The spectral-Galerkin approximation of nonlinear eigenvalue problems

Jing An a,b, Jie Shen c,d,*, Zhimin Zhang a,e

a Beijing Computational Science Research Center, Beijing 100193, China
b School of Mathematical Sciences, Guizhou Normal University, Guiyang 550025, China
c Fujian Provincial Key Laboratory on Mathematical Modeling and High Performance Scientific Computing and
School of Mathematical Science, Xiamen University, Xiamen 361005, China
d Department of Mathematics, Purdue University, West Lafayette, IN 47907, USA
e Department of Mathematics, Wayne State University, Detroit, MI 48202, USA

A R T I C L E   I N F O

Article history:
Received 3 June 2017
Received in revised form 20 February 2018
Accepted 19 April 2018
Available online 27 April 2018

Keywords:
Spectral-Galerkin approximation
Error estimation
Iteration algorithm
Nonlinear eigenvalue problems

A B S T R A C T

In this paper we present and analyze a polynomial spectral-Galerkin method for nonlinear elliptic eigenvalue problems of the form \(-\text{div}(AVu) + Vu + f(u^2)u = \lambda u, \|u\|_2 = 1\). We estimate errors of numerical eigenvalues and eigenfunctions. Spectral accuracy is proved under rectangular meshes and certain conditions of \(f\). In addition, we establish optimal error estimation of eigenvalues in some hypothetical conditions. Then we propose a simple iteration scheme to solve the underlying eigenvalue problem. Finally, we provide some numerical experiments to show the validity of the algorithm and the correctness of the theoretical results.

© 2018 IMACS. Published by Elsevier B.V. All rights reserved.

1. Introduction

Eigenvalue problems appear in many mathematical models for scientific and engineering applications, such as the calculation of the vibration modes of a mechanical structure in the framework of nonlinear elasticity, the Gross–Pitaevskii equation describing the steady states of Bose–Einstein condensates [16,2–4], and the Hartree–Fock and Kohn–Sham equations used to calculate ground state electronic structures of molecular systems in quantum chemistry and materials science [6,15,18].

However, most of the existing analysis for eigenvalue approximations are concerned with linear eigenvalue problems [1], and there are relatively few results concerning approximation of nonlinear eigenvalue problems [20,19,7,9–11,14], and most of them are based on finite element methods with an exception in [7,8] where an error estimate for Fourier spectral method to a periodic nonlinear eigenvalue problem is derived. To the best of our knowledge, there has no report on high order numerical methods for non-periodic nonlinear eigenvalue problems. Thus, the aim of this paper is to develop and analyze a spectral Galerkin method for a nonlinear elliptic eigenvalue problem. In particular, we shall extend the error estimates established in [7] for eigenvalues and eigenfunctions for periodic case with Fourier approximation to non-periodic case with polynomial approximation. We also describe an efficient implementation of the spectral-Galerkin method and present some numerical experiments to validate our analysis and to demonstrate the effectiveness of our algorithm.

* Corresponding author.
E-mail addresses: anjing@csrc.ac.cn, aj154@163.com (J. An), shen7@purdue.edu (J. Shen), zmzhang@csrc.ac.cn, zzhang@math.wayne.edu (Z. Zhang).

https://doi.org/10.1016/j.apnum.2018.04.012
0168-9274/© 2018 IMACS. Published by Elsevier B.V. All rights reserved.
The rest of this paper is organized as follows. In the next section, some preliminaries needed in this paper are presented. In §3, the error estimates of approximate eigenvalues and eigenfunctions are analyzed. In §4, we describe the details for an efficient implementation of the algorithm and we present several numerical experiments to demonstrate the accuracy and efficiency of our method. In addition, we give some concluding remarks.

2. Preliminaries

Following [7], we consider in this article a particular class of nonlinear eigenvalue problems arising in the study of variational models of the form

$$\inf\{E(v) : v \in X, \int_{\Omega} v^2 dx = 1\} , \tag{2.1}$$

where $X = H^1_0(\Omega)$ with $\Omega$ being a bounded domain, and the energy functional $E$ is of the form

$$E(v) = \frac{1}{2} a(v, v) + \frac{1}{2} \int_{\Omega} F(v^2(x)) dx \tag{2.2}$$

with

$$a(u, v) = \int_{\Omega} (A\nabla u) \cdot \nabla v dx + \int_{\Omega} Vu v dx. \tag{2.3}$$

We make the following assumptions:

1. $A(x)$ is symmetric, and $A \in (L^\infty(\Omega))^{d\times d}$; \tag{2.4}
2. $\exists \alpha > 0$ s.t. $\xi^T A(x) \xi \geq \alpha |\xi|^2$ for all $\xi \in \mathbb{R}^d$ and $x \in \Omega$; \tag{2.5}
3. $V \in L^2(\Omega)$; \tag{2.6}
4. $F \in C^1([0, +\infty), \mathbb{R}) \cap C^2((0, +\infty), \mathbb{R})$ and $F'' > 0$ on $(0, +\infty)$; \tag{2.7}
5. $\exists 0 \leq q < 2$ and $C \in \mathbb{R}_{+}$ s.t. $|F'(t)| \leq C(1 + t^q)$ $\forall t \geq 0$; \tag{2.8}
6. $F''(t)$, $F''(t)$ bounded in $[0, +\infty)$. \tag{2.9}

In order to simplify the notation, we let $f(t) = F'(t)$, and $\omega = v^2$. We can then reformulate (2.1) as

$$\inf\{\mathcal{G}(\omega) : \omega \geq 0, \sqrt{\omega} \in X, \int_{\Omega} \omega = 1\}, \tag{2.10}$$

where

$$\mathcal{G}(\omega) = \frac{1}{2} a(\sqrt{\omega}, \sqrt{\omega}) + \frac{1}{2} \int_{\Omega} F(\omega) dx.$$ 

It is shown in [7] that, under assumptions (2.4)-(2.8), (2.10) has a unique solution $\omega_0$ and (2.1) has exactly two solutions: $u = \sqrt{\omega_0}$ and $-u$. Moreover, $E$ is Gâteaux differentiable on $X$ for all $v \in X, E'(v) = A_v v$ with

$$A_v = -\text{div}(A\nabla v) + V + f(v^2)$$

being a self-adjoint operator on $L^2(\Omega)$. Thus, the function $u$ is a solution of the Euler–Lagrange equation

$$\langle A_u u - \lambda u, v \rangle_{X', X} = 0, \forall v \in X \tag{2.11}$$

for some $\lambda \in \mathbb{R}$.

We assume that $\{X_N\}$ is a sequence of approximation spaces for $X$ such that

$$\lim_{N \to \infty} \min_{v_N \in X_N} \|u - v_N\|_{H^1} = 0. \tag{2.12}$$

Then the discrete variational approximation of (2.1) is:

$$\inf\{E(v_N) : v_N \in X_N, \int_{\Omega} v_N^2 = 1\}. \tag{2.13}$$
Problem (2.13) has at least one minimizer \( u_N \), which satisfies
\[
(A_{uN}u_N - \lambda Nu_N, v_N)_{X',X} = 0, \quad \forall v_N \in X_N.
\]
for some \( \lambda_N \in \mathbb{R} \).

3. Error estimates

We will establish our main results in this section. We start with a basic error analysis for general approximation spaces, and then derive improved error estimates for the Legendre–Galerkin approximation.

3.1. Basic error analysis

Let \( u \) be the unique positive solution of (2.1) and let \( u_N \) be a minimizer of the discrete problem (2.13). Since if \( u_N \) is a minimizer of (2.13), so is \( -u_N \), we can assume that \( (u_N, u) \geq 0 \). We also introduce the bilinear form \( E'(u) \) defined on \( X \times X \) by
\[
(E'(u), w)_{X',X} = (A_{u} u, w)_{X',X} + 2 \int_{\Omega} f(u^2)u^2 vw.
\]

By using the same arguments as in the proof of Lemma 1 in [7], we can obtain the following Lemma:

**Lemma 3.1.** Under assumptions (2.4)–(2.9) and (2.12), there exist \( \beta > 0 \) and \( M \in \mathbb{R}_+ \) such that for all \( v \in X \),
\[
0 \leq (A_{u} - \lambda) v, v\rangle_{X',X} \leq M \|v\|_{H^1}^2, \\
\beta \|v\|_{H^1}^2 \leq (E'(u) - \lambda) v, v\rangle_{X',X} \leq M \|v\|_{H^1}^2.
\]

And there exists \( \gamma > 0 \) such that for all \( N > 0 \),
\[
\gamma \|u_N - u\|_{H^1}^2 \leq ((A_{u} - \lambda)(u_N - u), (u_N - u))_{X',X}.
\]

**Lemma 3.2.** Under assumptions (2.4)–(2.9) and (2.12), it holds that
\[
\lim_{N \to \infty} \|u_N - u\|_{H^1} = 0.
\]

**Proof.** Following [7], we can derive from the definition \( E(u) \) that
\[
E(u_N) - E(u) = \frac{1}{2} (A_{u} u_N, u_N)_{X',X} - \frac{1}{2} (A_{u} u, u)_{X',X} \\
+ \frac{1}{2} \int_{\Omega} (u_N^2 - u^2) - f(u^2)(u_N^2 - u^2) \\
= \frac{1}{2} ((A_{u} - \lambda)(u_N - u), (u_N - u))_{X',X} \\
+ \frac{1}{2} \int_{\Omega} (u_N^2 - f(u^2)(u_N^2 - u^2)).
\]

From (3.4) and the convexity of \( F \), we have
\[
E(u_N) - E(u) \geq \frac{\gamma}{2} \|u_N - u\|_{H^1}^2.
\]

Let \( \pi_N u \in X_N \) be such that
\[
\|u - \pi_N u\|_{H^1} = \min\{\|u - v_N\|_{H^1}, v_N \in X_N\}.
\]

We deduce from (2.12) that \( \pi_N u \) converges to \( u \) in \( X \) when \( N \to \infty \). The functional \( E \) being strongly continuous on \( X \), then when \( N \to \infty \) we have
\[
\|u_N - u\|_{H^1}^2 \leq \frac{2}{\gamma} (E(u_N) - E(u)) \leq \frac{2}{\gamma} (E(\pi_N u) - E(u)) \to 0.
\]

Thus, we have
\[
\lim_{N \to \infty} \|u_N - u\|_{H^1} = 0. \quad \square
\]
From Lemma 3.2 we know that there exist $N_1 > 0$ such that for all $N > N_1$,

$$
\|u_N\|_{H^1} \leq 2\|u\|_{H^1}, \quad \|u_N - u\|_{H^1} \leq \frac{1}{2}.
$$

(3.6)

Let $\Omega = \Omega_1 \cup \Omega_2$, and $\Omega_1 \cap \Omega_2 = \emptyset$, such that $\forall x \in \Omega_1, u(u_N - u) \geq 0$ and $\forall x \in \Omega_2, u(u_N - u) < 0$.

Since

$$
f(u_N^2) - f(u^2) = f'(\xi_N^2)(u_N^2 - u^2),
$$

where $\xi_N^2$ is between $u_N^2$ and $u^2$, and $f'(t)$ locally bounded in $[0, +\infty)$ and $f'(t) > 0$ on $(0, +\infty)$, then from (3.6) there exist non-negative constant $\alpha_N, \beta_N$ and $M > 0$ such that $\alpha_N < M, \beta_N < M$ and

$$
\int_{\Omega_1} 2f'(\xi_N^2)u^2(u(u_N - u)) = \alpha_N \int_{\Omega_1} u(u_N - u),
$$

$$
\int_{\Omega_2} 2f'(\xi_N^2)u^2(u(u_N - u)) = \beta_N \int_{\Omega_2} u(u_N - u).
$$

**Theorem 1.** Under assumptions (2.4)–(2.7), (2.9) and (2.12), it holds

$$
|\lambda_N - \lambda| \leq C(\|u_N - u\|_{H^1}^2 + \|u_N - u\|_{L^2}),
$$

(3.7)

$$
\|u_N - u\|_{H^1} \leq C \min_{\lambda_N \in \lambda_N} \|v_N - u\|_{H^1};
$$

(3.8)

In addition, if $\alpha_N \leq \beta_N$, we have

$$
|\lambda_N - \lambda| \leq C \|u_N - u\|_{H^1}^2,
$$

(3.9)

where $C$ is a constant independent of $N$.

**Proof.** We shall first prove (3.7).

Since $X_N \subset H^1_0(\Omega)$, we derive from (2.1), (2.11), (2.13) and (2.14) that $\lambda_N \geq \lambda$. On the other hand, by direct calculation, we have

$$
\lambda_N - \lambda = \langle A_N u_N, u_N \rangle_{X', X} - \langle A u, u \rangle_{X', X}
$$

$$
= (u_N - u) + \int_{\Omega} f(u_N^2)u_N^2 - \int_{\Omega} f(u^2)u^2
$$

$$
= \alpha(u_N - u, u_N - u) + 2\alpha(u, u_N - u) + \int_{\Omega} f(u_N^2)u_N^2 - \int_{\Omega} f(u^2)u^2
$$

$$
= \alpha(u_N - u, u_N - u) + 2\lambda \int_{\Omega} u(u_N - u) - 2\int_{\Omega} f(u^2)u(u_N - u)
$$

$$
+ \int_{\Omega} f(u_N^2)u_N^2 - \int_{\Omega} f(u^2)u^2
$$

$$
= \alpha(u_N - u, u_N - u) - \lambda \|u_N - u\|_{L^2}
$$

$$
- 2\int_{\Omega} f(u^2)u(u_N - u) + \int_{\Omega} f(u_N^2)u_N^2 - \int_{\Omega} f(u^2)u^2
$$

$$
= ((A_u - \lambda)(u_N - u), (u_N - u))_{X', X} + \int_{\Omega} u_N^2(f(u_N^2) - f(u^2)).
$$

We estimate below the two terms in the last line.

Since

$$
u_N^2(u_N^2 - u^2) = (2u^2 + 2uu_N + u_N^2)(u_N - u)^2 + 2u^2(u_N - u),
$$

we have
\[
\int_{\Omega} u_N^2 (f(u_N^2) - f(u^2))dx = \int_{\Omega} f'(\xi_N^2)(2u^2 + 2uu_N + u_N^2)(u_N - u)^2 dx \\
+ 2 \int_{\Omega} f'(\xi_N^2)u^2(u_N - u)dx.
\]

The above two terms can be estimated as follows:

\[
\int_{\Omega} f'(\xi_N^2)(2u^2 + 2uu_N + u_N^2)(u_N - u)^2 dx \\
\leq C \|(u^2 + u_N^2)(u_N - u)\|_{L^2} \|u_N - u\|_{L^2} \\
\leq C(\|u^2(u_N - u)\|_{L^2} + \|u_N^2(u_N - u)\|_{L^2}) \|u_N - u\|_{L^2} \\
\leq C(\|u\|^2_{L^6}\|u_N - u\|_{L^6} + \|u_N\|^2_{L^6}\|u_N - u\|_{L^6}) \|u_N - u\|_{L^2} \\
\leq C(c_6^2\|u\|^2_{H^1}\|u_N - u\|_{H^1} + 4c_6^2\|u\|^2_{H^1}\|u_N - u\|_{H^1}) \|u_N - u\|_{L^2} \\
\leq C\|u_N - u\|_{H^1},
\]

\[
2 \int_{\Omega} f'(\xi_N^2)u^3(u_N - u)dx \leq C \|u\|^3_{L^6}\|u_N - u\|_{L^2}
\]

\[
\leq C(c_6^3\|u\|^3_{H^1}) \|u_N - u\|_{L^2} \leq C\|u_N - u\|_{L^2}.
\]

Therefore, we have

\[
\int_{\Omega} u_N^2 (f(u_N^2) - f(u^2))dx \leq C(\|u_N - u\|^2_{H^1} + \|u_N - u\|_{L^2}).
\]

where \(c_6\) is the Sobolev constant in \(\|v\|_{L^6} \leq c_6\|v\|_{H^1}, \quad \forall v \in X\). Therefore, we obtain (3.7).

Next, we will evaluate the \(H^1\)-norm of the error \(u_N - v\). We first notice that for all \(v \in X_N\),

\[
\|u_N - u\|_{H^1} \leq \|u_N - v_N\|_{H^1} + \|v_N - u\|_{H^1}.
\]

From (3.3) of Lemma 3.1 we have

\[
\|u_N - v_N\|^2_{H^1} \leq \beta^{-1} \langle (E''(u) - \lambda)(u_N - v_N), (u_N - v_N) \rangle_{X',X} \\
= \beta^{-1} \langle (E''(u) - \lambda)(u - u), (u_N - v_N) \rangle_{X',X} \\
+ \beta^{-1} \langle (E''(u) - \lambda)(v_N - u), (u_N - v_N) \rangle_{X',X}.
\]

We proceed below in three steps.

**Step 1:** Estimation of \(\langle (E''(u) - \lambda)(u_N - u), (u_N - v_N) \rangle_{X',X}\).

Since

\[
\langle (E''(u) - \lambda)(u_N - u), (u_N - v_N) \rangle_{X',X} \\
= - \int_{\Omega} (f(u_N^2)u_N - f(u^2)u_N - 2f'(u^2)u^2(u_N - u))(u_N - v_N)dx \\
+ (\lambda_N - \lambda) \int_{\Omega} u_N(u_N - v_N)dx,
\]

we only need to estimate \(\int_{\Omega} (f(u_N^2)u_N - f(u^2)u_N - 2f'(u^2)u^2(u_N - u))(u_N - v_N)dx\) and \(\int_{\Omega} u_N(u_N - v_N)dx\).

For all \(v \in X_N\) such that \(\|v\|_{L^2} = 1\), we have

\[
\int_{\Omega} u_N(u_N - v_N)dx = 1 - \int_{\Omega} u_Nv_Ndx = \frac{1}{2}\|u_N - v_N\|^2_{L^2}.
\]

In addition, we have
\[
\int_{\Omega} (f(u_N^2) - f(u^2)u_N - 2f'(u^2)u_N^2(u_N - u))(u_N - v_N)dx \\
\leq \int_{\Omega} |f'(\xi_N^2)(u_N + u)u_N - 2f'(u^2)u_N| \cdot |u_N - v_N|dx \\
\leq \int_{\Omega} |(f'(\xi_N^2)(u_N + u)u_N - 2f'(\xi_N^2)u_N^2 + 2f'(\xi_N^2)u_N^2 - 2f'(u^2)u_N^2)| \cdot |u_N - v_N|dx \\
\leq \int_{\Omega} |(f'(\xi_N^2)(u_N - u)(u_N + 2u) - 2u^2(f'(\xi_N^2) - f'(u^2))(u_N - u)| \cdot |u_N - v_N|dx \\
\leq \int_{\Omega} \left| (f'(\xi_N^2)(u_N - u)(u_N + 2u) - 2u^2 f''(\xi_N^2)(\xi_N^2 - u^2))(u_N - u) \right| \cdot |u_N - v_N|dx \\
\leq \int_{\Omega} \left| (f'(\xi_N^2)(u_N + 2u)(u_N - u)) + |2u^2 f''(\xi_N^2)(\xi_N^2 - u^2)) \right| \cdot |u_N - u| \cdot |u_N - v_N|dx \\
\leq \int_{\Omega} \left| (f'(\xi_N^2)(u_N + 2u)(u_N - u)) + |2u^2 f''(\xi_N^2)(u_N^2 - u^2)) \right| \cdot |u_N - u| \cdot |u_N - v_N|dx \\
\leq \int_{\Omega} \left| (f'(\xi_N^2)(u_N + 2u) + |2u^2 f''(\xi_N^2)(u_N + u))) \cdot (u_N - u)^2 \right| \cdot |u_N - v_N|dx \\
\leq C \int_{\Omega} (|u_N + 2u| + |u^2(u_N + u)|) \cdot (u_N - u)^2 \cdot |u_N - v_N|dx,
\]

where \(\xi_N^2\) is between \(\xi_N^2\) and \(u^2\). Since for all \(N > N_1\) and all \(v_N \in X_N\),

\[
\int_{\Omega} |u_N + 2u| \cdot (u_N - u)^2 \cdot |u_N - v_N|dx \\
\leq \|u_N + 2u\|_{L^2} \|u_N - u\|^2 \|u_N - v_N\|_{L^2} \\
\leq (\|u_N\|_{L^2}^2 + 2\|u\|_{L^2}) \|u_N - u\|^2 \|u_N - v_N\|_{L^2} \\
\leq (C\|u_N\|_{L^2}^2 + 2\|u\|_{L^2}) \|u_N - u\|^2 \|u_N - v_N\|_{L^2} \\
\leq (Cc_6\|u_N\|_{H^1}^2 + 2\|u\|_{L^2})c_6 \|u_N - u\|^2 \|u_N - v_N\|_{H^1} \\
\leq (2Cc_6\|u\|_{H^1}^3 + 2\|u\|_{L^2})c_6 \|u_N - u\|^2 \|u_N - v_N\|_{H^1} \\
\leq C \|u_N - u\|^2 \|u_N - v_N\|_{H^1},
\]

we derive from (3.7) that
Therefore, and by direct calculations, we find

\[
\|(A\nu(u - v_N), u_N - v_N)_{X', X}\|
\]

\[
\leq |\int (\text{div}(A\nabla(u - v_N)) + V(u - v_N) + f(u^2)(u - v_N))(u_N - v_N)dx| \tag{2}
\]

\[
\leq |\int (A\nabla(u - v_N)\nabla(u_N - v_N) + V(u - v_N)(u_N - v_N) + f(u^2)(u - v_N))(u_N - v_N)dx| \tag{3}
\]

\[
\leq |A|_{L^\infty} \|\nabla(u - v_N)\|_{L^2} \|\nabla(u_N - v_N)\|_{L^2} + \int |V(u - v_N)(u_N - v_N)|dx \tag{4}
\]

\[
+ \int |f(u^2)(u - v_N)(u_N - v_N)|dx \tag{5}
\]

\[
\leq |A|_{L^\infty} \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1} + \|V\|_{L^2} \|(u - v_N)(u_N - v_N)\|_{L^2} \tag{6}
\]

\[
+ \|f(u^2)\|_{L^\infty} \|u - v_N\|_{L^2} \|u_N - v_N\|_{L^2} \tag{7}
\]

\[
\leq |A|_{L^\infty} \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1} + C \|V\|_{L^2} \|u - v_N\|_{L^2} \|u_N - v_N\|_{L^2} \tag{8}
\]

\[
+ \|f(u^2)\|_{L^\infty} \|u - v_N\|_{L^2} \|u_N - v_N\|_{L^2} \tag{9}
\]

\[
\leq \left( |A|_{L^\infty} + C \|V\|_{L^2} + \|f(u^2)\|_{L^\infty} \right) \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1} \tag{10}
\]

\[
\leq C \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1}. \tag{11}
\]

and

\[
2 \int |f'(u^2)u^2(u - v_N)(u_N - v_N)|dx + \lambda \int |(u - v_N)(u_N - v_N)|dx \tag{12}
\]

\[
\leq C \int |u^2(u - v_N)(u_N - v_N)|dx + \lambda \|u - v_N\|_{L^2} \|u_N - v_N\|_{L^2} \tag{13}
\]

\[
\leq C \|u\|_{L^6}^2 \left( \int (u - v_N)^2 \|u_N - v_N\|_{L^2} \right)^\frac{1}{2} \|u - v_N\|_{L^2} \|u_N - v_N\|_{L^2} \tag{14}
\]

\[
\leq C \|u\|_{L^6}^2 \|u - v_N\|_{L^2} \|u_N - v_N\|_{L^6} + \lambda \|u - v_N\|_{L^2} \|u_N - v_N\|_{L^2} \tag{15}
\]

\[
\leq C \|u\|_{H^1}^2 \|u - v_N\|_{L^2} \|u_N - v_N\|_{H^1} + \lambda \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1} \tag{16}
\]

\[
\leq C \left( C \|u\|_{H^1}^2 + \lambda \right) \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1} \tag{17}
\]

\[
\leq C \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1}. \tag{18}
\]

Therefore, we have

\[
\|(E''(u) - \lambda)(u - v_N), (u_N - v_N)\|_{X', X} \|
\]

\[
\leq |\|(A\nu(u - v_N), u_N - v_N)_{X', X}\| | + C \|u - v_N\|_{H^1} \|u_N - v_N\|_{H^1}. \tag{19}
\]
**Step 3:** Estimation of $\|u_N - u\|_{H^1}$.

From **Step 1** and **Step 2**, we have

$$
\|u_N - v_N\|_{H^1} \leq C(\|u_N - u\|_{H^1}^2 + \|u - v_N\|_{H^1}) + (\|u_N - u\|_{H^1} + \|u - N - u\|_{L^2})\|u_N - v_N\|_{H^1}.
$$

We derive from (3.5) and (3.6) that there exist $N_2 > N_1$ such that for all $N \geq N_2$,

$$
C(\|u_N - u\|_{H^1}^2 + \|u_N - u\|_{L^2}) \leq \gamma < \frac{1}{2},
$$

where $\gamma$ is a constant independent of $N$.

Thus, from (3.10) for all $N > N_2$ we have

$$
\|u_N - u\|_{H^1} \leq C \min_{v_N \in X_N, \|v_N\|_{L^2} = 1} \|v_N - u\|_{H^1}.
$$

We let $u_N^0$ be a minimizer of the following minimization problem

$$
\min_{v_N \in X_N} \|v_N - u\|_{H^1}.
$$

We know from (2.12) that $u_N^0$ converges to $u$ in $H^1$ when $N \to \infty$. In addition, we have

$$
\min_{v_N \in X_N, \|v_N\|_{L^2} = 1} \|v_N - u\|_{H^1} \leq \|\frac{u_N^0}{\|u_N^0\|_{L^2}} - u\|_{H^1}
$$

$$
\leq \|u_N^0 - u\|_{H^1} + \frac{\|u_N^0\|_{H^1}}{\|u_N^0\|_{L^2}}|1 - \|u_N^0\|_{L^2}|
$$

$$
\leq \|u_N^0 - u\|_{H^1} + \frac{\|u_N^0\|_{H^1}}{\|u_N^0\|_{L^2}}\|u_N^0 - u\|_{L^2}
$$

$$
\leq (1 + \frac{\|u_N^0\|_{H^1}}{\|u_N^0\|_{L^2}})\min_{v_N \in X_N} \|v_N - u\|_{H^1}.
$$

For $N > N_2 > N_1$, we have

$$
\|u_N^0 - u\|_{H^1} \leq \|u_N - u\|_{H^1} \leq \frac{1}{2},
$$

$$
\|u_N^0 - u\|_{L^2} \leq \|u_N - u\|_{H^1} \leq \frac{1}{2}.
$$

Then we have

$$
\|u_N^0\|_{H^1} \leq \|u_N^0 - u\|_{H^1} + \|u\|_{H^1} \leq \frac{1}{2} + \|u\|_{H^1}.
$$

Since

$$
1 = \|u\|_{L^2} \leq \|u_N^0 - u\|_{L^2} + \|u_N\|_{L^2} \leq \frac{1}{2} + \|u_N\|_{L^2},
$$

then we have

$$
\|u_N\|_{L^2} \geq \frac{1}{2}.
$$

Then we can get

$$
1 + \frac{\|u_N^0\|_{H^1}}{\|u_N^0\|_{L^2}} \leq 2(\|u\|_{H^1} + 1).
$$

Thus, (3.8) is proved.
Finally, if $\alpha_N \leq \beta_N$, then we have
\[
\int_\Omega 2f'(\xi_N^2)u^2(u(u_N - u)) = \int_\Omega 2f'(\xi_N^2)u^2(u(u_N - u)) + \int_\Omega 2f'(\xi_N^2)u^2(u(u_N - u))
\]
\[
= \alpha_N \int_\Omega u(u_N - u) + \beta_N \int_\Omega u(u_N - u)
\]
\[
\leq \beta_N(\int_\Omega u(u_N - u) + \int_\Omega u(u_N - u))
\]
\[
= \beta_N \int_\Omega u(u_N - u) = -\frac{1}{2} \beta_N \|u_N - u\|^2_{L^2}.
\]

Then from (3.2) in Lemma 3.1, we obtain (3.9). \(\square\)

**Remark 3.1.** The optimal eigenvalue error estimate (3.9) is proved without additional smoothness assumption but with the assumption $\alpha_N \leq \beta_N$ which is not easy to verify. However, the numerical results in §5 indicates that this optimal eigenvalue error estimate holds, at least for the tested cases.

### 3.2. Error estimates for Legendre–Galerkin method

The error estimates in Theorem 1 is proved under very general assumptions on the approximation space. In this subsection, we shall fix $\Omega = I^d$ ($d = 1, 2, 3$) with $I = (-1, 1)$, and consider the Legendre–Galerkin approximation with $X_N = P_N(I^d) \cap H^1_0(I^d)$ where $P_N$ stand for the set of all polynomials of at most degree $N$ in each direction. Then (2.12) is obviously satisfied.

We define the projection operator $\pi_N^{1,0} : H^1_0(I^d) \rightarrow X_N$ by
\[
(\nabla(\pi_N^{1,0} u - u), \nabla v) = 0, \quad v \in X_N,
\]
and recall the following result (cf. Remark 2.16 of [5]):

**Lemma 3.3.** Let $r \in \mathbb{N}_0$ with $r \geq 1$. If $f \in H^1_0(I^d) \cap H^r(I^d)$, then, for $N \geq 1$,
\[
\|f - \pi_N^{1,0} f\|_{H^1(I^d)} \leq C N^{-r+1} \|f\|_{H^r(I^d)},
\]
where $C$ is a constant independent of $N$.

Combining the above results with those in Theorem 1, we derive immediately the following:

**Theorem 2.** Under assumptions (2.4)–(2.7) and (2.9), if $u \in H^1_0(I^d) \cap H^m(I^d)$, then for $m \geq 1, N \geq 1$, we have
\[
|\lambda - \lambda_N| \leq C N^{-m+1} \|u\|_{H^m}, \hspace{1cm} (3.14)
\]
\[
\|u - u_N\|_{H^1} \leq C N^{-m+1} \|u\|_{H^m}; \hspace{1cm} (3.15)
\]

In addition, if $\alpha_N \leq \beta_N$, we have
\[
|\lambda - \lambda_N| \leq C N^{2(-m+1)}(\|u\|^2_{H^m} + \|u\|_{H^m}), \hspace{1cm} (3.16)
\]
where $C$ is a constant independent of $N$.

### 4. Implementation details and numerical results

In this section, we present an efficient implementation of the Legendre–Galerkin method for (2.14), and present some numerical experiments to validate our error analysis.
4.1. Implementation of the Legendre–Galerkin method

From equation (2.11) and the constraint \( \|u\|_{L^2}^2 = 1 \), we obtain an equivalent nonlinear eigenvalue problem
\[
- \text{div}(A \nabla u) + Vu + f(u^2)u = \lambda u, \quad \|u\|_{L^2}^2 = 1 \quad \text{in } \Omega,
\]
\( u\big|_{\partial \Omega} = 0. \) (4.1)

The weak form of (4.1) and (4.2) is: Find \((u, \lambda) \in X \times \mathbb{R}\) such that
\[
(A \nabla u, \nabla v) + (Vu, v) + (f(u^2)u, v) = \lambda (u, v), \quad \forall v \in X,
\]
\( \|u\|_{L^2}^2 = 1. \) (4.3)

Then the discrete form of (4.3) and (4.4) is: Find \((u_N, \lambda_N) \in X_N \times \mathbb{R}\) such that
\[
(A \nabla u_N, \nabla v_N) + (Vu_N, v_N) + (f(u_N^2)v_N, v_N) = \lambda_N (u_N, v_N), \quad \forall v_N \in X_N,
\]
\( \|u_N\|_{L^2}^2 = 1. \) (4.5)

To solve the above nonlinear eigenvalue problem, we use the Picard iteration as follows:
\[
(A \nabla u_N^p, \nabla v_N) + (Vu_N^p, v_N) + (f((u_N^{p-1})^2)u_N^p, v_N) = \lambda_N^p (u_N^p, v_N), \quad \forall v_N \in X_N,
\]
\( \|u_N^p\|_{L^2}^2 = 1. \) (4.7)

with initial guess determined by
\[
(A \nabla u_N^0, \nabla v_N) + (Vu_N^0, v_N) = \lambda_N^0 (u_N^0, v_N), \quad \forall v_N \in X_N,
\]
\( \|u_N^0\|_{L^2}^2 = 1. \) (4.9)

To simplify the presentation, we shall only consider the two-dimensional case although higher-dimensional case can be dealt with similarly. Let \( \phi_k = L_k(x) - L_{k+2}(x) \) \((k = 0, 1, \cdots, N-2)\), where \( L_k(x) \) denotes the Legendre polynomial of degree \( k \). Then, we have \( X_N = \text{span}\{\phi_i(x)\phi_j(y), i, j = 0, 1, \cdots, N-2\} \) [17].

We write
\[
u_N^p = \sum_{i,j=0}^{N-2} \nu_{ij}^p \phi_i(x)\phi_j(y). \] (4.11)

Then, we can reduce (4.7)–(4.8) and (4.9)–(4.10) to generalized eigenvalue problems:
\[
(\mathcal{S} + A(V) + M(\bar{u}^{p-1}))\bar{u}^p = \lambda_N^p \mathbb{B}\bar{u}^p, \quad p = 1, 2, \cdots, \] (4.12)

and
\[
(\mathcal{S} + A(V))\bar{u}^0 = \lambda_N^0 \mathbb{B}\bar{u}^0, \] (4.13)

respectively, where the stiff matrix \( \mathcal{S} \) and mass matrix \( \mathbb{B} \) are sparse [17]. The matrices \( M(V) \) and \( M(\bar{u}^{p-1}) \) are in general full but their matrix-vector product \((M(V))\bar{v}\) and \( M(\bar{u}^{p-1})\bar{v}\) can be efficiently computed by using a pseudo-spectral approach. Since only a few smallest eigenvalues are mostly interesting in real applications, it is most efficient to solve (4.12) (resp. (4.13)) using iterative eigen solvers such as shifted inverse power method (cf., for instance, [12]) which requires solving, repeatedly for different righthand side \( f \) (resp. \( \bar{f} \)),
\[
((\mathcal{S} + A(V) + M(\bar{u}^{p-1})) - \lambda_{ap}\mathbb{B})u_p = f \quad \text{(resp.} \quad (\mathcal{S} + A(V)) - \lambda_{a0}\mathbb{B})u_0 = \bar{f}, \] (4.14)

where \( \lambda_{ap} \) (resp. \( \lambda_{a0} \)) is some approximate value for the eigenvalue \( \lambda_N^p \) (resp. \( \lambda_N^0 \)). The above system can be efficiently solved by the Schur-complement approach, we refer to [13] for a detailed description on a related problem. In summary, the approximate eigenvalue problem (4.12) (resp. (4.13)) can be solved very efficiently.

4.2. Numerical experiments

We now perform some numerical tests to compute eigenvalues and eigenfunctions of (3.14)–(3.16). All numerical tests are performed using MATLAB 2015b.
Table 4.1
Numerical approximation to $\lambda_1$ for different $N$ and $L$ in 1D.

<table>
<thead>
<tr>
<th>$N$</th>
<th>$L = 10$</th>
<th>$L = 15$</th>
<th>$L = 20$</th>
<th>$L = 25$</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>2.038315754257193</td>
<td>2.038315754334226</td>
<td>2.038315754334225</td>
<td>2.038315754334222</td>
</tr>
<tr>
<td>15</td>
<td>2.038315716117307</td>
<td>2.038315716190542</td>
<td>2.038315716190543</td>
<td>2.038315716190548</td>
</tr>
<tr>
<td>20</td>
<td>2.038315716123191</td>
<td>2.038315716196435</td>
<td>2.038315716196433</td>
<td>2.038315716196430</td>
</tr>
<tr>
<td>25</td>
<td>2.038315716123191</td>
<td>2.038315716196430</td>
<td>2.038315716196430</td>
<td>2.038315716196434</td>
</tr>
</tbody>
</table>

Fig. 1. Errors between numerical solutions and the reference solution for $\lambda_1$ in 1D.

Fig. 2. The error figure of $\|u_{30} - u_{L,N,1}\|_2$ for different $N$ and $L$ in 1D.

4.2.1. One dimensional case

We take $A = \frac{1}{2}$, $V(x) = \frac{1}{2}x^2$, $f(u^2) = |u|^2$ and $\bar{\Omega} = [-1, 1]$. We take our first example. Numerical results for the first eigenvalue with different $N$ and iteration step $L$ are listed in Table 4.1. Since the exact eigen-pairs are unknown, we computed the reference solution, $\lambda_{L,N,1}$ and the associated numerical eigenfunction $u_{L,N,1}$, with $N = 40$ and iteration step $L = 30$.

We plot the error graphs of numerical approximations to the first eigenvalues $\lambda_1$ and the associated eigenfunction $u_1$ with different $N$ and $L$ in Fig. 1 and Fig. 3, respectively. In addition, in order to compare the convergence rates of $\|u_1 - u_{L,N,1}\|_{H^1}$ and $\|u_1 - u_{L,N,1}\|_2$, we also plot the error graphs of $\|u_{40,1} - u_{L,N,1}\|_2$ and $\|u_{40,1} - u_{L,N,1}\|_2/\|u_{40,1} - u_{L,N,1}\|_2$ for the first eigenvalue in Fig. 2, 4. We see from Table 4.1 and Fig. 1 that numerical eigenvalues achieve fifteen-digit accuracy with $N \geq 20$ and $L \geq 15$. We know from Fig. 1 that the numerical results are in agreement with the theoretical results, i.e., achieve spectral accuracy. From Fig. 4, the convergence rate of $\|u_1 - u_{L,N,1}\|_2$ is higher than that of $\|u_1 - u_{L,N,1}\|_2$. However, we know from Fig. 1 and Fig. 3 that the convergence rates of $|\lambda_{L,N,1} - \lambda_1|$ and $\|u_{L,N,1} - u_1\|_{H^1}$ are almost the same, which show that correctness of optimal error estimation in Theorem 1.
Fig. 3. The error figure of $|u_{40,1}^{30} - u_{N,1}^{L}|^2$ for different $N$ and $L$ in 1D.

Fig. 4. The error figure of $\|u_{40,1}^{30} - u_{N,1}^{L}\|_{H^1}^2 / \|u_{20,1}^{30} - u_{N,1}^{L}\|_{L^2}$ for different $N$ and $L$ in 1D.

Remark 4.1. For the non-periodic case, the numerical method in reference [7] is mainly based on the finite element discretization. The numerical method in this paper is based on spectral Galerkin approximation, thus, when the solution is smooth enough, the numerical solutions have spectral accuracy. In addition, we can observe from the numerical results in this paper that the numerical solutions have spectral accuracy. Compared with the numerical results in reference [7], the accuracy of the numerical solution in this paper is much higher in the same degree of freedom.

4.2.2. Two dimensional case

We take $A = \frac{1}{4} E$, $V(x_1, x_2) = \frac{1}{4}(x_1^2 + x_2^2)$, $f(u^2) = |u|^2$ and $\bar{\Omega} = [-1, 1]^2$ as our second example, where $E$ is the identity matrix. Numerical results for the first eigenvalue with different $N$ and iteration step $L$ are listed in Table 4.2. Since the exact eigen-pairs are unknown, we computed the reference solution, $\lambda_{N,1}^1$ and the associated numerical eigenfunction $u_{N,1}^1$, with $N = 40$ and iteration step $L = 30$.

We also plot the error graphs of numerical approximations to the first eigenvalue $\lambda_1$ and the associated eigenfunction $u_1$ with different $N$ and $L$ in Fig. 5 and Fig. 6, respectively. We see from Table 4.2 and Fig. 5 that numerical eigenvalues achieve at least thirteen-digit accuracy with $N \geq 15$ and $L \geq 12$. We know from Fig. 5 that the numerical results are in agreement with the theoretical results, i.e., achieve spectral accuracy. However, we know from Fig. 5 and Fig. 6 that the convergence rates of $|\lambda_{N,1}^1 - \lambda_1|$ and $\|u_{N,1}^1 - u_1\|_{H^1}$ are almost the same, which also show that correctness of optimal error estimation in Theorem 1. In order to further demonstrate the convergence of the approximate eigenfunctions, we plot the graphs of the first eigenfunction with $N = 15$ and $L = 18$ in Fig. 7 and the graphs of the first reference eigenfunction with $N = 40$ and $L = 30$ in Fig. 8, respectively.
4.3. Summary

We considered numerical approximations and error estimates for a nonlinear elliptic eigenvalue problem. Spectral accuracy error bounds are established for numerical eigenvalues and eigenfunctions. Numerical tests demonstrate that the method achieves high accuracy with relatively small number of unknowns. To simplify the analysis, we have restricted our analysis to rectangle and cubic domains. However, the approach presented in this paper can be extended to more general domains by using spectral-element method.
Fig. 7. The graphs of the first eigenfunction with $N = 15$ and $L = 18$ in 2D.

Fig. 8. The graphs of the first eigenfunction with $N = 40$ and $L = 30$ in 2D.

Acknowledgements

This work was supported in part by the National Natural Science Foundation of China (No. 11661022, 91630204, 11471031, 11371298, 11421110001, 51661135011), NASF U1530401, and the US National Science Foundation grant DMS-1419040.

References