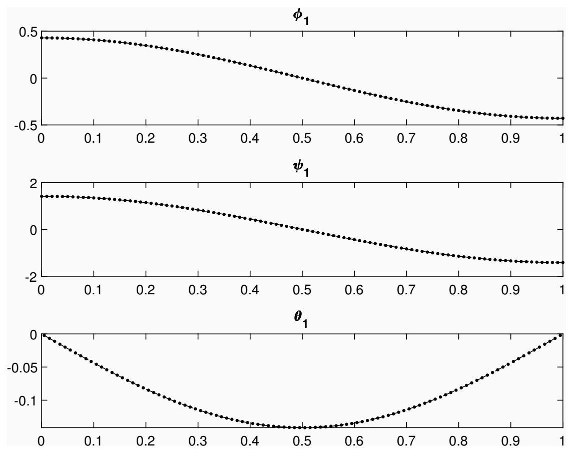

Figure 3.1: The computed functions $\phi_1^{(100)}$ (top), $\psi_1^{(100)}$ (middle), and $\theta_1^{(100)}$ (bottom) for Example 3, together with the corresponding exact functions ϕ_1, ψ_1, and θ_1. In each graph, the approximate function is the solid curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.

Extensive numerical testing suggests that

$$\left|a_k^{(j)} - a_k\right| = O(h^2) \text{ as } j \to \infty,$$
$$\left|b_k^{(j)} - b_k\right| = O(h^2) \text{ as } j \to \infty.$$

Each of the generalized singular spaces is one-dimensional and, therefore, we can compare the generalized singular functions directly rather than referring to the gap between subspaces (we just have to normalize the vectors and multiply by -1 when necessary so that the angle between each singular vector and its estimate is close to 0 rather than close to π). We observe

$$\left\|\phi_k^{(j)} - \phi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$
$$\left\|\psi_k^{(j)} - \psi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$

$$\left\|\theta_k^{(j)} - \theta_k\right\|_{L^2(0,1)} = O(h) \text{ as } j \to \infty.$$

In each case, the rate of convergence is optimal for the given discretization

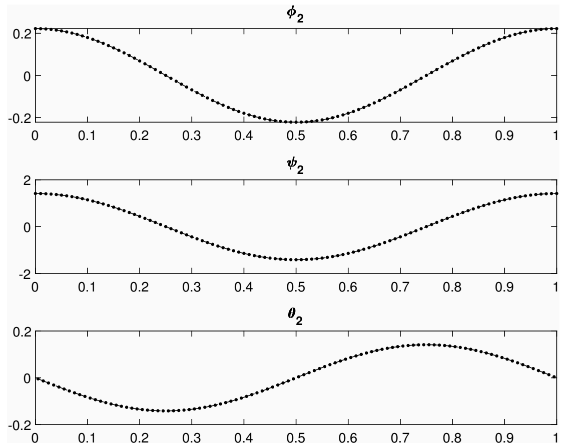

Figure 3.2: The computed functions $\phi_2^{(100)}$ (top), $\psi_2^{(100)}$ (middle), and $\theta_2^{(100)}$ (bottom) for Example 3, together with the corresponding exact functions ϕ_2, ψ_2, and θ_2. In each graph, the approximate function is the solid curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.

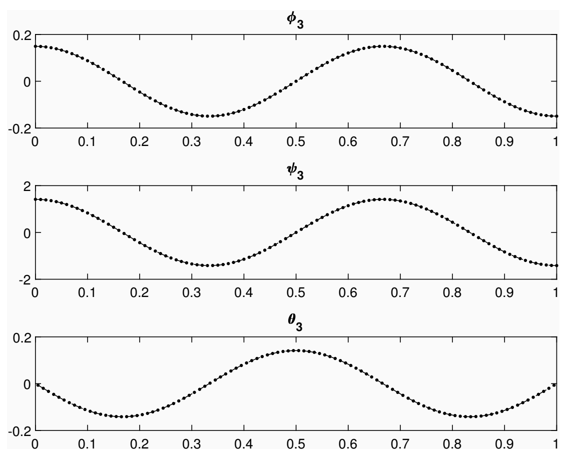

Figure 3.3: The computed functions $\phi_3^{(100)}$ (top), $\psi_3^{(100)}$ (middle), and $\theta_3^{(100)}$ (bottom) for Example 3, together with the corresponding exact functions ϕ_3, ψ_3, and θ_3. In each graph, the approximate function is the solid curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.

Chapter 4

Rates of convergence

In Example 3 of the previous chapter, we compared the generalized singular values and vectors of the operator pairs (T, L) and (T_j, L_j). In that example, we observed the following rates of convergence for the generalized singular values:

$$\left|a_k^{(j)} - a_k\right| = O(h^2) \text{ as } j \to \infty,$$
$$\left|b_k^{(j)} - b_k\right| = O(h^2) \text{ as } j \to \infty.$$

We also observed the following rates of convergence for the generalized singular functions:

$$\left\|\phi_k^{(j)} - \phi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$
$$\left\|\psi_k^{(j)} - \psi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$
$$\left\|\theta_k^{(j)} - \theta_k\right\|_{L^2(0,1)} = O(h) \text{ as } j \to \infty.$$

Here, $h = 1/j$ was the width of each interval of continuous piecewise linear finite elements in the discretization X_j of X. In this chapter, we analyze the rate of convergence of the generalized singular values and vectors of (T_j, L_j) to those of (T, L). We next consider a less trivial example that demonstrates the same rates of convergence observed in Example 3.

Example 4. *Let $X = L^2(0,1)$ and $Y = Z = L^2(0,1)$. Define operators $T : X \to Y$ and $L : D(L) \to Z$ by $Lx = x'$, and $Tx = \int_0^1 se^{st}x(t)\, dt$. We discretize (T, L) by defining $X_j = Y_j$ to be the space of continuous piecewise linear functions on a uniform mesh with j elements, and Z_j to be the set of piecewise constant functions defined on each subinterval of the mesh. Let $\{x_0, x_1, \cdots, x_j\}$ be the standard nodal basis. Define $T_j = P_{Y_j}T|_{X_j}$ and $L_j = P_{Z_j}L|X_j$. Using the method of computation for* *the GSVE from Chapter 2, We are able to compute the singular values and vectors for* (T_j, L_j). *The GSVE of* (T, L) *is unknown in this example, so convergence rates are approximated using Richardson extrapolation. In our discretization, we will have a total of 7 refinements of our finite element space. At each stage of refinement, the previous discretization is also interpolated into the new refinement in order to compare the functions from each refinement. Using Richardson extrapolation, we can use any 3 consecutive refinements of our discretization to produce a rate of approximation for each of the singular values and singular vectors.*

We begin with a discretization using 40 elements, and at each refinement of the discretization, we double the number of elements. Tables 4.1-4.4 give the rates of convergence for the first 5 generalized singular vectors and generalized singular values of T. *The numbers in Table 4.1 are the estimates for p using Richardson extrapolation such that*

$$\left|a_k^{(j)} - a_k\right| = Ch^p.$$

The numbers in Table 4.2 (and similarly in Tables 4.3 and 4.4) are estimates of p using Richardson extrapolation such that

$$\left\|\phi_k^{(j)} - \phi_k\right\|_{L^2(0,1)} = Ch^p.$$

The results of Table 4.1 suggest the following rates of convergence of the generalized singular values of (T_j, L_j) *to those of* (T, L).

$$\left|a_k^{(j)} - a_k\right| = O(h^2) \text{ as } j \to \infty,$$
$$\left|b_k^{(j)} - b_k\right| = O(h^2) \text{ as } j \to \infty.$$

The rate of convergence of $b_k^{(j)}$ *to* b_k *follows immediately from the equations*

$$a_k^2 + b_k^2 = 1,$$
$$\left(a_k^{(j)}\right)^2 + \left(b_k^{(j)}\right)^2 = 1.$$

Tables 4.2-4.4 suggest the following rates of convergence of the generalized singular vectors of (T_j, L_j) *to those of* (T, L).

$$\left\|\phi_k^{(j)} - \phi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$
$$\left\|\psi_k^{(j)} - \psi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$
$$\left\|\theta_k^{(j)} - \theta_k\right\|_{L^2(0,1)} = O(h) \text{ as } j \to \infty.$$

Table 4.1
Rate of convergence of $a_k^{(j)}$

k	$j = 160$	$j = 320$	$j = 640$	$j = 1240$	$j = 2480$
1	NaN	NaN	NaN	NaN	NaN
2	2.00065	2.00016	2.00004	2.00001	2.00000
3	2.00121	2.00031	2.00008	2.00002	2.00000
4	2.00339	2.00087	2.00022	2.00005	2.00001
5	2.00816	2.00214	2.00054	2.00013	2.00008

Table 4.2
Rate of convergence of $\phi_k^{(j)}$

k	$j = 160$	$j = 320$	$j = 640$	$j = 1240$	$j = 2480$
1	1.99998	2.00000	2.00000	2.00000	2.00000
2	2.00013	2.00003	2.00001	2.00000	2.00000
3	2.00020	2.00005	2.00001	2.00000	2.00000
4	1.99937	1.99984	1.99996	1.99999	2.00000
5	1.99644	1.99912	1.99978	1.99995	2.00000

In both examples, the same rates of convergence of the generalized singular values and vectors were observed when using continuous piecewise linear elements to discretize the space $D(L)$ and when using piecewise constant elements to discretize the space Y. In this chapter, we provide a theory and analysis to prove these rates of convergence. In our analysis, it will be important to consider the space $D(L^*L) \subseteq X$.

To be consistent with the notation in Chapter 5, we will denote $D(L^*L)$ by S_2. We define the bilinear form $\langle \cdot, \cdot \rangle_{S_2} : S_2 \times S_2 \to \mathbb{R}$ by

$$\langle x, y \rangle_{S_2} = \langle Mx, My \rangle_X \ , \ \forall x, y \in S_2,$$

where $M : S_2 \to X$ was defined by $M = T^*T + L^*L$ in Chapter 3.

By Theorem 14, M has bounded inverse and is, therefore, injective. Hence, for any $x \in S_2$,

$$\langle x, x \rangle_{S_2} = 0 \iff x = 0.$$

Therefore, $\langle \cdot, \cdot \rangle_{S_2}$ defines an inner product on S_2. The next two lemmas show that S_2 is a dense subspace of $D(L)$ and that S_2 is a Hilbert space with norm $\| \cdot \|_{S_2}$ defined

Table 4.3
Rate of convergence of $\theta_k^{(j)}$

k	$j = 160$	$j = 320$	$j = 640$	$j = 1240$	$j = 2480$
1	NaN	NaN	NaN	NaN	NaN
2	0.99971	0.99993	0.99998	1.00000	1.00000
3	0.99867	0.99967	0.99992	1.00000	1.00000
4	0.99605	0.99902	0.99975	0.99994	0.99998
5	0.99078	0.99771	0.99943	0.99986	0.99996

Table 4.4
Rate of convergence of $\psi_k^{(j)}$

k	$j = 160$	$j = 320$	$j = 640$	$j = 1240$	$j = 2480$
1	2.00007	2.00002	2.00000	2.00000	2.00000
2	2.00009	2.00002	2.00001	2.00000	2.00000
3	2.00058	2.00014	2.00005	2.00001	2.00000
4	2.00112	2.00028	2.00007	2.00002	2.00000
5	2.00150	2.00039	2.00010	2.00003	1.99999

by

$$\|x\|_{S_2} = \|Mx\|_X \ , \ \forall x \in X.$$

Lemma 23. *S_2 is dense in $D(L)$.*

Proof. It will suffice to prove that $S_2^{\perp_*} = \{0\}$ where $\perp_*$ denotes the orthogonal complement in $D(L)$ of S_2 with respect to the $*$-inner product. Let $w \in S_2^{\perp_*}$, and define $u = M^{-1}w$. Then $u \in D(L^*L)$, so by definition of w, $\langle w, u\rangle_* = 0$. Therefore,

$$\|w\|_X^2 = \langle w, w\rangle_X = \langle w, M^{-1}w\rangle_* = \langle w, u\rangle_* = 0.$$

Thus, $w = 0$ and the proof is complete. □

Lemma 24. *$\|\cdot\|_{S_2}$ is a stronger norm than $\|\cdot\|_*$. In particular, for every $x \in S_2$,*

$$\|x\|_* \leq (\gamma)^{-1/2}\|x\|_{S_2}.$$

Proof. Let $x \in S_2$. Then it follows that

$$\|x\|_*^2 = \langle x, x\rangle_*$$

$$
\begin{aligned}
&= \langle Tx, Tx\rangle_Y + \langle Lx, Lx\rangle_Z \\
&= \langle x, (T^*T + L^*L)x\rangle_X \\
&= \langle x, Mx\rangle_X \\
&\le \|x\|_X \|Mx\|_X \\
&\le (\gamma)^{-1/2}\|x\|_* \|Mx\|_X .
\end{aligned}
$$

After dividing each side of the inequality by $\|x\|_*$, we have

$$\|x\|_* \le (\gamma)^{-1/2}\|Mx\|_X = (\gamma)^{-1/2}\|x\|_{S_2},$$

and this completes the proof. □

We will now assume the following property of the discretization: There exists a sequence of positive real numbers $\{\varepsilon_j\}$ such that $\varepsilon_j \to 0$ as $j \to \infty$ and

$$\|\Pi_{X_j} x - x\|_* \le \varepsilon_j \|x\|_{S_2},\ \forall x \in S_2. \tag{4.1}$$

This assumption is consistent with finite element approximation results. For instance if X_j is the space of continuous piecewise linear finite elements in $L^2(0,1)$, and $L : H^1(0,1) \to L^2(0,1)$ is the derivative operator, then the $*$-norm is equivalent to the $H^1(0,1)$ norm, and the S_2-norm is related to the H^2-seminorm. In this case, we can take $\varepsilon_j = Ch_j$ for some constant $C > 0$, where h_j is the mesh size of the j-th mesh (see [14] Theorem 4.4.20).

Theorem 25. *For every $x \in S_2$,*

$$\|\Pi_{X_j} x - x\|_X \le \varepsilon_j^2 \|x\|_{S_2}. \tag{4.2}$$

Proof. Let $x \in S_2$, and define $\hat{x} = \Pi_{X_j} x$ and $w = M^{-1}(x - \hat{x})$. Then

$$
\begin{aligned}
\|x - \hat{x}\|_X^2 &= \langle x - \hat{x}, x - \hat{x}\rangle_X \\
&= \langle M^{-1}(x - \hat{x}), x - \hat{x}\rangle_* \\
&= \langle w, x - \hat{x}\rangle_* .
\end{aligned}
$$

Since $x - \hat{x}$ is orthogonal to the space X_j with respect to the $*$-inner product, it follows that $\langle \Pi_{X_j} w, x - \hat{x}\rangle_* = 0$. Therefore,

$$\langle w, x - \hat{x}\rangle_* = \langle w - \Pi_{X_j} w, x - \hat{x}\rangle_* .$$

Putting these results together, we have

$$
\begin{aligned}
\|x - \hat{x}\|_X^2 = \langle w, x - \hat{x}\rangle_* &= \langle w - \Pi_{X_j} w, x - \hat{x}\rangle_* \\
&\le \|w - \Pi_{X_j} w\|_* \|x - \hat{x}\|.
\end{aligned}
$$

Both w and x are in S_2, so by (4.1), it follows that

$$\begin{aligned} \|w - \Pi_{X_j} w\|_* &\le \varepsilon_j \|w\|_{S_2} \\ \|x - \hat{x}\|_* &\le \varepsilon_j \|x\|_{S_2}. \end{aligned}$$

From this we have

$$\begin{aligned} \|x - \hat{x}\|_X^2 \le \|w - \Pi_{X_j}\|_* \|x - \hat{x}\|_* &\le (\varepsilon_j \|w\|_{S_2}) (\varepsilon_j \|x\|_{S_2}) \\ &= \varepsilon_j^2 \|w\|_{S_2} \|x\|_{S_2}. \end{aligned}$$

Therefore,

$$\begin{aligned} \|x - \hat{x}\|_X^2 &\le \varepsilon_j^2 \|Mw\|_X \|x\|_{S_2} \\ &= \varepsilon_j^2 \|M \left(M^{-1}(x - \hat{x})\right) \|_X \|x\|_{S_2} \\ &= \varepsilon_j^2 \|x - \hat{x}\|_X \|x\|_{S_2}. \end{aligned}$$

After eliminating a factor of $\|x - \hat{x}\|_X$ from each side of the inequality, we have

$$\|x - \hat{x}\|_X \le \varepsilon_j^2 \|x\|_{S_2}.$$

□

It should be noted that a similar argument yields that

$$\|x - \Pi_{X_j} x\|_X \le \varepsilon_j \|x\|_* \ \forall x \in D(L). \tag{4.3}$$

It then follows that

$$\|\Pi_{X_j} x\|_X \le \|x\|_X + \varepsilon_j \|x\|_* \le (\gamma^{-1/2} + \varepsilon_j) \|x\|_* \ \forall x \in D(L). \tag{4.4}$$

The following lemma will be used in the analysis of the rate of convergence of the generalized singular values.

Lemma 26. *There exists a constant $C > 0$ such that*

$$\|T_j \Pi_{X_j} x - Tx\|_Y \le C(t_j \|x\|_* + \varepsilon_j^2 \|x\|_{S_2}) \ \forall x \in S_2. \tag{4.5}$$

Proof. For every $x \in S_2$, we have

$$\begin{aligned} \|T_j \Pi_{X_j} x - Tx\|_Y &\le \|T_j \Pi_{X_j} x - T \Pi_{X_j} x\|_Y + \|T \Pi_{X_j} x - Tx\|_Y \\ &\le \|(T - T_j) \Pi_{X_j}\|_Y + \|T(\Pi_{X_j} x - x)\|_Y \\ &\le t_j \|\Pi_{X_j} x\|_X + \|T\|_{\mathcal{L}(X,Y)} \|\Pi_{X_j} x - x\|_X \\ &\le t_j \gamma^{-1/2} \|x\|_* + \|T\|_{\mathcal{L}(X,Y)} \varepsilon_j^2 \|x\|_{S_2} \\ &\le C(t_j \|x\|_* + \varepsilon_j^2 \|x\|_{S_2}), \end{aligned}$$

where $C = \max\{\gamma^{-1/2}, \|T\|_{\mathcal{L}(X,Y)}\}$. $\square$

Recall that the GSVE of (T, L) can be expressed as

$$T = \sum_{k=1}^{\infty} a_k \psi_k \otimes_* \phi_k,$$
$$L = \sum_{k=1}^{\infty} b_k \theta_k \otimes_* \phi_k \ + \sum_{k \in M_a} b_k \theta_k \otimes_* \phi_k,$$

and similarly, the GSVE of (T_j, L_j) is given by

$$T_j = \sum_{k=1}^{\min\{m_j, n_j\}} a_k^{(j)} \psi_k^{(j)} \otimes_{*_j} \phi_k^{(j)},$$
$$L_j = \sum_{k=1}^{\min\{p_j, n_j\}} b_k^{(j)} \theta_k^{(j)} \otimes_{*_j} \phi_k^{(j)}.$$

Here, we order the generalized singular terms in the GSVE of (T, L) such that the generalized singular values of T are nonincreasing and the generalized singular values of L are nondecreasing, as was done in Chapter 3. To be precise, we assume that the index sets M_b and M_0, defined in Definition 1.11 of Chapter 3, are given by

$$M_b = \{1, 2, \cdots, N_b\}, \ M_0 = \{N_b + 1, N_b + 2, \cdots\},$$

where $N_b = \dim(\mathcal{N}(L))$. Since T is a compact operator, the null space of L must be finite-dimensional, hence making such an indexing possible. Similarly, the GSVE for (T_j, L_j), as given above, is such that the sequence of generalized singular values $\{a_k^{(j)}\}_{k=1}$ is nonincresing and the sequence of generalized singular values $\{b_k^{(j)}\}_{k=1}$ is nondecreasing for every $j \in \mathbb{Z}^+$.

The generalized singular values a_k of (T, L) can be characterized as

$$a_k = \max_{\substack{S \subseteq D(L) \\ \dim(S) = k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|Tx\|_Y}{\|x\|_*},$$

where this maximum is attained for each $k \in \mathbb{Z}^+$ by the space

$$S = \Phi_k = \operatorname{span}\{\phi_1, \phi_2, \cdots, \phi_k\}.$$

The corresponding singular values of T_j can be characterized as

$$a_k^{(j)} = \max_{\substack{S \subseteq X_j \\ \dim(S)=k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j x\|_Y}{\|x\|_{*_j}}.$$

We will often have to compare the norms $\|\cdot\|_*$ and $\|\cdot\|_{*_j}$ on the subspace X_j when analyzing the rates of convergence of the generalized singular vectors. Recall that $\eta_j = c_j^2 + 2(t_j + \ell_j)$. Then $\eta_j = O(c_j) \to 0$ as $j \to \infty$, and hence $0 \leq \eta_j < 1$ for all $j \in \mathbb{Z}^+$ that are sufficiently large. We will need the following fact about the generalized singular vectors of the operator pair (T, L).

Lemma 27. *For all $k \in \mathbb{Z}^+ = M_b \cup M_0$, $\phi_k \in S_2$.*

Proof. Let $k \in M_0 \cup M_b$. Then we have

$$T\phi_k = a_k \psi_k, \; T^{\#}\psi_k = a_k \phi_k,$$

where $a_k > 0$. Putting both of these things together, we have

$$T^{\#}T\phi_k = a_k^2 \phi_k.$$

Since $T^{\#} = M^{-1}T^*$ (see [9], Theorem 5.27), if follows that $\phi_k \in D(M) = S_2$. $\square$

Notice that when $a_k \in M_a$, the same argument does not hold since $T^{\#}T\phi_k = 0$. We will need a few more preliminary results in order to prove a particular rate of convergence.

Lemma 28. *For every $n \in \mathbb{Z}^+$,*

$$\max_{\substack{S \subseteq D(L) \\ \dim(S)=k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|Tx\|_Y}{\|x\|_*} = \max_{\substack{S \subseteq S_2 \\ \dim(S)=k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|Tx\|_Y}{\|x\|_*}.$$

Proof. This follows immediately from the fact that S_2 is dense in $D(L)$. $\square$

Lemma 29. *For every $k \in \mathbb{Z}^+$ and for every sufficiently large positive integer j,*

$$\max_{\substack{S \subseteq D(L) \\ \dim(S)=k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} = \max_{\substack{S \subseteq X_j \\ \dim(S)=k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*}. \tag{4.6}$$

Proof. We assume that j is sufficiently large such that $\dim(X_j) \geq k$. Clearly, the left-hand side of (4.6) is at least as big as the right-hand side. Therefore, we must prove that

$$\max_{\substack{S \subseteq D(L) \\ \dim(S)=k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} \leq \max_{\substack{S \subseteq X_j \\ \dim(S)=k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*}. \tag{4.7}$$

Let S be a k-dimensional subspace of X, and let $\hat{S} = \Pi_{X_j} S$. We consider two cases. If $S \cap X_j^{\perp_*}$ is nontrivial, where $X_j^{\perp_*}$ denotes the orthogonal complement of X_j in $D(L)$ with respect to the $*$-norm, then there exists $x \in S$ such that $x \neq 0$ and $T_j \Pi_{X_j} x = 0$. It then follows that

$$\begin{aligned} &\min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} = 0 \\ \implies &\min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} \leq \max_{\substack{K \subseteq X_j \\ \dim(K) = k}} \min_{\substack{x \in K \\ x \neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*} \end{aligned}$$

The second case is that $S \cap X_j^{\perp_*}$ is trivial. In this case, $\dim(\hat{S}) = k$, and there is a one-to-one correspondence between $x \in S$ and $\bar{x} \in \hat{S}$ ($\bar{x} = \Pi_{X_j} x$). For each such x and $\bar{x}$, we have $\bar{x} - x \in X_j^{\perp_*}$ and hence

$$T_j \bar{x} = T_j \Pi_{X_j} \bar{x} = T_j \Pi_{X_j} x.$$

Also, we have that $\|\bar{x}\|_* = \|\Pi_{X_j} x\|_* \leq \|x\|_*$. Therefore,

$$\frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} \leq \frac{\|T_j \bar{x}\|_Y}{\|\bar{x}\|_*}.$$

From this, it follows that

$$\min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} \leq \min_{\substack{\bar{x} \in \hat{S} \\ \bar{x} \neq 0}} \frac{\|T_j \bar{x}\|_Y}{\|\bar{x}\|_*} \leq \max_{\substack{\hat{S} \subseteq X_j \\ \dim(\hat{S}) = k}} \min_{\substack{\bar{x} \in \hat{S} \\ \hat{x} \neq 0}} \frac{\|T_j \bar{x}\|_Y}{\|\bar{x}\|_*},$$

and, thus, we have shown

$$\max_{\substack{S \subseteq D(L) \\ \dim(S) = k \\ S \cap X_j^{\perp_*} = \{0\}}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} \leq \max_{\substack{S \subseteq X_j \\ \dim(S) = k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*},$$

which completes the proof. □

We can now prove the desired result.

Theorem 30. *For each $k \in \mathbb{Z}^+$, there exists constants C_1 and C_2 such that for all sufficiently large positive integers j,*

$$a_k - C_1(c_j + \varepsilon_j^2) \leq a_k^{(j)} \leq a_k + C_2 c_j. \tag{4.8}$$

Proof. By Corollary 9 of [11],

$$\frac{\|x\|_*}{\|x\|_{*_j}} \leq \frac{1}{\sqrt{1-\eta_j}} \ \forall x \in X_j, \tag{4.9}$$

where $\eta_j = c_j^2 + 2(t_j + c_j)$. Using this result, we have

$$\begin{aligned}
a_k^{(j)} = \max_{\substack{S\subseteq X_j \\ \dim(S)=k}} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|T_j x\|_Y}{\|x\|_{*_j}} &= \max_{\substack{S\subseteq X_j \\ \dim(S)=k}} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*} \frac{\|x\|_*}{\|x\|_{*_j}} \\
&\leq \frac{1}{\sqrt{1-\eta_j}} \max_{\substack{S\subseteq X_j \\ \dim(S)=k}} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*} \\
&\leq \frac{1}{\sqrt{1-\eta_j}} \max_{\substack{S\subseteq X_j \\ \dim(S)=k}} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|Tx\|_Y + \|(T-T_j)x\|_Y}{\|x\|_*} \\
&\leq \frac{1}{\sqrt{1-\eta_j}} \left(\max_{\substack{S\subseteq X_j \\ \dim(S)=k}} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_*} + t_j \right) \\
&\leq \frac{1}{\sqrt{1-\eta_j}} \left(\max_{\substack{S\subseteq D(L) \\ \dim(S)=k}} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_*} + t_j \right) \\
&= \frac{1}{\sqrt{1-\eta_j}} (a_k + t_j).
\end{aligned}$$

Since

$$\frac{1}{\sqrt{1-\eta_j}} = 1 + \delta(\eta_j)\eta_j \text{ where } 0 < \delta(\eta_j) < 1,$$

for all η_j sufficiently small ($0 \leq \eta_j < 3/8$ suffices), and since $t_j = O(c_j)$ and $\eta_j = O(c_j)$, this establishes the upper bound in (4.8). To prove the lower bound, we define the subspace $\Phi_k = \text{span}\{\phi_1, \phi_2, \cdots, \phi_k\}$ and apply Lemma 26:

$$\begin{aligned}
a_k &= \max_{\substack{S\subseteq D(L) \\ \dim(S)=k}} \min_{\substack{x\in S \\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_*} \\
&= \min_{\substack{x\in \Phi_k \\ x\neq 0}} \frac{\|Tx\|_Y}{\|x\|_*} \\
&\leq \min_{\substack{x\in \Phi_k \\ x\neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y + C(t_j\|x\|_* + \varepsilon_j^2 \|x\|_{S_2})}{\|x\|_*}.
\end{aligned}$$

Since Φ_k is fixed and finite-dimensional, there exists $C_k > 0$ such that

$$\frac{\|x\|_{S_2}}{\|x\|_*} \leq C_k \ \forall x \in \Phi_k.$$

If we define $C' = \max\{C, CC_k\}$, we obtain

$$\begin{aligned} a_k &\leq \min_{\substack{x \in \Phi_k \\ x \neq 0}} \left(\frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} + C'(t_j + \varepsilon_j^2) \right) \\ &= \min_{\substack{x \in \Phi_k \\ x \neq 0}} \frac{\|T_j \Pi_{X_j} x\|_Y}{\|x\|_*} + C'(t_j + \varepsilon_j^2) \\ &\leq \max_{\substack{S \subseteq X_j \\ \dim(S) = k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*} + C'(t_j + \varepsilon_j^2). \end{aligned}$$

By Corollary 9 of [11],

$$\frac{\|x\|_{*_j}}{\|x\|_*} \leq \sqrt{1 + \eta_j} \ \forall x \in X_j.$$

Therefore,

$$\begin{aligned} a_k &\leq \sqrt{1 + \eta_j} \left(\max_{\substack{S \subseteq X_j \\ \dim(S) = k}} \min_{\substack{x \in S \\ x \neq 0}} \frac{\|T_j x\|_Y}{\|x\|_*} \right) + C'(t_j + \varepsilon_j^2) \\ &= \sqrt{1 + \eta_j} \left(a_k^{(j)} \right) + C'(t_j + \varepsilon_j^2) \\ &\implies \frac{1}{\sqrt{1 + \eta_j}} (a_k - C'(t_j + \varepsilon_j^2)) \leq a_k^{(j)}. \end{aligned}$$

Since

$$\frac{1}{\sqrt{1 + \eta_j}} = 1 - \varsigma(\eta_j)\eta_j, \ \text{where } 0 \leq \varsigma(\eta_j) \leq 1/2,$$

the lower bound in (4.8) follows (again using the fact that both t_j and η_j are $O(c_j)$), and the proof is complete. □

We now wish to analyze the convergence of the generalized singular vectors of (T_j, L_j) to the generalized singular vectors of (T, L). Recall that for each k, we define the spaces E_k and $E_k^{(j)}$ by

$$E_k = \text{span}\{\phi_i : a_i = a_k\}.$$

and

$$E_k^{(j)} = \text{span}\{\phi_i^{(j)} : a_i^{(\ell)} \to a_k \text{ as } \ell \to \infty\}.$$

We will need to prove that the subspace $E_k^{(j)}$ converges to the subspace E_k as $j \to \infty$ in the sense that the gap between E_k and $E_k^{(j)}$ converges to 0 as $j \to \infty$. Recall that the gap between subspaces U and V in H, denoted by $\hat{\delta}(U, V)$ is defined by

$$\begin{aligned}
\hat{\delta}(U, V) &= \max\{\delta(U, V), \delta(V, U)\}, \\
\delta(U, V) &= \sup_{\substack{u \in U \\ \|u\|_H = 1}} \inf_{v \in V} \|u - v\|_H.
\end{aligned}$$

We wish to derive estimates for $\hat{\delta}(E_k, E_k^{(j)})$, where gap is defined by either the X-norm or the $*$-norm. We will, therefore, write $\hat{\delta}_*$ and δ_* for the gap defined by the $*$-norm and $\hat{\delta}_X$ and δ_X for the gap defined by the X-norm.

It was shown in [11] that, under the assumptions made here, $\hat{\delta}_*(E_k, E_k^{(j)}) \to 0$ as $j \to \infty$. We will need the following results to conclude that $\delta_X(E_k, E_k^{(j)}) = \delta_X(E_k^{(j)}, E_k)$ for all positive integers j sufficiently large.

Theorem 31. *If U and V are k-dimensional subspaces of a Hilbert space H, where k is a positive integer, then $\delta(U, V) = \delta(V, U)$.*

Proof. Without loss of generality, let us assume that $\delta(U, V) \leq \delta(V, U)$. By Theorem 7, the result holds if $\delta(U, V)$ and $\delta(V, U)$ are both strictly less than 1. It suffices, therefore, to show that the assumption $\delta(U, V) < \delta(V, U) = 1$ produces a contradiction. Since V is finite dimensional,

$$\delta(V, U) = \max_{\substack{v \in V \\ \|v\| = 1}} \|P_U v - v\|.$$

Therefore, the assumption that $\delta(V, U) = 1$ implies that there exists $\hat{v} \in V$ such that $\|\hat{v}\| = 1$ and $\|P_U \hat{v} - \hat{v}\| = 1$. This is possible only if $P_U \hat{v} = 0$. That is, if $\hat{v} \in U^\perp$. On the other hand, the assumption that $\delta(U, V) < 1$ implies that $\|P_V u - u\| < 1$ for all $u \in U$ and hence that the null space of $P = P_V|_U$ (P_V restricted to U) is trivial. Since $\dim(U) = \dim(V) = k$, the fundamental theorem of linear algebra implies that P maps U onto V; thus, there exists $\hat{u} \in U$ such that $P\hat{u} = \hat{v}$. But then

$$\begin{aligned}
\left\| P\left(\frac{\hat{u}}{\|\hat{u}\|}\right) - \frac{\hat{u}}{\|\hat{u}\|} \right\| < 1 &\implies \|P\hat{u} - \hat{u}\| < \|\hat{u}\| \\
&\implies \|\hat{v} - \hat{u}\| < \|\hat{u}\| \\
&\implies \|\hat{v}\|^2 + \|\hat{u}\|^2 < \|\hat{u}\|^2
\end{aligned}$$

(where we used $\delta(U,V) < 1$ in the first step and $\hat{v} \in U^{\perp}$ in the last step). Since $\|\hat{v}\| = 1$, the last inequality is impossible, and the proof is complete. $\square$

The previous theorem implies that $\delta_X(E_k, E_k^{(j)}) = \delta_X(E_k^{(j)}, E_k)$. It follows, therefore, that $\hat{\delta}_X(E_k, E_k^{(j)}) = \delta(E_k, E_k^{(j)})$, and it, therefore, suffices to analyze the convergence of $\delta_X(E_k, E_k^{(j)})$ to zero. The same comments apply to $\hat{\delta}_*$: $\hat{\delta}_*(E_k, E_k^{(j)}) = \delta_*(E_k, E_k^{(j)})$. Specifically, we will show that

$$\delta_X(E_k, E_k^{(j)}) = O(c_j + \varepsilon_j^2),$$
$$\delta_*(E_k, E_k^{(j)}) = O(c_j + \varepsilon_j).$$

By definition of $\delta_X(E_k, E_k^{(j)})$, we must show that there exists $C = C_k > 0$ such that, for all $v \in E_k$ with $\|v\|_X = 1$, there exists $w \in E_k^{(j)}$ such that

$$\|v - w\|_X \leq C(c_j + \varepsilon_j^2). \tag{4.10}$$

We will show that the same vector w also satisfies

$$\|v - w\|_* \leq C(c_j + \varepsilon_j)$$

(albeit with a different constant for C).

We now proceed to show that there exists $C > 0$ such that given $v \in E_k$ with $\|v\|_* = 1$, one can define $w \in E_k^{(j)}$ such that inequality (4.10) holds. To do this, we will need some more notation and several preliminary results. Recall that the dimension of X_j is n_j; let r_j be the rank of T_j. Then there exist r_j generalized singular value/singular vector pairs $(a_i^{(j)}, \phi_i^{(j)})$ of T_j. If $r_j < n_j$, extend the set $\{\phi_i^{(j)} : i = 1, 2, \cdots, r_j\}$ to an orthonormal basis $\{\phi_i^{(j)} : i = 1, 2, \cdots, n_j\}$ for X_j, and define $a_i^{(j)} = 0$ for $i = r_j{+}1, r_j{+}2, \cdots, n_j$. It should be noted that $\{\phi_i^{(j)} : i = 1, 2, \cdots, n_j\}$ is orthonormal with respect to the $*_j$-inner product.

We will write

$$I_k = \{i \in \mathbb{Z}^+ : a_i = a_k\};$$

then

$$E_k = \mathrm{span}\{\phi_i : i \in I_k\},$$
$$E_k^{(j)} = \mathrm{span}\{\phi_i^{(j)} : i \in I_k\}.$$

We will also write $J_j = \{1, 2, \cdots, n_j\}$.

Recall that we have defined the operators $M : D(L^*L) \to X$ and $M_j : X_j \to X_j$ by

$$M = T^*T + L^*L,$$
$$M_j = T_j^*T_j + L_j^*L_j.$$

Both such operators are bijections and each has bounded inverse (see Chapter 3). For each $j \in \mathbb{Z}^+$, we define the operator $\Lambda_j : D(L^*L) \to X_j$ by

$$\Lambda_j = M_j^{-1} P_{X_j} M.$$

Here, again, P_{X_j} denotes the orthogonal projection onto the subspace X_j of X with respect to the X-norm. Notice that, for any $u, v \in X_j$,

$$\langle u, v\rangle_{*_j} = \langle T_j u, T_j v\rangle_Y + \langle L_j u, L_j v\rangle_Z = \langle (T_j^*T_j + L_j^*L_j)u, v\rangle_X = \langle M_j u, v\rangle_X.$$

Similarly, for any $u, v \in D(L)$ such that $u \in S_2$,

$$\langle u, v\rangle_* = \langle Mu, v\rangle_X.$$

Therefore, for any $u \in S_2$ and for any $v \in X_j$, we have

$$\begin{aligned}\langle \Lambda_j u, v\rangle_{*_j} &= \langle M_j^{-1} P_{X_j} M u, v\rangle_{*_j} = \langle P_{X_j} M u, v\rangle_X \\ &= \langle Mu, v\rangle_X = \langle u, v\rangle_*,\end{aligned}$$

that is,

$$\langle \Lambda_j u, v\rangle_{*_j} = \langle u, v\rangle_* \ \forall u \in S_2\ ,\ \forall v \in X_j. \tag{4.11}$$

The operator Λ_j approximates the operator Π_{X_j} in the following sense.

Lemma 32. *For every $j \in \mathbb{Z}^+$ and for every $u \in S_2$,*

$$\|(\Pi_{X_j} - \Lambda_j)u\|_* \le \frac{\eta_j}{(1-\eta_j)\sqrt{\gamma}} \|u\|_{S_2}.$$

Proof. By Theorem 16, for every $x \in X$,

$$\|(\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})x\|_* \le \frac{\eta_j}{(1-\eta_j)\sqrt{\gamma}} \|x\|_X.$$

Since $M : S_2 \to X$, it follows that for every $u \in S_2$,

$$\|(\Pi_{X_j} M^{-1} - M_j^{-1} P_{X_j})Mu\|_* \le \frac{\eta_j}{(1-\eta_j)\sqrt{\gamma}} \|Mu\|_X.$$

By definition, $\|u\|_{S_2} = \|Mu\|_X$. Thus we have

$$\begin{aligned}\|(\Pi_{X_j} - \Lambda_j)u\|_* = \|(\Pi_{X_j} - M_j^{-1}P_{X_j}M)u\|_* &= |(\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})Mu\|_* \\ &\leq \frac{\eta_j}{(1-\eta_j)\sqrt{\gamma}}\|u\|_{S_2}.\end{aligned}$$

$\square$

Lemma 26 and Lemma 32 yield the following estimate.

Lemma 33. *There exists a constant $C > 0$ such that for every $j \in \mathbb{Z}^+$,*

$$\|T - T_j\Lambda_j\|_{\mathcal{L}(S_2,Y)} \leq C(c_j + \varepsilon_j^2).$$

Proof. By Lemma 26, $\|\cdot\|_{S_2}$ is a stronger norm than $\|\cdot\|_*$ and there exists $C' > 0$ such that for every $j \in \mathbb{Z}^+$,

$$\begin{aligned}\|T - T_j\Lambda_j\|_{\mathcal{L}(S_2,Y)} &\leq \|T - T_j\Pi_{X_j}\|_{\mathcal{L}(S_2,Y)} + \|T_j(\Pi_{X_j} - \Lambda_j)\|_{\mathcal{L}(S_2,Y)} \\ &\leq C'(t_j + \varepsilon_j^2) + \|T_j\|_{\mathcal{L}(D(L),Y)}\|\Pi_{X_j} - \Lambda_j\|_{\mathcal{L}(S_2,D(L))}\end{aligned}$$

Since $\|T_j\|_{\mathcal{L}(D(L),Y)} \leq \|T\|_{\mathcal{L}(D(L),Y)} + t_j$ and $t_j, \eta_j = O(c_j)$, it follows that

$$\begin{aligned}\|T - T_j\Lambda_j\|_{\mathcal{L}(S_2,Y)} &\leq C'(t_j + \varepsilon_j^2) + \|T_j\|_{\mathcal{L}(D(L),Y)}\|\Pi_{X_j} - \Lambda_j\|_{\mathcal{L}(S_2,D(L))} \\ &\leq C'(t_j + \varepsilon_j^2) + (\|T\|_{\mathcal{L}(D(L),Y)} + t_j)\frac{\eta_j}{(1-\eta_j)\sqrt{\gamma}} = O(c_j + \varepsilon_j^2).\end{aligned}$$

This completes the proof. $\square$

Let $v \in E_k$ such that $\|v\|_X = 1$. We now define $w \in E_k^{(j)}$ by

$$w = \Pi^{(j)}_{E_k^{(j)}}\Lambda_j v,$$

where $\Pi^{(j)}_{E_k^{(j)}} : X_j \to E_k^{(j)}$ is defined to be the orthogonal projection of X_j onto $E_k^{(j)}$ with respect to the $*_j$-inner product defined on X_j. By the triangle inequality,

$$\|v - w\|_X \leq \|v - \Lambda_j v\|_X + \|\Lambda_j v - w\|_X. \tag{4.12}$$

We first consider the first term to the right of inequality (4.12) above. We already know that

$$\left\|v - \Pi_{X_j}v\right\|_X \leq \varepsilon_j^2\|v\|_{S_2} \ \forall v \in E_k \subseteq S_2$$

and

$$\left\|v - \Pi_{X_j} v\right\|_* \leq \varepsilon_j \|v\|_{S_2} \ \forall v \in E_k \subseteq S_2.$$

Therefore, since E_k is a finite-dimensional subspace of S_2, there exists a constant $C_k^{(1)} > 0$ such that

$$\begin{aligned} v \in E_k,\ \|v\|_X = 1 \quad &\Longrightarrow \quad \left\|v - \Pi_{X_j} v\right\|_X \leq C_k^{(1)} \varepsilon_j^2, \\ v \in E_k,\ \|v\|_* = 1 \quad &\Longrightarrow \quad \left\|v - \Pi_{X_j} v\right\|_* \leq C_k^{(1)} \varepsilon_j. \end{aligned}$$

Putting this together with Lemma 32, we have the following lemma.

Lemma 34. *For every $v \in E_k$ such that $\|v\|_X = 1$,*

$$\|v - \Lambda_j v\|_X = O(c_j + \varepsilon_j^2)\ ,\ \|v - \Lambda_j v\|_* = O(c_j + \varepsilon_j).$$

Proof. By the triangle inequality,

$$\|v - \Lambda_j v\|_X \leq \|v - \Pi_{X_j} v\|_X + \|(\Pi_{X_j} - \Lambda_j) v\|_X$$

From the above argument together with Lemma 32, we have

$$\begin{aligned} \|v - \Pi_{X_j} v\|_X + \|(\Pi_{X_j} - \Lambda_j) v\|_X &\leq \|v - \Pi_{X_j} v\|_X + \gamma^{-1/2} \|(\Pi_{X_j} - \Lambda_j) v\|_* \\ &\leq C_k^{(1)} \varepsilon_j^2 + \frac{\eta_j}{(1 - \eta_j)\gamma} \|v\|_{S_2} \\ &\leq C_k^{(1)} \varepsilon_j^2 + \frac{\eta_j}{(1 - \eta_j)\gamma} C_k^{(2)} \|v\|_X \\ &= C_k^{(1)} \varepsilon_j^2 + C_k^{(2)} \frac{\eta_j}{(1 - \eta_j)\gamma}. \end{aligned}$$

Here, such a constant $C_k^{(2)} > 0$ exists because X_j is a finite dimensional space and, therefore, all norms on X_j are equivalent. Since $\eta_j = O(c_j)$, it follows that

$$C_k^{(2)} \frac{\eta_j}{(1 - \eta_j)\gamma} = O(c_j).$$

Therefore, we have

$$\begin{aligned} \|v - \Lambda_j v\|_X &\leq C_k^{(1)} \varepsilon_j^2 + C_k^{(2)} \frac{\eta_j}{(1 - \eta_j)\gamma} \\ &= O(c_j + \varepsilon_j^2). \end{aligned}$$

Similarly, we have that

$$\|v - \Lambda_j v\|_* \leq \|v - \Pi_{X_j} v\|_* + \|(\Pi_{X_j} - \Lambda_j) v\|_*$$

$$\leq C_k^{(1)}\varepsilon_j + \frac{\eta_j}{(1-\eta_j)\gamma}\|v\|_{S_2}.$$

By exactly the same argument as above, it then follows that

$$\|v - \Lambda_j v\|_* = O(c_j + \varepsilon_j).$$

$\square$

We now consider the second term $\|\Lambda_j v - w\|_X$ to the right of inequality (4.12). By our coercivity condition (1.9), we have that

$$\|\Lambda_j v - w\|_X \leq \gamma^{-1/2} \|\Lambda_j v - w\|_* .$$

Therefore, it suffices to prove that

$$\|\Lambda_j v - w\|_* = O(c_j + \varepsilon_j^2).$$

Notice that $\Lambda_j v - w \in X_j$. Therefore, as a consequence of Corollary 13, we have

$$\|\Lambda_j v - w\|_*^2 \leq \frac{1}{1-\eta_j}\|\Lambda_j v - w\|_{*_j}^2.$$

Before we prove that $\|\Lambda_j v - w\|_{*_j} = O(c_j + \varepsilon_j^2)$, we need the following result.

Lemma 35. *For each $k \in \mathbb{Z}^+$, the quantity*

$$\rho_k^{(j)} = \max\left\{\frac{1}{a_k^2 - (a_i^j)^2} : i \in J_j \backslash I_k\right\}$$

is bounded by a constant C for all sufficiently large $j \in \mathbb{Z}^+$. In other words, there exists $C = C_k > 0$ such that for all $j \in \mathbb{Z}^+$ sufficiently large,

$$\rho_k^{(j)} \leq C.$$

Proof. The proof is similar to that of Lemma 7 in [15]. $\square$

The next result shows that the quantity $\|\Lambda_j v - w\|_{*_j}$ can be bounded by a constant times the quantity $\|(\Lambda_j T^\# T - T_j^{\#_j} T_j \Lambda_j)v\|_{*_j}^2$ for sufficiently large $j \in \mathbb{Z}^+$. Using previous results, we can then show that this upper bound converges to 0 at a rate at least as fast as $O(\varepsilon_j^2 + c_j)$.

Lemma 36. *For every $j \in \mathbb{Z}^+$,*

$$\|\Lambda_j v - w\|_{*_j} \leq \rho_k^{(j)}\|(\Lambda_j T^\# T - T_j^{\#_j} T_j \Lambda_j)v\|_{*_j}.$$

Proof. By definition, $w = \Pi^{(j)}_{E_k^{(j)}} \Lambda_j v$, where $\Pi^{(j)}_{E_k^{(j)}}$ is the orthogonal projection of X_j onto $E_k^{(j)}$ with respect to the $*_j$-inner product defined on X_j. Since $E_k^{(j)} = \text{span}\{\phi_i^{(j)} : i \in I_k\}$, we can express these vectors as

$$\Lambda_j v = \sum_{i \in J_j} \langle \phi_i^{(j)}, \Lambda_j v\rangle_{*_j} \phi_i^{(j)} \ , \ w = \sum_{i \in I_k} \langle \phi_i^{(j)}, \Lambda_j v\rangle_{*_j} \phi_i^{(j)}.$$

Taking the difference then gives

$$\|\Lambda_j v - w\|^2_{*_j} = \sum_{i \in J_j \backslash I_k} \langle \phi_i^{(j)}, \Lambda_j v\rangle^2_{*_j}.$$

For $i \in J_j \backslash I_k$, we have

$$\begin{aligned} a_k^2 \langle \phi_i^{(j)}, T_j^{\#_j} T_j \Lambda_j v\rangle_{*_j} = a_k^2 \left(a_i^{(j)}\right)^2 \langle \phi_i^{(j)}, \Lambda_j v\rangle_{*_j} &= a_k^2 \left(a_i^{(j)}\right)^2 \langle \phi_i^{(j)}, v\rangle_{*} \\ &= \left(a_i^{(j)}\right)^2 \langle \phi_i^{(j)}, T^{\#} T v\rangle_{*} \\ &= \left(a_i^{(j)}\right)^2 \langle \phi_i^{(j)}, \Lambda_j T^{\#} T v\rangle_{*_j}. \end{aligned}$$

Subtracting $\left(a_i^{(j)}\right)^2 \langle \phi_i^{(j)}, T_j^{\#_j} T_j \Lambda_j v\rangle_X$ from each side of the equation above, we have

$$\left(a_k^2 - (a_i^{(j)})^2\right) \langle \phi_i^{(j)}, T_j^{\#_j} T_j \Lambda_j v\rangle_{*_j} = \left(a_i^{(j)}\right)^2 \langle \phi_i^{(j)}, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v\rangle_{*_j}.$$

Therefore, for any $i \in J_j \backslash I_k$,

$$\frac{1}{\left(a_i^{(j)}\right)^2} \langle \phi_i^{(j)}, T_j^{\#_j} T_j \Lambda_j v\rangle_{*_j} = \frac{1}{a_k^2 - \left(a_i^{(j)}\right)^2} \langle \phi_i^{(j)}, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v\rangle_{*_j}.$$

Therefore, we have

$$\begin{aligned} \|\Lambda_j v - w\|^2_{*_j} &= \sum_{i \in J_j \backslash I_k} \langle \phi_i^{(j)}, \Lambda_j v\rangle^2_{*_j} \\ &= \sum_{i \in J_j \backslash I_k} \left(\frac{1}{\left(a_i^{(j)}\right)^2} \langle \phi_i^{(j)}, T_j^{\#_j} T_j \Lambda_j v\rangle_{*_j} \right)^2 \\ &= \sum_{i \in J_j \backslash I_k} \left(\frac{1}{a_k^2 - \left(a_i^{(j)}\right)^2} \langle \phi_i^{(j)}, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v\rangle_{*_j} \right)^2 \end{aligned}$$

$$\leq \sum_{i \in J_j \setminus I_k} \left(\rho_k^{(j)} \langle \phi_i^{(j)}, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v \rangle_{*_j} \right)^2$$
$$\leq \left(\rho_k^{(j)} \right)^2 \|(\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v\|_{*_j}.$$

This completes the proof. □

We want to show that $\|(\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v\|_{*_j} = O(c_j + \varepsilon_j^2)$. Let $y \in E_k$ such that $\|y\|_X = 1$, and let $x \in X_j$. Then

$$\begin{aligned}
&\langle x, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) y \rangle_{*_j} \\
&= \langle x, \Lambda_j T^{\#} T y \rangle_{*_j} - \langle x, T_j^{\#_j} T_j \Lambda_j y \rangle_{*_j} \\
&= \langle Tx, Ty \rangle_Y - \langle T_j x, T_j \Lambda_j y \rangle_Y = \langle (T - T_j) x, Ty \rangle_Y + \langle T_j x, (T - T_j \Lambda_j) y \rangle_Y.
\end{aligned}$$

Therefore,

$$\begin{aligned}
&\langle x, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) y \rangle_{*_j} \\
&= \langle (T - T_j) x, Ty \rangle_Y + \langle T_j x, (T - T_j \Lambda_j) y \rangle_Y \\
&\leq \|(T - T_j) x\|_Y \|Ty\|_Y + \|T_j x\|_Y \|(T - T_j \Lambda_j) y\|_Y.
\end{aligned}$$

As a consequence of Lemma 33, there exists a constant $C' > 0$ such that for every $j \in \mathbb{Z}^+$ and for every $u \in S_2$,

$$\|(T - T_j \Lambda_j) u\|_Y \leq C'(c_j + \varepsilon_j^2) \|u\|_{S_2}.$$

From this, it then follows that

$$\begin{aligned}
&\|(T - T_j) x\|_Y \|Ty\|_Y + \|T_j x\|_Y \|(T - T_j \Lambda_j) y\|_Y \\
&\qquad \leq t_j \|x\|_* \|T\|_{\mathcal{L}(X,Y)} \|y\|_X + (1 + t_j) \|x\|_* C'(c_j + \varepsilon_j^2) \|y\|_{S_2}.
\end{aligned}$$

Since E_k is finite dimensional, all norms on E_k are equivalent. Thus, there exists $C'' > 0$ such that for every $u \in E_k$, $C' \|u\|_{S_2} \leq C'' \|u\|_X$. Therefore, we have

$$\begin{aligned}
&t_j \|x\|_* \|T\|_{\mathcal{L}(X,Y)} \|y\|_X + C'(c_j + \varepsilon_j^2)(1 + t_j) \|x\|_* \|y\|_{S_2} \\
&\qquad \leq t_j \|x\|_* \|T\|_{\mathcal{L}(X,Y)} \|y\|_X + (1 + t_j) C''(c_j + \varepsilon_j^2) \|x\|_* \|y\|_X \\
&\qquad \leq \left(t_j \|T\|_{\mathcal{L}(X,Y)} + (1 + t_j) C''(c_j + \varepsilon_j^2) \right) \frac{1}{\sqrt{1 - \eta_j}} \|x\|_{*_j}.
\end{aligned}$$

Hence,

$$\langle x, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) y \rangle_{*_j} \leq \left[\left(t_j \|T\|_{\mathcal{L}(X,Y)} + (1+t_j) C''(c_j + \varepsilon_j^2) \right) \frac{1}{\sqrt{1-\eta_j}} \right] \|x\|_{*_j}. \tag{4.13}$$

Since $t_j = O(c_j)$ and $\eta_j = O(c_j)$ as $j \to \infty$, it follows that

$$\left[\left(t_j \|T\|_{\mathcal{L}(X,Y)} + (1+t_j) C''(c_j + \varepsilon_j^2) \right) \frac{1}{\sqrt{1-\eta_j}} \right] = O(c_j + \varepsilon_j^2).$$

Therefore, there exists $C > 0$ such that for all $j \in \mathbb{Z}^+$ sufficiently large,

$$\left[\left(t_j \|T\|_{\mathcal{L}(X,Y)} + (1+t_j) C''(c_j + \varepsilon_j^2) \right) \frac{1}{\sqrt{1-\eta_j}} \right] \leq C(c_j + \varepsilon_j^2).$$

We then combine this with inequality (4.13) to give

$$\langle x, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) y \rangle_{*_j} \leq C(c_j + \varepsilon_j^2) \|x\|_{*_j}. \tag{4.14}$$

Notice that the constant C does not depend on the choice of y or x from above. With inequality (4.14), we can prove the following theorem.

Lemma 37. *Let $k \in \mathbb{Z}^+$. Then there exists $C = C_k > 0$ and $j_k \in \mathbb{Z}^+$ such that for every $v \in E_k$ with $\|v\|_X = 1$ and for all $j \in \mathbb{Z}^+$ with $j \geq j_k$,*

$$\|(\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v\|_{*_j} \leq C(c_j + \varepsilon_j^2).$$

Proof. Let $v \in E_k$ such that $\|v\|_X = 1$, and let $z = (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v$. Then

$$\begin{aligned} \|z\|_{*_j} &= \sup_{\substack{x \in X_j \\ \|x\|_{*_j} = 1}} \langle x, z \rangle_{*_j} \\ &= \sup_{\substack{x \in X_j \\ \|x\|_{*_j} = 1}} \langle x, (\Lambda_j T^{\#} T - T_j^{\#_j} T_j \Lambda_j) v \rangle_{*_j}, \end{aligned}$$

and the previous theorem gives us the desired bound for sufficiently large j. □

From Lemmas 34, 36, and 37 it follows that there exists $C_1, C_2 > 0$ and $j_k \in \mathbb{Z}^+$ such that for all $j \in \mathbb{Z}^+$ with $j \geq j_k$,

$$\|v - w\|_X \leq \|v - \Lambda_j v\|_X + \|\Lambda_j v - w\|_X$$

$$
\begin{aligned}
&\leq \|v - \Lambda_j v\|_X + \frac{1}{\sqrt{(1-\eta_j)\gamma}} \|\Lambda_j v - w\|_{*_j} \\
&\leq C_1(\varepsilon_j^2 + c_j) + \frac{1}{\sqrt{(1-\eta_j)\gamma}} \left(\rho_k^{(j)} \|(\Lambda_j T^\# T - T_j^{\#_j} T_j \Lambda_j) v\|_{*_j} \right) \\
&\leq C_1(\varepsilon_j^2 + c_j) + \frac{1}{(1-\eta_j)\sqrt{\gamma}} \left(\rho_k^{(j)} C_2 (c_j + \varepsilon_j^2) \right) \\
&= \left(C_1 + \frac{\rho_k^{(j)} C_2}{\sqrt{(1-\eta_j)\gamma}} \right) (\varepsilon_j^2 + c_j) \\
&\leq C(c_j + \varepsilon_j^2),
\end{aligned}
$$

where $C > 0$ is a constant depending on k.

Since E_k is a finite-dimensional subspace, the preceding lemmas hold if we assume that $\|v\|_* = 1$ instead of $\|v\|_X = 1$. If the term $\|v - w\|_X$ is replaced with $\|v - w\|_*$, the same reasoning above using properties of norms and Lemma 34 shows that if we assume that v is chosen from E_k such that $\|v\|_* = 1$, then there exists a constant $C_3 > 0$ such that for all $j \in \mathbb{Z}^+$ sufficiently large,

$$\|v - w\|_* \leq C_3(c_j + \varepsilon_j).$$

In this bound, we only have ε_j, as oppose to ε_j^2. This is a consequence of Lemma 34 since the $*$-norm is stronger than X-norm from the coercivity condition (1.9). These results justify the following theorem about the rate of convergence of the gap between E_k and $E_k^{(j)}$ to 0 as $j \to \infty$.

Theorem 38. *Let $k \in \mathbb{Z}^+$ be given. Then there exists a constant $C = C_k > 0$ such that for any $v \in E_k$ with $\|v\|_X = 1$ and for all $j \in \mathbb{Z}^+$ sufficiently large, there exists $w \in E_k^{(j)}$ satisfying*

$$
\begin{aligned}
\|v - w\|_X &\leq C(c_j + \varepsilon_j^2), \\
\|v - w\|_* &\leq C(c_j + \varepsilon_j).
\end{aligned}
$$

In terms of gap, it follows that for all $j \in \mathbb{Z}^+$ sufficiently large,

$$
\begin{aligned}
\hat{\delta}_X(E_k, E_k^{(j)}) &\leq C(c_j + \varepsilon_j^2), \\
\hat{\delta}_*(E_k, E_k^{(j)}) &\leq C(c_j + \varepsilon_j).
\end{aligned}
$$

Using Theorem 38, we are able to derive the rates of convergence of the other generalized singular spaces. To do this, we define, for each $k \in \mathbb{Z}^+$, the following spaces:

$$F_k = \operatorname{span}\{\psi_i : i \in I_k\},$$

$$
\begin{aligned}
F_k^{(j)} &= \operatorname{span}\{\psi_i^{(j)} : i \in I_k\}, \\
G_k &= \operatorname{span}\{\theta_i : i \in I_k\}, \\
G_k^{(j)} &= \operatorname{span}\{\theta_i^{(j)} : i \in I_k\}.
\end{aligned}
$$

Using the GSVE's of (T, L) and (T_j, L_j), it is clear that for any $k \in \mathbb{Z}^+$,

$$
\begin{aligned}
T(E_k) &= F_k, \\
L(E_k) &= G_k, \\
T_j\left(E_k^{(j)}\right) &= F_k^{(j)}, \\
L_j\left(E_k^{(j)}\right) &= G_k^{(j)}.
\end{aligned}
$$

The next theorem gives a rate of convergence of $F_k^{(j)}$ to F_k and of $G_k^{(j)}$ to G_k as $j \to \infty$.

Theorem 39. *Let $k \in M_0$ be given. Then there exists a constant $C = C_k > 0$ such that*

$$
\hat{\delta}(F_k, F_k^{(j)}) \leq C(c_j + \varepsilon_j^2) \text{ and } \hat{\delta}(G_k, G_k^{(j)}) \leq C(c_j + \varepsilon_j).
$$

Proof. Let $k \in M_0$. Then $T|_{E_k} : E_k \to F_k$ and $L|_{E_k} : E_k \to G_k$ are bijections. In [11], it is proven that the space $F_k^{(j)}$ converges to F_k and $G_k^{(j)}$ converges to G_k as $j \to \infty$ when $k \in M_0$. Therefore, the respective gaps between these spaces is less than 1 for $j \in \mathbb{Z}^+$ sufficiently large. Thus, it follows that for sufficiently large $j \in \mathbb{Z}^+$,

$$
\begin{aligned}
\delta\left(F_k, F_k^{(j)}\right) &= \hat{\delta}\left(F_k, F_k^{(j)}\right), \\
\delta\left(G_k, G_k^{(j)}\right) &= \hat{\delta}\left(G_k, G_k^{(j)}\right).
\end{aligned}
$$

By definition of the asymmetric gap,

$$
\begin{aligned}
\delta\left(F_k, F_k^{(j)}\right) &= \max_{\substack{u \in F_k \\ \|u\|_Y = 1}} \min_{v \in F_k^{(j)}} \|u - v\|_Y \\
&= \max_{\substack{u \in E_k \\ \|Tu\|_Y = 1}} \min_{v \in F_k^{(j)}} \|Tu - v\|_Y.
\end{aligned}
$$

When $T : E_k \to F_k$ is understood to be restricted to E_k, we will just write T instead of $T|_{E_k}$. This map defines a bijection of finite dimensional spaces and therefore has a bounded inverse $T^{-1} : F_k \to E_k$. Suppose that $u \in E_k$ such that $\|Tu\|_Y = 1$. Then,

$$
\|u\|_X = \|T^{-1}Tu\|_X \leq \|T^{-1}\|_{\mathcal{L}(F_k, E_k)} \|Tu\|_Y
$$

$$= \|T^{-1}\|_{\mathcal{L}(F_k,E_k)}.$$

Then

$$1 \leq \frac{\|T^{-1}\|_{\mathcal{L}(F_k,E_k)}}{\|u\|_X}.$$

From this, it then follows that

$$\begin{aligned}
\max_{\substack{u\in E_k\\ \|Tu\|_Y=1}} \min_{v\in F_k^{(j)}} \|Tu-v\|_Y &\leq \max_{\substack{u\in E_k\\ \|Tu\|_Y=1}} \min_{v\in F_k^{(j)}} \frac{\|Tu-v\|_Y}{\|u\|_X}\|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \\
&= \|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \max_{\substack{u\in E_k\\ \|u\|_X=1}} \min_{v\in F_k^{(j)}} \|Tu-v\|_Y.
\end{aligned}$$

From Theorem 38, it follows that there exists a constant $C' > 0$ such that for any $u \in E_k$ with $\|u\| \leq \|T^{-1}\|_{\mathcal{L}(F_k,E_k)}$, there exists $w \in E_k^{(j)}$ such that for all $j \in \mathbb{Z}^+$ sufficiently large,

$$\begin{aligned}
\|u-w\|_X &\leq C'(c_j+\varepsilon_j^2), \\
\|u-w\|_* &\leq C'(c_j+\varepsilon_j).
\end{aligned}$$

Since $T_jw \in F_k^{(j)}$, we have

$$\begin{aligned}
\|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \max_{\substack{u\in E_k\\ \|u\|_X=1}} \min_{v\in F_k^{(j)}} \|Tu-v\|_Y &\leq \|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \max_{\substack{u\in E_k\\ \|u\|_X=1}} \|Tu-T_jw\|_Y \\
&\leq \|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \max_{\substack{u\in E_k\\ \|u\|_X=1}} \big(\|T\|_{\mathcal{L}(X,Y)}\|u-w\|_X + \|(T-T_j)w\|_Y\big) \\
&\leq \|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \left(\|T\|_{\mathcal{L}(X,Y)}C'(c_j+\varepsilon_j^2) + \frac{t_j}{\sqrt{1-\eta_j}}\|v\|_*\right).
\end{aligned}$$

The term $\dfrac{t_j}{\sqrt{1-\eta_j}}\|v\|_*$ at the end of this inequality follows from the fact

$$\begin{aligned}
\|(T-T_j)w\|_Y \leq t_j\|w\|_* &\leq \frac{t_j}{\sqrt{1-\eta_j}}\|w\|_{*_j} \\
&= \frac{t_j}{\sqrt{1-\eta_j}}\|\Pi^{(j)}_{E_k^{(j)}}\Lambda_j v\|_{*_j} \\
&\leq \frac{t_j}{\sqrt{1-\eta_j}}\|\Lambda_j v\|_{*_j} \\
&\leq \frac{t_j}{\sqrt{1-\eta_j}}\|v\|_*.
\end{aligned}$$

Since E_k is a finite dimensional space, it follows that all norms are equivalent on E_k. Also, $t_j(1-\eta_j)^{-1/2} = O(c_j)$ as $j \to \infty$. Hence, there exists a constant $C_1 > 0$ such that for any $v \in E_k$ with $\|v\|_X = 1$ and for all $j \in \mathbb{Z}^+$ sufficiently large,

$$\|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \left(\|T\|_{\mathcal{L}(X,Y)} C'(c_j + \varepsilon_j^2) + \frac{t_j}{\sqrt{1-\eta_j}} \|v\|_* \right) \leq C_1(\varepsilon_j^2 + c_j).$$

Thus, it follows that for all $j \in \mathbb{Z}^+$ sufficiently large,

$$\begin{aligned}
\delta\left(F_k, F_k^{(j)}\right) &= \max_{\substack{u \in F_k \\ \|u\|_Y = 1}} \min_{v \in F_k^{(j)}} \|u - v\|_Y = \max_{\substack{u \in E_k \\ \|Tu\|_Y = 1}} \min_{v \in F_k^{(j)}} \|Tu - v\|_Y \\
&\leq \|T^{-1}\|_{\mathcal{L}(F_k,E_k)} \left(\|T\|_{\mathcal{L}(X,Y)} C'(c_j + \varepsilon_j^2) + \frac{t_j}{\sqrt{1-\eta_j}} \|v\|_* \right) \\
&\leq C_1(c_j + \varepsilon_j^2).
\end{aligned}$$

We now prove a rate of convergence of the spaces $G_k^{(j)}$ to G_k as $j \to \infty$. Again considering only the asymmetric gap $\delta\left(G_k, G_k^{(j)}\right)$, the same argument above gives

$$\begin{aligned}
\delta\left(G_k, G_k^{(j)}\right) &= \max_{\substack{u \in G_k \\ \|u\|_Z = 1}} \min_{v \in G_k^{(j)}} \|u - v\|_Z \\
&= \max_{\substack{u \in E_k \\ \|Lu\|_Z = 1}} \min_{v \in G_k^{(j)}} \|Gu - v\|_Z \\
&\leq \|L^{-1}\|_{\mathcal{L}(G_k,E_k)} \max_{\substack{u \in E_k \\ \|u\|_X = 1}} \left(\|L\|_{\mathcal{L}(D(L),Y)} \|u - w\|_* + \|(L - L_j)w\|_Z \right),
\end{aligned}$$

where w is the same as above. By the same argument as above together with Lemma 34, it follows that there exists a constant $C_2 > 0$ such that

$$\begin{aligned}
\delta\left(G_k, G_k^{(j)}\right) &\leq \|L^{-1}\|_{\mathcal{L}(G_k,E_k)} \max_{\substack{u \in E_k \\ \|u\|_X = 1}} \left(\|L\|_{\mathcal{L}(D(L),Y)} \|u - w\|_* + \|(L - L_j)w\|_Z \right) \\
&\leq C_2(c_j + \varepsilon_j).
\end{aligned}$$

By letting $C = \max\{C_1, C_2\}$, the theorem is then proven. $\square$

In the previous theorem, the assumption that $k \in M_0$ is necessary for the spaces $G_k^{(j)}$ to converge to G_k at the rate provided in the theorem. When $k \in M_b = \mathcal{N}(L)$, the space G_k of left generalized singular vectors in Z is trivial, and the gap between $G_k^{(j)}$ and G_k is either 0 or 1, depending on whether $G_k^{(j)}$ is trivial or not. Therefore, a rate of convergence is not sensible when $k \in M_b$. Also, one should notice that the rate of convergence of $G_k^{(j)}$ to G_k as $j \to \infty$ is one order of ε_j worse than the convergence rate of the other singular spaces. This follows from the possibility of L being unbounded with respect to the weaker norm $\|\cdot\|_X$. Hence, in the derivation above, the quantity $\|L(v-w)\|_Y$ must be compared to $\|v-w\|_*$ instead of $\|v-w\|_X$, which yields a worse rate of convergence.

Chapter 5

Higher-order convergence and future work

The theory established in Chapter 4 proves the orders of convergence observed in Example 3 for the corresponding generalized singular values and vectors. In this example, X_j was the discretization of the space $X = H^1(0,1)$ using continuous piecewise linear finite elements on a mesh with elements of length $h = 1/j$. The rates of convergence for the generalized singular values were as follows:

$$\left|a_k^{(j)} - a_k\right| = O(h^2) \text{ as } j \to \infty,$$
$$\left|b_k^{(j)} - b_k\right| = O(h^2) \text{ as } j \to \infty.$$

Also, the following rates of convergence for the generalized singular functions were as follows:

$$\left\|\phi_k^{(j)} - \phi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$
$$\left\|\psi_k^{(j)} - \psi_k\right\|_{L^2(0,1)} = O(h^2) \text{ as } j \to \infty,$$
$$\left\|\theta_k^{(j)} - \theta_k\right\|_{L^2(0,1)} = O(h) \text{ as } j \to \infty.$$

If we consider this same example, but with X_j and Y_j the space of continuous piecewise quadratic finite elements, and Z_j the space of piecewise linear finite elements (not necessarily continuous), we then observe the following rates of convergence for the generalized singular values:

$$\left|a_k^{(j)} - a_k\right| = O(h^3) \text{ as } j \to \infty,$$
$$\left|b_k^{(j)} - b_k\right| = O(h^3) \text{ as } j \to \infty.$$

Also, the following rates of convergence of the generalized singular functions are observed to be

$$\begin{aligned}\left\|\phi_k^{(j)} - \phi_k\right\|_{L^2(0,1)} &= O(h^3) \text{ as } j \to \infty,\\ \left\|\psi_k^{(j)} - \psi_k\right\|_{L^2(0,1)} &= O(h^3) \text{ as } j \to \infty,\\ \left\|\theta_k^{(j)} - \theta_k\right\|_{L^2(0,1)} &= O(h^2) \text{ as } j \to \infty.\end{aligned}$$

Similarly, if we let X_j and Y_j be the space of continuous piecewise cubic finite elements, and if we let Z_j be the space of piecewise quadratic finite elements (not necessarily continuous), we observe for the generalized singular values,

$$\begin{aligned}\left|a_k^{(j)} - a_k\right| &= O(h^4) \text{ as } j \to \infty,\\ \left|b_k^{(j)} - b_k\right| &= O(h^4) \text{ as } j \to \infty,\end{aligned}$$

and for the generalized singular functions,

$$\begin{aligned}\left\|\phi_k^{(j)} - \phi_k\right\|_{L^2(0,1)} &= O(h^4) \text{ as } j \to \infty,\\ \left\|\psi_k^{(j)} - \psi_k\right\|_{L^2(0,1)} &= O(h^4) \text{ as } j \to \infty,\\ \left\|\theta_k^{(j)} - \theta_k\right\|_{L^2(0,1)} &= O(h^3) \text{ as } j \to \infty.\end{aligned}$$

The theory of Chapter 4 is based on the assumption that there exists a sequence of positive real numbers $\{\varepsilon_j\}$ such that $\varepsilon_j \to 0$ as $j \to \infty$ and

$$\|\Pi_{X_j} x - x\|_* \leq \varepsilon_j \|x\|_{S_2}, \ \forall x \in S_2.$$

A generalization of this assumption would make sense if we are using higher order finite elements. We begin by defining the sequence S_n of Hilbert spaces ordered by containment. Since $M : D(L^*L) \to X$ is densely defined, self-adjoint, and strictly positive, it has a square root $A = M^{1/2}$. We define

$$\mathcal{M} = \bigcap_{k=0}^{\infty} D(A^k).$$

Then, by Lemma 8.17 of [6], $\mathcal{M}$ is dense in X, and we define

$$\langle x, y\rangle_{S_k} = \langle A^k x, A^k y\rangle_X \ , \ \|x\|_{S_k} = \|A^k x\|_X \ , \ \forall x, y \in \mathcal{M}.$$

It is easy to see that $\langle \cdot, \cdot \rangle_{S_k}$ defines an inner product on $\mathcal{M}$, and we define S_k to be the completion of $\mathcal{M}$ with respect to the norm $\| \cdot \|_{S_k}$. We say that the collection of spaces $\{S_k : k \in \mathbb{Z}, k \geq 0\}$ is the *Hilbert scale* define by A. By definition, $S_0 = X$. we also have that $S_1 = D(L)$ and the definition of S_2 is consistent with that given in chapter 4. Further, the sequence of norms $\| \cdot \|_{S_k}$ is increasing in strength as $k \to \infty$. This is summarized in following theorem.

Theorem 40.

1.) $S_1 = D(L)$ and $\langle \cdot, \cdot \rangle_{S_1} = \langle \cdot, \cdot \rangle_*$.

2.) For every $n, m \in \mathbb{Z}$ such that $m > n \geq 0$, it follows that $\| \cdot \|_{S_m}$ is a stronger norm than $\| \cdot \|_{S_n}$, and S_m is a dense subspace of S_n with respect the norm $\| \cdot \|_{S_n}$.

Proof. To prove (1), let $x, y \in D(L^*L)$. Since M is self-adjoint with respect to the X-inner product, it follows that $A = M^{1/2}$ is also self adjoint with respect to the X-inner product, and

$$\langle x, y \rangle_{S_1} = \langle Ax, Ay \rangle_X = \langle M^{1/2}x, M^{1/2}y \rangle_X = \langle x, My \rangle_X = \langle x, y \rangle_*.$$

By Lemma 23 from Chapter 4, S_2 is dense in $D(L)$ with respect to the norm $\| \cdot \|_*$. Thus, it follows that for every $x, y \in D(L)$,

$$\langle x, y \rangle_{S_1} = \langle x, y \rangle_*.$$

For a proof of (2), see [6], Proposition 8.19. □

We now make the following assumption about the discretizations X_j: There exists a sequence of positive real numbers $\{\varepsilon_j\}$ and a positive integer n such that $\varepsilon_j \to 0$ as $j \to \infty$ and

$$\|\Pi_{X_j} x - x\|_* \leq \varepsilon_j^k \|x\|_{S_{k+1}}, \ \forall x \in S_{k+1} \ , \ \forall k \in \mathbb{Z} \text{ such that } 0 \leq k \leq n. \tag{5.1}$$

Theorem 41. *For every $x \in S_{k+1}$,*

$$\|\Pi_{X_j} x - x\|_X \leq \varepsilon_j^{k+1} \|x\|_{S_{k+1}} \ , \ \forall k \in \mathbb{Z}^+ \text{ such that } 0 \leq k \leq n. \tag{5.2}$$

Proof. Let $x \in S_{k+1}$, and define $\hat{x} = \Pi_{X_j} x$ and $w = M^{-1}(x - \hat{x})$. Then $w \in S_2$ and

$$\begin{aligned}
\|x - \hat{x}\|_X^2 &= \langle x - \hat{x}, x - \hat{x} \rangle_X \\
&= \langle M^{-1}(x - \hat{x}), x - \hat{x} \rangle_* \\
&= \langle w, x - \hat{x} \rangle_* \\
&= \langle w - \Pi_{X_j} w, x - \hat{x} \rangle_*
\end{aligned}$$

$$\leq \|w - \Pi_{X_j} w\|_* \|x - \hat{x}\|.$$

Since $w \in S_2$ and $x \in S_{k+1}$, it follows from (5.1), that

$$\|w - \Pi_{X_j} w\|_* \leq \varepsilon_j \|w\|_{S_2} \ , \ \|x - \hat{x}\|_* \leq \varepsilon_j^k \|x\|_{S_{k+1}}.$$

Therefore, we have

$$\begin{aligned}
\|x - \hat{x}\|_X &\leq \|w - \Pi_{X_j}\|_* \|x - \hat{x}\|_* \\
&\leq (\varepsilon_j \|w\|_{S_2}) \left(\varepsilon_j^k \|x\|_{S_{k+1}}\right) \\
&= \varepsilon_j^{k+1} \|w\|_{S_2} \|x\|_{S_{k+1}} \\
&= \varepsilon_j^{k+1} \|Mw\|_X \|x\|_{S_{k+1}} \\
&= \varepsilon_j^{k+1} \|M\left(M^{-1}(x - \hat{x})\right)\|_X \|x\|_{S_{k+1}} \\
&= \varepsilon_j^{k+1} \|x - \hat{x}\|_X \|x\|_{S_{k+1}}.
\end{aligned}$$

After eliminating a factor of $\|x - \hat{x}\|_X$ from each side of the inequality, we have

$$\|x - \hat{x}\|_X \leq \varepsilon_j^{k+1} \|x\|_{S_{k+1}}.$$

□

In the case that the generalized singular vectors of the operator pair (T, L) are in the space S_{n+1}, where n is defined in our discretization assumption above, the following theorems can be proven by following the analysis of Chapter 3.

Theorem 42. *For each $k \in \mathbb{Z}^+$, there exists constants C_1 and C_2 such that for all sufficiently large positive integers j,*

$$a_k - C_1(c_j + \varepsilon_j^{n+1}) \leq a_k^{(j)} \leq a_k + C_2 c_j. \tag{5.3}$$

Theorem 43. *Let $k \in \mathbb{Z}^+$ be given. Then there exists a constant $C = C_k > 0$ such that for any $v \in E_k$ with $\|v\|_X = 1$ and for all $j \in \mathbb{Z}^+$ sufficiently large, there exits $w \in E_k^{(j)}$ satisfying*

$$\begin{aligned}
\|v - w\|_X &\leq C(c_j + \varepsilon_j^{n+1}), \\
\|v - w\|_* &\leq C(c_j + \varepsilon_j^{n}).
\end{aligned}$$

In terms of gap, it follows that for all $j \in \mathbb{Z}^+$ sufficiently large,

$$\begin{aligned}
\hat{\delta}_X(E_k, E_k^{(j)}) &\leq C(c_j + \varepsilon_j^{n+1}), \\
\hat{\delta}_*(E_k, E_k^{(j)}) &\leq C(c_j + \varepsilon_j^{n}).
\end{aligned}$$

Theorem 44. *Let $k \in M_0$ be given. Then there exists a constant $C = C_k > 0$ such* *that*

$$\hat{\delta}(F_k, F_k^{(j)}) \leq C(c_j + \varepsilon_j^{n+1}),$$
$$\hat{\delta}(G_k, G_k^{(j)}) \leq C(c_j + \varepsilon_j^{n}).$$

The proofs for these theorems depend on the right generalized singular vectors ϕ_k being in S_{n+1} for each $k \in \mathbb{Z}^+$. Let ϕ_k be a right generalized singular vector of (T, L). Then

$$T\phi_k = a_k\psi_k,\ T^{\#}\psi_k = a_k\phi_k,$$

where $a_k > 0$. Putting both of these things together, we have

$$T^{\#}T\phi_k = a_k^2\phi_k.$$

Therefore, using the fact that $T^{\#}T = M^{-1}T^*T$, it follows that

$$\phi_k = \frac{1}{a_k^2}M^{-1}T^*T\phi_k.$$

It is clear that $\phi_k \in S_2$, but in order for $\phi_k \in S_n$, it must follows that $T^*T\phi_k \in S_{n-2}$. This assumption may be rather strong for a few reasons. One reason is that T^*T may not be smoothing with respect to the operator L^*L. In other words, it may be that T^*T does not map S_i back into S_i for some integer i. Another reason is that the functions in S_i may have to satisfy more and more boundary conditions as i becomes larger. Such boundary conditions arise, for example, when T is the solution operator to the Laplace equation.

The analysis for these observed rates of convergence is still incomplete. The above issues suggest that further assumptions may need to be made pertaining to the discretization. If this is so, then it should be possible to find a compact operator T such that even with higher order elements, convergence of the generalized singular values and vectors should be no better than $O(h^p)$ for some a fixed integer p, regardless the order of the finite elements used.

References

[1] Andreas Kirsch. *An Introduction to the Mathematical Theory of Inverse Problems*. Springer-Verlag New York, 2011.

[2] Victor Isakov. *Inverse Problems for Partial Differential Equations*. Springer International Publishing, 2017.

[3] Charles W. Groetsch. *Inverse Problems in the Mathematical Sciences*. Vieweg+Teubner Verlag, 1993.

[4] Mark S. Gockenbach. *Linear Inverse Problems and Tikhonov Regularization*. MAA Press, 2016.

[5] Charles W. Groetsch. Integral equations of the first kind, inverse problems and regularization: a crash course. *Journal of Physics: Conference Series*, 73:012001, 2007.

[6] M. Hanke H. W. Engl and A. Neubauer. *Regularization of Inverse Problems*. Kluwer Academic Publishers Group, Dordrecht, 1996.

[7] Kendall Atkinson and Weimin Han. *Theoretical Numerical Analysis*. Springer-Verlag New York, 2009.

[8] Ake Bjrck. Least squares methods. In *Handbook of Numerical Analysis*, Volume 1, North Holland, Amsterdam, 1990. P.G. Ciarlet and J.L. Lions.

[9] Mark S. Gockenbach. Generalizing the GSVD. *SIAM Journal on Numerical Analysis*, 54:2417–2540, 2016.

[10] Daniel Boffi. Finite element approximation of eigenvalue problems. *Acta Numerica*, 19:1–120, 2010.

[11] Mark S. Gockenbach and Matthew J. Roberts. Approximating the generalized singular value expansion. *SIAM Journal on Numerical Analysis*, 56(5):2776–2795, 2018.

[12] Ivo M. Babuska and John E. Osborn. Eigenvalue Problems. In *Handbook of Numerical Analysis*, Volume 2: Finite Element Methods: Part 1, pages 641–787, North Holland, Amsterdam, 1991. P.G. Ciarlet and J.L. Lions, eds., vol. I.

[13] P.M. Anselone. *Collectively Compact Operator Approximation Theory and Integral Equations*. Prentice Hall, Englewood Cliffs, 1971.

[14] Susanne Brenner and Ridgway Scott. *The Mathematical Theory of Finite Element Methods 2nd Edition*. Springer-Verlag New York, 2002.

[15] Mark S. Gockenbach Daniel K. Crane and Matthew J. Roberts. Approximating the singular value expansion of a compact operator. *Under Review*.