9-11-2018

Approximating the generalized singular value expansion

Mark Gockenbach
Michigan Technological University

Matthew J. Roberts
Michigan Technological University

Recommended Citation

Gockenbach, M., & Roberts, M. J. (2018). Approximating the generalized singular value expansion. *SIAM Journal on Numerical Analysis, 56*(5), 2776-2795. http://dx.doi.org/10.1137/18M1163713
Retrieved from: https://digitalcommons.mtu.edu/math-fp/11

Follow this and additional works at: https://digitalcommons.mtu.edu/math-fp
Part of the Mathematics Commons

SIAM J. NUMER. ANAL.
Vol. 56, No. 5, pp. 2776–2795

APPROXIMATING THE GENERALIZED SINGULAR VALUE EXPANSION*

MARK S. GOCKENBACH† AND MATTHEW J. ROBERTS†

Abstract. The generalized singular value expansion (GSVE) simultaneously diagonalizes a pair of operators on Hilbert space. From a theoretical point of view, the GSVE enables a straightforward analysis of, for example, weighted least-squares problems and the method of Tikhonov regularization with seminorms. When the operators are discretized, an approximate GSVE can be computed from the generalized singular value decomposition of a pair of Galerkin matrices. Unless the discretization is carefully chosen, spurious modes can appear, but a natural condition on the discretization guarantees convergence of the approximate GSVE to the exact one. Numerical examples illustrate the pitfalls of a poor discretization and efficacy of the convergence conditions.

Key words. singular value expansion, convergence, Galerkin discretization

AMS subject classifications. 65J22, 47A52

DOI. 10.1137/18M1163713

1. Introduction. Many problems in computational mathematics require the simultaneous analysis of two operators defined on a Hilbert space. Perhaps the most common example is a linear inverse problem $Tx = y$, in which it is desired to estimate the solution x from a (noisy) data vector y. A true inverse problem is unstable (that is, x does not depend continuously on y); for this reason, it is common to choose a regularization operator L with the property that Lx is small for reasonable solutions x and large for those dominated by undesirable features (most commonly, L is a derivative operator). The solution x is then estimated by solving

$$\min_x \|Tx - y\|^2 + \lambda \|Lx\|^2, \tag{1}$$

where $\lambda > 0$ is a constant. This approach is called *Tikhonov regularization with seminorms*; the idea is to choose x so that Tx is close to y, while simultaneously requiring that Lx is not too large.

Closely related to the problem just described is the following *weighted least-squares* problem:

$$\begin{aligned} &\min \|Lx\|_Z^2 \\ &\text{s.t. } x \text{ is a least-squares solution of } Tx = y. \end{aligned} \tag{2}$$

Indeed, the goal of solving (1) is to estimate the solution of (2) for $y = y^*$, where y^* is the exact data.

One more problem that involves the two operators T and L is the following equality-constrained least-squares problem:

$$\begin{aligned} &\min \|Tx - y\|_Y^2 \\ &\text{s.t. } Lx = z. \end{aligned} \tag{3}$$

*Received by the editors January 3, 2018; accepted for publication (in revised form) June 28, 2018; published electronically September 11, 2018.
http://www.siam.org/journals/sinum/56-5/M116371.html

†Department of Mathematical Sciences, Michigan Technological University, Houghton, MI 49931-1295 (msgocken@mtu.edu, majrober@mtu.edu).

The *generalized singular value expansion* (GSVE) of an operator pair (T, L), introduced in [6], allows the two operators to be simultaneously diagonalized and thereby makes much desired analysis relatively transparent. To describe the GSVE, we establish the following conditions on the operators T and L; these conditions will be assumed throughout the paper. 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

$$\langle Tx, Tx\rangle_Y + \langle Lx, Lx\rangle_Z \geq \gamma \|x\|_X^2 \text{ for all } x \in D(L). \tag{4}$$

Condition (4) is a common assumption that guarantees, among other things, that the regularized problem (1) has a unique solution for each $\lambda > 0$. It has been the basis for much analysis; see, for instance, [10, Chapter 1] (which refers to (4) as the *completion condition*), [7, Chapter 5], or [5, Chapter 8]. We define the inner product $\langle\cdot,\cdot\rangle_*$ on $D(L)$ by $\langle x, y\rangle_* = \langle Tx, Ty\rangle_Y + \langle Lx, Ly\rangle_Z$ and write $\|.\|_*$ for the corresponding norm. It is well known that $D(L)$ is a Hilbert space under the inner product $\langle\cdot,\cdot\rangle_*$ (see, for instance, [9] or section 5.2 of [7]). For future reference, we introduce the notation $T^\#$ for the adjoint of T with respect to the $*$-inner product:

$$\langle Tx, y\rangle_Y = \left\langle x, T^\# y\right\rangle_* \text{ for all } x \in D(L), y \in Y.$$

The following theorem expresses the GSVE of (T, L).

THEOREM 1 ([6, Theorem 4.2]). *There exist 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 $\{\psi_k : k \in M_0 \cup M_b\} \subset Y$, $\{\theta_k : k \in M_0 \cup M_a\} \subset Z$, and subsets $\{a_k : k \in I\}$ and $\{b_k : k \in I\}$ of $\mathbb{R}$ such that*

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

and $0 \leq a_k, b_k \leq 1$, $a_k^2 + b_k^2 = 1$ for all $k \in I$.

Here, $\otimes_*$ refers to the outer product with respect to the $*$-inner product.

As an example of the utility of the GSVE, we note that, for

$$x = \sum_{k\in I} \alpha_k \phi_k \in D(L),$$

we have

$$\begin{aligned} \|Tx - y\|_Y^2 + \lambda\|Lx\|_Z^2 = &\sum_{k\in M_0} \left\{(a_k\alpha_k - \langle\psi_k, y\rangle_Y)^2 + \lambda b_k^2\alpha_k^2\right\} \\ &+ \sum_{k\in M_b} (a_k\alpha_k - \langle\psi_k, y\rangle_Y)^2 + \sum_{k\in M_a} \lambda b_k^2\alpha_k^2 + \|\hat{y}\|_Y^2, \end{aligned} \tag{5}$$

where $\hat{y}$ is the projection of y onto $\mathcal{R}(T)^\perp$. It then follows easily that the solution of (1) is

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

The use of GSVE reduces (1) from a problem of minimizing over all $x \in D(L)$ to that of minimizing over each α_k individually, the key being that each α_k appears in only one term on the right-hand side of (5). From this formula for $x_{\lambda,y}$, various properties of the solution can be examined.

The purpose of this paper is to propose and analyze a general approach to estimating the GSVE of an operator pair (T, L). Two approaches were presented in [6]. The first is based on recognizing that the pairs α_k^2, ϕ_k, $k \in I$ (with $a_k = 0$ for $k \in M_a$), are the eigenpairs of the compact self-adjoint operator $T^\sharp T$. These eigenpairs can be estimated using the general theory for symmetric, variationally posed eigenvalue problems, as presented in [4]. However, this approach has two shortcomings. We must choose a finite-dimensional subspace $\hat{X}$ of $D(L)$ with basis $\{x_1, x_2, \ldots, x_n\}$ 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

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

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 ($\mathcal{R}(T)$ infinite-dimensional and not closed), M_0 has infinite cardinality and $a_k \to 0$. 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 [6], does have the advantage that its convergence follows directly from the theory of symmetric, variationally posed eigenvalue problems.

The second approach described in [6] is based on reducing the computation to that of a (matrix) generalized singular value decomposition (GSVD). Here is one version of the GSVD.

THEOREM 2 ([3, Theorem 22.2]). *Let $A \in \mathbb{R}^{m\times n}$ and $B \in \mathbb{R}^{p\times n}$ be given 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 $C \in \mathbb{R}^{n\times n}$ and $S \in \mathbb{R}^{p\times n}$ such that*

$$A = UCW^{-1} \text{ and } B = VSW^{-1}.$$

Moreover, the diagonal entries of C and S are nonnegative and satisfy

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

In the next section, we will show how to compute the GSVE of a discretization (T_j, L_j) of the operator pair (T, L) by computing the GSVD of a pair of matrices. This is essentially the second approach taken in [6], although here we allow for a more general discretization. In the following sections, we present conditions on the convergence of T_j to T and L_j to L under which the GSVE of (T_j, L_j) is guaranteed to converge to the GSVE of (T, L). We also give examples of both convergence and nonconvergence.

2. Discretization. To discretize (T, L), we assume that $\{X_j : j \in \mathbb{Z}^+\}$ is a sequence of finite-dimensional subspaces of $D(L)$ with the property that

$$\bigcup_{j=1}^{\infty} X_j \text{ is dense in } D(L).$$

For each j, we assume that $\{x_1^{(j)}, x_2^{(j)}, \ldots, x_{n_j}^{(j)}\}$ is a basis for X_j. Similarly, we assume that

$$Y_j = \mathrm{sp}\left\{y_1^{(j)}, y_2^{(j)}, \ldots, y_{m_j}^{(j)}\right\}$$

and

$$Z_j = \mathrm{sp}\left\{z_1^{(j)}, z_2^{(j)}, \ldots, z_{p_j}^{(j)}\right\}$$

are subspaces of Y and Z, respectively, for each j, and that

$$\bigcup_{j=1}^{\infty} Y_j \text{ is dense in } Y \quad \text{and} \quad \bigcup_{j=1}^{\infty} Z_j \text{ is dense in } Z.$$

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. In the next section, we present conditions on $\{T_j\}$ and $\{L_j\}$ that guarantee that the GSVE of (T_j, L_j) converges to the GSVE of (T, L) as $j \to \infty$. First, however, we show how to compute the GSVE of (T_j, L_j) by computing the GSVD of a related pair of matrices.

We will need the discrete version of the $*$-inner product defined by

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

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

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

The conditions that we impose on $\{T_j\}$ and $\{L_j\}$ in the next section will ensure that this holds at least for all j sufficiently large. It then follows that $\langle\cdot,\cdot\rangle_{*_j}$ defines an inner product on X_j for each $j \in \mathbb{Z}^+$.

The following result shows how to compute the GSVE of (T_j, L_j). The Gram matrix for a basis $\{x_1, x_2, \ldots, x_n\}$ is $G \in \mathbb{R}^{n\times n}$ defined by $G_{ij} = \langle x_i, x_j\rangle$, where we use the inner product from the space to which the linearly independent set $\{x_1, x_2, \ldots, x_n\}$ belongs.

THEOREM 3. *Define* $A_j \in \mathbb{R}^{m_j\times n_j}$ *and* $B_j \in \mathbb{R}^{p_j\times n_j}$ *by*

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

Let H_j *and* J_j *be the Gram matrices for* $\{y_1^{(j)}, y_2^{(j)}, \ldots, y_{m_j}^{(j)}\}$ *and* $\{z_1^{(j)}, z_2^{(j)}, \ldots, z_{p_j}^{(j)}\}$, *respectively, and let*

$$H_j^{-1/2} A_j = UCW^{-1}, \quad J_j^{-1/2} B_j = VSW^{-1}$$

be the GSVD of $(H_j^{-1/2} A_j, J_j^{-1/2} B_j)$. *Define*

$$\tilde{U} = H_j^{-1/2} U \quad \text{and} \quad \tilde{V} = J_j^{-1/2} V.$$

Then the GSVE of (T_j, L_j) *is*

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

where

$$\phi_k^{(j)} = \sum_{i=1}^{n_j} W_{ik} x_i^{(j)}, \quad \psi_k^{(j)} = \sum_{i=1}^{m_j} \tilde{U}_{ik} y_i^{(j)}, \quad \theta_k^{(j)} = \sum_{i=1}^{p_j} \tilde{V}_{ik} z_i^{(j)}$$

and $a_1^{(j)}, \ldots, a_{n_j}^{(j)}$ *and* $b_1^{(j)}, \ldots, b_{n_j}^{(j)}$ *are the diagonal entries of* C *and* S*, respectively. We assume that* $m_j \geq n_j$*. If* $p_j < n_j$*, then* $b_{p_j+1}^{(j)}, \ldots, b_{n_j}^{(j)}$ *are defined to be* 0.

The proof of Theorem 3 is similar to that of Theorem 4.4 of [6]. In that paper, we considered only discretization $T_j = P_{Y_j} T|_{X_j}$, $L_j = P_{Z_j} T|_{X_j}$, where P_{Y_j} and P_{Z_j} denote the orthogonal projection operators onto Y_j and Z_j, respectively. However, the derivation of the GSVE in the case of an arbitrary discretization T_j, L_j is similar to the special case covered in [6]. For this reason, the proof of Theorem 3 will be omitted.

In Theorem 3, the sets $\{\phi_k^{(j)}\}$, $\{\psi_k^{(j)}\}$, and $\{\theta_k^{(j)}\}$ are orthonormal with respect to the $*_j$-, Y-, and Z-inner products, respectively. To compute these orthonormal bases, it is necessary to compute $H^{1/2}$ and $J^{1/2}$. In discretizations based on finite element spaces, the Gram matrices H and J are tridiagonal or at least banded, and it is computationally feasible to compute their square roots using the spectral decomposition.

As noted above, our main goal is to analyze the convergence of the GSVE of a discretization (T_j, L_j) of (T, L). Before proceeding to the analysis, we present an example to show that a seemingly natural discretization need not lead to convergence.

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 $k \geq 1$,

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

and $\phi_0(t) = 1$, $\psi_0(t) = 1$, $a_0 = 1$, $b_0 = 0$. It can be verified that $\{\phi_k\}_{k=0}^{\infty}$, $\{\psi_k\}_{k=0}^{\infty}$, 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, we have $M_0 = \mathbb{Z}^+$, $M_a = \emptyset$, 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 with j elements. Let $\{x_0, x_1, \ldots, x_j\}$

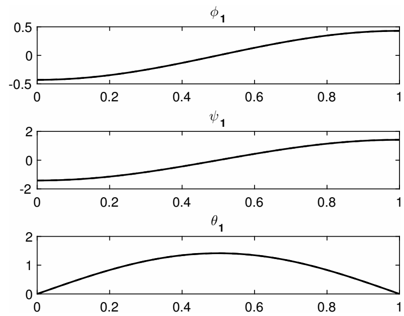

FIG. 1. *The computed values of $\phi_1^{(j)}$ (top), $\psi_1^{(j)}$ (middle), and $\theta_1^{(j)}$ (bottom) for Example 1, together with the corresponding exact functions ϕ_1, ψ_1, and θ_1. In each graph, the approximate function is the solid curve and the exact function is the dashed curve. The approximate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

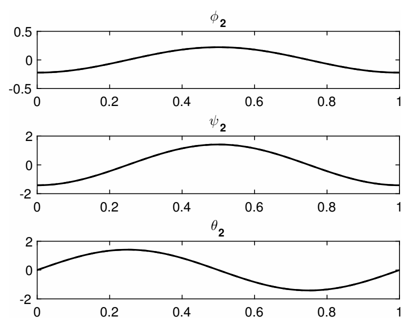

FIG. 2. *The computed values of $\phi_2^{(j)}$, $\psi_2^{(j)}$, and $\theta_2^{(j)}$ for Example 1, together with the corresponding exact functions ϕ_2, ψ_2, and θ_2. In each graph, the approximate function is the solid curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.*

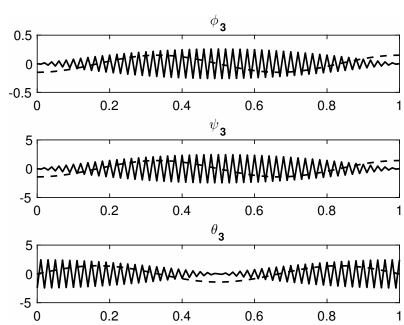

FIG. 3. *The computed values of $\phi_3^{(j)}$, $\psi_3^{(j)}$, and $\theta_3^{(j)}$ for Example 1, together with the corresponding exact functions ϕ_3, ψ_3, and θ_3. In each graph, the approximate function is the solid curve and the exact function is the dashed curve.*

be the standard nodal basis. Define T_j to be T restricted to X_j, and define L_j by $L_j = P_{Z_j} L|_{X_j}$, where P_{Z_j} is the orthogonal projection operator onto Z_j. We take $j = 100$ and compute the GSVE of (T_j, L_j) as described in Theorem 3 and graph $\phi_k^{(j)}$, $\psi_k^{(j)}$, and $\theta_k^{(j)}$ for $k = 1, 2, 3$ (see Figures 1–3).

We see that $\phi_1^{(j)}$, $\psi_1^{(j)}$, $\theta_1^{(j)}$ and $\phi_2^{(j)}$, $\psi_2^{(j)}$, $\theta_2^{(j)}$ are accurate approximations of the corresponding exact functions, but $\phi_3^{(j)}$, $\psi_3^{(j)}$, $\theta_3^{(j)}$ 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 [4]). The spurious mode persists as the mesh is refined.

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

In the next section, 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.

3. Convergence of the approximate GSVE.

3.1. Definition of convergence. Before we can analyze the convergence of the GSVE of (T_j, L_j) to that of (T, L), we must define what it means for a sequence of GSVEs to converge to a given GSVE. 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 [4] for a detailed discussion. In the case of eigenvalues and eigenvectors, we can expect that the eigenvalues of A_j 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 can be compared 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, there are probably k simple eigenvalues of A_j that converge to λ as $j \to \infty$. We will have to take this into account below.

When discussing the convergence of the GSVE, 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, inequality (4) 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

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

the singular values of T and L have the following properties:

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

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

With these definitions for M_b and M_a, we see that $\{a_k : k \in \mathbb{Z}^+\}$ is a nonincreasing sequence and $\{b_k : k \in \mathbb{Z}^+\}$ is a nondecreasing sequence. 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

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

For each j, let

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

be the GSVE of (T_j, L_j), written so that

$$a_1^{(j)} \geq a_2^{(j)} \geq \cdots \geq a_{n_j}^{(j)} \quad \text{and} \quad b_1^{(j)} \leq b_2^{(j)} \leq \cdots \leq b_{n_j}^{(j)}.$$

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 [4]).

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

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

It can be shown (in the Hilbert space setting, as we consider here) that

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

provided $\delta(U, V), \delta(V, U) < 1$ holds (see [8]).

Given the sequences $\{a_k\}$ and $\{b_k\}$ of singular values and the sequences $\{\phi_k\}$, $\{\psi_k\}$, and $\{\theta_k\}$ of singular vectors of (T, L), we make the following definitions:

$$S_k(\phi) = \mathrm{sp}\{\phi_i : a_i = a_k\},$$
$$S_k(\psi) = \mathrm{sp}\{\psi_i : a_i = a_k\},$$
$$S_k(\theta) = \mathrm{sp}\{\theta_i : a_i = a_k\}.$$

Typically, if a_k is a multiple singular value (that is, $\dim(S_k(\phi)) > 1$), say,

$$\{i \in \mathbb{Z}^+ : a_i = a_k\} = \{k_1, k_2, \ldots, k_q\},$$

then each $a_{k_r}^{(j)}$ will be a simple singular value of (T_j, L_j), meaning that

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

For this reason, we define

$$S_k^{(j)}(\phi) = \mathrm{sp}\{\phi_i^{(j)} : a_i^{(\ell)} \to a_k \text{ as } \ell \to \infty\},$$
$$S_k^{(j)}(\psi) = \mathrm{sp}\{\psi_i^{(j)} : a_i^{(\ell)} \to a_k \text{ as } \ell \to \infty\},$$
$$S_k^{(j)}(\theta) = \mathrm{sp}\{\theta_i^{(j)} : a_i^{(\ell)} \to a_k \text{ as } \ell \to \infty\}.$$

Note that because $a_k^2+b_k^2=1$ for all $k\in\mathbb{Z}^+$, $\{i\in\mathbb{Z}^+ : a_i=a_k\}=\{i\in\mathbb{Z}^+ : b_i=b_k\}$. 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).

Definition 5. *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 $\epsilon>0$, there exists j_0 such that all of the following are true for all $j\ge j_0$ and all $k=1,2,\dots,N$:*

$$\left|a_k^{(j)}-a_k\right|<\epsilon,\ \left|b_k^{(j)}-b_k\right|<\epsilon,$$
$$\hat{\delta}\left(S_k^{(j)}(\phi),S_k(\phi)\right)<\epsilon,\ \hat{\delta}\left(S_k^{(j)}(\psi),S_k(\psi)\right)<\epsilon,\ \hat{\delta}\left(S_k^{(j)}(\theta),S_k(\theta)\right)<\epsilon.$$

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

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

respectively.

Notice that Definition 5 does not refer to $\{\phi_k : k\in M_a\}$ or $\{\theta_k : k\in M_a\}$. Our theory will show that, in the representation

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

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 above, we see that the eigenpairs of $T^\#T$ are a_k^2,ϕ_k, $k=1,2,3,\dots$. Similarly, the eigenpairs for the operator $T_j^{\#_j}T_j$ are $a_{k,n}^2,\phi_{k,n}$, $k=1,2,\dots,n_j$. 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 $T_j^{\#_j}T_jP_{X_j} : X\to X$, 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 $T_j^{\#_j}T_jP_{X_j}$ to X_j, it is immediate that an eigenpair of $T_j^{\#_j}T_j$ is an eigenpair of $T_j^{\#_j}T_jP_{X_j}$. Conversely, if $T_j^{\#_j}T_jP_{X_j}x=\lambda x$, then since $T_j^{\#_j}T_jP_{X_j}$ 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 Babuška and Osborn ([2]; see also [4, sections 6 and 9]) shows that if a sequence $\{A_j\}$ of compact operators ($A_j : X\to X$ for all $j\in\mathbb{Z}^+$) converges in norm to the compact operator $A : X\to X$, then eigensystems of A_j converge to that of A as $j\to\infty$, provided we exclude the zero eigenvalues of A from consideration. Specifically, we have the following theorem [4, Theorem 9.1] (in which $\rho(A)$ denotes the resolvent set of A).

THEOREM 6. *Let $A : X \to X$ be a compact linear operator, and let $\{A_j\}$ be a sequence of compact operators from X to X such that*

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

Then for any compact set $K \subset \rho(A)$, there exists $j_0 \in \mathbb{Z}^+$ such that for every $j \geq j_0$, we have $K \subset \rho(A_j)$. If λ is a nonzero eigenvalue of A with algebraic multiplicity m, then there are m eigenvalues $\lambda_{1,j}, \lambda_{2,j}, \dots, \lambda_{m,j}$ of A_j, repeated according to their algebraic multiplicities, such that each $\lambda_{i,n}$ converges to λ as $j \to \infty$. Moreover, if we define $E_j(\lambda)$ to be the sum of the eigenspaces of $\lambda_{1,j}, \lambda_{2,j}, \dots, \lambda_{m,j}$, then the gap between $E_j(\lambda)$ and the eigenspace $E(\lambda)$ of λ tends to zero 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 norm, then it will follow that the eigensystem of $T_j^{\#_j} T_j P_{X_j}$ converges to the eigensystem of $T^{\#}T$.

3.2. Preliminary results. We will use the following fundamental result.

THEOREM 7 ([1]). *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 $M_j \to M$ 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. For each $j \in \mathbb{Z}^+$, we define

$$\begin{aligned} t_{j,1} &= \max_{\substack{x \in X_j \\ x \neq 0}} \frac{\|(T - T_j)x\|_Y}{\|x\|_X}, \; t_{j,2} = \max_{\substack{x \in X_j \\ x \neq 0}} \frac{\|(T - T_j)x\|_Y}{\|x\|_*}, \; t_j = \max\{t_{j,1}, t_{j,2}\}, \\ \ell_j &= \max_{\substack{x \in X_j \\ x \neq 0}} \frac{\|(L - L_j)x\|_Z}{\|x\|_*}, \; c_j = \sqrt{t_j^2 + \ell_j^2}. \end{aligned} \tag{6}$$

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 (6), we have

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

and

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

Therefore, for all $x \in X_j$,

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

Therefore, we have

$$\|T_j x\|_Y \leq t_j \|x\|_* + \|Tx\|_Y \leq (1 + t_j)\|x\|_* \text{ for all } x \in X_j \tag{7}$$

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

$$\|L_j x\|_Z \leq (1+\ell_j)\|x\|_* \text{ for all } x \in X_j. \tag{8}$$

We now derive several preliminary results that are needed for our main convergence theorem (Theorem 18). These results concern the relationship between the $*_j$- and $*$-inner products and norms (Lemma 8 and Corollary 9), the basic properties of the operators $M = T^*T + L^*L$ and $M_j = T_j^*T_j + L_j^*L_j$ (Theorems 10 and 11), the relationship between M and M_j (Theorems 12 and 13), a property of P_{X_j} (Lemma 15), and various results about the convergence of T_j to T (Theorems 14, 16, and 17).

LEMMA 8.

$$|\langle x,y\rangle_* - \langle x,y\rangle_{*_j}| \leq (c_j^2 + 2(t_j+\ell_j))\|x\|_*\|y\|_* \textit{ for all } x,y \in X_j.$$

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

$$\begin{aligned}
&|\langle x,y\rangle_* - \langle x,y\rangle_{*_j}| \\
&= |\langle Tx,Ty\rangle_Y + \langle Lx,Ly\rangle_Z - \langle T_jx,T_jy\rangle_Y - \langle L_jx,L_jy\rangle_Z| \\
&\leq |\langle Tx,Ty\rangle_Y + \langle Lx,Ly\rangle_Z - \langle Tx,T_jy\rangle_Y - \langle Lx,L_jy\rangle_Z + \langle Tx,T_jy\rangle_Y + \langle Lx,L_jy\rangle_Z \\
&\quad - \langle T_jx,T_jy\rangle_Y - \langle L_jx,L_jy\rangle_Z| \\
&= |\langle Tx,(T-T_j)y\rangle_Y + \langle Lx,(L-L_j)y\rangle_Z + \langle (T-T_j)x,T_jy\rangle_Y + \langle (L-L_j)x,L_jy\rangle_Z| \\
&\leq t_j\|Tx\|_Y\|y\|_* + \ell_j\|Lx\|_Z\|y\|_* + t_j\|x\|_*\|T_jy\|_Y + \ell_j\|x\|_*\|L_jy\|_Z \\
&\leq (t_j+\ell_j)\|x\|_*\|y\|_* + t_j(1+t_j)\|x\|_*\|y\|_* + \ell_j(1+\ell_j)\|x\|_*\|y\|_* \\
&= (t_j^2+\ell_j^2+2(t_j+\ell_j))\|x\|_*\|y\|_* \\
&= (c_j^2+2(t_j+\ell_j))\|x\|_*\|y\|_*. \qquad \square
\end{aligned}$$

Since $c_j, t_j, \ell_j \to 0$ as $j \to \infty$, we see that $1-c_j^2-2(t_j+\ell_j) > 0$ for all j sufficiently large, say $j \geq \bar{j}$.

COROLLARY 9. *For all* $j \geq \bar{j}$,

$$(1-c_j^2-2(t_j+\ell_j))\|x\|_*^2 \leq \|x\|_{*_j}^2 \leq (1+c_j^2+2(t_j+\ell_j))\|x\|_*^2 \textit{ for all } x \in X_j, \tag{9}$$

$$\frac{1}{1+c_j^2+2(t_j+\ell_j)}\|x\|_{*_j}^2 \leq \|x\|_*^2 \leq \frac{1}{1-c_j^2-2(t_j+\ell_j)}\|x\|_{*_j}^2 \textit{ for all } x \in X_j. \tag{10}$$

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

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

These operators will be central to our analysis; the following three results describe some properties of M and M_j that will be needed.

THEOREM 10. *The operator M is a bijection with bounded inverse, and*

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

that is,

$$\|M^{-1}x\|_* \leq \frac{\|x\|_X}{\sqrt{\gamma}} \textit{ for all } x \in X,$$

where γ is the constant from condition (4).

Proof. See [7, Theorem 5.25]. ◻

THEOREM 11. *For all $j \geq \bar{j}$, the operator M_j is invertible, with*

$$\|M_j^{-1}\|_{\mathcal{L}(X,D(L))} \leq \frac{1}{(1-c_j^2-2(t_j+\ell_j))\sqrt{\gamma}}.$$

That is,

$$\|M_j^{-1}x\|_* \leq \frac{1}{(1-c_j^2-2(t_j+\ell_j))\sqrt{\gamma}}\|x\|_X \text{ for all } x \in X_j.$$

Proof. Let $x \in X_j$. By Corollary 9, we have

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

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

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

THEOREM 12. *For all $x \in X$ and all $j \geq \bar{j}$,*

$$\|\Pi_{X_j}M^{-1}x - M_j^{-1}P_{X_j}x\|_* \leq \frac{c_j^2+2(t_j+\ell_j)}{(1-c_j^2-2(t_j+\ell_j))\sqrt{\gamma}}\|x\|_X.$$

Proof. Let $x \in X$. Then

$$\begin{aligned}
&\|\Pi_{X_j}M^{-1}x - M_j^{-1}P_{X_j}x\|_*^2 \\
&= \left\langle \Pi_{X_j}M^{-1}x - M_j^{-1}P_{X_j}x, (\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\right\rangle_* \\
&\leq \left\langle M^{-1}x, (\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\right\rangle_* - \left\langle M_j^{-1}P_{X_j}x, (\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\right\rangle_{*_j} \\
&\qquad + (c_j^2+2(t_j+\ell_j))\|M_j^{-1}P_{X_j}x\|_*\|(\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\|_*,
\end{aligned}$$

where we have applied Lemma 8 at the last step. Moreover,

$$\left\langle M^{-1}x, (\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\right\rangle_* = \left\langle x, (\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\right\rangle_X$$

and

$$\left\langle M_j^{-1}P_{X_j}x, (\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\right\rangle_{*_j} = \left\langle x, (\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\right\rangle_X.$$

It follows that

$$\|\Pi_{X_j}M^{-1}x - M_j^{-1}P_{X_j}x\|_*^2 \leq (c_j^2+2(t_j+\ell_j))\|M_j^{-1}P_{X_j}x\|_*\|(\Pi_{X_j}M^{-1} - M_j^{-1}P_{X_j})x\|_*$$

and hence that

$$\|\Pi_{X_j} M^{-1}x - M_j^{-1}P_{X_j}x\|_* \le (c_j^2 + 2(t_j+\ell_j))\|M_j^{-1}P_{X_j}x\|_*.$$

Applying Theorem 11 (and the fact that $\|P_{X_j}x\|_X \le \|x\|_X$), we obtain

$$\|\Pi_{X_j} M^{-1}x - M_j^{-1}P_{X_j}x\|_* \le \frac{c_j^2 + 2(t_j+\ell_j)}{(1-c_j^2-2(t_j+\ell_j))\sqrt{\gamma}}\|x\|_X,$$

as desired. ∎

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

THEOREM 13. *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

$$\begin{aligned}\|M^{-1}x - M_j^{-1}P_{X_j}x\|_* &\le \|M^{-1}x - \Pi_{X_j}M^{-1}x\|_* + \|\Pi_{X_j}M^{-1}x - M_j^{-1}P_{X_j}x\|_* \\ &\le \|(I-\Pi_{X_j})M^{-1}x\|_* + \frac{c_j^2 + 2(t_j+\ell_j)}{(1-c_j^2-2(t_j+\ell_j))\sqrt{\gamma}}\|x\|_X.\end{aligned}$$

By assumption, $\cup_{j=1}^{\infty}X_j$ is a dense subset of $D(L)$ with respect to the $*$-norm, and $t_j, \ell_j, c_j \to 0$ as $j\to\infty$. The desired result follows. ∎

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

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

Also,

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

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

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

Similarly,

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

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

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

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

$$\|P_{X_j}T^* - T_j^*P_{Y_j}\|_{\mathcal{L}(Y,X_j)} \to 0 \text{ as } j\to\infty. \tag{11}$$

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

THEOREM 14. *$T_j^{\#_j} P_{Y_j} T \to T^{\#} T$ in the $\mathcal{L}(D(L), D(L))$-norm.*

Proof. By definition,

$$T^{\#} = (T^*T + L^*L)^{-1}T^* = M^{-1}T^*.$$

We also have

$$T_j^{\#_j} = (T_j^*T_j + L_j^*L_j)^{-1}T_j^* = M_j^{-1}T_j^*.$$

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 7, it suffices to prove that $M_j^{-1}T_j^* P_{Y_j} - M^{-1}T^* \to 0$ pointwise on Y. Let $y \in Y$. Then

$$\begin{aligned}
&\|M^{-1}T^*y - M_j^{-1}T_j^*P_{Y_j}y\|_* \\
&\le \|(M^{-1} - M_j^{-1}P_{X_j})T^*y\|_* + \|M_j^{-1}P_{X_j}T^*y - M_j^{-1}T_j^*P_{Y_j}y\|_* \\
&= \|(M^{-1} - M_j^{-1}P_{X_j})T^*y\|_* + \|M_j^{-1}(P_{X_j}T^*y - T_j^*P_{Y_j}y)\|_* \\
&\le \|(M^{-1} - M_j^{-1}P_{X_j})T^*y\|_* + \frac{1}{(1 - c_j^2 - 2(t_j + \ell_j))\sqrt{\gamma}}\|(P_{X_j}T^* - T_j^*P_{Y_j})y\|_X.
\end{aligned}$$

It now follows from Theorem 13 and (11) that $\|M^{-1}T^*y - M_j^{-1}T_j^*P_{Y_j}y\|_* \to 0$ as $j \to \infty$. ∎

We need two more results.

LEMMA 15. *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

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

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

THEOREM 16. *$T_jP_{X_j} \to T$ in the $\mathcal{L}(X, Y)$-norm.*

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

$$\|v_{j_k}\|_X = 1 \quad \text{and} \quad \|T_{j_k}P_{X_{j_k}}v_{j_k} - Tv_{j_k}\|_Y \ge \epsilon. \tag{12}$$

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

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

($\|(T_{j_k} - T)P_{X_{j_k}}v_{j_k}\|_Y \le t_j \to 0$, and $TP_{X_{j_k}}v_{j_k} \to Tv$ because $P_{X_{j_k}}v_{j_k} \to v$ weakly 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 (12). The contradiction completes the proof. ∎

We have been working toward the following result.

THEOREM 17. *$T_j^{\#_j} T_j P_{X_j} \to T^\# T$ 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 14. Therefore, it suffices to show that the first term goes to 0. Applying Theorem 11, we have

$$\begin{aligned}
&\|M_j^{-1} T_j^* P_{Y_j}(T_j P_{X_j} - T)\|_{\mathcal{L}(D(L),D(L))} \\
&\le \frac{1}{(1 - c_j^2 - 2(t_j + \ell_j))\sqrt{\gamma}} \|T_j^* P_{Y_j}(T_j P_{X_j} - T)\|_{\mathcal{L}(D(L),X_j)} \\
&\le \frac{\|T_j^*\|_{\mathcal{L}(Y_j,X_j)}}{(1 - c_j^2 - 2(t_j + \ell_j))\sqrt{\gamma}} \|T_j P_{X_j} - T\|_{\mathcal{L}(D(L),Y)} \\
&= \frac{\|T_j\|_{\mathcal{L}(X_j,Y_j)}}{(1 - c_j^2 - 2(t_j + \ell_j))\sqrt{\gamma}} \|T_j P_{X_j} - T\|_{\mathcal{L}(D(L),Y)} \\
&\le \frac{t_j + \|T\|_{\mathcal{L}(X,Y)}}{(1 - c_j^2 - 2(t_j + \ell_j))\sqrt{\gamma}} \|T_j P_{X_j} - T\|_{\mathcal{L}(D(L),Y)} \\
&\le \frac{t_j + \|T\|_{\mathcal{L}(X,Y)}}{(1 - c_j^2 - 2(t_j + \ell_j))\gamma} \|T_j P_{X_j} - T\|_{\mathcal{L}(X,Y)}.
\end{aligned}$$

Since t_j goes to 0 and T is a bounded operator, it follows from Theorem 16 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. ∎

3.3. Convergence theorem. Theorem 17, together with Theorem 6, shows 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 our main theorem.

THEOREM 18. *Assuming that $c_j \to 0$ as $j \to \infty$ (where c_j is defined by (6)), the GSVE of (T_j, L_j) converges to the GSVE of (T, L) in the sense of Definition 5.*

Proof. Since $a_k^{(j)}, \phi_k^{(j)}$, $k = 1, 2, \dots, n_j$, are the eigenpairs of $T_j^{\#_j} T_j P_{X_j}$, a_k, ϕ_k are the eigenpairs of $T^\# T$, and $T_j^{\#_j} T_j P_{X_j} \to T^\# T$ in norm, it follows from Theorem 6 that $\{a_k^{(j)}\}$ converges to $\{a_k\}$ and $\{\phi_k^{(j)}\}$ converges to $\{\phi_k\}$ in the manner described by Definition 5. Moreover, since $(a_k^{(j)})^2 + (b_k^{(j)})^2 = 1$ for all $k = 1, 2, \dots, n_j$ and $a_k^2 + b_k^2 = 1$ for all $k \in \mathbb{Z}^+$, it follows that $\{b_k^{(j)}\}$ also converges to $\{b_k\}$, consistent with Definition 5.

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

$$j \ge j_0 \ \Rightarrow \ \max\left\{\delta\left((S_k(\phi), S_k^{(j)}(\phi)\right), \delta\left(S_k^{(j)}(\phi), S_k(\phi)\right)\right\} < \epsilon.$$

First, we show that $j_0 \in \mathbb{Z}^+$ can be chosen such that $\delta(S_k(\phi), S_k^{(j)}(\phi)) < \epsilon$ for all

$j \geq j_0$, that is, that

$$j \geq j_0 \;\Rightarrow\; \sup_{\substack{y\in S_k(\psi)\\ \|y\|_Y=1}} \inf_{v\in S_k^{(j)}(\phi)} \|y-v\|_Y < \epsilon. \tag{13}$$

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

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

We will show that this value of j_0 satisfies (13). It suffices to show that, for any $j \geq j_0$ and any $y \in S_k(\psi)$ satisfying $\|y\|_Y = 1$, there exists $v \in S_k^{(j)}(\psi)$ such that $\|y-v\|_Y < \epsilon$. Suppose

$$S_k(\psi) = \mathrm{sp}\{\psi_{k_1}, \psi_{k_2}, \ldots, \psi_{k_q}\}.$$

There exist real numbers $\alpha_1, \alpha_2, \ldots, \alpha_q$ such that

$$y = \sum_{i=1}^{q} \alpha_i \psi_{k_i} \text{ and } \sum_{i=1}^{q} \alpha_i^2 = 1.$$

But then, since $T\phi_{k_i} = \alpha_k \psi_{k_i}$ for $i = 1, 2, \ldots, q$, we have

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

where

$$x = \sum_{i=1}^{q} \alpha_i \phi_{k_1}.$$

Moreover, since $\{\phi_{k_1}, \phi_{k_2}, \ldots, \phi_{k_1}\}$ (like $\{\psi_{k_1}, \psi_{k_2}, \ldots, \psi_{k_1}\}$) is orthonormal, we see that $\|x\|_* = \|y\|_Y = 1$. Hence, there exists $u \in S_k^{(j)}(\phi)$ such that

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

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

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

Here we have used (7) to bound $\|T_j\|_{\mathcal{L}(D(L),Y)}$.

This concludes the proof that j_0 satisfies (13), and hence that $\delta(S_k(\phi), S_k^{(j)}(\phi)) < \epsilon$ for all $j \geq j_0$. The proof that $j_0 \in \mathbb{Z}^+$ can be chosen so that also $\delta(S_k^{(j)}(\phi), S_k(\phi)) < \epsilon$ for all $j \geq j_0$ is similar. Thus we have shown that $\{\psi_k^{(j)}\}$ converges to $\{\psi_k\}$ in the sense of Definition 5.

The proof that $\{\theta_k^{(j)}\}$ converges to $\{\theta_k\}$ in the sense of Definition 5 is exactly the same, and the proof of the theorem is complete. ∎

4. Further numerical examples. Since we observed nonconvergence in Example 1, the discretization (T_j, L_j) must fail to satisfy the hypotheses of Theorem 18.

Example 1 (continued). In this example, $T_j = T|_{X_j}$ and hence $t_j = 0$ for all j. However (recalling that x_i is the ith standard nodal basis function), a direct calculation shows that

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

(where $h = 1/j$), and hence ℓ_j is bounded away from 0. Therefore, Theorem 18 does not apply to this example.

We now present a discretization of the operators of Example 1 that satisfies the hypotheses of Theorem 18 and hence leads to convergence.

Example 2. Let T, L, X_j, and Y_j be as defined 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 defined to be $T|_{X_j}$, 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 $t_j = \ell_j = 0$ for all j, and hence Theorem 18 guarantees that the GSVE of (T_j, L_j) converges to the GSVE of (T, L) (in the sense of Definition 5) as $j \to \infty$.

Figures 4–6 show the approximate and exact singular functions for $k = 1, 2, 3$ (analogous to Figures 1–3). 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.

Extensive numerical testing suggests that

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

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

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

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

We close this paper with another example.

Example 3. Define $\hat{X} = \{x \in L^2(0,1) : \int_0^1 x = 0\}$ and define $T : \hat{X} \to \hat{X}$ by the

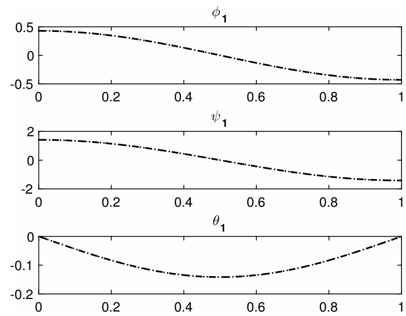

FIG. 4. *The computed values of* $\phi_1^{(j)}$ *(top),* $\psi_1^{(j)}$ *(middle), and* $\theta_1^{(j)}$ *(bottom) for Example* 2*, together with the corresponding exact functions* ϕ_1*,* ψ_1*, and* θ_1*. In each graph, the approximate function is the dotted curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.*

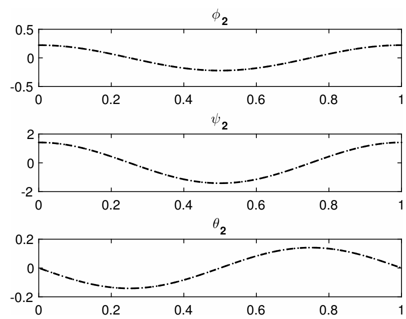

FIG. 5. *The computed values of* $\phi_2^{(j)}$*,* $\psi_2^{(j)}$*, and* $\theta_2^{(j)}$ *for Example* 2*, together with the corresponding exact functions* ϕ_2*,* ψ_2*, and* θ_2*. In each graph, the approximate function is the dotted curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.*

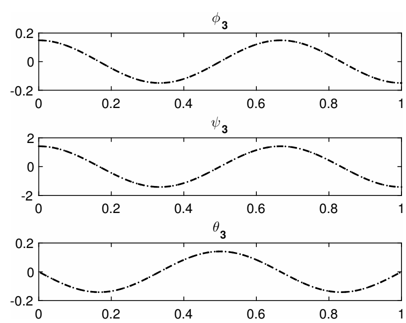

FIG. 6. *The computed values of* $\phi_3^{(j)}$*,* $\psi_3^{(j)}$*, and* $\theta_3^{(j)}$ *for Example* 2*, together with the corresponding exact functions* ϕ_1*,* ψ_1*, and* θ_1*. In each graph, the approximate function is the dotted curve and the exact function is the dashed curve. The approximate and exact curves are indistinguishable at this scale.*

condition that $u = Tf$ is the solution of the weak form of the boundary value problem

$$\begin{aligned} -u'' &= f \text{ in } (0,1), \\ u'(0) &= u'(1) = 0, \\ \int_0^1 u &= 0. \end{aligned}$$

(Note that this BVP has a unique solution in $\hat{X}$ for each $f \in \hat{X}$.) Thus $u = Tf$ is the solution of the variational problem

$$u \in \hat{X}_1,\ \langle u', v' \rangle_{L^2(0,1)} = \langle f, v \rangle_{L^2(0,1)} \text{ for all } v \in \hat{X}_1,$$

where $\hat{X}_1 = \hat{X} \cap H^1(0,1)$. We also define $L : \hat{X}_1 \to L^2(0,1)$ by $Lx = x'$ (thus $D(L) = \hat{X}_1$). It is straightforward to derive the GSVE of (T, L) using Fourier analysis, specifically the Fourier cosine series. The result is

$$T = \sum_{n=1}^{\infty} a_n \psi_n \otimes_* \phi_n, \quad L = \sum_{n=1}^{\infty} b_n \theta_n \otimes_* \phi_n,$$

where

$$\phi_n(t) = \frac{n^2\pi^2\sqrt{2}}{\sqrt{n^6\pi^6+1}} \cos(n\pi t),\ \psi_n(t) = \sqrt{2}\cos(n\pi t),\ \theta_n(t) = -\sqrt{2}\sin(n\pi t),$$
$$a_n = \frac{1}{\sqrt{n^6\pi^6+1}},\ b_n = \frac{n^3\pi^3}{\sqrt{n^6\pi^6+1}}.$$

In terms of the notation of this paper, we have $X = Y = \hat{X}$ and $Z = L^2(0,1)$. We impose the L^2-norm on all three spaces X, Y, and Z.

To discretize this example, we define $X_j = Y_j$ to be the space of continuous piecewise linear functions with mean 0 on a uniform mesh of j elements and Z_j to be the space of piecewise constant functions defined on the same mesh. We define $T_j : X_j \to X_j$ by the condition that, for $f \in X_j$, $u = T_j f$ is the solution of

$$u \in X_j,\ \langle u', v' \rangle_{L^2(0,1)} = \langle f, v \rangle_{L^2(0,1)} \text{ for all } v \in X_j.$$

We also define $L_j : X_j \to Z_j$ by $L_j x = x'$ for all $x \in X_j$.

By standard finite element analysis, we have $t_j = O(h^2)$ and, since $L_j = L|_{X_j}$, $\ell_j = 0$ for all j. It follows from Theorem 18 that the GSVE of (T_j, L_j) is guaranteed to converge, in the sense of Definition 5, to the GSVE of (T, L).

Tables 1 and 2 show the errors in the computed estimates of $\phi_0, \phi_1, \phi_2, \phi_3, \phi_4$ and $\theta_1, \theta_2, \theta_3, \theta_4$, respectively. These errors are consistent with

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

TABLE 1
Example 3: Errors in $\phi_k^{(j)}$.

k	$n = 40$	$n = 80$	$n = 160$	$n = 320$	$n = 640$
0	$1.914 \cdot 10^{-15}$	$1.840 \cdot 10^{-15}$	$2.891 \cdot 10^{-15}$	$7.660 \cdot 10^{-15}$	$1.233 \cdot 10^{-14}$
1	$1.3104 \cdot 10^{-4}$	$3.2756 \cdot 10^{-5}$	$8.1887 \cdot 10^{-6}$	$2.0472 \cdot 10^{-6}$	$5.1179 \cdot 10^{-7}$
2	$2.7773 \cdot 10^{-4}$	$6.9408 \cdot 10^{-5}$	$1.7350 \cdot 10^{-5}$	$4.3375 \cdot 10^{-6}$	$1.0844 \cdot 10^{-6}$
3	$4.2161 \cdot 10^{-4}$	$1.0532 \cdot 10^{-4}$	$2.6325 \cdot 10^{-5}$	$6.5809 \cdot 10^{-6}$	$1.6452 \cdot 10^{-6}$
4	$5.6487 \cdot 10^{-4}$	$1.4103 \cdot 10^{-4}$	$3.5244 \cdot 10^{-5}$	$8.8103 \cdot 10^{-6}$	$2.2025 \cdot 10^{-6}$

TABLE 2
Example 3: Errors in $\theta_k^{(j)}$.

k	$j=40$	$j=80$	$j=160$	$j=320$	$j=640$
1	$2.2671\cdot10^{-2}$	$1.1336\cdot10^{-2}$	$5.6681\cdot10^{-3}$	$2.8341\cdot10^{-3}$	$1.4170\cdot10^{-3}$
2	$4.5333\cdot10^{-2}$	$2.2671\cdot10^{-2}$	$1.1336\cdot10^{-2}$	$5.6681\cdot10^{-3}$	$2.8341\cdot10^{-3}$
3	$6.7978\cdot10^{-2}$	$3.4004\cdot10^{-2}$	$1.7004\cdot10^{-2}$	$8.5021\cdot10^{-3}$	$4.2511\cdot10^{-3}$
4	$9.0597\cdot10^{-2}$	$4.5333\cdot10^{-2}$	$2.2671\cdot10^{-2}$	$1.1336\cdot10^{-2}$	$5.6681\cdot10^{-3}$

To save space, we do not display the errors in $\psi_k^{(j)}$, $a_k^{(j)}$, and $b_k^{(j)}$, but these errors are consistent with

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

Our analysis does not predict any particular rate of convergence. In Examples 2 and 3, we observed the expected (optimal) rates of convergence: $O(h^2)$ for $\{\phi_k^{(j)}\}$ and $\{\psi_k^{(j)}\}$, and $O(h)$ for $\{\theta_k^{(j)}\}$. Referring to the analogous analysis for variationally posed eigenvalue problems, Boffi [4, p. 50] has observed that the

> definition of convergence (7.7) does not give any indication of the approximation rate. It is indeed quite common to separate the convergence analysis for eigenvalue problems into two steps: firstly, the convergence and the absence of spurious modes is investigated in the spirit of (7.7), then suitable approximation rates are proved.

We intend to present an analysis of the rate of convergence in a future paper.

REFERENCES

[1] P. M. ANSELONE, *Collectively Compact Operator Approximation Theory and Applications to Integral Equations*, Prentice-Hall, Englewood Cliffs, NJ, 1971.

[2] I. BABUŠKA AND J. OSBORN, *Eigenvalue problems*, in Handbook of Numerical Analysis. Volume 2: Finite Element Methods: Part 1, P. G. Ciarlet and J. L. Lions, eds., Elsevier, Amsterdam, 1991, pp. 641–787.

[3] Å. BJÖRCK, *Least squares methods*, in Handbook of Numerical Analysis, Volume 1, P. G. Ciarlet and J. L. Lions, eds., North–Holland, Amsterdam, 1990.

[4] D. BOFFI, *Finite element approximation of eigenvalue problems*, Acta Numer., 19 (2010), pp. 1–120.

[5] H. W. ENGL, M. HANKE, AND A. NEUBAUER, *Regularization of Inverse Problems*, Math. Appl. 375, Kluwer Academic, Dordrecht, 1996.

[6] M. S. GOCKENBACH, *Generalizing the GSVD*, SIAM J. Numer. Anal., 54 (2016), pp. 2517–2540, https://doi.org/10.1137/15M1019453.

[7] M. S. GOCKENBACH, *Linear Inverse Problems and Tikhonov Regularization*, MAA Press, Washington, DC, 2016.

[8] T. KATO, *Perturbation theory for nullity, deficiency and other quantities of linear operators*, J. Analyse Math., 6 (1958), pp. 261–322.

[9] J. LOCKER AND P. M. PRENTER, *Regularization with differential operators.* I. *General theory*, J. Math. Anal. Appl., 74 (1980), pp. 503–529.

[10] V. A. MOROZOV, *Methods for Solving Incorrectly Posed Problems*, Springer-Verlag, New York, 1984.