ICES REPORT 13-24

August 2013

Finite Element and NURBS Approximations of Eigenvalue, Boundary-value, and Initial-value Problems

by

Thomas J.R. Hughes, John A. Evans, Alessandro Reali

The Institute for Computational Engineering and Sciences
The University of Texas at Austin
Austin, Texas 78712

Reference: Thomas J.R. Hughes, John A. Evans, Alessandro Reali, Finite Element and NURBS Approximations of Eigenvalue, Boundary-value, and Initial-value Problems, ICES REPORT 13-24, The Institute for Computational Engineering and Sciences, The University of Texas at Austin, August 2013.

Finite Element and NURBS Approximations of Eigenvalue, Boundary-value, and Initial-value Problems

Thomas J.R. Hughes [a], John A. Evans [a,b,*], and Alessandro Reali [c]

[a] *Institute for Computational Engineering and Sciences, The University of Texas at Austin, USA*

[b] *Department of Aerospace Engineering Sciences, University of Colorado Boulder, USA*

[c] *Department of Civil Engineering and Architecture, University of Pavia, and IMATI-CNR, Pavia, Italy*

[*] Corresponding author. *E-mail address:* john.a.evans@colorado.edu

Abstract

We study the spectral approximation properties of finite element and NURBS spaces from a global perspective. We focus on eigenfunction approximations and discover that the L^2-norm errors for finite element eigenfunctions exhibit pronounced "spikes" about the transition points between branches of the eigenvalue spectrum. This pathology is absent in NURBS approximations. By way of the Pythagorean eigenvalue error theorem, we determine that the squares of the energy-norm errors of the eigenfunctions are the sums of the eigenvalue errors and the squares of the L^2-norm eigenfunction errors. The spurious behavior of the higher eigenvalues for standard finite elements is well-known and therefore inherited by the energy-norm errors along with the spikes in the L^2-norm of the eigenfunction errors. The eigenvalue pathology is absent for NURBS. The implications of these results to the corresponding elliptic boundary-value problem and parabolic and hyperbolic initial-value problems are discussed.

Keywords: Spectral approximation properties, global analysis, finite elements, isogeometric analysis, eigenvalues and eigenfunctions, error analysis

Contents

1 Introduction

We study the spectral approximation properties of finite elements and NURBS. Our work builds on earlier investigations, namely, Cottrell et al. [7], Reali [15], Hughes et al. [13], and Evans and Hughes [9]. See also Cottrell et al. [6]. In these works, we focused primarily on the approximations of eigenvalues from a *global* perspective. That is, we investigated a number of modal problems and compared finite elements with NURBS in their ability to approximate eigenvalues for *all* modes. In the case of [13], we also compared discrete wave number approximations for the Helmholtz equation. Our primary analytical tool in these studies was discrete Fourier/von Neumann analysis (see Richtmyer and Morton [16]). Our main findings were that, per degree-of-freedom and for the same polynomial order p, NURBS were much more accurate than finite elements. With the exception of a few "outlier" modes, in many cases the NURBS eigenvalues enjoyed almost spectral accuracy, and the so-called "optical" branches of finite element eigenvalue approximations diverged with p. This last fact was a surprising finding and one that caused considerable interest and some consternation. A question that naturally arises is "What are the consequences of these errors in the spectrum to various boundary- and initial-value problems?" An initial answer to this question, in the context of the Helmholtz equation, was provided in [13], in which it was shown that in certain wave-number intervals propagating waves become evanescent waves, a completely spurious phenomenon.

We approach this question more systematically herein. To that end we review the eigenvalue problem and its approximation by finite elements. We note that it has been well-known in engineering and mathematics circles that the higher modes are poorly approximated (see, e.g., Hughes [12] and Strang and Fix [17]). In fact, it is well-known that they are not approximated at all. This is explicit in the standard finite element error estimates. However, in consideration of various boundary- and initial-value problems, the eigenvalue approximation is not the only issue. The eigenfunction approximation is fundamental as well. The errors in the discrete approximations of linear boundary- and initial-value problems can be completely expressed in terms of the eigenvalue and eigenfunction errors of the corresponding eigenproblem using modal analysis techniques. For this purpose we need to know the errors in the eigenvalues and the eigenfunctions for *all* the modes, again necessitating a global analysis approach, in contrast with the standard asymptotic finite element error estimates which tell us that the low modes are accurate and converge at specific rates. There is one very interesting result for the Galerkin formulation of the eigenvalue problem that is global, that is, it pertains to each and every mode. We refer to it as the "Pythagorean eigenvalue error theorem"; roughly speaking, it says that, for each mode, the error in the eigenvalue and the

square of the error in the eigenfunction (in say the L^2-norm[1]) sum to the square of the error in the energy-norm. This constitutes an *error budget* that will influence the errors in the corresponding boundary- and initial-value problems. We focus in this work on the simplest elliptic operator, the second derivative in one space dimension. We determine the eigenvalue and eigenfunction errors, in the L^2- and energy-norm, and plot the budgets for all modes for C^0-continuous finite elements and C^{p-1}- continuous NURBS for $p = 2, 3$, and 4. There is nothing new to say about the eigenvalue errors. However, there is quite a surprise in the eigenfunction results for finite elements. Specifically, large error "spikes" appear at the transition points between the acoustic (lowest modes) and optical branches. We note that there are $p-1$ transition points in the spectrum for order-p approximations. Furthermore, these spikes are not present in L^2-best approximations of the eigenfunctions. These results are contrasted with those for NURBS in which case there are no error spikes whatsoever and the approximation quality of the eigenfunctions in L^2 is indistinguishable from the L^2-best approximation. Overall, the spectral accuracy differences between finite elements and NURBS is, to say the least, striking.

We then return to the original question of the implications of spectral errors on those of boundary- and initial-value problem errors. To do this we construct expressions for the errors in the boundary- and initial-value problems in terms of the eigenvalues and eigenfunctions of the discrete and continuous (i.e., exact) problems. Specifically, we look at the elliptic boundary-value problem and parabolic and hyperbolic initial-value problems. The different mathematical structures of these cases lead to different conclusions. We argue that the inaccuracy of the higher finite element modes is not a serious issue for the elliptic boundary-value problem; it is of some, but small, concern for the parabolic initial-value problem; and it is a significant concern for the hyperbolic initial-value problem. We illustrate this point with numerical calculations. We further discuss this issue from the standpoint of current industrial practice in solving hyperbolic initial-value problems, specifically structural dynamics problems. In this discipline, it is apparent that the preferred time-integration procedures are designed to compensate for the inaccuracy of the higher finite element modes. This has also been known for some time in the computational structural dynamics community. Nevertheless, the magnitude of the problem, that the current results dramatize for higher-order finite elements, seems to be a new finding. A positive result is that NURBS elements are not subject to the pathologies noted for C^0-continuous finite elements and produce accurate results in the same circumstances.

A summary of the remainder of the paper follows. We begin with preliminaries and notational conventions in Section 2. In Section 3 we present the elliptic eigenvalue problem, standard finite element error estimates, and the aforementioned spectral results. In Section 4 we assess the elliptic boundary-value problem in light of the spectral results, and we do the same for the parabolic and hyperbolic initial-value problems in Sections 5 and 6, respectively. We draw conclusions in Section 7. In Appendix A we perform an analytical investigation to verify some of the numerical results presented in Section 3.

2 Preliminaries

As a prelude to our study, we mention that we have certain specific problem classes in mind. For example, our results will apply to the Laplace operator and to the classical linear elasticity operator,

[1]For purposes of this discussion we are assuming the eigenvalue problem involves a mass matrix and a stiffness matrix. A typical case would be linear elasticity, in which case the mass and stiffness matrices would be associated with the L^2- and H^1-norms. In buckling problems of thin structures the mass matrix is replaced by an initial-stress stiffness matrix, the L^2-norm is replaced by the H^1-norm, and the stiffness matrix is associated with the H^2-norm.

and to the various types of boundary- and initial-value problems that can be formulated in terms of these operators. The abstract variational framework that pertains to these problems includes many other cases of practical interest as well.

Let Ω be a bounded and connected domain in $\mathbb{R}^d$, where $d \in \mathbb{Z}^+$ is the number of space dimensions. We assume Ω has a Lipschitz boundary $\partial\Omega$. Let $H^m(\Omega)$ denote the Sobolev space of functions with square-integrable generalized derivatives of degree r, where $r = 0, 1, \ldots, m \in \mathbb{Z}^+$. Let $\mathcal{V}$ be a closed subspace of $(H^m(\Omega))^n$, where $n \in \mathbb{Z}^+$. In the typical case under consideration, the functions in $\mathcal{V}$ will satisfy some type of homogenous boundary conditions, appropriate to the problem being considered. We will not need to be more specific about this for what we have in mind. Let $(\cdot, \cdot)$ and $a(\cdot, \cdot)$ be symmetric bilinear forms. We assume both are continuous and coercive in the following sense: For all $v, w \in \mathcal{V}$,

$$a(v, w) \leq \|v\|_E \|w\|_E \tag{1}$$

$$\|w\|_E^2 = a(w, w) \tag{2}$$

$$(v, w) \leq \|v\| \|w\| \tag{3}$$

$$\|w\|^2 = (v, w) \tag{4}$$

where $\|\cdot\|_E$ is the energy-norm which is assumed equivalent to the $(H^m(\Omega))^n$-norm on $\mathcal{V}$ and $\|\cdot\|$ is the $(L^2(\Omega))^n = (H^0(\Omega))^n$ norm.

3 Elliptic Eigenvalue Problem

The elliptic eigenvalue problem is fundamental to our investigation. It is of course of engineering interest itself, in that within the context of elasticity it describes important cases of free vibration and linearized buckling. As we shall see later on, the errors in the eigenvalue problem engendered by the finite-dimensional space provide insight into potential errors in the elliptic boundary-value problem and the parabolic and hyperbolic initial-value problems. The elliptic eigenvalue problem is stated as follows: Find eigenvalues $\lambda_l \in \mathbb{R}^+$ and eigenfunctions $u_l \in \mathcal{V}$, for $l = 1, 2, \ldots, \infty$, such that, for all $w \in \mathcal{V}$,

$$\lambda_l(w, u_l) = a(w, u_l) \tag{5}$$

It is well-known that $0 < \lambda_1 \leq \lambda_2 \leq \lambda_3 \leq \ldots$, and that the eigenfunctions are $(L^2(\Omega))^n$-orthonormal, that is, $(u_k, u_l) = \delta_{kl}$ where δ_{kl} is the Kronecker delta, for which $\delta_{kl} = 1$ if $k = l$ and $\delta_{kl} = 0$ otherwise. The normalization of the eigenfunctions is actually arbitrary. We have assumed without loss of generality that $\|u_l\| = 1$, for all $l = 1, 2, \ldots, \infty$. It follows from (5) that

$$\|u_l\|_E^2 = a(u_l, u_l) = \lambda_l \tag{6}$$

and $a(u_k, u_l) = 0$ for $k \neq l$.

3.1 Galerkin Formulation

Now let us introduce a finite-dimensional approximation space $\mathcal{V}^h \subset \mathcal{V}$. We think of $\mathcal{V}^h$ as either a standard finite element space or a space of maximally smooth B-splines or NURBS[2]. There are

[2]Throughout, we usually favor using the terminology NURBS over B-splines for generality because the conventional view is that B-splines just represent the special case of NURBS that are "unweighted" in the sense that the weights are all equal to 1. However, if a study was performed specifically with unweighted NURBS, that is, B-splines, we will so specify. On the other hand, NURBS are simply B-splines in a space of dimension $d + 1$, but this only creates confusion in the specificity of the terminologies.

obviously many other possibilities but we will focus on these herein. The discrete counterpart of (5) is: Find $\lambda_l^h \in \mathbb{R}^+$ and $u_l^h \in \mathcal{V}^h$ such that for all $w^h \in \mathcal{V}^h$,

$$\lambda_l^h(w^h, u_l^h) = a(w^h, u_l^h) \tag{7}$$

The solution of (7) has similar properties to the solution of (5). Specifically, $0 < \lambda_1^h \leq \lambda_2^h \leq \ldots \leq \lambda_N^h$, where N is the dimension of $\mathcal{V}^h$, $(u_k^h, u_l^h) = \delta_{kl}$, $\|u_l^h\|_E^2 = a(u_l^h, u_l^h) = \lambda_l^h$, and $a(u_k^h, u_l^h) = 0$ if $k \neq l$.

The comparison of $\left\{\lambda_l^h, u_l^h\right\}$ to $\{\lambda_l, u_l\}$ for all $l = 1, 2, \ldots, N$ is the key to gaining insight into the errors of the discrete approximations to the elliptic boundary-value problem and the parabolic and hyperbolic initial-value problems. Most of the classical functional analysis results describing error estimates for the eigenvalues and eigenfunctions are *asymptotic*, that is, for each fixed N they are applicable only to the errors of the lowest modes of the finite-dimensional space $\mathcal{V}^h$. In general, a large portion of the eigenvalue/eigenfunction spectrum, the so-called "higher modes," are not approximations of their exact counterparts in any meaningful sense. A nice analytical study of a finite element method utilizing Hermite cubic basis functions is contained in Strang and Fix [17], Chapter 6, pages 226–228, where it is shown that the upper part of the spectrum has *no* approximability whatsoever. It is further stated therein on page 227 that these higher modes can be simply ignored. We take issue with this. In fact, in certain circumstances we are in almost complete disagreement with it. Nevertheless, it is well-known in the structural engineering discipline that the higher modes are grossly inaccurate, but the precise point in the spectrum where the eigenvalues and eigenfunctions cease to approximate their corresponding exact counterparts is never known in realistic engineering situations. Engineers typically make extremely conservative estimates based on experience, often only "accepting" a minute fraction of N, the number of modes represented by the entire discrete system. However, the situation is potentially very different for the discrete approximations of elliptic boundary-value problems and parabolic and hyperbolic problems. In these cases, all modes may participate in the solution to some extent and, as we will see later, inaccurate higher modes may not always be simply ignored.

Asymptotic error estimates may be very optimistic in that they ignore the deleterious effects of the higher modes. The situation is even more disconcerting in nonlinear analyses in that there is modal coupling between inaccurate higher modes and accurate lower modes. In these cases the ostensible accuracy of the lower modes can be vitiated by the large errors in the higher modes. For example, insufficiently refined discretizations of turbulent flows are susceptible to this fate [3].

In what follows we will investigate the errors caused by *all* modes in the discrete approximation. The results will provide insight into the very different behavior of C^0-continuous finite elements in elliptic boundary-value problems compared with parabolic problems compared with hyperbolic problems. Our approach is *not* asymptotic, rather it may be characterized as a *global analysis* approach.

The fundamental global error analysis result for elliptic eigenvalue problems is the *Pythagorean eigenvalue error theorem*. It is simply derived and is done so on page 233 of Strang and Fix [17]. The theorem is global in that it is applicable to each and every mode in the discrete approximation. Provided that $\|u_l^h\| = \|u_l\|$,

$$\frac{\lambda_l^h - \lambda_l}{\lambda_l} + \frac{\|u_l^h - u_l\|^2}{\|u_l\|^2} = \frac{\|u_l^h - u_l\|_E^2}{\|u_l\|_E^2}, \qquad \forall l = 1, 2, \ldots, N \tag{8}$$

Note that the relative error in the l^{th} eigenvalue and the square of the $(L^2(\Omega))^n$-norm error in the l^{th} eigenfunction sum to equal the square of the relative energy-norm error in the l^{th} eigenfunction.

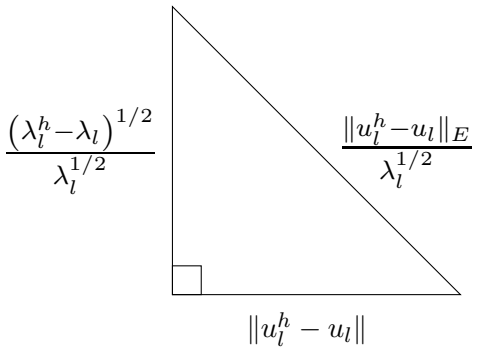


Figure 1: Graphical representation of the Pythagorean eigenvalue error theorem.

Due to the normalizations introduced earlier, (8) can also be written as

$$\frac{\lambda_l^h-\lambda_l}{\lambda_l}+\|u_l^h-u_l\|^2=\frac{\|u_l^h-u_l\|_E^2}{\lambda_l},\qquad \forall l=1,2,\ldots,N \tag{9}$$

See Figure 1. We note that the first term in (9) is always non-negative as $\lambda_l^h \geq \lambda_l$, a consequence of the "minimax" characterization of eigenvalues (see [17], page 223). It also immediately follows from (9) that

$$\lambda_l^h-\lambda_l \leq \|u_l^h-u_l\|_E^2 \tag{10}$$

$$\|u_l^h-u_l\|^2 \leq \frac{\|u_l^h-u_l\|_E^2}{\lambda_l} \tag{11}$$

3.2 Summary of Standard Finite Element Estimates for the Eigenvalue Problem

For a given mode, finite element error estimates can be derived provided the mesh size h satisfies a smallness condition of the form $h\lambda_l^{1/(2m)} \leq \epsilon$ where $2m$ is the order of the differential operator and $\epsilon < 1$ is a constant independent of h and λ_l. The main results, from Strang and Fix [17], pages 228–236, are

$$\frac{\lambda_l^h-\lambda_l}{\lambda_l} \leq c\left(h\lambda_l^{1/(2m)}\right)^{2(p+1-m)} \tag{12}$$

$$\frac{\|u_l^h-u_l\|_E}{\lambda_l^{1/2}} \leq c\left(h\lambda_l^{1/(2m)}\right)^{p+1-m} \tag{13}$$

$$\|u_l^h-u_l\| \leq c\left(h\lambda_l^{1/(2m)}\right)^{\sigma} \tag{14}$$

where

$$\sigma=\min\left\{p+1, 2\left(p+1-m\right)\right\} \tag{15}$$

in which p is the degree of complete polynomial exactly representable by the finite element space $\mathcal{V}^h$ and c is a constant independent of h and λ_l. Note that the dimensionless quantity $h\lambda_l^{1/(2m)}$ is a measure of how well the discrete finite element space $\mathcal{V}^h$ resolves the continuous eigenfunction u_l. Consequently, as the continuous eigenvalue λ_l increases in size, a more refined mesh must be utilized in order to obtain discrete eigenvalues and eigenfunctions satisfying a prescribed accuracy.

Moreover, it should be noted that the estimates given by (12)-(14) only hold for the lowest modes of $\mathcal{V}^h$ due to the smallness condition $h\lambda_l^{1/(2m)} \leq \epsilon$. This smallness condition also suggests that the lowest modes converge exponentially as the polynomial order p is increased, but this exponential convergence becomes less pronounced as $h\lambda_l^{1/(2m)} \to \epsilon$.

Analogous estimates for B-splines and NURBS, to the best of our knowledge, have not yet been obtained, but we conjecture that similar results can be obtained. Indeed, the estimates given by (12)-(14) are largely a consequence of best approximation error. As such, we expect the lowest modes for NURBS to be more accurate than the lowest modes for finite elements. This expectation is based on prior work which compared the best approximation properties of finite elements and NURBS using the framework of Kolmogorov n-widths [8]. In Subsection 3.3, we will zoom in on the lowest modes for both finite element and NURBS spaces to examine their corresponding accuracy.

Unfortunately, the estimates given by (12)-(14) do not apply to the highest modes of the discrete space $\mathcal{V}^h$. Instead, other approaches such as discrete Fourier/von Neumann analysis (see Richtmyer and Morton [16]) must be employed to analyze these often troublesome modes. This was precisely the approach taken in investigations by Cottrell et al. [7], Reali [15], and Hughes et al. [13] where the approximation of eigenvalues was conducted from a global perspective. The main conclusions from these works were that (i) the "eigenfrequencies" (i.e., square roots of the eigenvalues) were significantly more accurate for NURBS than for finite elements for a fixed number of degrees-of-freedom; (ii) finite element spectra are characterized by spurious "optical branches" while NURBS spectra are not; (iii) for a fixed number of degrees-of-freedom, the "optical branches" of the finite element spectra diverge with increasing polynomial order p; and (iv) for a fixed number of degrees-of-freedom, almost the entire NURBS spectrum converges with increasing polynomial order, achieving almost spectral accuracy. The one caveat to (iv) is so-called "outlier" frequencies, but their number is very small and independent of N. The divergent optical branches of finite element spectra correspond to a fixed but rather large fraction of N. These are the modes that offer no approximability of their exact counterparts. For example, for the standard, second-derivative operator in one-dimension, one-half the modes, that is, $N/2$ modes, occupy the optical branch for $p = 2$. That the behavior of the optical branches is poor for finite elements has been known since the earliest numerical studies of finite element methods (see, e.g., Hughes [12]). However, that the situation actually worsened with p was a surprise. This created some consternation within the finite element research community. One response was that since the lowest modes were, at the same time, becoming more accurate with p, the higher modes could be ignored. Concerning the eigenproblem per se, we view this response as reasonable and consistent with our previous discussion. Another response was that the inaccuracy of the higher modes was irrelevant for other problems of interest. We accept this response if we consider only the elliptic boundary-value problem. In this case, the best approximation property in the energy-norm is the trump card. The participation of the inaccurate higher modes in the solution is suppressed thereby. However, we feel this response is only partially appropriate for the parabolic initial-value problem and false for the hyperbolic initial-value problem, as well as for the corresponding time harmonic case (i.e., the generalized Helmholtz equation).

We previously analyzed the classical Helmholtz equation case in Hughes et al. [13]. Our results indicated that within certain bands of wave numbers, the finite element solutions would be completely spurious in that propagating waves would be transmogrified into evanescent waves. This was investigated thoroughly by Thompson and Pinsky in [18] and verified numerically for cubic finite elements in [13]. This pathology does not occur for NURBS.

One limitation of our previous work was that we focused our attention exclusively on eigenfrequency errors. As we shall see shortly, it is also important to investigate eigenfunction errors. In

fact, another surprise is in the offing.

3.3 Numerical Investigation of Spectral Approximations

We consider the elliptic eigenvalue problem for the second-order differential operator in one-dimension with homogeneous Dirichlet boundary conditions. The variational form of the problem is given by (5), in which

$$a(w, u_l) = \int_0^1 \frac{dw}{dx}\frac{du_l}{dx}dx \tag{16}$$

$$(w, u_l) = \int_0^1 w u_l dx \tag{17}$$

The eigenvalues are $\lambda_l = \pi^2 l^2$ and the eigenfunctions are $u_l = \sqrt{2}\sin(l\pi x)$, $l = 1, 2, \ldots, \infty$. This may be interpreted physically as the eigenvalue problem associated with a linear elastic rod undergoing axial deformations, or a string under constant tension undergoing transverse deformation, etc. In contrast with our previous studies, we will present the eigenvalue errors, rather than the eigenfrequency errors, and, in addition, $L^2(0,1)$- and energy-norm eigenfunction errors. We will plot the various errors in a format that represents the Pythagorean eigenvalue error theorem budget. We will restrict our study to quadratic, cubic, and quartic finite elements and B-splines. In all cases, we assume linear geometric parametrizations and uniform meshes. In the case of B-splines this means the Bézier elements all have the same size, namely, $h = 1/(N - (p-2))$. See Cottrell et al. [7]. Our results constitute portraits of all the errors for discrete spaces of any dimension N. Strictly speaking, for B-splines the results are only true for sufficiently large N, due to the use of open knot vectors, but in this case "sufficiently large" is not very large at all, say $N > 30$. For smaller spaces, the results change slightly. The results that we present here were computed using $N \approx 1,000$ and validated using a mesh convergence study and by comparing our results to analytical computations (see Appendix A). Moreover, we confirmed that the Pythagorean eigenvalue error theorem was preserved to high precision in all of our numerical experiments.

Let us begin with results for C^1-continuous quadratic B-splines, presented in Figure 2(a). The results for the relative eigenvalue errors (red curve) follow the usual pattern that has been seen before. The squares of the eigenfunction errors in $L^2(0,1)$ are also well-behaved (blue curve) with virtually no discernible error until about $l/N = 0.6$, and then monotonically increasing errors in the highest modes. The sums of the errors produce the squares of the relative energy-norm errors (black curve), as per the Pythagorean eigenvalue error theorem budget. There are no surprises here.

Next we compare with C^0-continuous quadratic finite elements in Figure 2(b). The pattern of eigenvalue errors (red curve), consisting of two branches, the acoustic branch for $l/N < 1/2$, and the optical branch for $l/N \geq 1/2$, is well-known for this case. However, the eigenfunction error in $L^2(0,1)$ (blue curve) represents a surprise in that there is a large spike about $l/N = 1/2$, the transition point between the acoustic and optical branches. Again, the square of the energy-norm eigenfunction error term (black curve) is the sum, as per the budget. This is obviously not a happy result. It suggests that if modes in the neighborhood of $l/N = 1/2$ are participating in the solution of a boundary-value or initial-value problem, the results will be in significant error. The two unpleasant features of this result are (i) the large magnitude of the eigenfunction errors about $l/N = 1/2$ and (ii) the fact that they occur at a relatively low mode number. That the highest modes are significantly in error is well-established for C^0-continuous finite elements, but that there are potential danger zones much earlier in the spectrum has not been recognized previously. The

midpoint of the spectrum in one-dimension corresponds to the quarter point in two dimensions and the eighth point in three dimensions, and so one must be aware of the fact that the onset of inaccurate modes occurs much earlier in higher dimensions.

The spikes in the eigenfunction error spectrum for C^0-finite elements raise the question as to whether or not the eigenfunctions are representative of the best approximation to eigenfunctions in the vicinity of $l/N = 1/2$. To answer this question, we computed the $L^2(0,1)$ best approximations of some of the exact eigenfunctions and plotted them in Figure 8(b). (They are indicated by $\times$.) The case for C^1-continuous quadratic B-splines is presented in Figure 8(a) for comparison. For this case there are almost no differences between the best approximation of the exact eigenfunctions and the computed eigenfunctions. However, for the C^0-continuous quadratic finite elements, the differences between the computed eigenfunctions and the $L^2(0,1)$ best approximations of the exact eigenfunctions are significant, as can be seen in Figure 8(b). The spike is nowhere to be seen in the best approximation results. We conclude that the Galerkin formulation of the eigenvalue problem is simply not producing good approximations to the exact eigenfunctions about $l/N = 1/2$ in the finite element case.[3]

We observe from Figures 2(a) and (b) that the initial portion of the spectrum is well-resolved for both C^1-continuous quadratic B-splines and C^0-continuous quadratic finite elements. This is expected due to the error estimates that were discussed in Subsection 3.2. In order to better compare the initial portion of the error spectrum for B-splines and finite elements, detailed views of the first 10% of the modes (i.e., $l/N \leq 1/10$) are presented in Figures 5(a) and (b). Note that while all of the errors are of small magnitude in this portion of the spectrum, the eigenvalue and eigenfunction errors are much smaller for C^1-continuous quadratic B-splines than they are for C^0-continuous quadratic finite elements. Indeed, the eigenvalue error at $l/N \approx 1/10$ (and the square of the energy-norm eigenfunction error) is more than fifteen times smaller for C^1-continuous quadratic B-splines than for C^0-continuous quadratic finite elements. This is consistent with available analytical results (see Table 1, Appendix A). Also note that in this portion of the spectrum, the eigenvalue errors dominate $L^2(0,1)$ eigenfunction errors, and thus the energy-norm errors are almost entirely attributable to the eigenvalue errors.

The comparison of C^2-continuous cubic B-splines and C^0-continuous cubic finite elements is presented in Figures 3(a) and (b), respectively. See Table 1, Appendix A, for the analytical comparison. As we would anticipate, the C^2-continuous B-spline results are uniformly more accurate than the C^1-continuous quadratic B-spline results (cf. Figure 2(a)). The eigenvalue error results for the C^0-continuous cubic finite elements are divided among three branches (see Figure 3(b)). The separation points are at $l/N = 1/3$ and $2/3$. (This follows from the general fact that there are p equal-length branches for p^{th}-degree C^0-continuous finite elements.) The eigenvalue and eigenfunction errors in this case for the acoustic branch, that is the low modes $l/N < 1/3$, are smaller than the corresponding ones for the C^0-continuous quadratic finite elements (cf. Figure 2(a)). However, for the branch beyond $l/N = 2/3$ they are worse. The results for the region between, $1/3 \leq l/N \leq 2/3$, are a bit mixed due to the fact that the locations of branch separation are not coincident. (This becomes clearer when we consider the case $p = 4$ and compare it with $p = 2$, for which we have coincident separation points.) Note that the spikes in the eigenfunction errors are again located at the separation points. We have plotted the $L^2(0,1)$ best approximations of some of the exact eigenfunctions in Figures 9(a) and (b) and, as in the quadratic case, the spikes are nowhere to be seen in the best approximation results. In Figures 6(a) and (b), detailed views of the first 10% of

[3]In contrast with the elliptic boundary-value problem, there is no guarantee that the Galerkin formulation of the eigenvalue problem will produce best approximations of the eigenfunctions in any norm. This is apparent here for the $L^2(0,1)$- and energy-norm for C^0-continuous finite elements. Curiously, the results for B-splines seem to be at least very close to best approximations in the $L^2(0,1)$-norm.

the modes in the error spectrum are presented, and like before, the eigenvalue and eigenfunction errors are far smaller for C^2-continuous cubic B-splines than they are for C^0-continuous cubic finite elements.

The comparison of C^3-continuous quartic B-splines and C^0-continuous quartic finite elements is presented in Figures 4(a) and (b), respectively. Again, the improvement of all the results for the quartic B-splines is uniform (see Figure 4(a)). There are no discernible errors of any kind up to about $l/N = 0.7$. For the C^0-continuous quartic finite elements there are four branches (see Figure 4(b)). The first, $l/N < 1/4$, is extremely accurate with respect to all quantities. There is also no discernible spike in eigenfunction errors at $l/N = 1/4$, but there are large spikes about $l/N = 1/2$ and $3/4$. As in the previous cases, the spikes are absent in the $L^2(0,1)$-best approximations of the functions, as displayed in Figures 10(a) and (b). The uppermost optical branch errors are growing compared with the lower-order finite element cases (cf. Figures 2(b) and 3(b)). In Figures 7(a) and (b), detailed views of the first 10% of the modes in the error spectrum are provided, and even though C^0-continuous quartic finite elements are extremely accurate in the first portion of the spectrum, they are markedly less accurate than C^3-continuous B-splines. At $l/N \approx 1/10$, the eigenvalue error for C^0-continuous quartic finite elements is nearly three thousand times as large as the eigenvalue error for C^3-continuous B-splines, as confirmed analytically (see Table 1, Section A), and the energy-norm eigenfunction error is more than fifty times as large. Consequently, even for problems in which the higher modes are not participating, C^3-continuous quartic B-splines hold a distinct advantage over C^0-continuous quartic finite elements in terms of approximability.

One question that remains is what is the behavior of discrete eigenfunctions near the transition points between the branches of the error spectrum. To answer this question, we have visualized discrete eigenfunctions for C^3-continuous quartic B-splines and C^0-continuous finite elements near the transition points $l/N = 1/4, 1/2$, and $3/4$ (for $N = 99$) in Figures 12 and 13 and the corresponding exact eigenfunctions in Figure 11. We immediately observe that all of the discrete eigenfunctions for C^3-continuous quartic B-splines are qualitatively and quantitatively accurate, though less so for the discrete eigenfunctions near the last transition point $l/N = 3/4$. This is consistent with our prior observations. On the other hand, while the discrete eigenfunctions for C^0-continuous quartic B-splines are accurate near the first transition point $l/N = 1/4$, they are exceptionally inaccurate near the transition points $l/N = 1/2$ and $3/4$, both qualitatively and quantitatively. In fact, the discrete eigenfunctions near these last two points appear to be the superposition of several waves with different frequencies and wavelengths as opposed to just a single wave. We have observed similar phenomena for C^0-continuous finite elements of different degree $p \geq 2$.

Engineering folklore regards the accuracy of the eigenvalues as a surrogate for the accuracy of the eigenfunctions. The justification for this is the Rayleigh quotient,

$$\lambda_l^h = \frac{a(u_l^h, u_l^h)}{(u_l^h, u_l^h)} \tag{18}$$

in which the eigenvalue is determined by the corresponding eigenfunction. Clearly, if the eigenfunction is accurate, so will be the eigenvalue. However, it is not clear at all that an accurate eigenvalue implies an accurate eigenfunction, and our spectral results demonstrate that such an implication is incorrect. For example, in Figure 4(a), as we approach $l/N = 1$, the eigenfunction errors are increasing when the eigenvalue errors are decreasing. Likewise, in Figure 4(b), in the acoustic branch as we approach $l/N = 1/2$, the eigenvalue error is small and diminishing at the same time as the eigenfunction error is increasing rapidly and becoming quite large.

As has been mentioned in previous works [7, 13], there are a collection of so-called "outlier" modes in the NURBS spectrum that appear to offer no approximability when a uniform mesh is

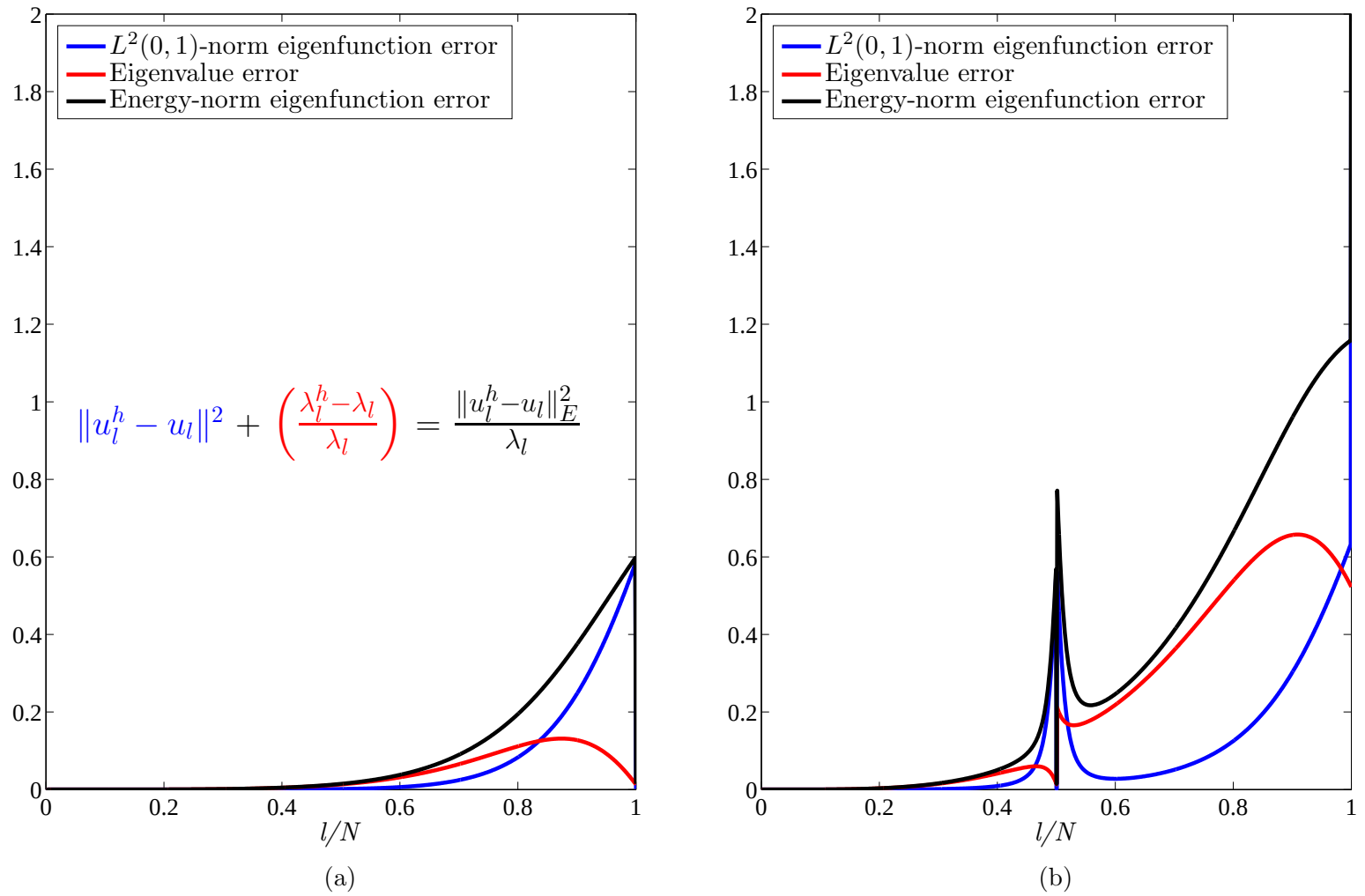


Figure 2: Pythagorean eigenvalue error theorem budget for quadratic elements. (a) C^1-continuous B-splines; (b) C^0-continuous finite elements. The blue curves are $\|u_l^h - u_l\|^2$, the red curves are $(\lambda_l^h - \lambda_l)/\lambda_l$, and the black curves are $\|u_l^h - u_l\|_E^2/\lambda_l$. Note that $\|u_l\| = \|u_l^h\| = 1$, $\|u_l\|_E^2 = \lambda_l$, and $\|u_l^h\|_E^2 = \lambda_l^h$.

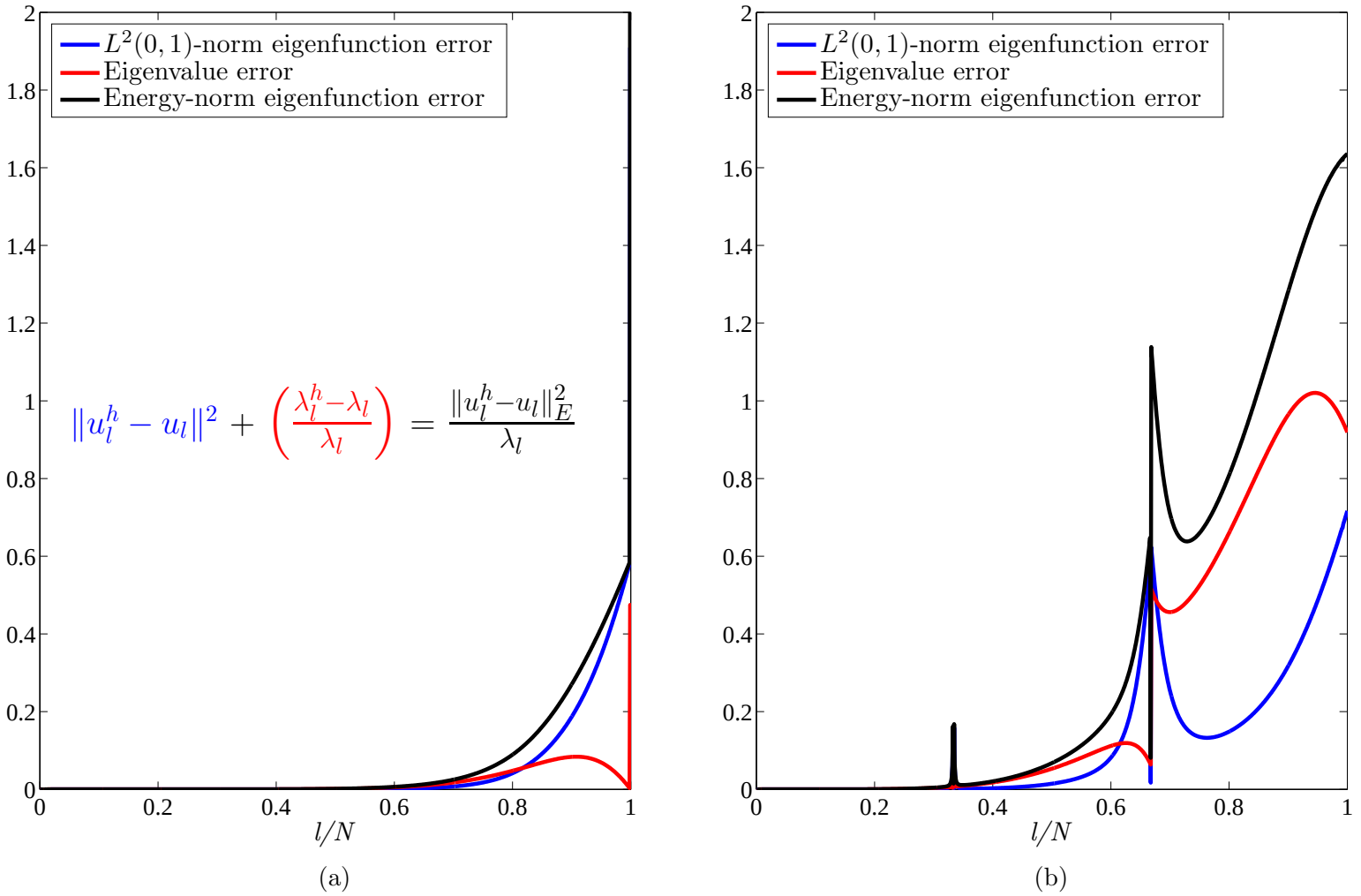


Figure 3: Pythagorean eigenvalue error theorem budget for cubic elements. (a) C^2-continuous B-splines; (b) C^0-continuous finite elements. The blue curves are $\|u_l^h - u_l\|^2$, the red curves are $(\lambda_l^h - \lambda_l)/\lambda_l$, and the black curves are $\|u_l^h - u_l\|_E^2/\lambda_l$. Note that $\|u_l\| = \|u_l^h\| = 1$, $\|u_l\|_E^2 = \lambda_l$, and $\|u_l^h\|_E^2 = \lambda_l^h$.

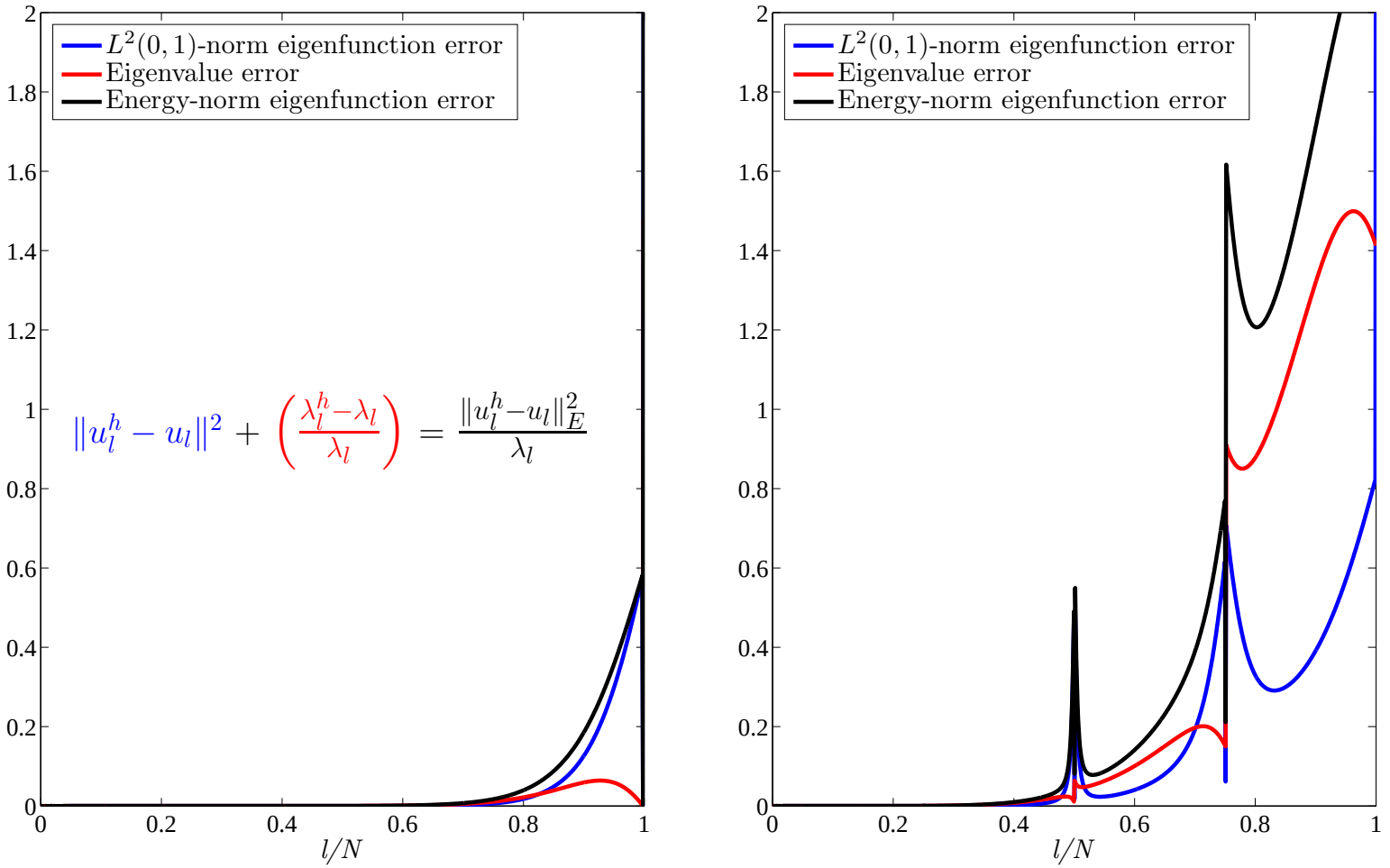


Figure 4: Pythagorean eigenvalue error theorem budget for quartic elements. (a) C^3-continuous B-splines; (b) C^0-continuous finite elements. The blue curves are $\|u_l^h - u_l\|^2$, the red curves are $(\lambda_l^h - \lambda_l)/\lambda_l$, and the black curves are $\|u_l^h - u_l\|_E^2/\lambda_l$. Note that $\|u_l\| = \|u_l^h\| = 1$, $\|u_l\|_E^2 = \lambda_l$, and $\|u_l^h\|_E^2 = \lambda_l^h$.

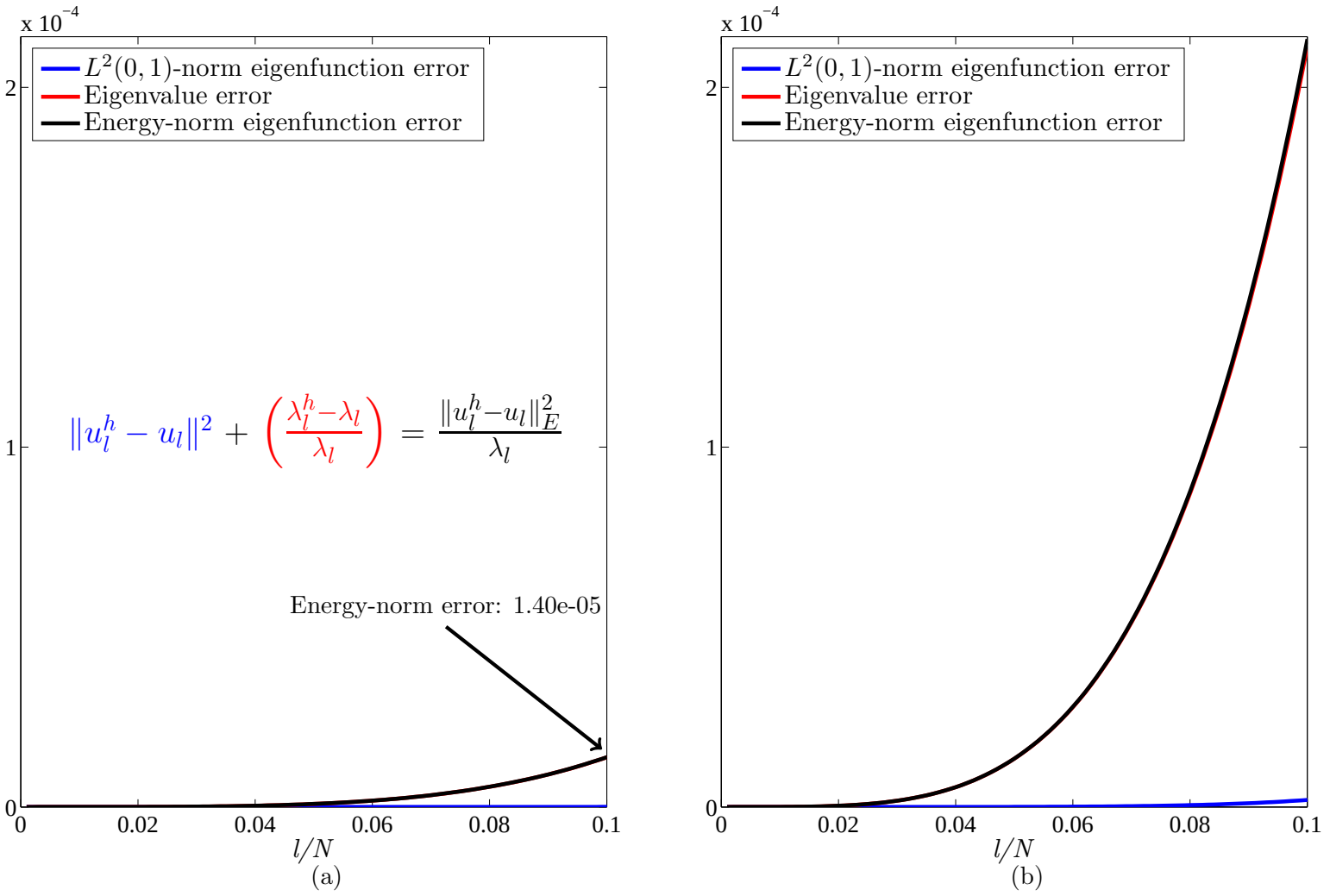


Figure 5: Pythagorean eigenvalue error theorem budget for quadratic elements and $l/N \leq 1/10$. (a) C^1-continuous B-splines; (b) C^0-continuous finite elements. The blue curves are $\|u_l^h - u_l\|^2$, the red curves are $(\lambda_l^h - \lambda_l)/\lambda_l$, and the black curves are $\|u_l^h - u_l\|_E^2/\lambda_l$. Note that $\|u_l\| = \|u_l^h\| = 1$, $\|u_l\|_E^2 = \lambda_l$, and $\|u_l^h\|_E^2 = \lambda_l^h$.

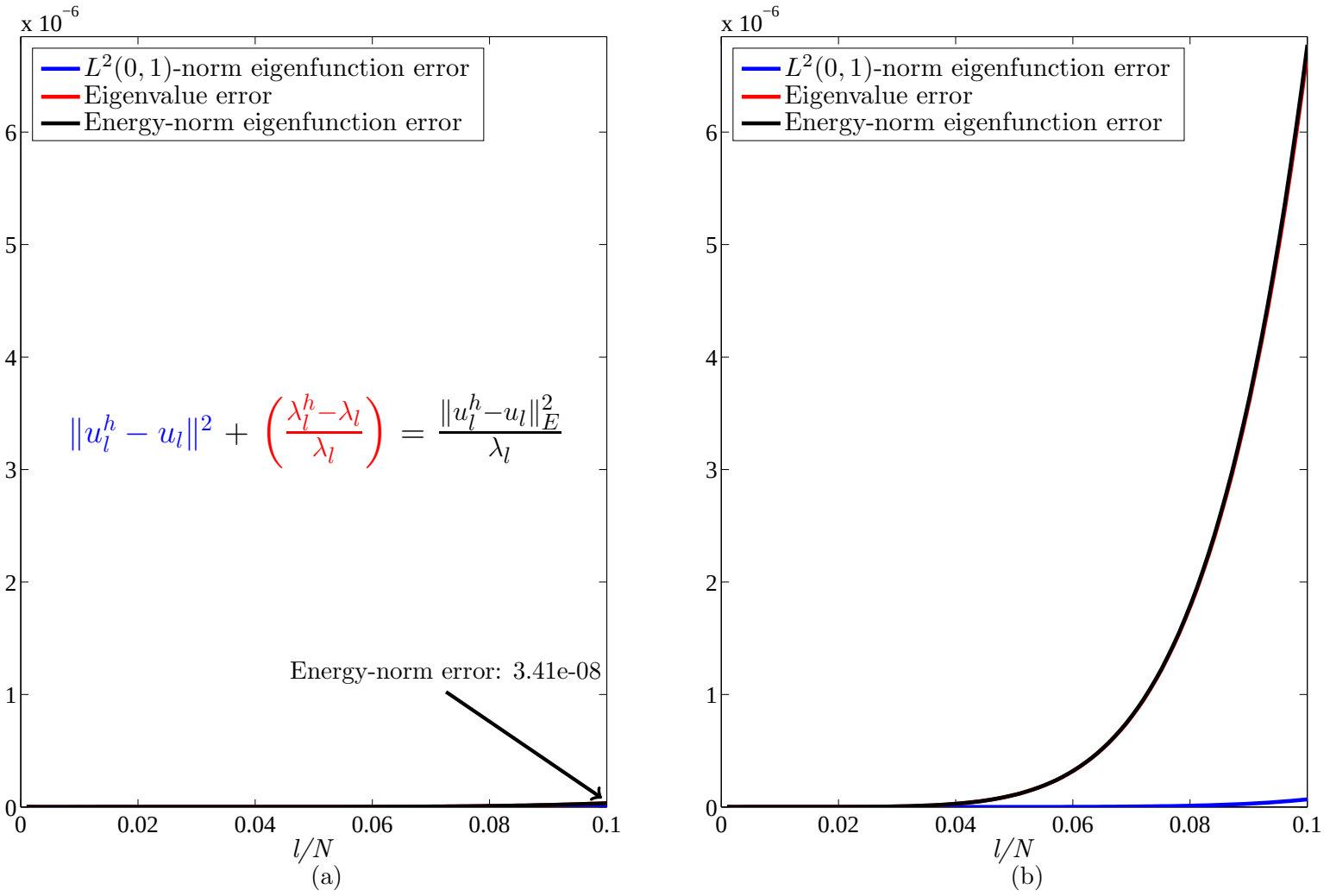


Figure 6: Pythagorean eigenvalue error theorem budget for cubic elements and $l/N \leq 1/10$. (a) C^2-continuous B-splines; (b) C^0-continuous finite elements. The blue curves are $\|u_l^h - u_l\|^2$, the red curves are $(\lambda_l^h - \lambda_l)/\lambda_l$, and the black curves are $\|u_l^h - u_l\|_E^2/\lambda_l$. Note that $\|u_l\| = \|u_l^h\| = 1$, $\|u_l\|_E^2 = \lambda_l$, and $\|u_l^h\|_E^2 = \lambda_l^h$.

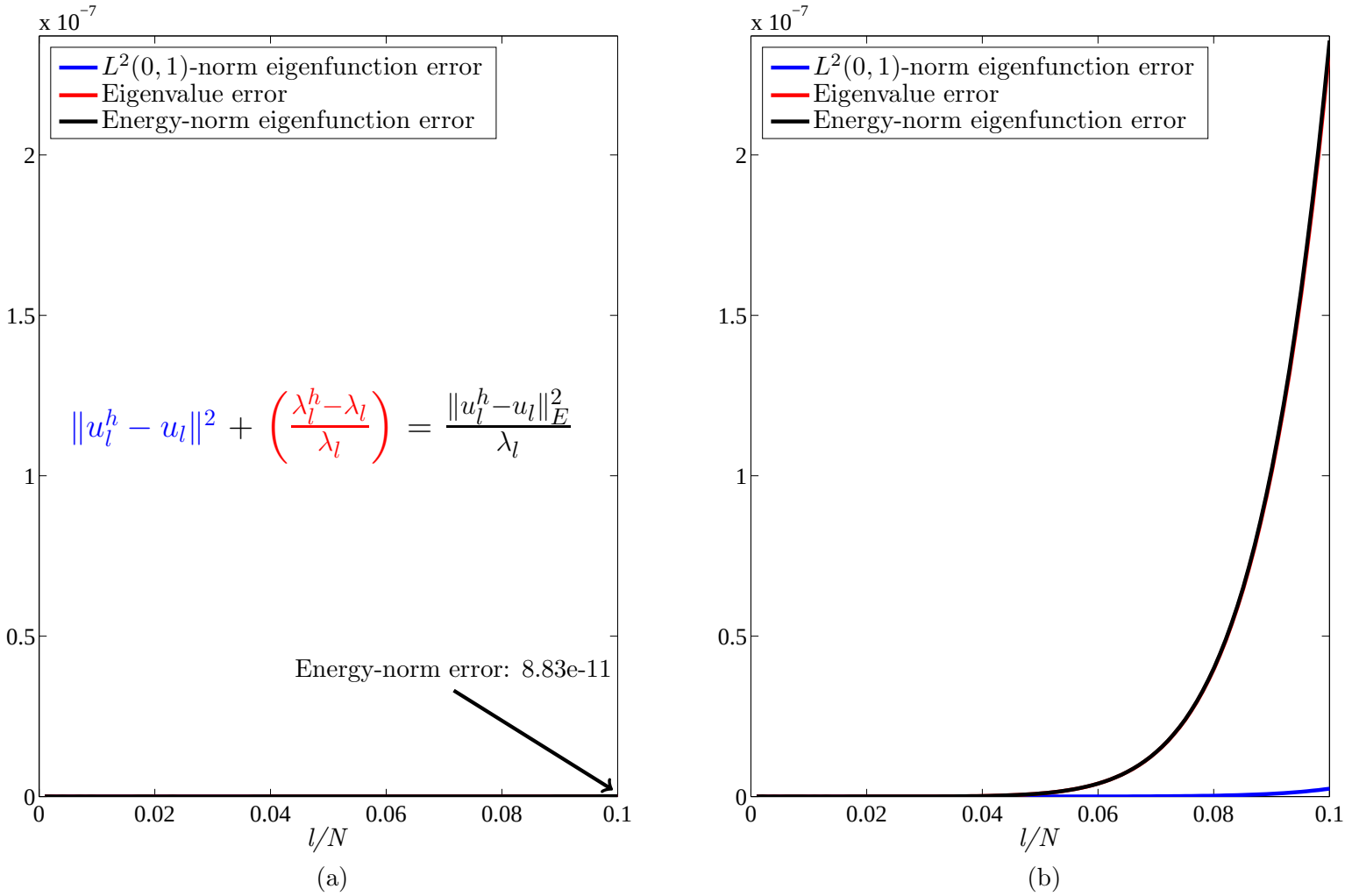
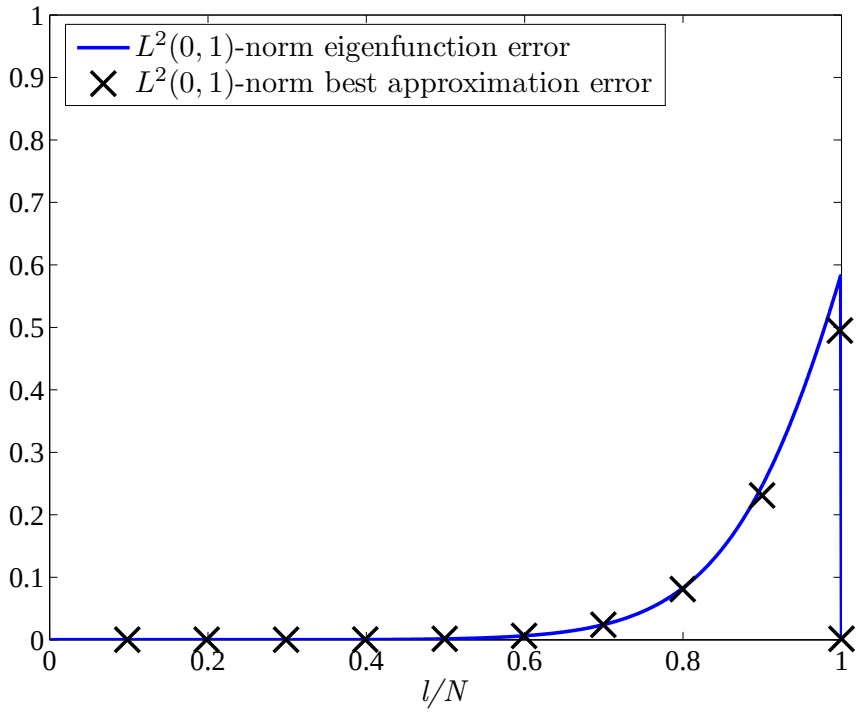
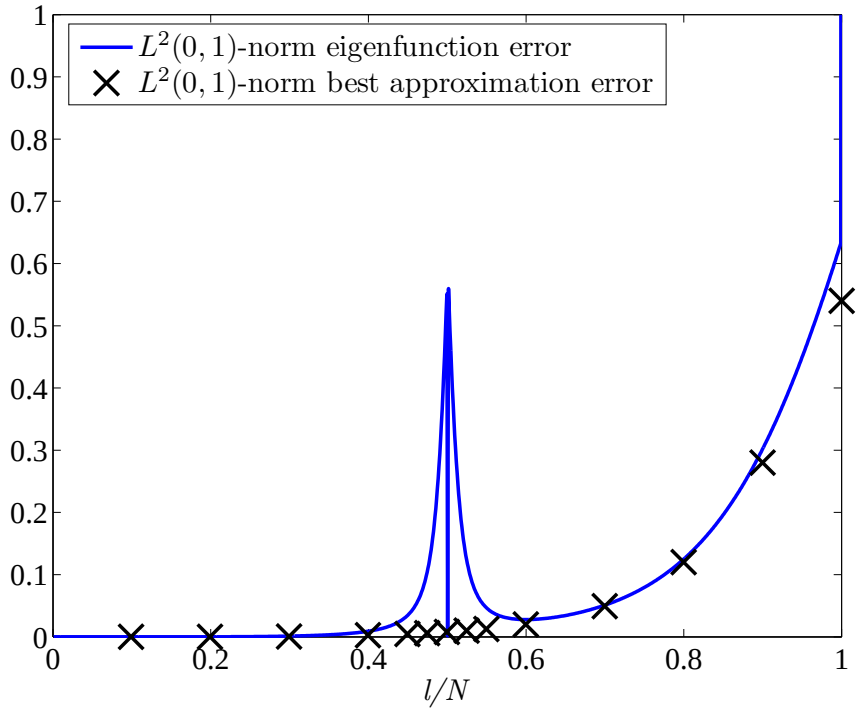


Figure 7: Pythagorean eigenvalue error theorem budget for quartic elements and $l/N \leq 1/10$. (a) C^3-continuous B-splines; (b) C^0-continuous finite elements. The blue curves are $\|u_l^h - u_l\|^2$, the red curves are $(\lambda_l^h - \lambda_l)/\lambda_l$, and the black curves are $\|u_l^h - u_l\|_E^2/\lambda_l$. Note that $\|u_l\| = \|u_l^h\| = 1$, $\|u_l\|_E^2 = \lambda_l$, and $\|u_l^h\|_E^2 = \lambda_l^h$.

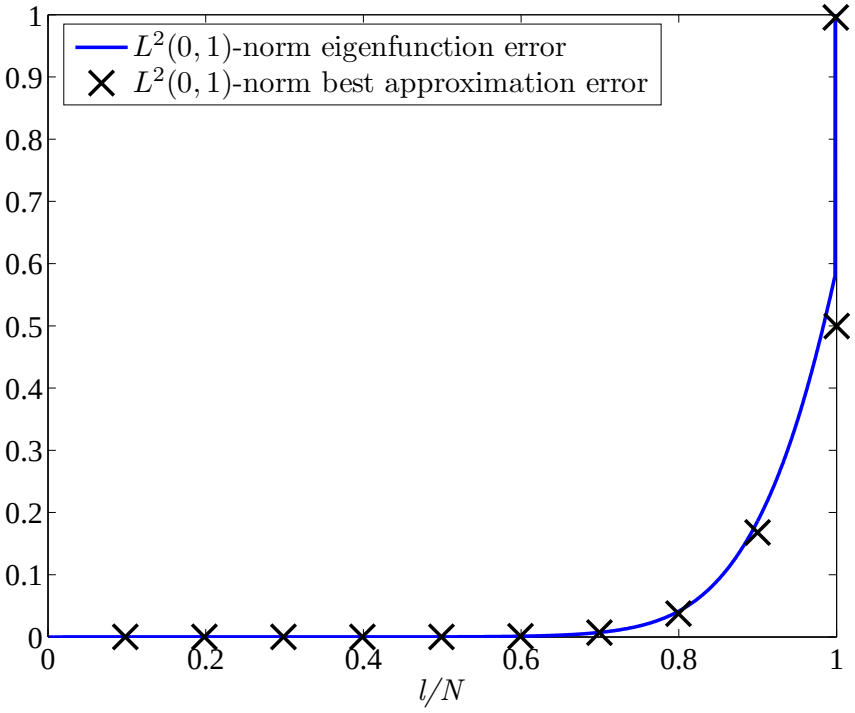


(a)

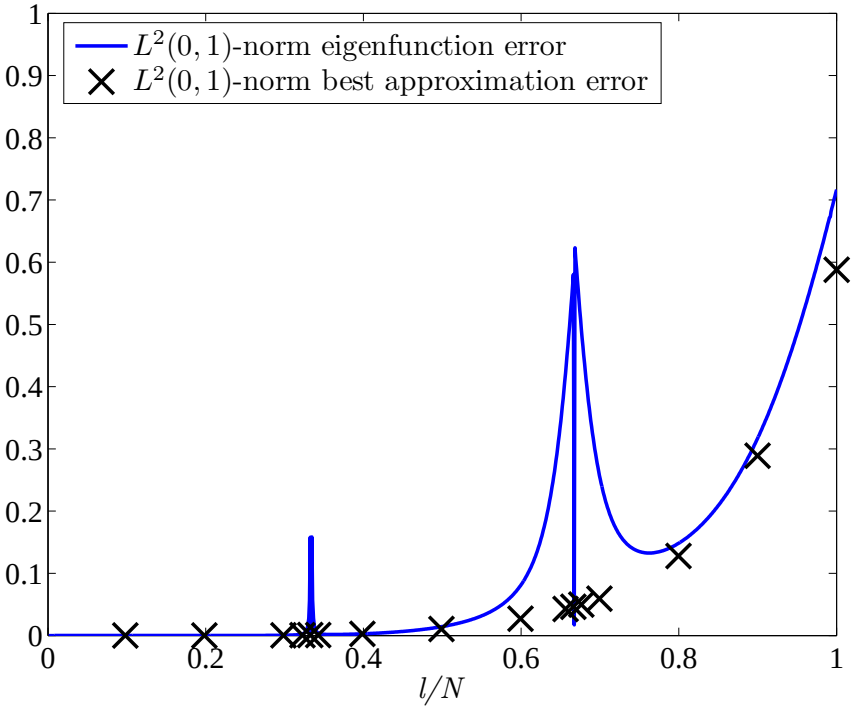


(b)

Figure 8: Comparisons of eigenfunctions computed by the Galerkin method with $L^2(0,1)$ best approximations of the exact eigenfunctions. (a) C^1-continuous quadratic B-splines; (b) C^0-continuous quadratic finite elements. The blue curves are $\|u_l^h - u_l\|^2$, where u_l^h is the Galerkin approximation of u_l, and the $\times$'s are $\|\tilde{u}_l^h - u_l\|^2$, where $\tilde{u}_l^h$ is the $L^2(0,1)$ best approximation of u_l.

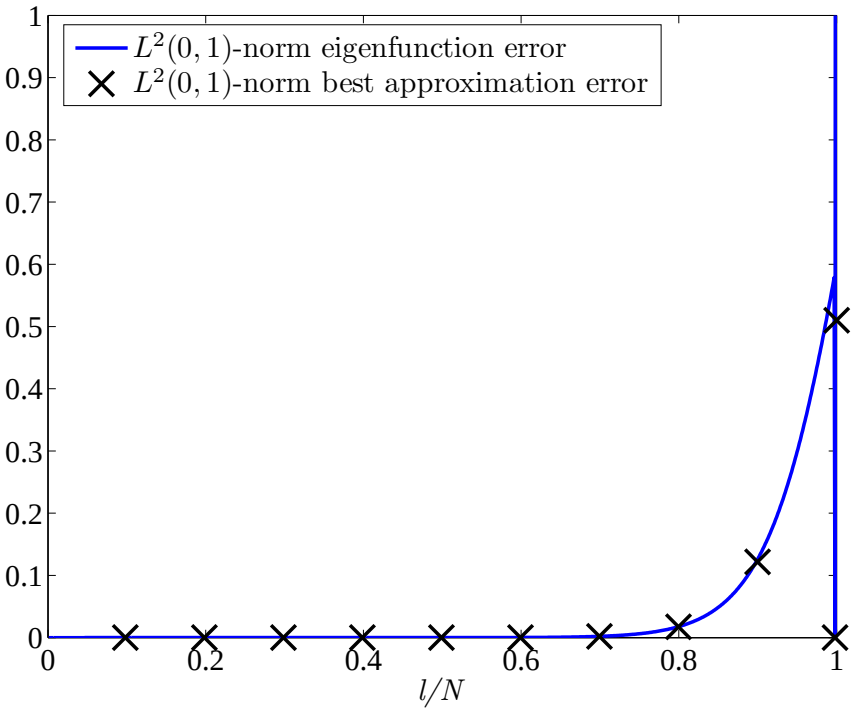


(a)

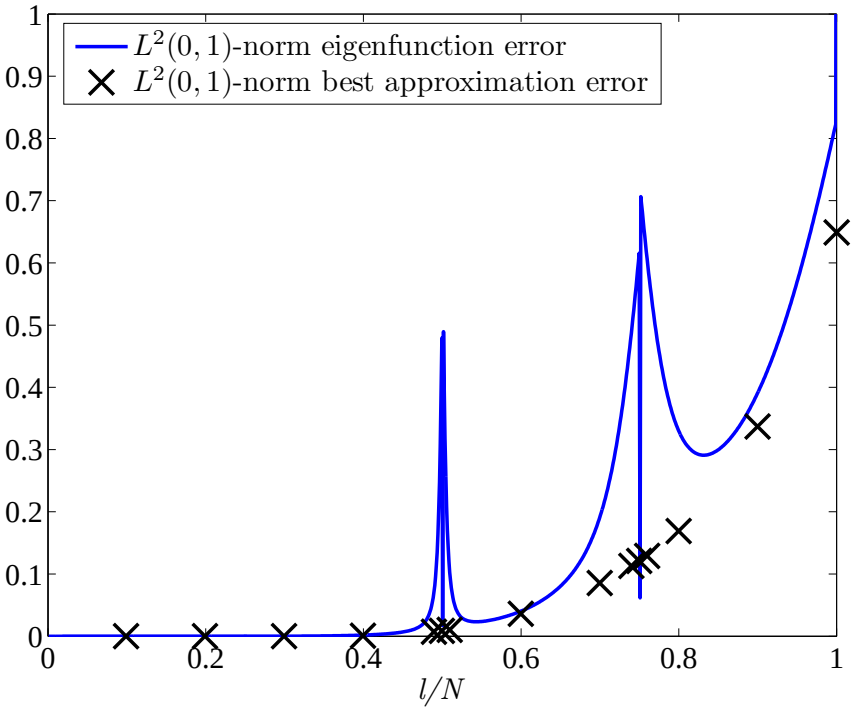


(b)

Figure 9: Comparisons of eigenfunctions computed by the Galerkin method with $L^2(0,1)$ best approximations of the exact eigenfunctions. (a) C^2-continuous cubic B-splines; (b) C^0-continuous cubic finite elements. The blue curves are $\|u_l^h - u_l\|^2$, where u_l^h is the Galerkin approximation of u_l, and the $\times$'s are $\|\tilde{u}_l^h - u_l\|^2$, where $\tilde{u}_l^h$ is the $L^2(0,1)$ best approximation of u_l.



(a)



(b)

Figure 10: Comparisons of eigenfunctions computed by the Galerkin method with $L^2(0,1)$ best approximations of the exact eigenfunctions. (a) C^3-continuous quartic B-splines; (b) C^0-continuous quartic finite elements. The blue curves are $\|u_l^h - u_l\|^2$, where u_l^h is the Galerkin approximation of u_l, and the $\times$'s are $\|\tilde{u}_l^h - u_l\|^2$, where $\tilde{u}_l^h$ is the $L^2(0,1)$ best approximation of u_l.

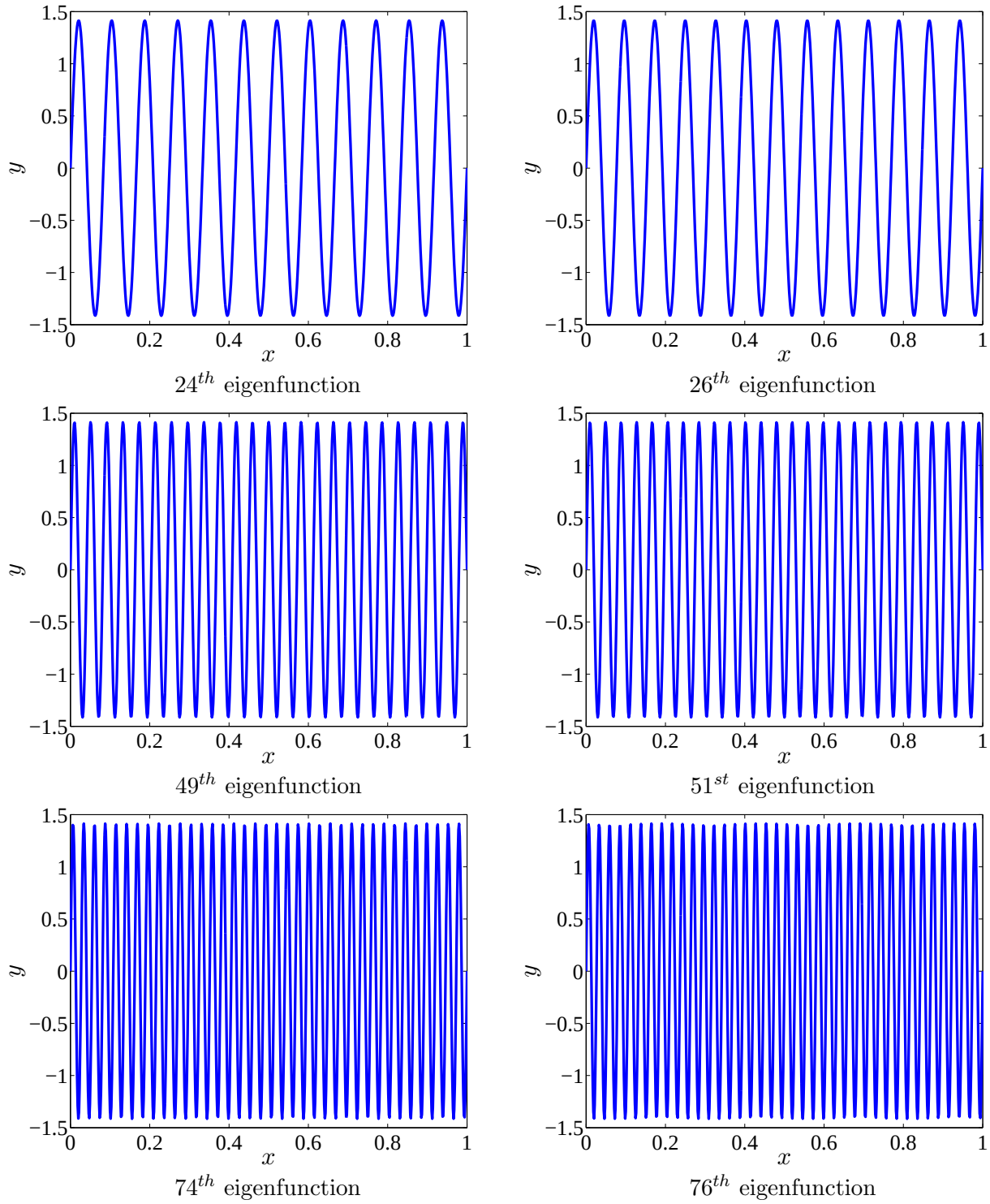


Figure 11: Visualization of the exact 24^{th}, 26^{th}, 49^{th}, 51^{st}, 74^{th}, and 76^{th} eigenfunctions.

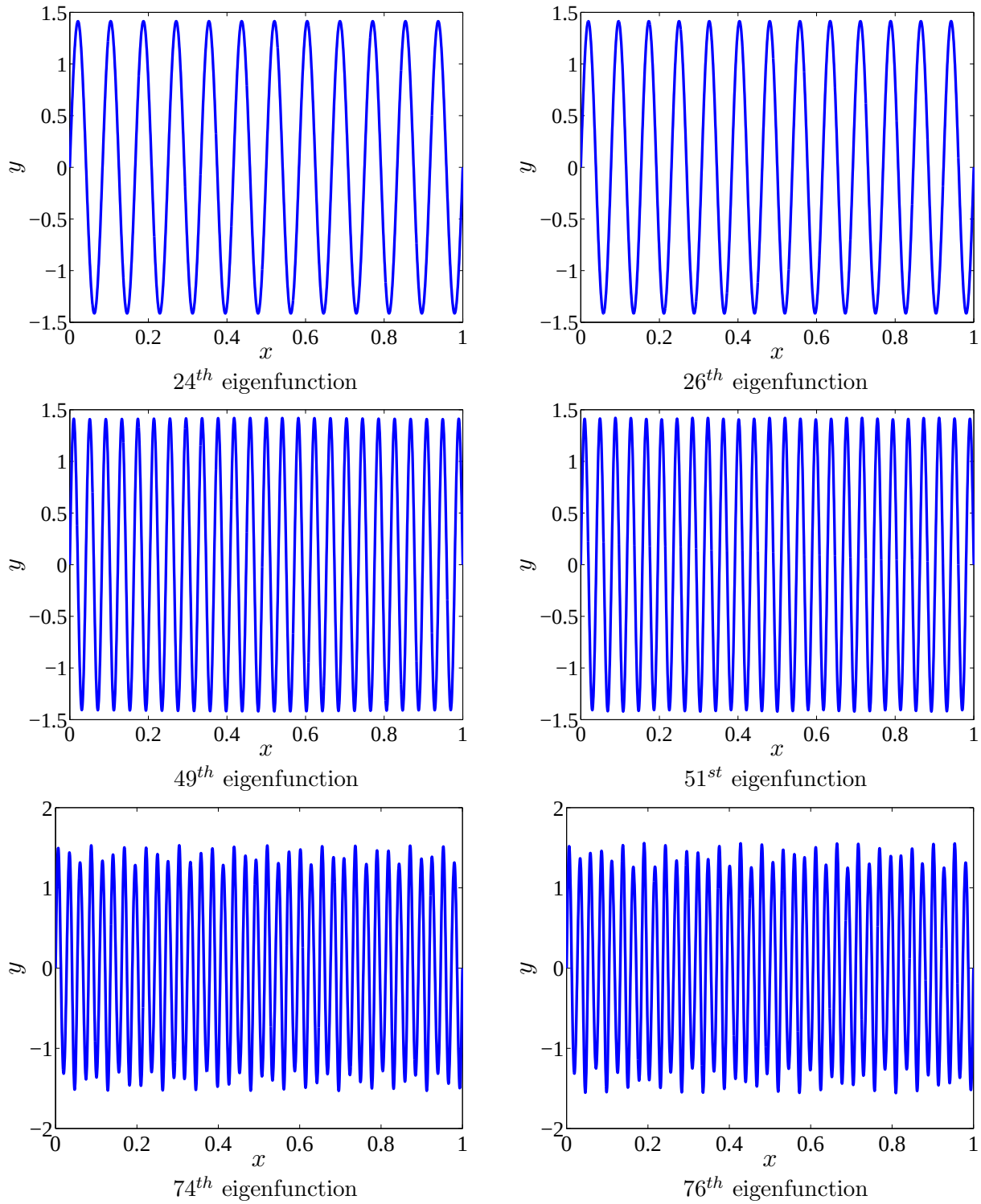


Figure 12: Visualization of the discrete 24^{th}, 26^{th}, 49^{th}, 51^{st}, 74^{th}, and 76^{th} eigenfunctions for a C^3-continuous quartic B-spline approximation where the total number of discrete modes is $N = 99$.

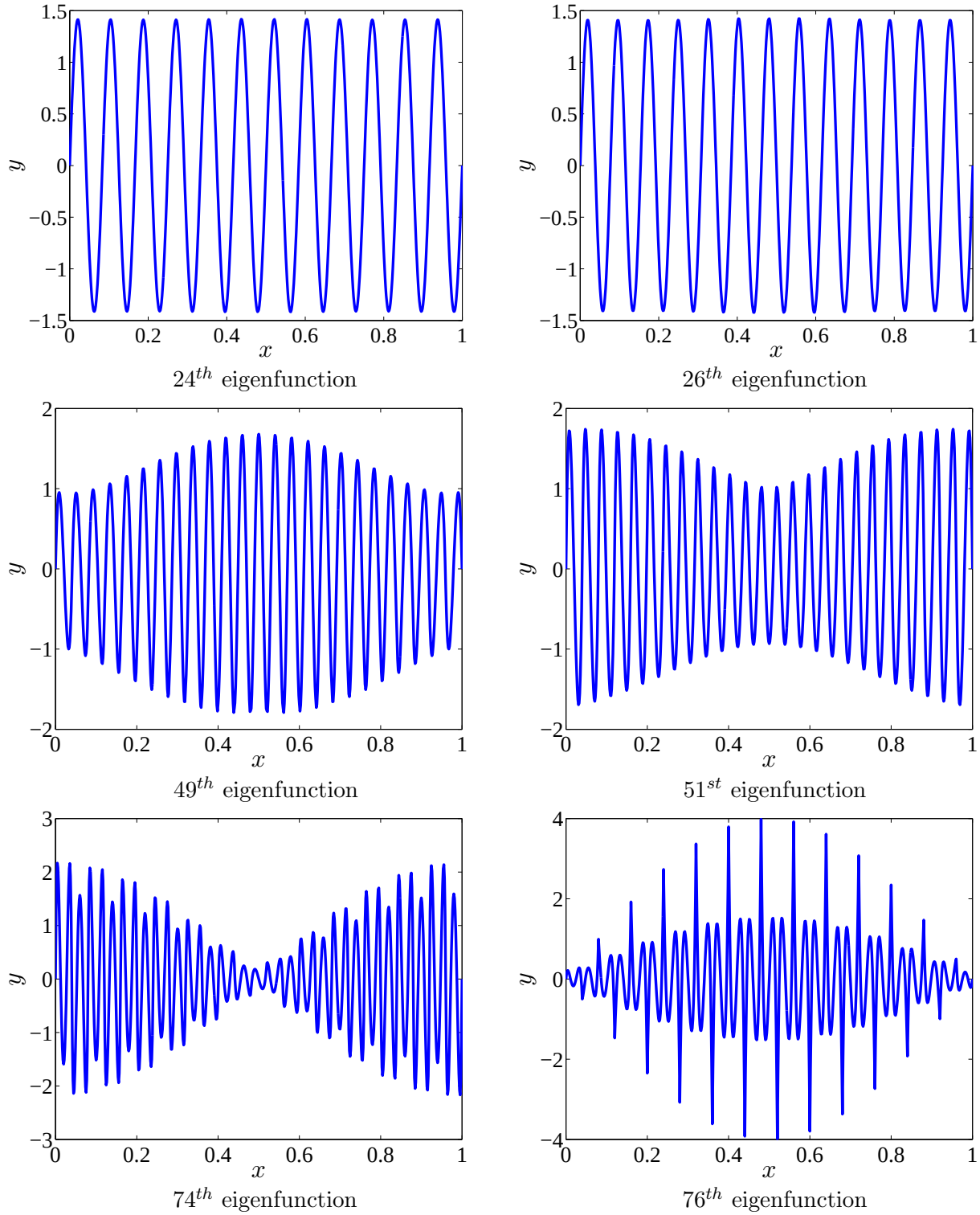
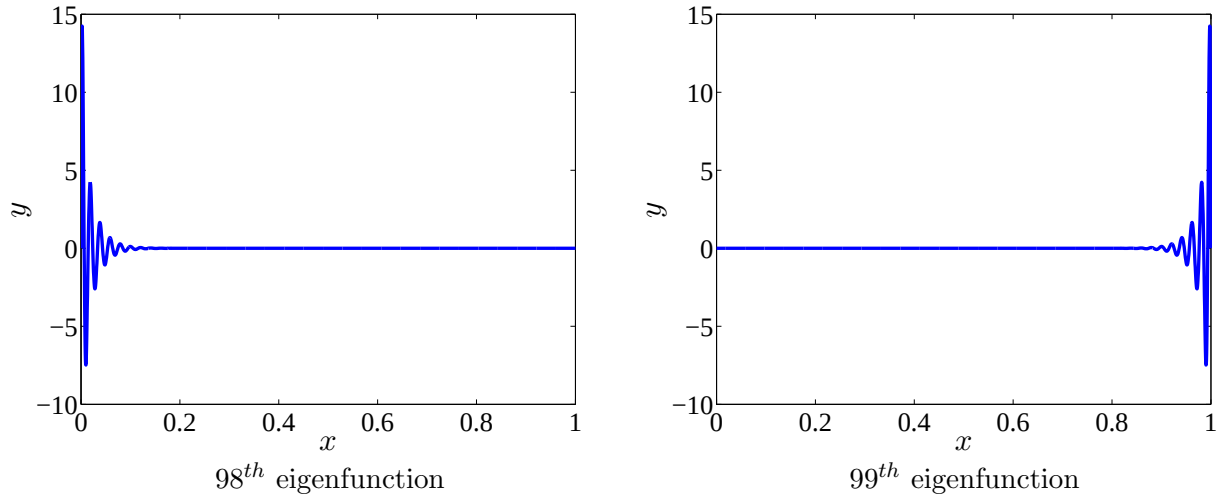


Figure 13: Visualization of the discrete 24^{th}, 26^{th}, 49^{th}, 51^{st}, 74^{th}, and 76^{th} eigenfunctions for a C^0-continuous quartic finite element approximation where the total number of discrete modes is $N = 99$. These eigenfunctions correspond to modes near the transition points between branches in the error spectrum.



98^{th} eigenfunction 99^{th} eigenfunction

Figure 14: Visualization of the discrete eigenfunctions corresponding to "outlier" modes for a C^3-continuous quartic B-spline approximation where the total number of discrete modes is $N = 99$. Only the last two modes are significantly in error. This holds for all significantly large values of N.

utilized. These modes lie at the end of the spectrum and are finite in number. Numerical evidence has revealed that, in the context of the problem given by (5), there are precisely $p-1$ outlier modes when p is odd and $p-2$ outlier modes when p is even. In Figures 14(a) and (b), the discrete eigenfunctions corresponding to the two outlier modes for C^3-continuous quartic B-splines are presented for $N = 99$. It is apparent from the figures that the $(N-1)^{st}$ and N^{th} discrete eigenfunctions are very inaccurate. In fact, they are mirror images of each other, zero throughout most of the domain, and only nonzero near the domain boundaries. While this spurious qualitative behavior may introduce error for problems in which the outlier modes are participating, the fact that the eigenfunctions hold support only near the boundaries of the domain indicates that their presence will not spoil the accuracy of numerical approximations in the interior of the domain, at least in the context of linear elliptic, parabolic, and hyperbolic problems. It is worth noting that we have observed that the outlier modes for other polynomial degrees are also only nonzero near the domain boundaries.

We think the results comparing the accuracy of eigenvalues and eigenfunctions for equal-order B-spline and finite element cases speak for themselves. Maximally smooth B-splines are simply much more accurate than equal-order C^0-continuous finite elements based on equal numbers of degrees-of-freedom.

4 Elliptic Boundary-value Problem

Let $f \in (L^2(\Omega))^n$ be a given function, referred to as the forcing. The continuous elliptic boundary-value problem is stated as follows: Find $u \in \mathcal{V}$ such that for all $w \in \mathcal{V}$

$$a(w,u) = (w,f) \tag{19}$$

The discrete counterpart is: Find $u^h \in \mathcal{V}^h$ such that for all $w^h \in \mathcal{V}^h$

$$a(w^h,u^h) = (w^h,f) \tag{20}$$

The usual approach to obtaining error estimates for (20) is to utilize the fact that u^h is the best approximation to u in $\|\cdot\|_E$, that $\|\cdot\|_E$ is equivalent to the $(H^m(\Omega))^n$-norm, and standard approximation estimates [17], from which it immediately follows that

$$\|u^h - u\|_E \leq ch^{p+1-m}\|u\|_{p+1} \tag{21}$$

Error estimates in $(L^2(\Omega)^n)$ can be obtained using the Aubin-Nitsche method, viz.,

$$\|u^h - u\| \leq ch^{\sigma}\|u\|_{p+1} \tag{22}$$

where $\sigma = \min\{p+1, 2(p+1-m)\}$. These results hold for both finite elements and B-splines. All standard mathematical texts on finite elements may be consulted for background. Equation (21) is proved for B-splines in Bazilevs et al. [4], and (22) is a simple exercise.

We now investigate the approximation of (19) by (20) from the point of view of eigenfunction expansions. Due to the completeness of the eigenfunctions, we can expand the solutions of (19) and (20) in terms of the eigenfunctions of the continuous and discrete eigenvalue problems,

$$u = \sum_{l=1}^{\infty} d_l u_l \tag{23}$$

$$u_l = \sum_{l=1}^{N} d_l^h u_l^h \tag{24}$$

where d_l and d_l^h are the Fourier coefficients of the continuous and discrete solutions, respectively. We can solve for the Fourier coefficients by substituting (23) and (24) into (19) and (20), respectively, and by replacing w in (19) with u_k and w^h in (20) by u_k^h, and by using the orthogonality properties and normalizations described earlier. We have

$$\lambda_l d_l = f_l \stackrel{\text{def}}{=} (u_l, f) \tag{25}$$

$$\lambda_l^h d_l^h = f_l^h \stackrel{\text{def}}{=} (u_l^h, f) \tag{26}$$

and so

$$u(x) = \sum_{l=1}^{\infty} \frac{f_l}{\lambda_l} u_l(x) \tag{27}$$

$$u^h(x) = \sum_{l=1}^{N} \frac{f_l^h}{\lambda_l^h} u_l^h(x) \tag{28}$$

Consequently, the error in the solution is

$$\begin{aligned} e(x) &= u^h(x) - u(x) \\ &= \bar{e}(x) + e'(x) \\ \bar{e}(x) &= \sum_{l=1}^{N} \bar{e}_l(x) \\ &= \sum_{l=1}^{N} \left(\frac{f_l^h}{\lambda_l^h} u_l^h(x) - \frac{f_l}{\lambda_l} u_l(x) \right) \end{aligned} \tag{29}$$

$$e'(x) = \sum_{l=N+1}^{\infty} e'_l(x)$$
$$= \sum_{l=N+1}^{\infty} \left(-\frac{f_l}{\lambda_l} u_l(x) \right) \tag{30}$$

We refer to $\bar{e}$ as the *resolved* error and e' as the *unresolved* error. Clearly, nothing can be done about e' within the Galerkin method. It is what it is. However, $\bar{e}$ is completely determined by the errors in the discrete eigenvalues and eigenfunctions. If the projection of the forcing onto the eigenfunctions is confined to the eigenmodes that are grossly inaccurate, it follows from (29) that u^h will likely be significantly in error.[4] Let us assume that the l^{th} eigenmode is one such mode and estimate $\bar{e}_l$ in the $\left(L^2(\Omega)\right)^n$-norm and energy-norm. To this end, it is convenient to introduce the temporary notations

$$a = \frac{1}{\lambda_l} \tag{31}$$
$$b = f_l = (u_l, f) \tag{32}$$
$$c = u_l \tag{33}$$
$$a^h = \frac{1}{\lambda_l^h} \tag{34}$$
$$b^h = f_l^h = (u_l^h, f) \tag{35}$$
$$c^h = u_l^h \tag{36}$$
$$\varepsilon_a = a^h - a = \frac{1}{\lambda_l^h} - \frac{1}{\lambda_l} \tag{37}$$
$$\varepsilon_b = b^h - b = \left(u_l^h - u_l, f \right) = (e_l, f) \tag{38}$$
$$\varepsilon_c = c^h - c = u_l^h - u_l = e_l \tag{39}$$

With these, we can write

$$\begin{aligned} \bar{e}_l &= a^h b^h c^h - abc \\ &= (a + \varepsilon_a)(b + \varepsilon_b)(c + \varepsilon_c) - abc \\ &= (ab\varepsilon_c + a\varepsilon_b c + ab\varepsilon_c) + (a\varepsilon_b\varepsilon_c + \varepsilon_a b\varepsilon_c + \varepsilon_a\varepsilon_b c) + \varepsilon_a\varepsilon_b\varepsilon_c \end{aligned} \tag{40}$$

This is a cubic polynomial in the errors ε_a, ε_b, and ε_c, with coefficients depending on the exact data represented by a, b, and c. Note that this is an *exact* expression for the modal error, $\bar{e}_l$. Now we wish to estimate the norms of $\bar{e}_l$. To do this we first need to estimate ε_a, ε_b, ε_c, a, b, and c:

$$|a| = \frac{1}{\lambda_l} \tag{41}$$
$$|b| = |(u_l, f)| \leq \|u_l\| \|f\| = \|f\| \tag{42}$$
$$\|c\| = \|u_l\| = 1 \tag{43}$$
$$\|c\|_E = \|u_l\|_E = \lambda_l^{1/2} \tag{44}$$

[4] In the elliptic boundary-value problem setting, there is the possibility of cancellation of modal errors as each error component $\bar{e}_l$ can be either positive or negative. Cancellation of modal errors will be discussed further in Subsection 4.1.

$$
\begin{aligned}
|\varepsilon_a| &= \left|\frac{1}{\lambda_l^h} - \frac{1}{\lambda_l}\right| \\
&= \frac{1}{\lambda_l^h}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right) \\
&\leq \frac{1}{\lambda_l}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right) \qquad (45)\\
|\varepsilon_b| &= |(e_l, f)| \\
&\leq \|e_l\|\|f\| \qquad (46)\\
\|\varepsilon_c\| &= \|e_l\| \qquad (47)\\
\|\varepsilon_c\|_E &= \|e_l\|_E \qquad (48)
\end{aligned}
$$

Employing these in (40), we have

$$
\begin{aligned}
\|\bar{e}_l\| \leq & \left(\frac{1}{\lambda_l}\|f\|\|e_l\| + \frac{1}{\lambda_l}\|e_l\|\|f\| + \frac{1}{\lambda_l}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|f\|\right) \\
& + \left(\frac{1}{\lambda_l}\|e_l\|^2\|f\| + \frac{1}{\lambda_l}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|f\|\|e_l\| + \frac{1}{\lambda_l}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|e_l\|\|f\|\right) \\
& + \frac{1}{\lambda_l}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|e_l\|^2\|f\| \\
= & \frac{\|f\|}{\lambda_l}\left(2\|e_l\| + \left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\left(1 + 2\|e_l\| + \|e_l\|^2\right) + \|e_l\|^2\right) \\
= & \frac{\|f\|}{\lambda_l}\left(\left(2\|e_l\| + \|e_l\|^2\right)\left(1 + \left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\right) + \left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\right) \\
= & 2\frac{\|f\|}{\lambda_l}\left(\|e_l\|\left(1 + \frac{1}{2}\|e_l\|\right)\left(1 + \left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\right) + \frac{1}{2}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\right) \qquad (49)
\end{aligned}
$$

It is apparent from (49) that $\|\bar{e}_l\|$ is bounded by terms involving $\|e_l\|$ and $\left(\lambda_l^h - \lambda_l\right)/\lambda_l$. By (10) and (11) these can both be estimated in terms of $\|e_l\|_E$:

$$
\|e\|_l \leq 2\frac{\|f\|}{\lambda_l}\left(\frac{\|e_l\|_E}{\lambda_l^{1/2}}\left(1 + \frac{1}{2}\frac{\|e_l\|_E}{\lambda_l^{1/2}}\right)\left(1 + \frac{\|e_l\|_E^2}{\lambda_l}\right) + \frac{1}{2}\frac{\|e_l\|_E^2}{\lambda_l}\right) \qquad (50)
$$

As is evident, the exact eigenvalues in the denominators tend to diminish the influence of the higher modes.

Proceeding as before, but this time for the energy-norm, we have

$$
\begin{aligned}
\|\bar{e}_l\|_E \leq & \left(\frac{1}{\lambda_l}\|f\|\|e_l\|_E + \frac{1}{\lambda_l^{1/2}}\|e_l\|\|f\| + \frac{1}{\lambda_l^{1/2}}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|f\|\right) \\
& + \left(\frac{1}{\lambda_l}\|e_l\|\|f\|\|e_l\|_E + \frac{1}{\lambda_l}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|f\|\|e_l\|_E + \frac{1}{\lambda_l^{1/2}}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|e_l\|\|f\|\right) \\
& + \frac{1}{\lambda_l}\left(\frac{\lambda_l^h - \lambda_l}{\lambda_l}\right)\|e_l\|\|f\|\|e_l\|_E \qquad (51)
\end{aligned}
$$

This expression reveals that the eigenfunction errors in the $(L^2(\Omega))^n$-norm and the energy-norm, and the eigenvalue errors, influence $\|\bar{e}_l\|_E$. As before, thanks to (10) and (11), we can express $\|e_l\|$

and $\left(\lambda_l^h - \lambda_l\right)/\lambda_l$ in terms of $\|e_l\|_E$:

$$\begin{aligned}\|\bar{e}_l\|_E \leq & \left(\frac{1}{\lambda_l}\|f\|\|e_l\|_E + \frac{1}{\lambda_l}\|e_l\|_E\|f\| + \frac{1}{\lambda_l^{3/2}}\|e_l\|_E^2\|f\|\right) \\ & + \left(\frac{1}{\lambda_l^{3/2}}\|f\|\|e_l\|_E^2 + \frac{1}{\lambda_l^2}\|f\|\|e_l\|_E^3 + \frac{1}{\lambda_l^2}\|e_l\|_E^3\|f\|\right) \\ & + \frac{1}{\lambda_l^{5/2}}\|e_l\|_E^4\|f\| \end{aligned} \tag{52}$$

Simplifying the above expression:

$$\begin{aligned}\|\bar{e}_l\|_E &\leq 2\frac{\|f\|\|e_l\|_E}{\lambda_l}\left(1 + \frac{\|e_l\|_E}{\lambda_l^{1/2}} + \frac{\|e_l\|_E^2}{\lambda_l} + \frac{1}{2}\frac{1}{\lambda_l^{3/2}}\|e_l\|_E^3\right) \\ &= 2\frac{\|f\|\|e_l\|_E}{\lambda_l}\left(1 + \frac{\|e_l\|_E}{\lambda_l^{1/2}}\left(1 + \frac{\|e_l\|_E}{\lambda_l^{1/2}} + \frac{1}{2}\frac{\|e_l\|_E^2}{\lambda_l}\right)\right)\end{aligned} \tag{53}$$

Once again, the beneficial effect of the exact eigenvalues is apparent.

4.1 Reconciliation with the "Best Approximation Property"

From (29) it seems apparent that a discrete solution to the elliptic boundary-value problem has the potential to be significantly in error. On the other hand, it is well-known that the Galerkin solution to the elliptic boundary-value problem possesses the so-called best approximation property in the energy-norm. The best approximation property is indeed a very strong and important result. It guarantees that we get the best possible fit of u by u^h in the energy-norm. Optimal error estimates ensue from this property. However, it is not in conflict with the observations made about (29). If the discrete eigenvalues and eigenfunctions that are participating in the solution are grossly in error, the "best approximation" will most likely not be very good. There are well-known examples of this phenomenon. The phenomenon of "locking" may be mentioned, of which there are several different manifestations, such as, volumetric locking, shear locking, membrane locking, etc. Volumetric locking fits precisely into the present elliptic boundary-value problem format. For example, consider a plane strain elasticity problem in which the ratio of the Lamé parameters, λ and μ, is very large, say $\lambda/\mu \gg 1$. Solve a typical problem, such as the classical benchmark of a plane strain beam in bending (see, e.g., Figures 4.4.1 and 4.4.2 on page 220 in Hughes [12]), for which there is an exact Airy stress function solution. Utilize standard linear triangular or bilinear quadrilateral finite elements. The results are grossly inaccurate (see, e.g., Figure 4.7.1 on page 249 in Hughes [12]). Nevertheless, they represent the best approximation in the energy-norm for the function spaces employed. Over the years, deficiencies of this kind have been circumvented through a variety of alternative finite elements and variational formulations. However, the type of pathology identified in our eigenvalue and eigenfunction calculations, that is, Figures 2-4, represents a quite different problem, but one that is not a major concern for the elliptic boundary-value problem. The reason for this is that higher modes usually do not participate significantly in solutions, except possibly locally. This is referred to as elliptic regularity in mathematics and Saint-Venant's principle in mechanics. The manifestation of this in the modal result is that eigenvalues are in the *denominators* of the discrete and exact solutions and they significantly attenuate the participation of the higher modes. Consequently, the standard elliptic boundary-value problems are usually forgiving, despite the potential for error expressed in Figures 2-4.

There is also a possibility that the overall solution error may be small even if the modal errors are large due to cancellation of error. Each modal component $\bar{e}_l$ of the resolved error $\bar{e}$ may be positive or negative. If two of the modal errors are large but equal in magnitude and opposite in sign, then they will cancel each other out when summed together in (29). However, this cancellation phenomena is entirely limited to the elliptic setting. In the hyperbolic setting, for example, modal errors oscillate at different frequencies and the consequences of our global error estimates are fully realized at later times even if there is an initial cancellation of errors. In Subsection 6.1, we will present a numerical example illustrating this phenomenon.

5 Parabolic Initial-value Problem

Let

$$\mathcal{V}_T = \left\{ v \in L^2\left(0,T;\mathcal{V}\right) : \frac{\partial v}{\partial t} \in L^2\left((0,T);\mathcal{V}^*\right) \right\} \tag{54}$$

where $T \in \mathbb{R}^+$ is the final time of interest and $\mathcal{V}^*$ is the dual space of $\mathcal{V}$, and let $f \in L^2\left((0,T);(L^2(\Omega))^n\right)$, the forcing, and $U \in (L^2(\Omega))^n$, the initial condition, be given functions. The continuous, parabolic initial-value problem is stated as follows: Find $u \in \mathcal{V}_T$ such that for all $w \in \mathcal{V}$ and almost every $t \in (0,T)$,

$$\left\langle w, \frac{\partial u}{\partial t}(t) \right\rangle + a(w, u(t)) = (w, f(t)) \tag{55}$$

and

$$(w, u(0)) = (w, U) \tag{56}$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing. We remark that (56) makes sense as functions in $\mathcal{V}_T$ also lie in $C\left([0,T];(L^2(\Omega))^n\right)$, the space of continuous functions $u : [0,T] \to (L^2(\Omega))^n$ (see Theorem 3 in Section 5.9.2 of Evans [10]). Let

$$\mathcal{V}_T^h = \left\{ v \in L^2\left((0,T);\mathcal{V}^h\right) : \frac{\partial v}{\partial t} \in L^2\left((0,T);\mathcal{V}^h\right) \right\} \tag{57}$$

The semi-discrete counterpart of the parabolic initial-value problem is: Find $u^h \in \mathcal{V}_T^h$ such that for all $w^h \in \mathcal{V}$ and almost every $t \in (0,T)$,

$$\left\langle w^h, \frac{\partial u^h}{\partial t}(t) \right\rangle + a(w^h, u^h(t)) = (w^h, f(t)) \tag{58}$$

and

$$(w^h, u^h(0)) = (w^h, U) \tag{59}$$

The standard finite element error estimate for the parabolic case is

$$\|e(t)\| \leq ch^\sigma \left(\|u(t)\|_{(H^{p+1}(\Omega))^n)} + \exp\left(-\lambda_1^h t\right) \|u(0)\|_{(H^{p+1}(\Omega))^n} + \int_0^t \exp\left(-\lambda_1^h (t-\tau)\right) \left\| \frac{\partial u}{\partial t}(\tau) \right\|_{(H^{p+1}(\Omega))^n} \mathrm{d}\tau \right) \tag{60}$$

where σ is the constant defined in (15) and $\lambda_1^h > 0$ is the smallest eigenvalue of the discretized eigenproblem [17].

Employing eigenfunction expansions, we can express the solutions of (55) and (58) as

$$u(t) = \sum_{l=1}^{\infty} d_l(t) u_l \tag{61}$$

$$u^h(t) = \sum_{l=1}^{N} d_l^h(t) u_l^h \tag{62}$$

where $d_l(t)$ and $d_l^h(t)$ are the Fourier coefficients of the continuous and discrete solutions, respectively. To generate uncoupled first-order ordinary differential equations governing the Fourier coefficients, we substitute (61) and (62) into (55) and (58), respectively, and replace w in (55) by u_k and w^h in (58) by u_k^h, and employ the orthogonality properties of the eigenfunctions, to arrive at

$$\dot{d}_l(t) + \lambda_l d_l(t) = f_l(t) \stackrel{\text{def}}{=} (u_l, f(t)) \quad \forall l = 1, 2, \ldots, \infty \tag{63}$$

and

$$\dot{d}_l^h(t) + \lambda_l^h d_l^h(t) = f_l^h(t) \stackrel{\text{def}}{=} (u_l^h, f(t)) \quad \forall l = 1, 2, \ldots, N \tag{64}$$

where a superposed dot denotes time differentiation. Proceeding analogously with (56) and (59), we determine the corresponding initial conditions

$$d_l(0) = U_l \stackrel{\text{def}}{=} (u_l, U) \quad \forall l = 1, 2, \ldots, \infty \tag{65}$$

and

$$d_l^h(0) = U_l^h \stackrel{\text{def}}{=} (u_l^h, U) \quad \forall l = 1, 2, \ldots, N \tag{66}$$

Solving the initial-value problems for the ordinary differential equations, we obtain

$$d_l(t) = U_l \exp(-\lambda_l t) + \int_0^t \exp(-\lambda_l (t-\tau)) f_l(\tau) \mathrm{d}\tau \tag{67}$$

and

$$d_l^h(t) = U_l^h \exp\left(-\lambda_l^h t\right) + \int_0^t \exp\left(-\lambda_l^h (t-\tau)\right) f_l^h(\tau) \mathrm{d}\tau \tag{68}$$

and so

$$u(x,t) = \sum_{l=1}^{\infty} \left(U_l \exp(-\lambda_l t) + \int_0^t \exp(-\lambda_l (t-\tau)) f_l(\tau) \mathrm{d}\tau \right) u_l(x) \tag{69}$$

and

$$u^h(x,t) = \sum_{l=1}^{N} \left(U_l^h \exp\left(-\lambda_l^h t\right) + \int_0^t \exp\left(-\lambda_l^h (t-\tau)\right) f_l^h(\tau) \mathrm{d}\tau \right) u_l^h(x) \tag{70}$$

The error can then be expressed as

$$\begin{aligned}
e(x,t) =& u^h(x,t) - u(x,t) \\
=& \bar{e}(x,t) + e'(x,t) \\
\bar{e}(x,t) =& \sum_{l=1}^{N} \bar{e}_l(x,t) \\
=& \sum_{l=1}^{N} \left(U_l^h \exp\left(-\lambda_l^h t\right) u_l^h(x) - U_l \exp(-\lambda_l t) u_l(x) \right. \\
&+ \left. \int_0^t \exp\left(-\lambda_l^h (t-\tau)\right) f_l^h(\tau) \mathrm{d}\tau u_l^h(x) - \int_0^t \exp(-\lambda_l (t-\tau)) f_l(\tau) \mathrm{d}\tau u_l(x) \right)
\end{aligned} \tag{71}$$

$$
\begin{aligned}
e'(x,t) &= \sum_{l=N+1}^{\infty} e'_l(x,t) \\
&= \sum_{l=N+1}^{\infty} \left(-U_l \exp\left(-\lambda_l t\right) u_l(x) - \int_0^t \exp\left(-\lambda_l (t-\tau)\right) f_l(\tau) \mathrm{d}\tau u_l(x) \right)
\end{aligned} \tag{72}
$$

Again, there is nothing more to be said about e'. For $\bar{e}$ the error is due to the errors in eigenvalues and eigenfunctions. The errors in decay rates are due exclusively to the errors in eigenvalues. The initial condition error is due to the projection error of U onto the eigenfunctions. This decays rapidly for the higher modes but is important at early times say up to $t = O(\lambda_l^{-1})$ for the error in U_l. The situation is similar for the forcing. If f is concentrated around time $\bar{t}$, then errors may be appreciable for short time intervals thereafter, say for $(t-\bar{t})^+ = O(\lambda_l^{-1})$ where $(\cdot)^+$ denotes the Macaulay bracket, that is, $(\tau)^+ = \tau$ if $\tau > 0$, and is zero otherwise. The rationale for this is provided by assuming the forcing to be an impulse at time $\bar{t}$, that is, a Dirac distribution with respect to time

$$
f(x,t) = \bar{f}(x)\delta(t-\bar{t}) \tag{73}
$$

Then

$$
\begin{aligned}
f_l(t) &= \left(u_l(x), f(x,t)\right) \\
&= \left(u_l(x), \bar{f}\right) \delta(t-\bar{t}) \\
&= \bar{f}_l \delta(t-\tau)
\end{aligned} \tag{74}
$$

Likewise,

$$
\begin{aligned}
f_l^h(t) &= \left(u_l^h(x), f(x,t)\right) \\
&= \left(u_l^h(x), \bar{f}\right) \delta(t-\bar{t}) \\
&= \bar{f}_l^h \delta(t-\tau)
\end{aligned} \tag{75}
$$

Substituting (74) and (75) into (71) results in

$$
\begin{aligned}
\bar{e}(x,t) &= \sum_{l=1}^{N} \left(U_l^h \exp\left(-\lambda_l^h t\right) u_l^h(x) - U_l \exp\left(-\lambda_l t\right) u_l(x) \right. \\
&\quad \left. + \int_0^t \left(\exp\left(-\lambda_l^h (t-\tau)\right) \bar{f}_l^h u_l^h(x) - \exp\left(-\lambda_l (t-\tau)\right) \bar{f}_l u_l(x) \right) \delta(t-\tau) \mathrm{d}\tau \right) \\
&= \sum_{l=1}^{N} \left(U_l^h \exp\left(-\lambda_l^h t\right) u_l^h(x) - U_l \exp\left(-\lambda_l t\right) u_l(x) \right. \\
&\quad \left. + \bar{f}_l^h \exp\left(-\lambda_l^h \left(t-\bar{t}\right)^+\right) u_l^h(x) - \bar{f}_l \exp\left(-\lambda_l \left(t-\bar{t}\right)^+\right) u_l(x) \right)
\end{aligned} \tag{76}
$$

From (76) we see that the error due to an impulse at time $\bar{t}$ behaves very much like the error in initial data. If we view the forcing to be decomposed into a sequence of impulses, then each impulse would create an error similar in form to the error created by the initial data. Consequently, we can simplify our discussion of errors by considering only those caused by initial data.

As in the case of the elliptic boundary-value problem, it is revealing to investigate the contribution of a typical modal error. However, the analysis is similar to that performed for the elliptic

boundary-value problem and is tedious, so we will content ourselves with presenting only final results, neglecting forcing contributions. For the $\left(L^2(\Omega)\right)^n$-norm, the estimate is

$$\begin{aligned}\|\bar{e}_l(t)\| &\leq \exp\left(-\lambda_l t\right)\|U\|\left\{2\|e_l\|\left(2+\|e_l\|\right)+\min\left\{\|e_l\|_E^2 t,1\right\}\right\}\\ &\leq \exp\left(-\lambda_l t\right)\|U\|\left\{2\frac{\|e_l\|_E}{\lambda_l^{1/2}}\left(2+\frac{\|e_l\|_E}{\lambda_l^{1/2}}\right)+\min\left\{\|e_l\|_E^2 t,1\right\}\right\}\end{aligned} \tag{77}$$

Clearly for large times, the term $\exp\left(-\lambda_l t\right)$ dominates, and $\|\bar{e}_l\| \to 0$ as $-\lambda_l t \to -\infty$. For short times, that is $\|e_l\|_E^2 t < 1$, we have

$$\|\bar{e}_l(t)\| \leq \exp\left(-\lambda_l t\right)\|U\|\left\{2\frac{\|e_l\|_E}{\lambda_l^{1/2}}\left(2+\frac{\|e_l\|_E}{\lambda_l^{1/2}}\right)+\|e_l\|_E^2 t\right\} \tag{78}$$

Likewise, for the energy-norm, we have

$$\|\bar{e}_l(t)\|_E \leq \exp\left(-\lambda_l t\right)\|U\|\left\{2\|e_l\|_E\left(2+\frac{\|e_l\|_E}{\lambda_l^{1/2}}\right)+\lambda_l^{1/2}\min\left\{\|e_l\|_E^2 t,1\right\}\right\} \tag{79}$$

As for the $(L^2(\Omega))^n$-norm, $\|\bar{e}_l\|_E \to 0$ as $-\lambda t \to -\infty$. For short times,

$$\|\bar{e}_l(t)\|_E \leq \exp\left(-\lambda_l t\right)\|U\|\left\{2\|e_l\|_E\left(2+\frac{\|e_l\|_E}{\lambda_l^{1/2}}\right)+\lambda_l^{1/2}\min\left\{\|e_l\|_E^2 t,1\right\}\right\} \tag{80}$$

Examination of (77)-(80) again confirms that modal errors may be explicitly expressed in terms of errors of the eigenfunctions in the $(L^2(\Omega))^n$-norm and the energy-norm, and errors in the eigenvalues.

The fact that higher modes decay exponentially fast suggests that the parabolic problem is also very forgiving, at least after a short time interval of $O(\lambda_k^{-1})$, where k here denotes the lowest mode number for which the discrete eigenvalue and/or eigenfunction are/is significantly in error.

6 Hyperbolic Initial-value Problem

Let

$$\mathcal{V}_T = \left\{v \in L^2\left(0,T;\mathcal{V}\right) : \frac{\partial v}{\partial t} \in L^2\left((0,T);\left(L^2(\Omega)\right)^n\right) \text{ and } \frac{\partial^2 v}{\partial t^2} \in L^2\left((0,T);\mathcal{V}^*\right)\right\} \tag{81}$$

where, as before, $T \in \mathbb{R}^+$, $\mathcal{V}^*$ is the dual of $\mathcal{V}$ and $f \in L^2\left((0,T);\left(L^2(\Omega)\right)^n\right)$ is the forcing. Let $U \in \left(L^2(\Omega)\right)^n$ and $V \in \mathcal{V}^*$ be the initial data, referred to as the initial displacement and initial velocity, respectively. The continuous hyperbolic initial-value problem is: Find $u \in \mathcal{V}_T$ such that for all $w \in V$ and almost every time $t \in (0,T)$,

$$\left\langle w, \frac{\partial^2 u}{\partial t^2}(t)\right\rangle + a(w,u(t)) = (w,f(t)) \tag{82}$$

$$(w,u(0)) = (w,U) \tag{83}$$

and

$$\left\langle w, \frac{\partial u}{\partial t}(0)\right\rangle = \langle w,V\rangle \tag{84}$$

where $\langle \cdot, \cdot \rangle$ is the duality pairing. We remark that (83) and (84) make sense because functions in $\mathcal{V}_T$ lie in $C\left([0,T];(L^2(\Omega))^n\right)$ and their time derivatives lie in $C\left([0,T];\mathcal{V}^*\right)$ (see Theorem 2 in Section 5.9.2 of Evans [10]). Let

$$\mathcal{V}_T^h = \left\{ v \in L^2\left(0,T;\mathcal{V}^h\right) : \frac{\partial v}{\partial t} \in L^2\left((0,T);\mathcal{V}^h\right) \text{ and } \frac{\partial^2 v}{\partial t^2} \in L^2\left((0,T);\mathcal{V}^h\right) \right\} \tag{85}$$

The semi-discrete counterpart of (82)-(84) is: Find $u^h \in \mathcal{V}_T^h$ such that for all $w^h \in V^h$ and almost every time $t \in (0,T)$,

$$\left\langle w^h, \frac{\partial^2 u^h}{\partial t^2}(t) \right\rangle + a(w^h, u^h(t)) = (w^h, f(t)) \tag{86}$$

$$\left(w^h, u^h(0)\right) = \left(w^h, U\right) \tag{87}$$

and

$$\left(w^h, \frac{\partial u^h}{\partial t}(0)\right) = \left\langle w^h, V \right\rangle \tag{88}$$

Let

$$E\left(u, \frac{\partial u}{\partial t}\right) = \frac{1}{2}\left(\left(\frac{\partial u}{\partial t}, \frac{\partial u}{\partial t}\right) + a(u,u)\right) \tag{89}$$

denote the *total energy*. The square root of E is a norm on $\left(L^2(\Omega)\right)^n \times \mathcal{V}$, equivalent to the $\left(L^2(\Omega)\right)^n \times (H^m(\Omega))^n$-norm. The standard finite element error estimate for the hyperbolic problem in the energy-norm is

$$\begin{aligned} E\left(e, \frac{\partial e}{\partial t}\right)^{1/2} \leq & c\left(h^\nu \left(\|u(0)\|_{(H^{p+1}(\Omega))^n} + \|u(t)\|_{(H^{p+1}(\Omega))^n} \right) \right. \\ & \left. + h^\sigma \left(\left\| \frac{\partial u}{\partial t}(0) \right\|_{(H^{p+1}(\Omega))^n} + \left\| \frac{\partial u}{\partial t}(t) \right\|_{(H^{p+1}(\Omega))^n} + \int_0^t \left\| \frac{\partial^2 u}{\partial t^2}(\tau) \right\|_{(H^{p+1}(\Omega))^n} \mathrm{d}\tau \right) \right) \end{aligned} \tag{90}$$

where σ is defined in (15) and $\nu = p + 1 - m$ [17]. Note that $\nu \leq \sigma$. Consequently, for early times the rate of convergence of the energy-norm is ν. At times of order h^{-m}, the σ term begins to dominate in the sense that the error becomes $O(h^\sigma t)$ for $t \geq O(h^{-m})$ due to the time integral in (90).

Proceeding as for the parabolic case (see (61)-(66)), we arrive at the uncoupled second-order, ordinary differential equations, and corresponding initial conditions, governing the Fourier coefficients:

$$\ddot{d}_l(t) + \lambda_l d_l(t) = f_l(t) \quad \forall l = 1, 2, \ldots, \infty \tag{91}$$

with

$$d_l(0) = U_l \quad \forall l = 1, 2, \ldots, \infty \tag{92}$$

$$\dot{d}_l(0) = V_l \stackrel{\text{def}}{=} \langle u_l, V \rangle \quad \forall l = 1, 2, \ldots, \infty \tag{93}$$

and

$$\ddot{d}_l^h(t) + \lambda_l^h d_l^h(t) = f_l^h(t) \quad \forall l = 1, 2, \ldots, N \tag{94}$$

with

$$d_l^h(0) = U_l^h \quad \forall l = 1, 2, \ldots, N \tag{95}$$

$$\dot{d}_l^h(0) = V_l^h \stackrel{\text{def}}{=} \langle u_l^h, V \rangle \quad \forall l = 1, 2, \ldots, N \tag{96}$$

Solving these ordinary differential equations and defining the two frequencies $\omega_l = (\lambda_l)^{1/2}$ and $\omega_l^h = \left(\lambda_l^h\right)^{1/2}$, we obtain

$$d_l(t) = U_l \cos(\omega_l t) + \frac{V_l}{\omega_l} \sin(\omega_l t) + \frac{1}{\omega_l} \int_0^t \sin(\omega_l (t-\tau))\, f_l(\tau) \mathrm{d}\tau \tag{97}$$

$$d_l^h(t) = U_l^h \cos\left(\omega_l^h t\right) + \frac{V_l^h}{\omega_l^h} \sin\left(\omega_l^h t\right) + \frac{1}{\omega_l^h} \int_0^t \sin\left(\omega_l^h (t-\tau)\right) f_l^h(\tau) \mathrm{d}\tau \tag{98}$$

and so

$$u(x,t) = \sum_{l=1}^{\infty} \left(U_l \cos(\omega_l t) + \frac{V_l}{\omega_l} \sin(\omega_l t) + \frac{1}{\omega_l} \int_0^t \sin(\omega_l (t-\tau))\, f_l(\tau) \mathrm{d}\tau \right) u_l(x) \tag{99}$$

$$u^h(x,t) = \sum_{l=1}^{N} \left(U_l^h \cos\left(\omega_l^h t\right) + \frac{V_l^h}{\omega_l^h} \sin\left(\omega_l^h t\right) + \frac{1}{\omega_l^h} \int_0^t \sin\left(\omega_l^h (t-\tau)\right) f_l^h(\tau) \mathrm{d}\tau \right) u_l^h(x) \tag{100}$$

The error is then expressed in the now familiar way as

$$e(x,t) = u^h(x,t) - u(x,t) = \bar{e}(x,t) + e'(x,t)$$

$$\begin{aligned} \bar{e}(x,t) = \sum_{l=1}^{N} \Bigg(& U_l^h \cos\left(\omega_l^h t\right) u_l^h(x) - U_l \cos(\omega_l t)\, u_l(x) + \frac{V_l^h}{\omega_l^h} \sin\left(\omega_l^h t\right) u_l^h(x) - \frac{V_l}{\omega_l} \sin(\omega_l t)\, u_l(x) \\ & + \frac{1}{\omega_l^h} \int_0^t \sin\left(\omega_l^h (t-\tau)\right) f_l^h(\tau) \mathrm{d}\tau u_l^h(x) - \frac{1}{\omega_l} \int_0^t \sin(\omega_l (t-\tau))\, f_l(\tau) \mathrm{d}\tau u_l(x) \Bigg) \end{aligned} \tag{101}$$

$$e'(x,t) = \sum_{l=N+1}^{\infty} \left(-U_l \cos(\omega_l t) - \frac{V_l}{\omega_l} \sin(\omega_l t) - \frac{1}{\omega_l} \int_0^t \sin(\omega_l (t-\tau))\, f_l(\tau) \mathrm{d}\tau \right) u_l(x) \tag{102}$$

By assuming the forcing is a Dirac distribution with respect to time, as in the parabolic case (see (73)-(76)), we can simplify (101) to

$$\begin{aligned} \bar{e}(x,t) = \sum_{l=1}^{N} \Bigg(& U_l^h \cos\left(\omega_l^h t\right) u_l^h(x) - U_l \cos(\omega_l t)\, u_l(x) + \frac{V_l^h}{\omega_l^h} \sin\left(\omega_l^h t\right) u_l^h(x) - \frac{V_l}{\omega_l} \sin(\omega_l t)\, u_l(x) \\ & + \frac{\bar{f}_l^h}{\omega_l^h} \sin\left(\omega_l^h \left(t-\bar{t}\right)^+\right) u_l^h(x) - \frac{\bar{f}_l}{\omega_l} \sin\left(\omega_l \left(t-\bar{t}\right)^+\right) u_l(x) \Bigg) \end{aligned} \tag{103}$$

Thus, we see that an impulse at $\bar{t}$ behaves similarly to the initial velocity, but at times later than $\bar{t}$. Therefore, we can simplify our discussion of errors to consideration of those induced by the initial data.

As before, the analysis of modal errors is tedious. Consequently, we will just present final

results. The errors due to initial displacement are

$$\begin{aligned}\|\bar{e}_l(t)\| &\le \|U\| \left\{ \|e_l\| \left(2 + \|e_l\|\right) |\cos\left(\omega_l t\right)| + \left(1 + \|e_l\| \left(2 + \|e_l\|\right)\right) \left|\cos\left(\omega_l^h t\right) - \cos\left(\omega_l t\right)\right| \right\} \\ &\le \|U\| \left\{ \frac{\|e_l\|_E}{\omega_l} \left(2 + \frac{\|e_l\|_E}{\omega_l}\right) |\cos\left(\omega_l t\right)| + \left(1 + \frac{\|e_l\|_E}{\omega_l} \left(2 + \frac{\|e_l\|_E}{\omega_l}\right)\right) \left|\cos\left(\omega_l^h t\right) - \cos\left(\omega_l t\right)\right| \right\} \end{aligned} \tag{104}$$

$$\begin{aligned}\|\bar{e}_l(t)\|_E &\le \|U\| \left\{ \|e_l\|_E \left(2 + \|e_l\|\right) |\cos\left(\omega_l t\right)| + \left(\omega_l + \|e_l\|_E \left(2 + \|e_l\|\right)\right) \left|\cos\left(\omega_l^h t\right) - \cos\left(\omega_l t\right)\right| \right\} \\ &\le \|U\| \left\{ \|e_l\|_E \left(2 + \frac{\|e_l\|_E}{\omega_l}\right) |\cos\left(\omega_l t\right)| + \left(\omega_l + \|e_l\|_E \left(2 + \frac{\|e_l\|_E}{\omega_l}\right)\right) \left|\cos\left(\omega_l^h t\right) - \cos\left(\omega_l t\right)\right| \right\} \end{aligned} \tag{105}$$

The errors due to initial velocity may be obtained from the above by replacing $\|U\|$ with $\|V\|/\omega_l$, $\cos\left(\omega_l^h t\right)$ by $\sin\left(\omega_l^h t\right)$, and $\cos\left(\omega_l t\right)$ by $\sin\left(\omega_l t\right)$, viz.,

$$\begin{aligned}\|\bar{e}_l(t)\| &\le \frac{\|V\|}{\omega_l} \left\{ \|e_l\| \left(2 + \|e_l\|\right) |\sin\left(\omega_l t\right)| + \left(1 + \|e_l\| \left(2 + \|e_l\|\right)\right) \left|\sin\left(\omega_l^h t\right) - \sin\left(\omega_l t\right)\right| \right\} \\ &\le \frac{\|V\|}{\omega_l} \left\{ \frac{\|e_l\|_E}{\omega_l} \left(2 + \frac{\|e_l\|_E}{\omega_l}\right) |\sin\left(\omega_l t\right)| + \left(1 + \frac{\|e_l\|_E}{\omega_l} \left(2 + \frac{\|e_l\|_E}{\omega_l}\right)\right) \left|\sin\left(\omega_l^h t\right) - \sin\left(\omega_l t\right)\right| \right\} \end{aligned} \tag{106}$$

$$\begin{aligned}\|\bar{e}_l(t)\|_E &\le \frac{\|V\|}{\omega_l} \left\{ \|e_l\|_E \left(2 + \|e_l\|\right) |\sin\left(\omega_l t\right)| + \left(\omega_l + \|e_l\|_E \left(2 + \|e_l\|\right)\right) \left|\sin\left(\omega_l^h t\right) - \sin\left(\omega_l t\right)\right| \right\} \\ &\le \frac{\|V\|}{\omega_l} \left\{ \|e_l\|_E \left(2 + \frac{\|e_l\|_E}{\omega_l}\right) |\sin\left(\omega_l t\right)| + \left(\omega_l + \|e_l\|_E \left(2 + \frac{\|e_l\|_E}{\omega_l}\right)\right) \left|\sin\left(\omega_l^h t\right) - \sin\left(\omega_l t\right)\right| \right\} \end{aligned} \tag{107}$$

As for the elliptic boundary-value problem and the parabolic initial-value problem, we see from (104)-(107) that the modal errors may be bounded by the eigenfunction errors in the $(L^2(\Omega))^n$-norm and energy-norm, and the eigenvalue errors.

As we can observe from the initial displacement terms, any projection error in initial displacement will manifest itself immediately, as $\cos 0 = 1$, and will oscillate in time thereafter unattenuated due to the presence of the cosine behavior in time. There are no eigenvalues in the denominator, as in the elliptic boundary-value problem, and no exponentially decaying terms, as in the parabolic initial-value problem, to mitigate high modal errors. The errors in initial velocity start out as zero but quickly build up to their full values and oscillate in time thereafter undiminished in magnitude. Forcing errors behave similarly. As these errors will occur in the higher and intermediate modes, their impact will be felt almost instantaneously compared with the periods of the accurately represented lower modes. Nevertheless, there is some help from the eigenfrequencies in the denominators, but it is not as strong as for the eigenvalues in the denominators for the elliptic boundary-value problem because $\omega_l = (\lambda_l)^{1/2}$.

Both displacement and velocity phase errors, $\left|\cos\left(\omega_l^h t\right) - \cos\left(\omega_l t\right)\right|$ and $\left|\sin\left(\omega_l^h t\right) - \sin\left(\omega_l t\right)\right|$, eventually vitiate the solution no matter how accurate the projection errors.

The main conclusion for the hyperbolic case is that any modal error that is severe, combined with any initial data or forcing that non-negligibly projects onto that mode, will create significant errors that will appear virtually instantaneously and will persist thereafter for all time. In the hyperbolic case, we conclude that one cannot simply ignore the higher modes if they are significantly in error, and for the case of C^0-continuous finite elements the higher-modes, and even modes that might be categorized as intermediate, are significantly in error. The latter case might be of the most concern.

Nonlinear cases involving sudden loading, such as occurs in physical impact of elastic structures, would seem to fall into the category of hyperbolic problems in which errors of the kind anticipated might be severe. We believe that there is overwhelming practical evidence to support this. In engineering practice one does not exactly solve the t-continuous semi-discrete Galerkin equations. Time-stepping algorithms are utilized that introduce strong damping of the higher modes, suppressing their detrimental influence. In this regard, we may mention the HHT algorithm of Hilber et al. [11] and the generalized α algorithm of Chung and Hulbert [5], that are employed in implicit codes, such as Abaqus [1], and the artificial viscosities employed in explicit codes such as LS-DYNA [14].[5] Thus, in suppressing the influence of erroneous higher modes, these strategies result in solutions that are closer to the projection of the exact solutions. Nevertheless, it would seem that the more accurate higher-mode behavior of spline-based approximations would diminish, but not entirely eliminate, the need for algorithmic damping mechanisms and would produce significantly more accurate response in hyperbolic cases. An example illustrating this follows in the next subsection.

6.1 Numerical Investigation of Hyperbolic Approximations

To further elucidate the implications of error estimate (101), let us consider a very simple numerical test problem. Given $f(t) = 0$, $U = \sin(51\pi x)$, and $V = 0$, the exact solution to the hyperbolic initial-value problem given by (82)-(84) is

$$u(x,t) = \sin(51\pi x)\cos(51\pi t)$$

Note that this is a single wave solution, with solution coefficients (see (99))

$$U_l = \begin{cases} 1/\sqrt{2} & \text{if } l = 51 \\ 0 & \text{otherwise} \end{cases}, \quad V_l = 0, \quad f_l = 0$$

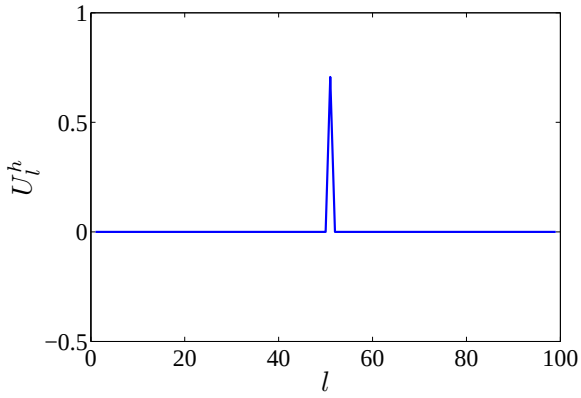
B-spline and finite element approximations to the hyperbolic initial-value problem are explicitly given by formula (100), wherein

$$U_l^h = \left(u_l^h, \sin(51\pi x)\right), \quad V_l^h = 0, \quad f_l^h = 0 \tag{108}$$

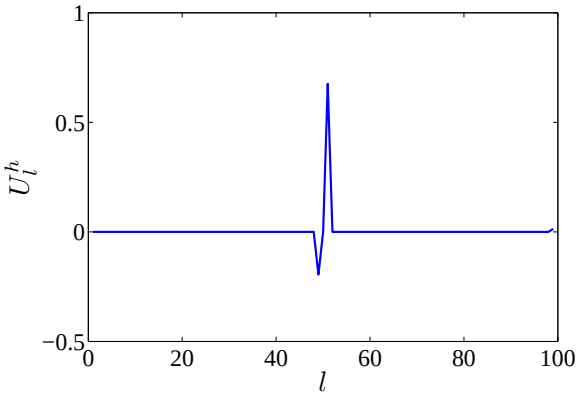
In what follows, we consider C^3-continuous quartic B-spline and C^0-continuous quartic finite element approximation spaces of dimension $N = 99$. We plot the solution coefficients U_l^h for these two approximation spaces in Figure 15 and note that $U_l^h \approx U_l$ for all l in the case of C^3-continuous quartic B-splines. Consequently, we expect the quartic B-spline approximation to the hyperbolic initial-value problem to resemble a single wave solution. In the case of C^0-continuous quartic finite elements, U_l^h is of appreciable magnitude for both $l = 49$ and $l = 51$, and thus we expect the quartic finite element approximation to be composed of two different waves. Moreover, finite element modes 49 and 51 are both located near the transition point between the second and third branches in the error spectrum, and as illustrated in Figure 13, the eigenfunctions corresponding to these modes are very inaccurate. For all of the above reasons, and in light of the error estimate provided by (101), we anticipate that the quartic finite element approximation to the hyperbolic initial-value problem has no approximability whatsoever, especially at later times t. This runs counter-intuitive to what one might normally anticipate, as modes with high wavenumber are not participating in the numerical solution of the hyperbolic problem.

We present the exact solution $u(x,t) = \sin(51\pi x)\cos(51\pi t)$ in conjunction with the C^3-continuous quartic B-spline and C^0-continuous quartic finite element approximations over the first half-cycle

[5] Both Abaqus and LS-DYNA now provide both implicit and explicit algorithmic options.



C^3 quartic B-spline initial condition



C^0 quartic finite element initial condition

Figure 15: Plots of the initial condition coefficients U_l^h for C^3-continuous quartic B-spline and C^0-continuous quartic finite element solutions of the hyperbolic initial-value problem with exact solution $u(x,t) = \sin(51\pi x)\cos(51\pi t)$. The total number of discrete modes for the numerical solutions is $N = 99$.

$t \in [0, 1/51]$ in Figure 16. The B-spline and finite element approximations were obtained by evaluating formula (100) using the coefficients given by (108). The B-spline approximation is visually exact over the first half-cycle and the finite element approximation is also fairly accurate. However, the finite element approximation does exhibit some spurious composite wave structure. It should be remarked that while the discrete initial condition for the finite element approximation is composed of modes 49 and 51, both of which are highly inaccurate, the initial condition is actually well-resolved. This enhanced resolution is due to cancellation of modal errors. In Figure 17, we present the exact solution and quartic B-spline and finite element approximations over the sixth half-cycle $t \in [5/51, 6/51]$. We immediately note that the B-spline approximation is nearly exact in the "eyeball norm", but the finite element approximation is both quantitatively and qualitatively inaccurate. The two-wave structure of the finite element approximation is visually apparent over this time period, as the influence of both modes 49 and 51 is observable in the plots. Lastly, in Figure 18, we present the exact solution and B-spline and finite element approximations over the eleventh half-cycle $t \in [10/51, 11/51]$. The B-spline approximation is still very accurate over this time period, but dispersion error has started to affect the solution to a minor degree. The B-spline wave approximation is oscillating at a slightly higher frequency than the exact wave solution due to the minimax principle, and hence the B-spline approximation will eventually lose approximability at sufficiently later times t. In contrast with the B-spline approximation, the finite element approximation has lost all approximability by the eleventh half-cycle and bears no resemblance whatsoever to the exact solution.

In addition to directly evaluating formula (100), we also solved the semi-discrete hyperbolic problem given by (86)-(88) numerically in time using the classical Runge-Kutta method as well as the generalized α algorithm [5]. We obtained very similar results to those reported here provided sufficiently small time-steps were employed to resolve the frequencies of the semi-discrete approximations. Moreover, we found that algorithmic damping was unable to restore qualitative accuracy for the quartic finite element approximation. In particular, it was discovered that algorithmic damping either did not significantly affect the quartic approximation or damped both participating modes. This is due to the fact that both participating finite element modes are medium modes.

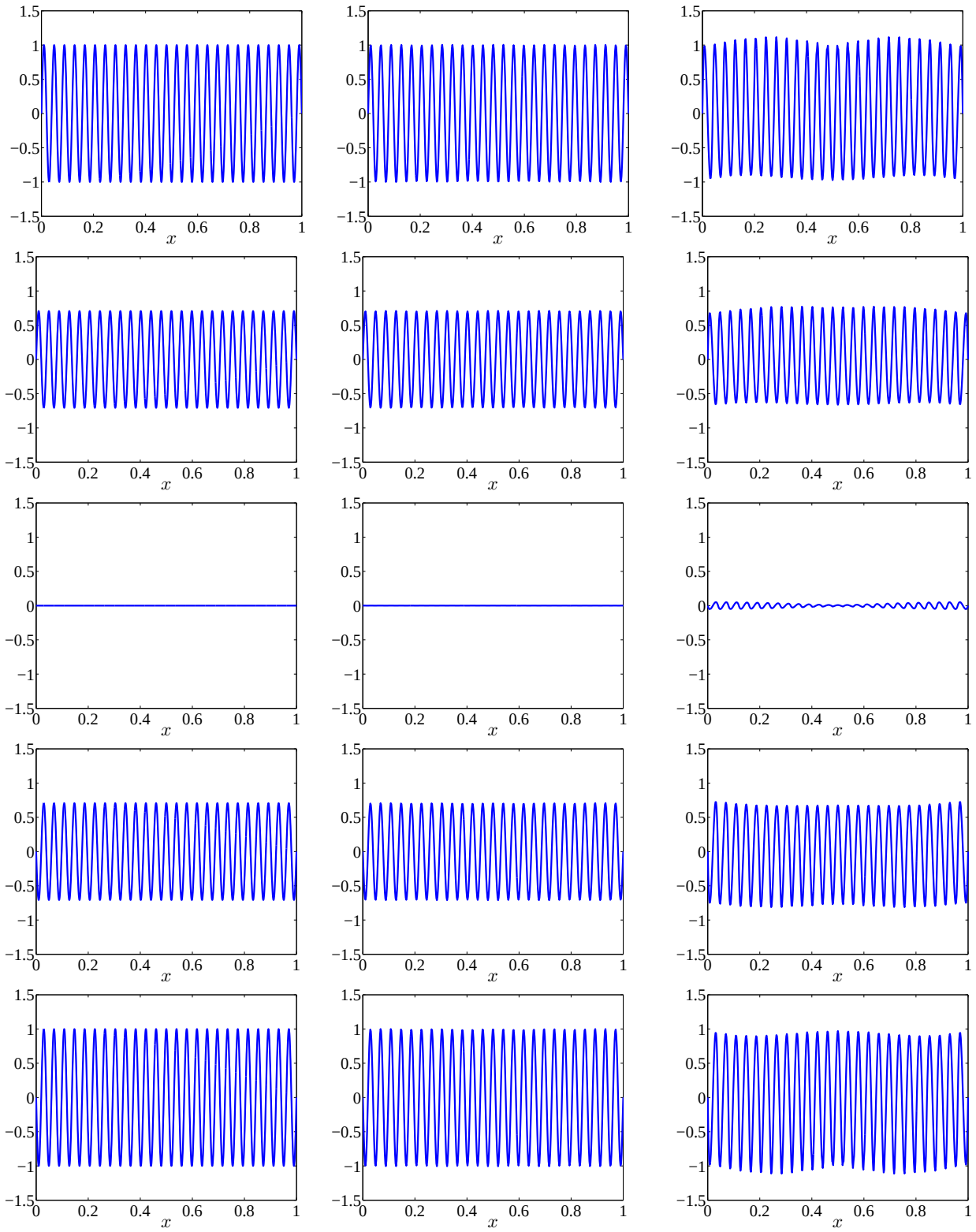


Figure 16: Plots of the exact and numerical solutions to the hyperbolic initial-value problem with exact solution $u(x,t) = \sin(51\pi x)\cos(51\pi t)$ over the first half-cycle $t \in [0, 1/51]$. The total number of discrete modes for the displayed numerical solutions is $N = 99$. Snapshots from top to bottom correspond to times $t = 0, 0.25/51, 0.5/51, 0.75/51,$ and $1/51$ respectively.

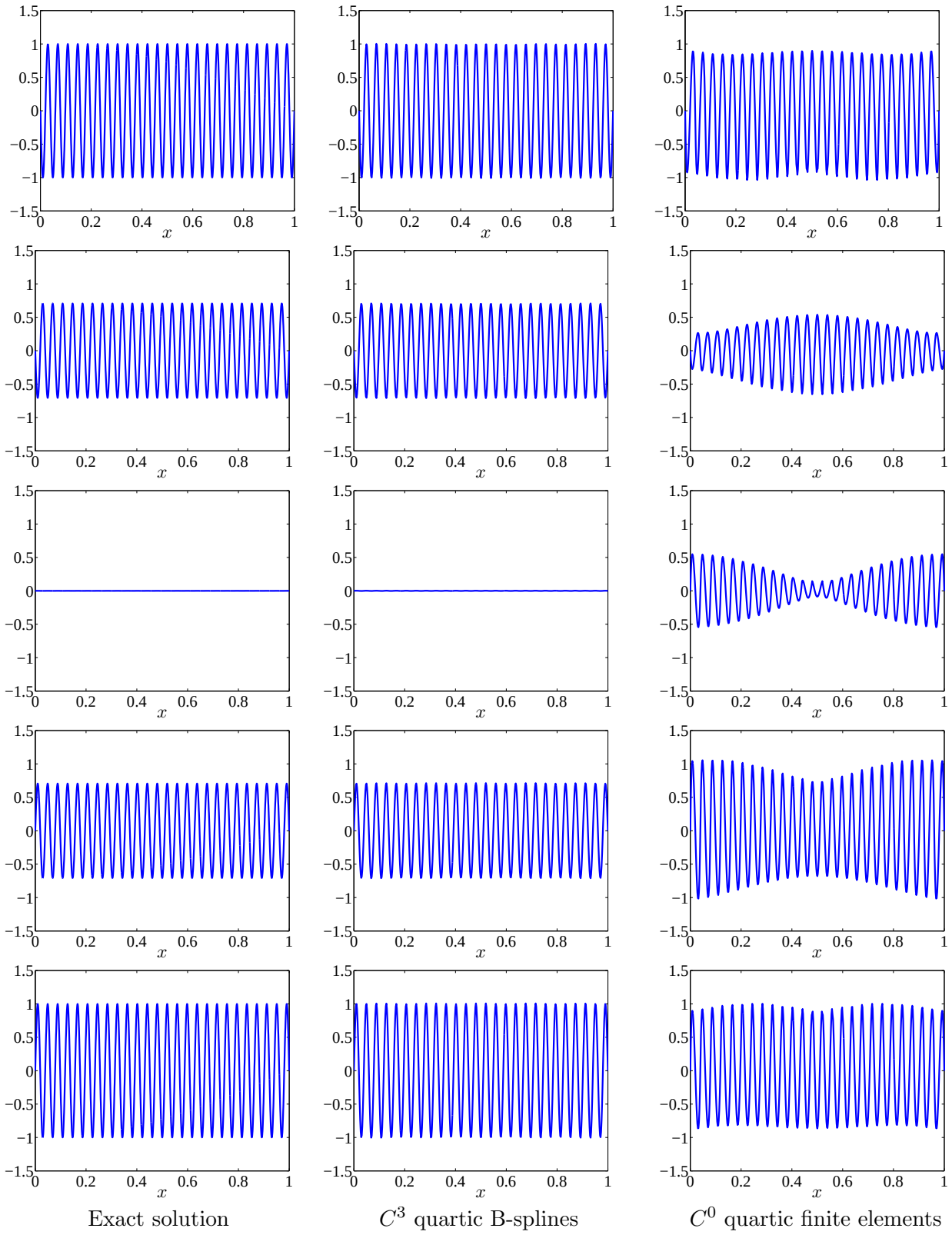


Figure 17: Plots of the exact and numerical solutions to the hyperbolic initial-value problem with exact solution $u(x,t) = \sin(51\pi x)\cos(51\pi t)$ over the sixth half-cycle $t \in [5/51, 6/51]$. The total number of discrete modes for the displayed numerical solutions is $N = 99$. Snapshots from top to bottom correspond to times $t = 5/51, 5.25/51, 5.5/51, 5.75/51,$ and $6/51$ respectively.

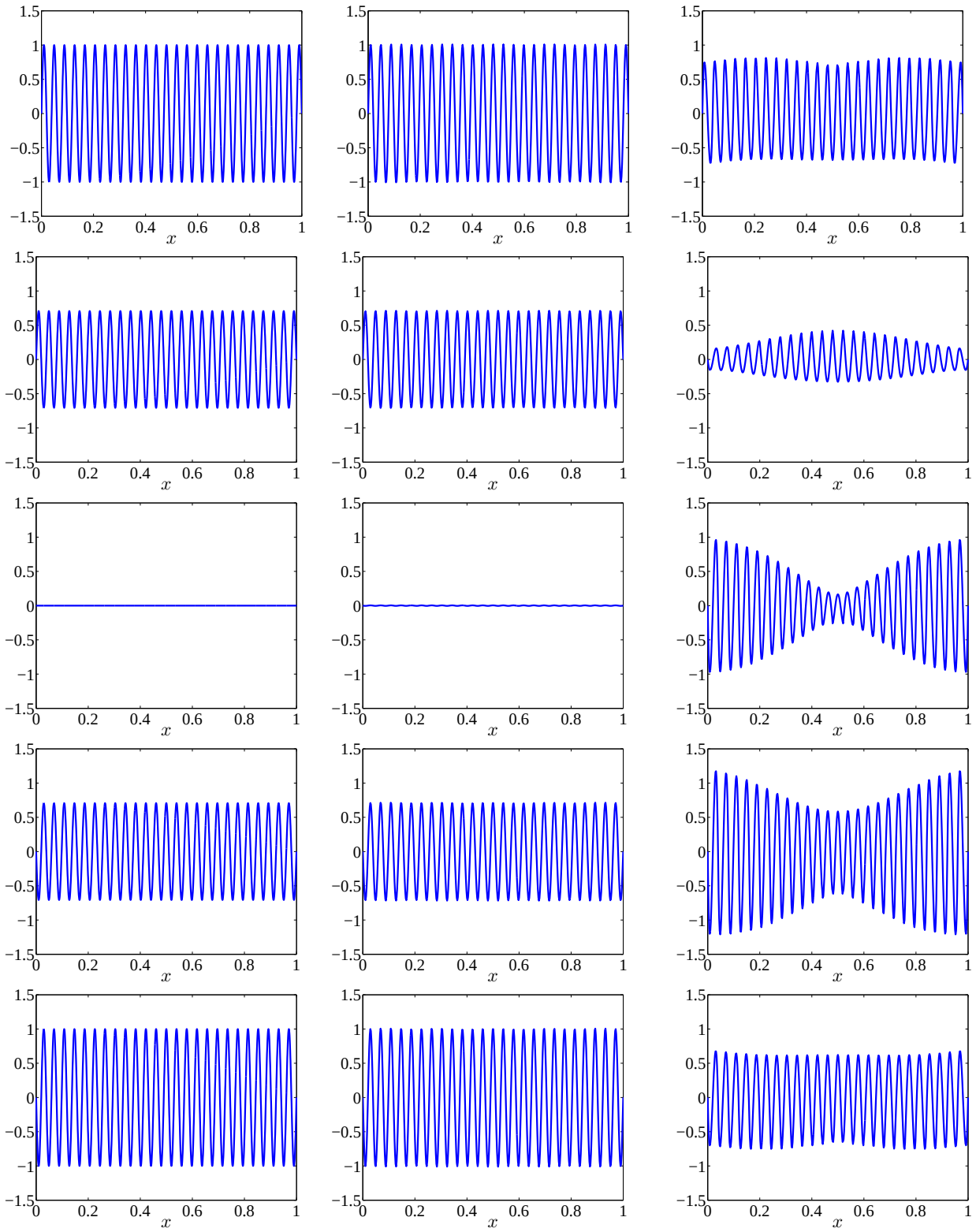


Figure 18: Plots of the exact and numerical solutions to the hyperbolic initial-value problem with exact solution $u(x,t) = \sin(51\pi x)\cos(51\pi t)$ over the eleventh half-cycle $t \in [10/51, 11/51]$. The total number of discrete modes for the displayed numerical solutions is $N = 99$. Snapshots from top to bottom correspond to times $t = 10/51, 10.25/51, 10.5/51, 10.75/51,$ and $11/51$ respectively.

We have conducted a collection of other numerical tests, and our conclusion has always been the same: if any of the "spurious" finite element eigenfunctions corresponding to spikes in the error spectrum participate in the numerical solution of a hyperbolic initial-value problem, then the corresponding finite element approximation has no qualitative or quantitative accuracy. On the other hand, NURBS approximations of hyperbolic initial-value problems are well-behaved in all modes (except a finite number of "outlier" modes) and are limited only by dispersion error for sufficiently smooth solutions. Additionally, NURBS approximations exhibit much less dispersion error than finite element approximations, even for the lowest modes of the spectrum (recall Figures 5-7). We simply conclude that, for the same polynomial order and an equal number of degrees-of-freedom, NURBS approximations of hyperbolic initial-value problems are much more robust and accurate than finite element approximations.

7 Conclusions

We have studied the accuracy of finite elements and NURBS approximations to the elliptic eigenvalue problem, and the implications of these results to the corresponding elliptic boundary-value problem and the parabolic and hyperbolic initial-value problems. One observation made herein is that the errors in the solutions to these problems can be expressed entirely in terms of the eigenfunction and eigenvalue errors. The inaccuracy of the higher modal eigenvalues in finite element approximations has been known for quite some time. The results herein have revealed a new phenomenon: There are large error spikes in the L^2-norms of the eigenfunction errors centered about the transitions between branches of the finite element eigenvalue spectrum. The squares of the energy-norm eigenfunction errors are sums of squares of L^2-norm eigenfunction errors and eigenvalue errors, which exhibit thereby the worst behavior of both. The NURBS errors are better behaved in every respect. The L^2-norms of the eigenfunction errors are indistinguishable from the L^2 best approximation errors of the eigenfunctions. The eigenvalue spectra do not have optical branches and, with the exception of a few "outlier modes", exhibit convergence of almost the entire spectrum. The upshot is that this is a demonstrably serious problem in the hyperbolic case, much less so in the parabolic case, and not a significant concern for the elliptic boundary-value problem. The situation is much more positive for NURBS elements in all cases.

Acknowledgements

Thomas J.R. Hughes and John A. Evans were partially supported by the Office of Naval Research under contract number N00014-08-0992. John A. Evans was additionally partially supported by the Department of Energy Computational Science Graduate Fellowship, provided under grant number DE-FG02-97ER25308, and the ICES Postdoctoral Fellowship at the Institute for Computational Engineering and Sciences. Alessandro Reali was partially supported by the European Research Council through the FP7 Ideas Starting Grant no. 259229 "ISOBIO". This support is gratefully acknowledged.

References

[1] Abaqus Unified FEA. `http://www.3ds.com/products/simulia/portfolio/abaqus`. Online; accessed 30-May-2013.

[2] M Ainsworth. Discrete dispersion relation for *hp*-version finite element approximation at high wave-number. *SIAM Journal on Numerical Analysis*, 42:553–575, 2004.

[3] I Akkerman, Y Bazilevs, V M Calo, T J R Hughes, and S Hulshoff. The role of continuity in residual-based variational multiscale modeling of turbulence. *Computational Mechanics*, 41:371–378, 2008.

[4] Y Bazilevs, L Beirao da Veiga, J A Cottrell, T J R Hughes, and G Sangalli. Isogeometric analysis: Approximation, stability and error estimates for *h*-refined meshes. *Mathematical Models and Methods in Applied Sciences*, 16:1–60, 2006.

[5] J. Chung and G. M. Hulbert. A time integration algorithm for structural dynamics with improved numerical dissipation: The generalized-α method. *Journal of Applied Mechanics*, 60:371–75, 1993.

[6] J A Cottrell, T J R Hughes, and Y Bazilevs. *Isogeometric Analysis: Toward Integration of CAD and FEA*. John Wiley and Sons, 2009.

[7] J A Cottrell, A Reali, Y Bazilevs, and T J R Hughes. Isogeometric analysis of structural vibrations. *Computer Methods in Applied Mechanics and Engineering*, 195:5257–5296, 2006.

[8] J A Evans, Y Bazilevs, I Babuška, and T J R Hughes. *n*-widths, sup-infs, and optimality ratios for the *k*-version of the isogeometric finite element method. *Computer Methods in Applied Mechanics and Engineering*, 198:1726–1741, 2009.

[9] J A Evans and T J R Hughes. Discrete spectrum analyses for various mixed discretizations of the Stokes eigenproblem. *Computational Mechanics*, 50:667–674, 2012.

[10] L C Evans. *Partial Differential Equations*. American Mathematical Society, Providence, Rhode Island, 1998.

[11] H Hilber, T J R Hughes, and R Taylor. Improved numerical dissipation for time integration algorithms in structural dynamics. *Earthquake Engineering and Structural Dynamics*, 5:283–292, 1977.

[12] T J R Hughes. *The Finite Element Method: Linear Static and Dynamic Finite Element Analysis*. Dover Publications, Inc., New York, 2000.

[13] T J R Hughes, A Reali, and G Sangalli. Duality and unified analysis of discrete approximations in structural dynamics and wave propagation: Comparison of *p*-method finite elements with *k*-method NURBS. *Computer Methods in Applied Mechanics and Engineering*, 197:4104–4124, 2008.

[14] LS-DYNA. `http://www.lstc.com/products/ls-dyna`. Online; accessed 30-May-2013.

[15] A Reali. An isogeometric analysis approach for the study of structural vibrations. Master's thesis, University of Pavia, Pavia, Italy, 2004.

[16] R D Richtmyer and K W Morton. *Difference Methods for Initial-value Problems*. John Wiley and Sons, second edition, 1967.

[17] G Strang and G Fix. *An Analysis of the Finite Element Method*. Wellesley-Cambridge Press, second edition, 2008.

[18] L T Thompson and P M Pinsky. Complex wavenumber Fcourier analysis of the p-version finite element method. *Computational Mechanics*, 13:255–275, 1994.

A Analytical Computation of Eigenvalue Errors

In this appendix, we analytically compute the eigenvalue errors for the one-dimensional Laplace eigenproblem subject to homogeneous Dirichlet boundary conditions and discretized by B-splines and finite elements. We begin by deriving an expression for the eigenvalue error for linear approximations (where the two approaches coincide), and we then extend our results to higher-order B-spline and finite element approximations.

A.1 Eigenvalue Errors for Linear Approximations

In what follows, consider a uniform mesh $0 = x_0 < x_1 < \dots x_A < \dots < x_{N+1} = 1$, where the number of elements is $n_{el} = N + 1$ and the mesh-size is $h = 1/n_{el}$. The system obtained by discretizing the homogeneous Laplace eigenproblem by linear finite elements on the uniform mesh is

$$\frac{\lambda^h h}{6}(u_{A-1} + 4u_A + u_{A+1}) + \frac{1}{h}(u_{A-1} - 2u_A + u_{A+1}) = 0, \qquad A = 1, \dots, N \tag{109}$$

$$u_0 = u_{N+1} = 0 \tag{110}$$

where N is the total number of degrees-of-freedom, $u_A = u^h(x_A)$ is the nodal value of the discrete normal mode u^h at node x_A, and λ^h is the eigenvalue corresponding to mode u^h. We write (109) in the compact form

$$\left(\frac{\lambda^h h^2}{c} + \beta\right) u_A = 0, \qquad A = 1, \dots, N \tag{111}$$

where c is a constant and α and β are operators, defined in this case to be

$$c = 6, \tag{112}$$

$$\alpha u_A = u_{A-1} + 4u_A + u_{A+1} \tag{113}$$

$$\beta u_A = u_{A-1} - 2u_A + u_{A+1} \tag{114}$$

Assuming a sinusoidal discrete normal mode which satisfies $u_A = C\sin(\omega x_A)$ for $A = 1, \dots, N$ and C a normalization constant (see Hughes [12], Chapter 9), we obtain

$$\left(\frac{(\omega^h h)^2}{c}\alpha + \beta\right) \sin(A\omega h) = 0, \qquad A = 1, \dots, N \tag{115}$$

where $\omega = \sqrt{\lambda}$ and $\omega^h = \sqrt{\lambda^h}$ are the exact and discrete eigenfrequencies, respectively. Substituting expressions (112)-(114) for c, α, and β and using the trigonometric identity $\sin(a+b) + \sin(a-b) = 2\sin(a)\sin(b)$ yields

$$\frac{(\omega^h h)^2}{6}(2 + \cos(\omega h)) - 1 + \cos(\omega h) = 0 \tag{116}$$

The eigenvalue error $\dfrac{\lambda^h - \lambda}{\lambda} = \left(\dfrac{\omega^h}{\omega}\right)^2 - 1$ is then easily expressed as

$$\frac{\lambda^h - \lambda}{\lambda} = \frac{6(1 - \cos(\omega h))}{(\omega h)^2(2 + \cos(\omega h))} - 1 \tag{117}$$

This formula can now be evaluated to obtain the induced error in approximating any eigenvalue $\lambda_l = (\omega_l)^2 = \pi^2 l^2$ using linear finite elements.

A.2 Eigenvalue Errors for Higher-Order B-spline Approximations

The computations carried out above for linear elements can largely be repeated for C^{p-1}-continuous B-spline approximations of degree $p > 1$. This is due to the fact that smooth B-spline approximations inherit a uniform stencil equation for all control points except the first and last p points when defined on a uniform mesh. For example, the system obtained by discretizing the homogeneous Laplace eigenproblem by C^1-continuous quadratic B-splines on the uniform mesh $0 = x_0 < x_1 < \ldots x_A < \ldots < x_N = 1$ is

$$\frac{\lambda^h h}{120}(14u_0 + 40u_1 + 25u_2 + u_3) + \frac{1}{6h}(6u_0 - 8u_1 + u_2 + u_3) = 0 \tag{118}$$

$$\frac{\lambda^h h}{120}(2u_0 + 25u_1 + 66u_2 + 26u_3 + u_4) + \frac{1}{6h}(2u_0 + u_1 - 6u_2 + 2u_3 + u_4) = 0 \tag{119}$$

$$\begin{aligned}\frac{\lambda^h h}{120}&(u_{A-2} + 26u_{A-1} + 66u_A + 26u_{A+1} + u_{A+2}) \\ &+\frac{1}{6h}(u_{A-2} + 2u_{A-1} - 6u_A + 2u_{A+1} + u_{A+2}) = 0, \qquad A = 3, \ldots, N-2\end{aligned} \tag{120}$$

$$\begin{aligned}\frac{\lambda^h h}{120}&(u_{N-3} + 26u_{N-2} + 66u_{N-1} + 25u_N + 2u_{N+1}) \\ &+\frac{1}{6h}(u_{N-3} + 2u_{N-2} - 6u_{N-1} + u_N + 2u_{N+1}) = 0\end{aligned} \tag{121}$$

$$\begin{aligned}\frac{\lambda^h h}{120}&(u_{N-2} + 25u_{N-1} + 40u_N + 14u_{N+1}) \\ &+\frac{1}{6h}(u_{N-2} + u_{N-1} - 8u_N + 6u_{N+1}) = 0\end{aligned} \tag{122}$$

$$u_0 = u_{N+1} = 0 \tag{123}$$

where $h = 1/n_{el}$ and the u_A's are the control variables of mode u^h. We refer to (118), (119), (121), and (122) as the boundary stencil equations and (120) as the interior stencil equations. In the following, rather than find the exact eigenpair (u^h, λ^h) which satisfies both the boundary and interior stencil equations, we seek an admissible normal mode and associated eigenvalue which only satisfy the interior stencil equations and homogeneous Dirichlet boundary conditions. We expect that this approximate eigenpair will converge to the exact discrete eigenpair as $N \to \infty$.[6] If we assume that the admissible normal mode takes the form

$$u_A = C \sin\left(\omega h \left(A - 1/2\right)\right), \qquad A = 1, \ldots, N \tag{124}$$

$$u_0 = u_{N+1} = 0 \tag{125}$$

where $\omega = \sqrt{\lambda}$ is the exact eigenfrequency, we obtain the simplified interior stencil equations

$$\left(\frac{(\omega^h h)^2}{c}\alpha + \beta\right)\sin((A - 1/2)\,\omega h) = 0, \qquad A = 3, \ldots, N-2 \tag{126}$$

[6] It should be noted that the admissible normal mode and its associated eigenvalue will not converge to the exact discrete eigenpair (u^h, λ^h) if the discrete normal mode u^h is an outlier mode. However, there are no outlier modes in the quadratic setting.

where $\omega^h = \sqrt{\lambda^h}$ is the discrete eigenfrequency and

$$c = 20 \tag{127}$$

$$\alpha u_A = u_{A-2} + 26u_{A-1} + 66u_A + 26u_{A+1} + u_{A+2} \tag{128}$$

$$\beta u_A = u_{A-2} + 2u_{A-1} - 6u_A + 2u_{A+1} + u_{A+2} \tag{129}$$

Invoking the trigonometric identity $\sin(a+b) + \sin(a-b) = 2\sin(a)\sin(b)$ in (126) yields the following expression for the eigenvalue error:

$$\frac{\lambda^h - \lambda}{\lambda} = \frac{20(3 - 2\cos(\omega h) - \cos(2\omega h))}{(\omega h)^2(33 + 26\cos(\omega h) + \cos(2\omega h))} - 1 \tag{130}$$

Note that the above expression is valid only for the admissible normal mode and not the exact discrete normal mode. Nonetheless, we expect that it yields a correct expression for the eigenvalue error in the limit of mesh refinement.

We now repeat the above procedure for cubic and quartic B-splines, but we omit details for brevity. The interior stencil equations for a uniform C^2-continuous cubic B-spline approximation take the form

$$\left(\frac{\lambda^h h^2}{c} + \beta\right) u_A = 0, \qquad A = 4, \ldots, N-3 \tag{131}$$

where

$$c = 42 \tag{132}$$

$$\alpha u_A = u_{A-3} + 120u_{A-2} + 1191u_{A-1} + 2416u_A + 1191u_{A+1} + 120u_{A+2} + u_{A+3} \tag{133}$$

$$\beta u_A = u_{A-3} + 24u_{A-2} + 15u_{A-1} - 80u_A + 15u_{A+1} + 24u_{A+2} + u_{A+3} \tag{134}$$

If we assume a sinusoidal admissible normal mode and follow the same steps as in the quadratic case, we arrive at the following expression for the eigenvalue error:

$$\frac{\lambda^h - \lambda}{\lambda} = \frac{42(40 - 15\cos(\omega h) - 24\cos(2\omega h) - \cos(3\omega h))}{(\omega h)^2(1208 + 1191\cos(\omega h) + 120\cos(2\omega h) + \cos(3\omega h))} - 1 \tag{135}$$

The interior stencil equations for a uniform C^3-continuous quartic B-spline approximation are

$$\left(\frac{\lambda^h h^2}{c} + \beta\right) u_A = 0, \qquad A = 5, \ldots, N-4 \tag{136}$$

where

$$c = 72 \tag{137}$$

$$\begin{aligned}\alpha u_A = {} & u_{A-4} + 502u_{A-3} + 14608u_{A-2} + 88234u_{A-1} + 156190u_A \\ & + 88234u_{A+1} + 14608u_{A+2} + 502u_{A+3} + u_{A+4}\end{aligned} \tag{138}$$

$$\begin{aligned}\beta u_A = {} & u_{A-4} + 118u_{A-3} + 952u_{A-2} + 154u_{A-1} - 2450u_A \\ & + 154u_{A+1} + 952u_{A+2} + 118u_{A+3} + u_{A+4}\end{aligned} \tag{139}$$

and assuming a sinusoidal admissible normal yields the following expression for the eigenvalue error:

$$\frac{\lambda^h - \lambda}{\lambda} = \frac{72(1225 - 154\cos(\omega h) - 952\cos(2\omega h) - 118\cos(3\omega h) - \cos(4\omega h))}{(\omega h)^2(78095 + 88234\cos(\omega h) + 14608\cos(2\omega h) + 502\cos(3\omega h) + \cos(4\omega h))} - 1 \tag{140}$$

	$p=2$	$p=3$	$p=4$
C^{p-1}	$1.382{\cdot}10^{-5}$	$3.338{\cdot}10^{-8}$	$8.542{\cdot}10^{-11}$
C^0	$2.121{\cdot}10^{-4}$	$6.740{\cdot}10^{-6}$	$2.347{\cdot}10^{-7}$
C^0 to C^{p-1} ratio	$1.535{\cdot}10^{1}$	$2.019{\cdot}10^{2}$	$2.748{\cdot}10^{3}$

Table 1: Analytically computed eigenvalue errors at 10% of the spectrum (i.e., for $l/N = 1/10$) for C^{p-1}-continuous B-spline and C^0-continuous finite element approximations of degree $p = 2, \ldots, 4$.

In the first row of Table 1, we have reported the analytically computed eigenvalue errors at 10% of the spectrum (i.e., for $l/N = 1/10$) for C^{p-1}-continuous B-spline approximations of degree $p = 2, \ldots, 4$ in the limit $N \to \infty$. It should be repeated that these errors correspond to the admissible normal modes and not the exact discrete normal modes, but we expect these errors to coincide in the limit of mesh refinement.

A.3 Eigenvalue Errors for Higher-Order Finite Element Approximations

The extension of our previous results to C^0-continuous finite elements of degree $p > 1$ is a delicate procedure as multiple stencils are present even away from the domain boundaries. Namely, there is a stencil associated with the "element edges" and a set of $p-1$ stencils associated with the internal nodes (i.e., the nodes internal to an element). In the context of dispersion analysis, this issue has been circumvented by means of a static condensation procedure which solves for the internal "bubble" degrees of freedom in terms of the edge nodes. By employing this technique, Ainsworth [2] derived a collection of discrete dispersion relations for higher-order finite elements of the form $\cos(k^h h) = f(kh)$ where $h = 1/n_{el}$ and k and k^h are the exact and discrete wave numbers, respectively. In virtue of the duality principle between dispersion and spectrum analysis (see [13]), we equivalently have the relationship $\cos(\omega h) = f(\omega^h h)$ where ω and ω^h are the exact and discrete eigenfrequencies. For quadratic finite elements, this relationship takes the form

$$\cos(\omega h) = \frac{3(\omega^h h)^4 - 104(\omega^h h)^2 + 240}{(\omega^h h)^4 + 16(\omega^h h)^2 + 240} \tag{141}$$

The above equation can be solved for the discrete eigenvalues $\lambda^h = (\omega^h)^2$, giving rise to the following two solutions for the eigenvalue error $\dfrac{\lambda^h - \lambda}{\lambda} = \left(\dfrac{\omega^h}{\omega}\right)^2 - 1$:

$$\frac{\lambda^h_{1,2} - \lambda}{\lambda} = \frac{52 + 8\cos(\omega h) \mp \sqrt{-176\cos^2(\omega h) + 1792\cos(\omega h) + 1984}}{(\omega h)^2(3 - \cos(\omega h))} - 1 \tag{142}$$

These two solutions represent two distinct branches of the eigenvalue spectrum. The first branch, the so-called acoustical branch, is valid for $\omega h \in [0, \pi]$, while the second branch, the optical branch, is valid for $\omega h \in [\pi, 2\pi]$. By substituting in the expression $\omega_l h = \pi l/n_{el}$ into (142) and noting which branch the eigenfrequency belongs to, one can evaluate the error for any discrete eigenvalue λ^h_l.

The relation between the exact and discrete eigenfrequencies in the cubic finite element setting is

$$\cos(\omega h) = \frac{-4(\omega^h h)^6 + 540(\omega^h h)^4 - 11520(\omega^h h)^2 + 25200}{(\omega^h h)^6 + 30(\omega^h h)^4 + 1080(\omega^h h)^2 + 25200} \tag{143}$$

For cubic finite elements, there are three eigenvalue branches (one acoustical and two optical, valid respectively for $\omega h \in [0, \pi]$, $\omega h \in [\pi, 2\pi]$, and $\omega h \in [2\pi, 3\pi]$) which correspond to the three solutions

$\lambda^h_{1,2,3}$ of the cubic equation:

$$(\cos(\omega h)+4)(\lambda^h h)^3+(30\cos(\omega h)-540)(\lambda^h h)^2+(1080\cos(\omega h)+1520)\lambda^h h+ \\ +25200(\cos(\omega h)-1)=0 \tag{144}$$

Similarly, there are four eigenvalue branches (one acoustical and three optical, valid respectively for $\omega h\in[0,\pi]$, $\omega h\in[\pi,2\pi]$, $\omega h\in[2\pi,3\pi]$, and $\omega h\in[3\pi,4\pi]$) for a quartic finite element approximation. These can be obtained by starting from the relation

$$\cos(\omega h)=\frac{5(\omega^h h)^8-1800(\omega^h h)^6+134064(\omega^h h)^4-2378880(\omega^h h)^2+5080320}{(\omega^h h)^8+48(\omega^h h)^6+3024(\omega^h h)^4+161280(\omega^h h)^2+5080320} \tag{145}$$

given in [2], replacing $\omega^h=\sqrt{\lambda^h}$, and then finding the four solutions $\lambda^h_{1,2,3,4}$ of the quartic equation:

$$(\cos(\omega h)-5)(\lambda^h h)^4+(48\cos(\omega h)+1800)(\lambda^h h)^3+(3024\cos(\omega h)-134064)(\lambda^h h)^2+ \\ +(161280\cos(\omega h)+2378880)\lambda^h h+5080320(\cos(\omega h)-1)=0 \tag{146}$$

In the second row of Table 1, we have reported the analytically computed eigenvalue errors at 10% of the spectrum (i.e., for $l/N=1/10$) for C^0-continuous finite elements of degree $p=2,\dots,4$ in the limit $N\to\infty$, and in the third row of the table, we have reported the ratio between the eigenvalue errors of the finite element and B-spline approximations.