Multigrid Method for Nonlinear Eigenvalue Problems Based on Newton Iteration

Fei Xu[1] · **Manting Xie**[2] · **Meiling Yue**[3]

Abstract

In this paper, a novel multigrid method based on Newton iteration is proposed to solve nonlinear eigenvalue problems. Instead of handling the eigenvalue λ and eigenfunction u separately, we treat the eigenpair (λ, u) as one element in a product space $\mathbb{R} \times H_0^1(\Omega)$. Then in the presented multigrid method, only one discrete linear boundary value problem needs to be solved for each level of the multigrid sequence. Because we avoid solving large-scale nonlinear eigenvalue problems directly, the overall efficiency is significantly improved. The optimal error estimate and linear computational complexity can be derived simultaneously. In addition, we also provide an improved multigrid method coupled with a mixing scheme to further guarantee the convergence and stability of the iteration scheme. More importantly, we prove convergence for the residuals after each iteration step. For nonlinear eigenvalue problems, such theoretical analysis is missing from the existing literatures on the mixing iteration scheme.

Keywords Multigrid method · Nonlinear eigenvalue problems · Newton iteration

1 Introduction

Solving large-scale nonlinear eigenvalue problems is a basic and challenging task in the field of scientific and engineering computing. Various practical problems reduce to solve

✉ Fei Xu
 xufei@lsec.cc.ac.cn

 Manting Xie
 mtxie@tju.edu.cn

 Meiling Yue
 yuemeiling@lsec.cc.ac.cn

1   Institute of Computational Mathematics, Department of Mathematics, Faculty of Science, Beijing University of Technology, Beijing 100124, China

2   Center for Applied Mathematics, Tianjin University, Tianjin 300072, China

3   School of Mathematics and Statistics, Beijing Technology and Business University, Beijing 100048, China

🖄 Springer

nonlinear eigenvalue problems. For example, the Schrödinger-Newton equation which has been widely used to model the quantum state reduction [19], the Gross-Pitaevskii equation which describes Bose-Einstein condensates [4], and the Kohn-Sham equation which is used to describe the electronic structure in quantum chemistry [10, 11]. In this paper, we will study a class of nonlinear eigenvalue models with convex energy functional. It is quite difficult to solve such nonlinear eigenvalue problems whose computational work grows exponentially as the problem size increases. Additionally, there are far fewer available numerical methods for nonlinear eigenvalue problems than those for the boundary value problems.

In this study, we resort to the multigrid method to improve the solving efficiency for nonlinear eigenvalue problems. The multigrid method was first proposed by Fedorenko [17] in 1961. Later, Brandt [6] demonstrated the efficiency of the multigrid method, which was then a widely examined subject in the computational mathematics field. Subsequently, Hackbusch, Xu, and many others began to use the tool of functional analysis to analyze this algorithm [18, 29, 31], which made the multigrid method experience rapid development. As we know, the multigrid method is able to derive an approximate solution possessing the optimal error estimates with the linear computational complexity. For eigenvalue problems, the study on the multigrid method is relatively limited. Xu and Zhou [32] first designed a two-grid method for the linear eigenvalue problem. This algorithm requires solving a small-scale eigenvalue problem on a coarse mesh and a large-scale boundary value problem on a fine mesh. When the mesh sizes of the coarse mesh (H) and fine mesh (h) have an appropriate proportional relation ($H = \sqrt{h}$), the optimal estimate for the approximate solution can be derived. Later, Chen and Liu et al. [13] extended the two-grid method to solve nonlinear eigenvalue problems using specially selected coarse and fine meshes. However, owing to the strict constrains on the ratio (i.e., $H = \sqrt{h}$), the two-grid method only performs on two mesh levels and cannot be used in the multigrid sequence.

To solve the nonlinear eigenvalue problems via the multigrid technique, the Newton iteration is adopted. In this paper, the eigenvalue λ and eigenfunction u are treated as one element in a product space $\mathbb{R} \times H_0^1(\Omega)$. Then, the nonlinear eigenvalue equation can be viewed as a special nonlinear equation defined in $\mathbb{R} \times H_0^1(\Omega)$. Next, based on the Newton iteration technique, we simply solve a linear boundary value problem in each layer of the multigrid sequence. Solving the large-scale nonlinear eigenvalue problem is avoided in our algorithm. Besides, the involved linearized boundary value problems can be solved quickly by many mature numerical algorithms, such as the classical multigrid method. Hence, the presented algorithm can significantly improve the solving efficiency for nonlinear eigenvalue problems. The well-posedness for these linearized boundary value problems and the convergence for our entire algorithm have been rigorously analyzed in this paper.

For the above-mentioned multigrid method, we find that the Newton iteration scheme may fail to converge for some complicated nonlinear eigenvalue models. In fact, this problem exists widely during the self-consistent field iteration for nonlinear eigenvalue problems. To derive a convergence result, the mixing scheme (or damping method) is introduced. The most widely used mixing scheme is Anderson acceleration, which is used to improve the convergence rate for fixed-point iterations. Anderson acceleration was first designed by D. G. Anderson in 1965 to solve the integral equation [2]. Subsequently, this technique was used to solve various models, such as, the self-consistent field iteration in electronic structure computations [16], flow problems [25, 27], molecular interaction [28], etc. Anderson acceleration adopts the mixing scheme to make the fixed-point iterations converge for general equations. Despite its widespread applications, the first mathematical convergence result for linear and nonlinear problems did not appear until 2015 in [30]. For nonlinear eigenvalue problems, there is still no strictly theoretical analysis. In general, the mixing scheme is used in fixed-point

iterations to accelerate the convergence rate. Thus, we use the mixing scheme in the Newton iteration scheme to design a more efficient multigrid method. By treating the nonlinear eigenvalue problem as a special nonlinear equation, we just solve a series of linearized boundary value problems derived from the Newton iteration scheme defined on a multigrid space sequence. Next, we mix the approximate solution in each iteration step to generate a more accurate approximation. More importantly, we can rigorously prove convergence for the residuals of the nonlinear eigenvalue problem after a one-time iteration step. This may provide some inspiration to prove the theoretical conclusions of the self-consistent field iteration for nonlinear eigenvalue problems.

The remainder of this paper is organized as follows. In Sect. 2, some preliminaries about the nonlinear eigenvalue problems and the corresponding finite element method are presented. In Sect. 3, we use the Newton iteration to solve the nonlinear eigenvalue problems and provide several useful error estimates. In Sect. 4, the multigrid method based on Newton iteration is introduced. In Sect. 5, we adopt the mixing scheme to provide an improved multigrid algorithm. Numerical experiments are presented in Sect. 6 to verify the theoretical results derived in this paper. Some concluding remarks are presented in the last section.

2 Preliminaries of nonlinear eigenvalue problems

In this section, we first introduce some standard notations for Sobolev space $W^{s,p}(\Omega)$ and the associated norm $\| \cdot \|_{s,p,\Omega}$ on a bounded domain $\Omega \in \mathbb{R}^d$ ($d \geq 1$). For $p = 2$, we denote $\| \cdot \|_{s,\Omega} = \| \cdot \|_{s,2,\Omega}$, $H^s(\Omega) = W^{s,2}(\Omega)$ and $H_0^1(\Omega) = \{v \in H^1(\Omega) : v|_{\partial\Omega} = 0\}$. Hereafter, we use $\lesssim, \gtrsim, \approx$ to denote $\leq, \geq, =$ with a mesh-independent coefficient for simplicity.

In this paper, we focus on the nonlinear eigenvalue problems arising from the following variational problem:

$$I = \inf\{E(v), v \in H_0^1(\Omega), \int_\Omega v^2 d\Omega = 1\}, \tag{1}$$

where the energy functional E is defined by

$$E(v) = \frac{1}{2}a(v, v) + \frac{1}{2}\int_\Omega F(v^2)d\Omega,$$

with

$$a(u, v) = \int_\Omega (A\nabla u \cdot \nabla v + Vuv)d\Omega.$$

To simplify the notation, let us define $f(t) = F'(t)$. Then, for all $v \in H_0^1(\Omega)$, $E'(v) = A_v v$ holds with

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

Next, applying the Lagrange multiplier method to (1) yields the following nonlinear eigenvalue problem:

$$\begin{cases} A_u u = \lambda u, & \text{in } \Omega, \\ u = 0, & \text{on } \partial\Omega, \\ \int_\Omega u^2 d\Omega = 1, \end{cases} \tag{3}$$

where λ comes from the Lagrange multiplier of the constraint $\|u\|_0^2 = 1$.

For future theoretical analysis, we assume that the following conditions hold true (see [10, 12]):

- $\mathcal{A}$ is a symmetric positive-definite matrix;
- $V \in L^p(\Omega)$ for some $p > \max(1, d/2)$;
- $F \in C^1([0, \infty), \mathbb{R}) \cap C^2((0, \infty), \mathbb{R})$ and $F'' > 0$ on $(0, \infty)$;
- $\exists\, 0 \leq m < 2$ and $C \in \mathbb{R}_+$, s.t. $\forall t \geq 0$, $|F'(t)| \leq C(1 + t^m)$;
- $F''(t)t$ is locally bounded in $[0, \infty)$;
- $\exists\, 1 < r \leq 2$ and $0 \leq s \leq 5-r$ such that $\forall a > 0$, $\exists\, C \in \mathbb{R}_+$, s.t. $\forall\, 0 < t_1 \leq a$, $\forall t_2 \in \mathbb{R}$,

$$|f(t_2^2)t_2 - f(t_1^2)t_1 - 2f'(t_1^2)t_1^2(t_2 - t_1)| \leq C(1 + |t_2|^s)|t_2 - t_1|^r;$$

- $\exists\, 0 \leq q < 1$ and $C \in \mathbb{R}_+$, s.t. $\forall t \geq 0$, $|f'(t)| + |f''(t)t| \leq C(1 + t^q)$.

In the remainder of this section, we present two lemmas for the nonlinear eigenvalue problem (3), and the detailed proof can be found in [10].

Lemma 1 *There exist $M \in \mathbb{R}_+$ and $\beta > 0$ such that for all $v \in H_0^1(\Omega)$, there holds*

$$0 \leq ((A_u - \lambda)v, v) \leq M\|v\|_1^2, \tag{4}$$

and

$$\beta\|v\|_1^2 \leq ((E''(u) - \lambda)v, v) \leq M\|v\|_1^2, \tag{5}$$

where $E''(u)$ has the following form

$$(E''(u)w, v) = (A_u w, v) + 2(f'(u^2)u^2 w, v). \tag{6}$$

Now, we introduce the finite element method for the nonlinear eigenvalue problem (3). To use the finite element method, we define the variational form for (3) as follows: Find $(\lambda, u) \in \mathbb{R} \times H_0^1(\Omega)$ such that $\|u\|_0 = 1$ and

$$(\mathcal{A}\nabla u, \nabla v) + (Vu + f(u^2)u, v) = \lambda(u, v), \quad \forall v \in H_0^1(\Omega). \tag{7}$$

Next, we define the finite element space V_h on a triangulation $\mathcal{T}_h$. The finite element space V_h is composed of piecewise polynomials such that $V_h \subset H_0^1(\Omega)$ and

$$\lim_{h \to 0} \inf_{v_h \in V_h} \|w - v_h\|_1 = 0, \quad \forall w \in H_0^1(\Omega). \tag{8}$$

Based on the finite element space V_h, we can derive the finite element solution for (7) by solving the discrete nonlinear eigenvalue problem as follows: Find $(\lambda_h, u_h) \in \mathbb{R} \times V_h$ such that $\|u_h\|_0 = 1$ and

$$(\mathcal{A}\nabla u_h, \nabla v_h) + (Vu_h + f(u_h^2)u_h, v_h) = \lambda_h(u_h, v_h), \quad \forall v_h \in V_h. \tag{9}$$

The standard error estimates for the finite element approximate eigenpair (λ_h, u_h) are described in the following lemma.

Lemma 2 *There exists $h_0 > 0$, such that for all $0 < h < h_0$, the following error estimates hold true*

$$\|u - u_h\|_1 \lesssim \delta_h(u), \tag{10}$$

and

$$|\lambda - \lambda_h| \lesssim \|u - u_h\|_1^2 + \|u - u_h\|_{L^{6/(5-2m)}}, \tag{11}$$

where

$$\delta_h(u) := \inf_{v_h \in V_h} \|u - v_h\|_1. \tag{12}$$

3 Newton iteration for nonlinear eigenvalue problems

To introduce Newton iteration for nonlinear eigenvalue problem (7), we first give some symbols to simplify the description. Let us define

$$\langle \mathcal{F}(\lambda, u), v \rangle = (\mathcal{A}\nabla u, \nabla v) + (Vu + f(u^2)u - \lambda u, v), \quad \forall v \in H_0^1(\Omega). \tag{13}$$

The Fréchet derivation of $\mathcal{F}$ with respect to u at (λ, u) is denoted by

$$\langle \mathcal{F}_u'(\lambda, u)w, v \rangle = (\mathcal{A}\nabla w, \nabla v) + ((V + f(u^2) - \lambda)w, v) + 2(f'(u^2)u^2 w, v)$$
$$= (E''(u)w, v) - \lambda(w, v). \tag{14}$$

Next, we denote the product space $\mathbb{R} \times H_0^1(\Omega)$ by X. Let us define the norms in X as

$$\|(\lambda, u)\|_X = |\lambda| + \|u\|_1 \quad \text{and} \quad \|(\lambda, u)\|_0 = |\lambda| + \|u\|_0, \quad \forall(\lambda, u) \in X. \tag{15}$$

Let us define

$$\langle \mathcal{G}(\lambda, u), (\mu, v) \rangle = \langle \mathcal{F}(\lambda, u), v \rangle + \frac{1}{2}\mu(1 - \int_\Omega u^2 d\Omega), \quad \forall(\mu, v) \in X. \tag{16}$$

Obviously, the nonlinear eigenvalue problem (7) is equivalent to the following nonlinear equation defined in the product space X: Find $(\lambda, u) \in X$ such that

$$\langle \mathcal{G}(\lambda, u), (\mu, v) \rangle = 0, \quad \forall(\mu, v) \in X. \tag{17}$$

Similarly, to solve (17) using the finite element method, we also introduce a finite dimensional space $X_h := \mathbb{R} \times V_h$ as the approximation for X. Then based on X_h, the finite element equation (9) is equivalent to the following nonlinear equation defined in the tensor product space X_h: Find $(\lambda_h, u_h) \in X_h$ such that

$$\langle \mathcal{G}(\lambda_h, u_h), (\mu, v_h) \rangle = 0, \quad \forall(\mu, v_h) \in X_h. \tag{18}$$

The Fréchet derivation of $\mathcal{G}$ at (λ, u) is denoted by:

$$\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, v) \rangle = \langle \mathcal{F}_u'(\lambda, u)w, v \rangle - \gamma(u, v) - \mu(u, w), \quad \forall(\mu, v) \in X. \tag{19}$$

Next, we begin to introduce the Newton iteration for (18). Given an initial value $(\lambda^{(0)}, u^{(0)}) \in X_h$, the one step of the Newton iteration is defined as follows, which will generate a more accurate approximate solution: Find $(\lambda^{(1)}, u^{(1)}) \in X_h$ such that

$$\langle \mathcal{G}'(\lambda^{(0)}, u^{(0)})(\lambda^{(1)} - \lambda^{(0)}, u^{(1)} - u^{(0)}), (\mu, v_h) \rangle = -\langle \mathcal{G}(\lambda^{(0)}, u^{(0)}), (\mu, v_h) \rangle,$$
$$\forall(\mu, v_h) \in X_h. \tag{20}$$

Based on the definitions (16) and (19), the linearized boundary value problem (20) can be rewritten as follows: Find $(\lambda^{(1)}, u^{(1)}) \in \mathbb{R} \times V_h$ such that for any $(\mu, v_h) \in \mathbb{R} \times V_h$, there holds

$$\begin{cases} a(\lambda^{(0)}, u^{(0)}; u^{(1)}, v_h) + b(u^{(0)}; v_h, \lambda^{(1)}) = \langle \mathcal{F}_u'(\lambda^{(0)}, u^{(0)})u^{(0)}, v_h \rangle \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad - \langle \mathcal{F}(\lambda^{(0)}, u^{(0)}), v_h \rangle - \lambda^{(0)}(u^{(0)}, v_h), \\ b(u^{(0)}; u^{(1)}, \mu) \qquad\qquad\qquad = -\mu/2 - \mu(u^{(0)}, u^{(0)})/2, \end{cases} \tag{21}$$

where

$$a(\lambda^{(0)}, u^{(0)}; u^{(1)}, v_h) = \langle \mathcal{F}'_u(\lambda^{(0)}, u^{(0)})u^{(1)}, v_h \rangle, \quad b(u^{(0)}; u^{(1)}, \mu) = -\mu(u^{(0)}, u^{(1)}).$$

To guarantee the well-posedness of (20) and (21), an isomorphism property is proved in Theorem 1:

Theorem 1 *Let (λ, u) be the exact solution of (17). When the mesh size h is sufficiently small, we can derive the following isomorphism properties for the operator $\mathcal{G}'$:*

$$\sup_{(\mu,v)\in X} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, v) \rangle}{\|(\mu, v)\|_X} \gtrsim \|(\gamma, w)\|_X, \quad \forall (\gamma, w) \in X, \tag{22}$$

and

$$\sup_{(\mu,v)\in X_h} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, v) \rangle}{\|(\mu, v)\|_X} \gtrsim \|(\gamma, w)\|_X, \quad \forall (\gamma, w) \in X_h. \tag{23}$$

For any $(\widehat{\lambda}, \widehat{u}) \in X$ such that $\|(\widehat{\lambda} - \lambda, \widehat{u} - u)\|_X$ is sufficiently small, we can also derive

$$\sup_{(\mu,v)\in X_h} \frac{\langle \mathcal{G}'(\widehat{\lambda}, \widehat{u})(\gamma, w), (\mu, v) \rangle}{\|(\mu, v)\|_X} \gtrsim \|(\gamma, w)\|_X, \quad \forall (\gamma, w) \in X_h. \tag{24}$$

Proof Firstly, proving (22) is equivalent to prove that for any $(\tau, f) \in \mathbb{R} \times H^{-1}(\Omega)$, the following equation is uniquely solvable: Find $(\gamma, w) \in \mathbb{R} \times H^1_0(\Omega)$ such that

$$\begin{cases} a(\lambda, u; w, v) + b(u; v, \gamma) = (f, v), & \forall v \in H^1_0(\Omega), \\ b(u; w, \mu) = \mu\tau, & \forall \mu \in \mathbb{R}. \end{cases} \tag{25}$$

According to the Brezzi theory [7], we just need to prove the following two conclusions:
(1) The following equation

$$a(\lambda, u; w, v) = (f, v), \quad \forall v \in V_0, \tag{26}$$

is uniquely solvable in $V_0 := \{v \in H^1_0(\Omega) : b(u; v, \mu) = 0, \ \forall \mu \in \mathbb{R}\}$.
(2) The following inf-sup condition holds true

$$\inf_{\mu\in\mathbb{R}} \sup_{v\in H^1_0(\Omega)} \frac{b(u; v, \mu)}{\|v\|_1 |\mu|} \geq k_b, \tag{27}$$

for some constant $k_b > 0$.

Firstly, from (5), we can easily find that (26) is uniquely solvable. Next, since $\|u\|_0 = 1$, taking $v = -\mu u$ leads to

$$\inf_{\mu\in\mathbb{R}} \sup_{v\in H^1_0(\Omega)} \frac{b(u; v, \mu)}{\|v\|_1 |\mu|} = \frac{1}{\|u\|_1} =: k_b.$$

Then we derive the inf-sup condition (27). Thus, we can derive the well-posedness of (25) based on Brezzi theory, which further yields (22).

Next, we begin to prove the second formula (23). Based on (5), let us define the projection $P_h : H^1_0(\Omega) \to V_h$ by

$$a(\lambda, u; w, v - P_h v) = 0, \quad \forall w \in V_h, \ \forall v \in H^1_0(\Omega). \tag{28}$$

Using (5) and (28), we can derive

$$\|P_h v\|_1 \lesssim \|v\|_1, \quad \forall v \in H^1_0(\Omega), \tag{29}$$

and

$$\|v - P_h v\|_0 \lesssim h\|v - P_h v\|_1 \lesssim h\|v\|_1, \quad \forall v \in H_0^1(\Omega). \tag{30}$$

Combining (22), (28), (29) and (30) leads to

$$
\begin{aligned}
\|(\gamma, w)\|_X &\lesssim \sup_{(\mu, v) \in X} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, v) \rangle}{\|(\mu, v)\|_X} \\
&= \sup_{(\mu, v) \in X} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (0, v - P_h v) \rangle + \langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, v)\|_X} \\
&= \sup_{(\mu, v) \in X} \frac{a(\lambda, u; w, v - P_h v) - \gamma(u, v - P_h v) + \langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, v)\|_X} \\
&= \sup_{(\mu, v) \in X} \frac{-\gamma(u, v - P_h v) + \langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, v)\|_X} \\
&\lesssim \sup_{(\mu, v) \in X} \frac{\gamma\|u\|_0\|v - P_h v\|_0 + \langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, v)\|_X} \\
&\lesssim \sup_{(\mu, v) \in X} \frac{\gamma h\|v\|_1 + \langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, v)\|_X} \\
&\lesssim h\|(\gamma, w)\|_X + \sup_{(\mu, v) \in X} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, v)\|_X}. \tag{31}
\end{aligned}
$$

Thus, when the mesh size h is small enough, we can obtain

$$
\begin{aligned}
\|(\gamma, w)\|_X &\lesssim \sup_{(\mu, v) \in X} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, v)\|_X} \\
&\lesssim \sup_{(\mu, v) \in X} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, P_h v) \rangle}{\|(\mu, P_h v)\|_X} \\
&\lesssim \sup_{(\mu, v) \in X_h} \frac{\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, v) \rangle}{\|(\mu, v)\|_X}, \tag{32}
\end{aligned}
$$

which is just the second formula (23). For the last formula (24), we first prove a Lipschitz continuity for $\mathcal{G}'$ based on the assumptions presented in Sect. 2, Hölder inequality and the imbedding theorem (see [1]):

$$
\begin{aligned}
&\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, v) \rangle - \langle \mathcal{G}'(\widehat{\lambda}, \widehat{u})(\gamma, w), (\mu, v) \rangle \\
&\quad = \left((f(u^2) - f(\widehat{u}^2))w, v \right) + ((\widehat{\lambda} - \lambda)w, v) \\
&\qquad + 2\left((f'(u^2)u^2 - f'(\widehat{u}^2)\widehat{u}^2)w, v \right) - \gamma(\widehat{u} - u, v) - \mu(\widehat{u} - u, w) \\
&\quad = 2\left(f'(\xi^2)\xi(u - \widehat{u})w, v \right) + ((\widehat{\lambda} - \lambda)w, v) \\
&\qquad + 4\left((f''(\xi^2)\xi^3 + f'(\xi^2)\xi)(u - \widehat{u})w, v \right) - \gamma(\widehat{u} - u, v) - \mu(\widehat{u} - u, w) \\
&\quad \lesssim \left((1 + \xi^{2q})\xi(u - \widehat{u})w, v \right) + ((\widehat{\lambda} - \lambda)w, v) \\
&\qquad + \left((1 + \xi^{2q})\xi(u - \widehat{u})w, v \right) - \gamma(\widehat{u} - u, v) - \mu(\widehat{u} - u, w) \\
&\quad \lesssim \|\xi\|_{0,4}\|u - \widehat{u}\|_{0,4}\|w\|_{0,4}\|v\|_{0,4} + \|\xi^{2q}\|_{0,3/q}\|\xi\|_{0,6}\|u - \widehat{u}\|_{0,6}\|w\|_{0,6}\|v\|_{0,6/(3-2q)} \\
&\qquad + |\widehat{\lambda} - \lambda|\|w\|_0\|v\|_0 + |\gamma|\|\widehat{u} - u\|_0\|v\|_0 + |\mu|\|\widehat{u} - u\|_0\|w\|_0 \\
&\quad \lesssim \|\xi\|_1\|u - \widehat{u}\|_1\|w\|_1\|v\|_1 + \|\xi\|_1^{2q}\|\xi\|_1\|u - \widehat{u}\|_1\|w\|_1\|v\|_1 \\
&\qquad + |\widehat{\lambda} - \lambda|\|w\|_0\|v\|_0 + |\gamma|\|\widehat{u} - u\|_0\|v\|_0 + |\mu|\|\widehat{u} - u\|_0\|w\|_0,
\end{aligned}
$$

where $\xi = \tau u + (1 - \tau)\widehat{u}$ with $\tau \in (0, 1)$.

Thus, we can derive

$$
\begin{aligned}
&\langle \mathcal{G}'(\lambda, u)(\gamma, w), (\mu, v)\rangle - \langle \mathcal{G}'(\widehat{\lambda}, \widehat{u})(\gamma, w), (\mu, v)\rangle \\
&\lesssim \|(\lambda - \widehat{\lambda}, u - \widehat{u})\|_X \|(\gamma, w)\|_X \|(\mu, v)\|_X.
\end{aligned} \tag{33}
$$

Let ε be a sufficiently small constant and $\|(\lambda - \widehat{\lambda}, u - \widehat{u})\|_X \leq \varepsilon$. Based on (23) and (33), for any $(\gamma, w) \in X_h$, we have

$$
\begin{aligned}
\|(\gamma, w)\|_X &\lesssim \sup_{(\mu, v) \in X_h} \frac{\varepsilon \|(\gamma, w)\|_X \|(\mu, v)\|_X + \langle \mathcal{G}'(\widehat{\lambda}, \widehat{u})(\gamma, w), (\mu, v)\rangle}{\|(\mu, v)\|_X} \\
&\lesssim \varepsilon \|(\gamma, w)\|_X + \sup_{(\mu, v) \in X_h} \frac{\langle \mathcal{G}'(\widehat{\lambda}, \widehat{u})(\gamma, w), (\mu, v)\rangle}{\|(\mu, v)\|_X}.
\end{aligned}
$$

Then (24) can be derived using the same way as the proof process in (32). This completes the proof.   □

Next, in Theorem 2, we provide the estimate for the residual of the Newton iteration scheme (20).

Theorem 2 *For any $(\lambda^{(0)}, u^{(0)}) \in X$, $(\lambda^{(1)}, u^{(1)}) \in X$, we can derive the following expansion*

$$
\begin{aligned}
\langle \mathcal{G}(\lambda^{(1)}, u^{(1)}), (\sigma, \eta)\rangle = {}&\langle \mathcal{G}(\lambda^{(0)}, u^{(0)}), (\sigma, \eta)\rangle + \langle \mathcal{G}'(\lambda^{(0)}, u^{(0)})(\lambda^{(1)} - \lambda^{(0)}, u^{(1)} - u^{(0)}), (\sigma, \eta)\rangle \\
&+ R\big((\lambda^{(0)}, u^{(0)}), (\lambda^{(1)}, u^{(1)}), (\sigma, \eta)\big), \qquad \forall (\sigma, \eta) \in X,
\end{aligned} \tag{34}
$$

where the residual $R\big((\lambda^{(0)}, u^{(0)}), (\lambda^{(1)}, u^{(1)}), (\sigma, \eta)\big)$ satisfies the following estimate

$$
\big| R\big((\lambda^{(0)}, u^{(0)}), (\lambda^{(1)}, u^{(1)}), (\sigma, \eta)\big) \big| \lesssim \|(\lambda^{(0)} - \lambda^{(1)}, u^{(0)} - u^{(1)})\|_X^2 \|(\sigma, \eta)\|_X.
$$

Proof We prove this theorem through the Taylor expansion. Let us define

$$
\eta(t) = \big\langle \mathcal{G}\big((\lambda^{(0)}, u^{(0)}) + t\{(\lambda^{(1)}, u^{(1)}) - (\lambda^{(0)}, u^{(0)})\}\big), (\sigma, \eta)\big\rangle, \quad t \in [0, 1]. \tag{35}
$$

Obviously,

$$
\eta(0) = \langle \mathcal{G}(\lambda^{(0)}, u^{(0)}), (\sigma, \eta)\rangle \quad \text{and} \quad \eta(1) = \langle \mathcal{G}(\lambda^{(1)}, u^{(1)}), (\sigma, \eta)\rangle. \tag{36}
$$

The first order derivative and the second order derivative of $\eta(t)$ with respect to t are listed as follows:

$$
\begin{aligned}
\eta'(t) = {}&\big\langle \mathcal{G}'((\lambda^{(0)}, u^{(0)}) + t\{(\lambda^{(1)}, u^{(1)}) - (\lambda^{(0)}, u^{(0)})\})((\lambda^{(1)}, u^{(1)}) - (\lambda^{(0)}, u^{(0)})), (\sigma, \eta)\big\rangle \\
= {}&\big\langle \mathcal{F}'_u((\lambda^{(0)}, u^{(0)}) + t\{(\lambda^{(1)}, u^{(1)}) - (\lambda^{(0)}, u^{(0)})\})(u^{(1)} - u^{(0)}), \eta\big\rangle \\
&- (\lambda^{(1)} - \lambda^{(0)})(u^{(0)} + t(u^{(1)} - u^{(0)}), \eta) - \sigma(u^{(0)} + t(u^{(1)} - u^{(0)}), u^{(1)} - u^{(0)}) \\
= {}&(\mathcal{A}\nabla(u^{(1)} - u^{(0)}), \nabla\eta) + \big((u^{(0)} + f((u^{(0)} + t(u^{(1)} - u^{(0)}))^2))(u^{(1)} - u^{(0)}), \eta\big) \\
&- \big((\lambda^{(0)} + t(\lambda^{(1)} - \lambda^{(0)}))(u^{(1)} - u^{(0)}), \eta\big) \\
&+ 2\big(f'((u^{(0)} + t(u^{(1)} - u^{(0)}))^2)(u^{(0)} + t(u^{(1)} - u^{(0)}))^2(u^{(1)} - u^{(0)}), \eta\big) \\
&- (\lambda^{(1)} - \lambda^{(0)})(u^{(0)} + t(u^{(1)} - u^{(0)}), \eta) - \sigma(u^{(0)} + t(u^{(1)} - u^{(0)}), u^{(1)} - u^{(0)}), \tag{37}
\end{aligned}
$$

and

$$
\begin{aligned}
\eta''(t) = {}&-2(\lambda^{(1)} - \lambda^{(0)})(u^{(1)} - u^{(0)}, \eta) - \sigma(u^{(1)} - u^{(0)}, u^{(1)} - u^{(0)}) \\
&+ 6\big(f'((u^{(0)} + t(u^{(1)} - u^{(0)}))^2)(u^{(0)} + t(u^{(1)} - u^{(0)}))(u^{(1)} - u^{(0)})^2, \eta\big) \\
&+ 4\big(f''((u^{(0)} + t(u^{(1)} - u^{(0)}))^2)(u^{(0)} + t(u^{(1)} - u^{(0)}))^3(u^{(1)} - u^{(0)})^2, \eta\big). \tag{38}
\end{aligned}
$$

For the nonlinear terms involved in (38), we have the following estimates

$$
\begin{aligned}
&\left|\left(f'((u^{(0)}+t(u^{(1)}-u^{(0)}))^2)(u^{(0)}+t(u^{(1)}-u^{(0)}))(u^{(1)}-u^{(0)})^2,\eta\right)\right| \\
&\quad\lesssim\left|\left((1+(u^{(0)}+t(u^{(1)}-u^{(0)}))^{2q})(u^{(0)}+t(u^{(1)}-u^{(0)}))(u^{(1)}-u^{(0)})^2,\eta\right)\right| \\
&\quad\lesssim\left|\left((u^{(0)}+t(u^{(1)}-u^{(0)}))(u^{(1)}-u^{(0)})^2,\eta\right)\right. \\
&\qquad+\left((u^{(0)}+t(u^{(1)}-u^{(0)}))^{2q}(u^{(0)}+t(u^{(1)}-u^{(0)}))(u^{(1)}-u^{(0)})^2,\eta\right)\Big| \\
&\quad\lesssim\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_{0,4}\|u^{(1)}-u^{(0)}\|_{0,4}^2\|\eta\|_{0,4} \\
&\qquad+\|(u^{(0)}+t(u^{(1)}-u^{(0)}))^{2q}\|_{0,3/q}\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_{0,6}\|u^{(1)}-u^{(0)}\|_{0,6}^2\|\eta\|_{0,6/(3-2q)} \\
&\quad\lesssim\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_{0,4}\|u^{(1)}-u^{(0)}\|_{0,4}^2\|\eta\|_{0,4} \\
&\qquad+\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_{0,6}^{2q}\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_{0,6}\|u^{(1)}-u^{(0)}\|_{0,6}^2\|\eta\|_{0,6/(3-2q)} \\
&\quad\lesssim\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_1\|u^{(1)}-u^{(0)}\|_1^2\|\eta\|_1 \\
&\qquad+\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_1^{2q}\|u^{(0)}+t(u^{(1)}-u^{(0)})\|_1\|u^{(1)}-u^{(0)}\|_1^2\|\eta\|_1 \\
&\quad\lesssim\|u^{(1)}-u^{(0)}\|_1^2\|\eta\|_1,
\end{aligned}
\tag{39}
$$

and

$$
\begin{aligned}
&\left|\left\langle f''((u^{(0)}+t(u^{(1)}-u^{(0)}))^2)(u^{(0)}+t(u^{(1)}-u^{(0)}))^3(u^{(1)}-u^{(0)})^2,\eta\right\rangle\right| \\
&\quad\lesssim\left|\left((1+(u^{(0)}+t(u^{(1)}-u^{(0)}))^{2q})(u^{(0)}+t(u^{(1)}-u^{(0)}))(u^{(1)}-u^{(0)})^2,\eta\right)\right| \\
&\quad\lesssim\|u^{(1)}-u^{(0)}\|_1^2\|\eta\|_1,
\end{aligned}
\tag{40}
$$

where the Hölder inequality and the imbedding theorem are used.

Based on (38), (39) and (40), we can derive

$$
\begin{aligned}
|\eta''(t)|&\lesssim|\lambda^{(1)}-\lambda^{(0)}|\|u^{(1)}-u^{(0)}\|_0\|\eta\|_0+|\sigma|\|u^{(1)}-u^{(0)}\|_0^2+\|u^{(1)}-u^{(0)}\|_1^2\|\eta\|_1 \\
&\lesssim\|(\lambda^{(0)}-\lambda^{(1)},u^{(0)}-u^{(1)})\|_X^2\|(\sigma,\eta)\|_X.
\end{aligned}
\tag{41}
$$

Thus, using (41) and the following Taylor expansion:

$$
\eta(1)=\eta(0)+\eta'(0)+\int_0^1\eta''(t)(1-t)dt,
$$

we can derive

$$
\begin{aligned}
\langle\mathcal{G}(\lambda^{(1)},u^{(1)}),(\sigma,\eta)\rangle=&\langle\mathcal{G}(\lambda^{(0)},u^{(0)}),(\sigma,\eta)\rangle+\langle\mathcal{G}'(\lambda^{(0)},u^{(0)})((\lambda^{(1)},u^{(1)}) \\
&-(\lambda^{(0)},u^{(0)})),(\sigma,\eta)\rangle \\
&+R((\lambda^{(0)},u^{(0)}),(\lambda^{(1)},u^{(1)}),(\sigma,\eta))
\end{aligned}
$$

with the residual

$$
|R((\lambda^{(0)},u^{(0)}),(\lambda^{(1)},u^{(1)}),(\sigma,\eta))|\lesssim\|(\lambda^{(0)}-\lambda^{(1)},u^{(0)}-u^{(1)})\|_X^2\|(\sigma,\eta)\|_X.
$$

Then the proof is completed. $\qquad\square$

4 Multigrid method for nonlinear eigenvalue problems based on Newton iteration

In this section, a novel multigrid method is proposed for solving the nonlinear eigenvalue problems based on the Newton iteration. To design the multigrid method, let us construct a

nested sequence of finite element spaces based on a nested multigrid sequence in the first step. The sequence of meshes is denoted by $\{\mathcal{T}_{h_k}\}_{k=1}^n$. For any two consecutive meshes $\mathcal{T}_{h_k}$ and $\mathcal{T}_{h_{k-1}}(k \geq 2)$, $\mathcal{T}_{h_k}$ is produced through a one-time uniform refinement form $\mathcal{T}_{h_{k-1}}$. The corresponding mesh sizes satisfy the following conditions:

$$h_k \approx \frac{1}{\beta} h_{k-1}, \quad k = 2, \cdots, n, \tag{42}$$

where the refinement index $\beta > 1$. Meanwhile, the following approximate relationship holds true:

$$\delta_{h_k}(u) \approx \frac{1}{\beta} \delta_{h_{k-1}}(u), \quad k = 2, \cdots, n. \tag{43}$$

Based on the multigrid sequence $\{\mathcal{T}_{h_k}\}_{k=1}^n$, we denote the corresponding sequence of finite element spaces by:

$$V_{h_1} \subset V_{h_2} \subset \cdots \subset V_{h_n} \subset H_0^1(\Omega), \tag{44}$$

and denote the sequence of tensor product finite element spaces by:

$$X_{h_1} \subset X_{h_2} \subset \cdots \subset X_{h_n} \subset X. \tag{45}$$

4.1 One step of the multigrid method

To describe the multigrid method more clearly, we first introduce how to perform one step of the Newton iteration.

For two consecutive finite element spaces X_{h_k} and $X_{h_{k+1}}$, assume that we have obtained an approximate solution $(\lambda^{h_k}, u^{h_k}) \in X_{h_k}$, Algorithm 1 shows the procedure for obtaining a new approximate solution $(\lambda^{h_{k+1}}, u^{h_{k+1}}) \in X_{h_{k+1}}$.

Algorithm 1 *One step of the multigrid method*

1. *Assume that we have obtained an approximate solution $(\lambda^{h_k}, u^{h_k}) \in X_{h_k}$.*
2. *Solve the linearized equation derived from Newton iteration in the finite element space $X_{h_{k+1}}$: Find $(\lambda^{h_{k+1}}, u^{h_{k+1}}) \in X_{h_{k+1}}$ such that for any $(\mu, v_{h_{k+1}}) \in X_{h_{k+1}}$, there holds*

$$\langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle$$
$$= -\langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle + \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle. \tag{46}$$

The equation (46) can be solved by the mixed finite element method in the following form: Find $(\lambda^{h_{k+1}}, u^{h_{k+1}}) \in \mathbb{R} \times V_{h_{k+1}}$ such that for any $(\mu, v_{h_{k+1}}) \in \mathbb{R} \times V_{h_{k+1}}$, there holds

$$\begin{cases} a(\lambda^{h_k}, u^{h_k}; u^{h_{k+1}}, v_{h_{k+1}}) + b(u^{h_k}; v_{h_{k+1}}, \lambda^{h_{k+1}}) = \langle \mathcal{F}_u'(\lambda^{h_k}, u^{h_k}) u^{h_k}, v_{h_{k+1}} \rangle \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad -\langle \mathcal{F}(\lambda^{h_k}, u^{h_k}), v_{h_{k+1}} \rangle - \lambda^{h_k}(u^{h_k}, v_{h_{k+1}}), \\ b(u^{h_k}; u^{h_{k+1}}, \mu) = -\mu/2 - \mu(u^{h_k}, u^{h_k})/2. \end{cases}$$

In order to simplify the notation, we use the following symbol to denote the above two solving steps:

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) = Newton\_Iteration(\lambda^{h_k}, u^{h_k}, X_{h_{k+1}}).$$

Next, we can prove that the new approximate solution $(\lambda^{h_{k+1}}, u^{h_{k+1}}) \in X_{h_{k+1}}$ derived by Algorithm 1 has a better accuracy than $(\lambda^{h_k}, u^{h_k}) \in X_{h_k}$. The detailed conclusion is presented in Theorem 3.

Theorem 3 *After implementing Algorithm* 1, *the new approximate solution* $(\lambda^{h_{k+1}}, u^{h_{k+1}})$ *has the following error estimate*

$$\|(\lambda - \lambda^{h_{k+1}}, u - u^{h_{k+1}})\|_X \lesssim \delta_{h_{k+1}}(u) + \|(\lambda - \lambda^{h_k}, u - u^{h_k})\|_X^2.$$

Proof We use $(\lambda_{h_{k+1}}, u_{h_{k+1}})$ to denote the standard finite element solution of (9), that is $(\lambda_{h_{k+1}}, u_{h_{k+1}})$ satisfies

$$(\mathcal{A}\nabla u_{h_{k+1}}, \nabla v_{h_{k+1}}) + (V u_{h_{k+1}}, v_{h_{k+1}}) + (f(u_{h_{k+1}}^2)u_{h_{k+1}}, v_{h_{k+1}})$$
$$= \lambda_{h_{k+1}}(u_{h_{k+1}}, v_{h_{k+1}}), \quad \forall v_{h_{k+1}} \in V_{h_{k+1}},$$

or equivalently

$$\langle \mathcal{G}(\lambda_{h_{k+1}}, u_{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle = 0, \quad \forall (\mu, v_{h_{k+1}}) \in X_{h_{k+1}}. \tag{47}$$

Combining (46) and (47) leads to

$$\langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda_{h_{k+1}} - \lambda^{h_{k+1}}, u_{h_{k+1}} - u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle$$
$$= \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle + \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda_{h_{k+1}} - \lambda^{h_k}, u_{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle$$
$$= -\langle \mathcal{G}(\lambda_{h_{k+1}}, u_{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle + \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle$$
$$+ \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda_{h_{k+1}} - \lambda^{h_k}, u_{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle$$
$$= -R((\lambda^{h_k}, u^{h_k}), (\lambda_{h_{k+1}}, u_{h_{k+1}}), (\mu, v_{h_{k+1}})). \tag{48}$$

Combining (24), Theorem 2 and (48), we can derive

$$\|(\lambda_{h_{k+1}} - \lambda^{h_{k+1}}, u_{h_{k+1}} - u^{h_{k+1}})\|_X \lesssim \|(\lambda_{h_{k+1}} - \lambda^{h_k}, u_{h_{k+1}} - u^{h_k})\|_X^2. \tag{49}$$

Using the standard error estimates in Lemma 2 and (49), we can obtain

$$\|(\lambda - \lambda^{h_{k+1}}, u - u^{h_{k+1}})\|_X \lesssim \|(\lambda - \lambda_{h_{k+1}}, u - u_{h_{k+1}})\|_X + \|(\lambda_{h_{k+1}} - \lambda^{h_{k+1}}, u_{h_{k+1}} - u^{h_{k+1}})\|_X$$
$$\lesssim \delta_{h_{k+1}}(u) + \|(\lambda - \lambda^{h_k}, u - u^{h_k})\|_X^2.$$

Then we complete the proof. □

4.2 Multigrid method based on Newton iteration

In this subsection, we introduce the multigrid method to solve the nonlinear eigenvalue problems. We use the following nested sequence of finite element spaces:

$$V_{h_1} \subset V_{h_2} \subset \cdots \subset V_{h_n} \subset H_0^1(\Omega), \tag{50}$$

and

$$X_{h_1} \subset X_{h_2} \subset \cdots \subset X_{h_n} \subset X. \tag{51}$$

The multigrid method is presented in Algorithm 42. The new algorithm requires solving the nonlinear eigenvalue problem directly in the initial space to obtain an initial value; then performing Algorithm 1 in the subsequent spaces. The approximate solution derived from the last space is used as the initial iterative value in the current space.

Algorithm 42 *Multigrid method based on Newton iteration*

1. *Solve the following nonlinear eigenvalue problem directly in the initial space V_{h_1}: Find $(\lambda^{h_1}, u^{h_1}) \in \mathbb{R} \times V_{h_1}$ such that $\|u^{h_1}\|_0 = 1$ and*

$$(\mathcal{A}\nabla u^{h_1}, \nabla v_{h_1}) + (V u^{h_1}, v_{h_1}) + (f((u^{h_1})^2)u^{h_1}, v_{h_1}) = \lambda^{h_1}(u^{h_1}, v_{h_1}), \quad \forall v_{h_1} \in V_{h_1}.$$

2. *For $k = 1, \cdots, n-1$, obtain a new approximate solution $(\lambda^{h_{k+1}}, u^{h_{k+1}}) \in X_{h_{k+1}}$ through*

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) = Newton\_Iteration(\lambda^{h_k}, u^{h_k}, X_{h_{k+1}}),$$

   End for.
3. *Finally, we obtain the approximate solution $(\lambda^{h_n}, u^{h_n}) \in X_{h_n}$.*

The error estimate for the final approximate solution $(\lambda^{h_n}, u^{h_n}) \in X_{h_n}$ derived by Algorithm 42 is presented in Theorem 4.

Theorem 4 *After performing Algorithm 42, the final approximate solution $(\lambda^{h_n}, u^{h_n}) \in X_{h_n}$ satisfies the following error estimate*

$$\|(\lambda - \lambda^{h_n}, u - u^{h_n})\|_X \lesssim \delta_{h_n}(u). \tag{52}$$

Proof We prove this theorem by the mathematical induction method. Since we solve the nonlinear eigenvalue problem directly in the initial space, using Lemma 2 leads to

$$\|(\lambda - \lambda^{h_1}, u - u^{h_1})\|_X \lesssim \delta_{h_1}(u).$$

Assume that (52) holds true for $(\lambda^{h_{n-1}}, u^{h_{n-1}}) \in X_{h_{n-1}}$. According to Theorem 3, we can derive

$$\begin{aligned}
&\|(\lambda - \lambda^{h_n}, u - u^{h_n})\|_X \\
&\lesssim \delta_{h_n}(u) + \|(\lambda - \lambda^{h_{n-1}}, u - u^{h_{n-1}})\|_X^2 \\
&\lesssim \delta_{h_n}(u) + \delta_{h_{n-1}}^2(u) \\
&\lesssim \delta_{h_n}(u).
\end{aligned}$$

So (52) also holds true for $(\lambda^{h_n}, u^{h_n}) \in X_{h_n}$. Then the proof is completed. $\qquad \square$

4.3 Work estimate of Algorithm 42

In this subsection, we end Sect. 4 by estimating the computational work of Algorithm 42 briefly. Let us use N_k to denote the dimensions of V_{h_k}. The dimensions of the sequence of finite element spaces satisfy

$$N_k \approx \beta^{d(k-n)} N_n, \quad k = 1, 2, \cdots, n.$$

Then the computational work of Algorithm 42 is presented in Theorem 5, which shows that Algorithm 42 has linear computational complexity.

Theorem 5 *Assume that solving the nonlinear eigenvalue problem directly in the initial space V_{h_1} needs work $\mathcal{O}(M_{h_1})$, and solving the linearized boundary value problem (46) in V_{h_k} needs work $\mathcal{O}(N_k)$ for $k = 2, 3, \cdots, n$. Then the total computational work of Algorithm 42 is $\mathcal{O}(N_n + M_{h_1})$. Furthermore, the linear computational complexity $\mathcal{O}(N_n)$ can be derived provided $M_{h_1} \leq N_n$.*

Proof Let us use W_k to denote the computational work of V_{h_k}. Then the computational work W of Algorithm 42 can be estimated by:

$$W = \sum_{k=1}^{n} W_k = \mathcal{O}(M_{h_1} + \sum_{k=2}^{n} N_k)$$

$$= \mathcal{O}(\sum_{k=2}^{n} \beta^{d(k-n)} N_n + M_{h_1})$$

$$= \mathcal{O}(N_n + M_{h_1}).$$

Then we get the desired conclusion $\mathcal{O}(N_n + M_{h_1})$, and when $M_{h_1} \leq N_n$, the total computational work changes to $\mathcal{O}(N_n)$. □

Remark 1 The linearized boundary value problem (46) can be solved efficiently by the multigrid method with linear computational complexity $\mathcal{O}(N_k)$ (see e.g., [5, 29]). Because the dimension of the initial space V_{h_1} is small and independent of the final finite element space, it is easy to get $M_{h_1} \leq N_n$. Thus, the linear computational complexity for Algorithm 42 can be obtained.

5 Multigrid method for nonlinear eigenvalue problems based on mixing scheme

In this section, an improved multigrid method is designed. The motivation is that when solving the nonlinear equation (18) by Newton iteration, we may encounter the non-convergence issue for some complicated models. This issue is the same as that of the classical self-consistent field iteration for the nonlinear eigenvalue problems. In order to overcome this difficulty, the mixing theme (damping method) is introduced to solve the nonlinear eigenvalue problems. The most widely used mixing scheme is Anderson acceleration [2], which is used to improve the convergence rate for fixed-point iterations. Although Anderson acceleration was widely used for solving various models, the first mathematical convergence result for linear and nonlinear problems did not appear until 2015 in [30]. For nonlinear eigenvalue problems, there is still no strictly theoretical analysis. In general, the mixing scheme is used in fixed-point iterations to accelerate the convergence rate. In this section, we use the mixing scheme in Newton iteration to design an improved multigrid method for nonlinear eigenvalue problems. Most importantly, we can prove that the norm of the residual decreases in each step of the Newton iteration. This may provide some inspiration to prove the theoretical conclusions for the more general mixing scheme for nonlinear eigenvalue problems.

Assume that we have obtained an approximate solution $(\lambda^{h_k}, u^{h_k}) \in X_{h_k}$, we introduce a novel iteration step in Algorithm 51 to obtain a better approximation $(\lambda^{h_{k+1}}, u^{h_{k+1}}) \in X_{h_{k+1}}$ on the basis of Newton iteration and the mixing scheme.

Algorithm 51 *One step of the mixing scheme*

1. *Given a parameter $\theta_{k+1} \in (0, 1)$ and an approximate solution $(\lambda^{h_k}, u^{h_k}) \in X_{h_k}$.*
2. *Solve the linearized equation derived from Newton iteration in the finite element space $X_{h_{k+1}}$: Find $(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}) \in X_{h_{k+1}}$ such that for any $(\mu, v_{h_{k+1}}) \in X_{h_{k+1}}$, there holds*

$$\langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle$$

$$= -\langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle + \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle. \quad (53)$$

The equation (53) *can be solved by the mixed finite element method in the following form:*
Find $(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}) \in \mathbb{R} \times V_{h_{k+1}}$ *such that for any* $(\mu, v_{h_{k+1}}) \in \mathbb{R} \times V_{h_{k+1}}$, *there holds*

$$\begin{cases} a(\lambda^{h_k}, u^{h_k}; \widehat{u}^{h_{k+1}}, v_{h_{k+1}}) + b(u^{h_k}; v_{h_{k+1}}, \widehat{\lambda}^{h_{k+1}}) = \langle \mathcal{F}'_u(\lambda^{h_k}, u^{h_k}) u^{h_k}, v_{h_{k+1}} \rangle \\ \qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad -\langle \mathcal{F}(\lambda^{h_k}, u^{h_k}), v_{h_{k+1}} \rangle - \lambda^{h_k}(u^{h_k}, v_{h_{k+1}}), \\ b(u^{h_k}; \widehat{u}^{h_{k+1}}, \mu) \qquad\qquad\qquad\qquad = -\mu/2 - \mu(u^{h_k}, u^{h_k})/2. \end{cases}$$

3. *Set*

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) = (1 - \theta_{k+1})(\lambda^{h_k}, u^{h_k}) + \theta_{k+1}(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}).$$

Next, we can prove that the value of the residual $|\langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle|$ monotonically decreases after performing Algorithm 51.

Theorem 6 *Assume that we have obtained an approximate solution* (λ^{h_k}, u^{h_k}), *then, there exists a sufficiently small* θ_{k+1} *such that the new approximation* $(\lambda^{h_{k+1}}, u^{h_{k+1}})$ *derived by Algorithm 51 has the following convergence*

$$|\langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle| \le (1 - \frac{\theta_{k+1}}{2})|\langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle|, \ \ \forall (\mu, v_{h_{k+1}}) \in X_{h_{k+1}}.$$

Proof From the following mixing scheme

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) = (1 - \theta_{k+1})(\lambda^{h_k}, u^{h_k}) + \theta_{k+1}(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}),$$

used in Algorithm 51, we can derive

$$(\lambda^{h_{k+1}} - \lambda^{h_k}, u^{h_{k+1}} - u^{h_k}) = \theta_{k+1}(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}). \tag{54}$$

Combining (54) and the mean value theorem leads to

$$\begin{aligned} &\langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle - \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \\ &\quad -\langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_{k+1}} - \lambda^{h_k}, u^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \\ &= \langle \mathcal{G}'(\widetilde{\lambda}^{h_k}, \widetilde{u}^{h_k})(\lambda^{h_{k+1}} - \lambda^{h_k}, u^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \\ &\quad -\langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_{k+1}} - \lambda^{h_k}, u^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \\ &= \theta_{k+1}\langle \mathcal{G}'(\widetilde{\lambda}^{h_k}, \widetilde{u}^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \\ &\quad -\theta_{k+1}\langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle, \end{aligned} \tag{55}$$

where $(\widetilde{\lambda}^{h_k}, \widetilde{u}^{h_k}) = (\alpha\lambda^{h_{k+1}} + (1 - \alpha)\lambda^{h_k}, \alpha u^{h_{k+1}} + (1 - \alpha)u^{h_k})$ with $\alpha \in (0, 1)$.
From (54), we know that when $\theta_{k+1} \to 0$, there holds

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) \to (\lambda^{h_k}, u^{h_k}),$$

and thus

$$(\widetilde{\lambda}^{h_k}, \widetilde{u}^{h_k}) \to (\lambda^{h_k}, u^{h_k}),$$

which also indicates

$$\begin{aligned} &\langle \mathcal{G}'(\widetilde{\lambda}^{h_k}, \widetilde{u}^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \\ &\quad \to \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle. \end{aligned}$$

Hence, we can choose a sufficiently small θ_{k+1}, such that

$$
\begin{aligned}
&\left| \langle \mathcal{G}'(\widetilde{\lambda}^{h_k}, \widetilde{u}^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right. \\
&\quad \left. - \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&\leq \frac{1}{2} \left| \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&= \frac{1}{2} \left| \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right|.
\end{aligned}
\tag{56}
$$

Based on (55) and (56), we can derive

$$
\begin{aligned}
&\left| \langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle - \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right. \\
&\quad \left. - \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_{k+1}} - \lambda^{h_k}, u^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&\leq \frac{\theta_{k+1}}{2} \left| \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right|.
\end{aligned}
\tag{57}
$$

Besides, using (53) and (54), we can also derive

$$
\begin{aligned}
&\left| \langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle - \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right. \\
&\quad \left. - \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_{k+1}} - \lambda^{h_k}, u^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&= \left| \langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle - \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right. \\
&\quad \left. - \theta_{k+1} \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\widehat{\lambda}^{h_{k+1}} - \lambda^{h_k}, \widehat{u}^{h_{k+1}} - u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&= \left| \langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle - \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right. \\
&\quad \left. + \theta_{k+1} \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&= \left| \langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle - (1 - \theta_{k+1}) \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right|.
\end{aligned}
\tag{58}
$$

Combining (57) and (58) leads to

$$
\begin{aligned}
&\left| \langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle - (1 - \theta_{k+1}) \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&\leq \frac{\theta_{k+1}}{2} \left| \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right|.
\end{aligned}
$$

That is

$$
\begin{aligned}
&\left| \langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle \right| \\
&\leq (1 - \frac{\theta_{k+1}}{2}) \left| \langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle \right|.
\end{aligned}
$$

Then we complete the proof. $\square$

In order to give a practical algorithm, there are two remaining problems. The first one is to choose an appropriate parameter θ_{k+1}. The second one is to compute the residual after each iteration step.

For the first problem, we provide an adaptive strategy for choosing θ_{k+1} in Algorithm 52, which gradually reduces the value of θ_{k+1}. For the second problem, in order to compare the following two values $|\langle \mathcal{G}(\lambda^{h_{k+1}}, u^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle|$ and $|\langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle|$ for any $(\mu, v_{h_{k+1}}) \in X_{h_{k+1}}$, we define the residual for the nonlinear eigenvalue problem (3) directly in the following way:

$$
Resi(\lambda^{h_k}, u^{h_k}) = \| A_{u^{h_k}} u^{h_k} - \lambda^{h_k} u^{h_k} \|_1 + \frac{1}{2} |1 - \| u^{h_k} \|_0^2|.
$$

Then next, we use $Resi(\lambda^{h_k}, u^{h_k})$ to judge the error reduction.

The modified mixing scheme and the corresponding multigrid method are presented in Algorithms 52 and 53, respectively.

Algorithm 52 *One step of the modified mixing scheme*

1. *Assume that we have obtained an approximate solution* $(\lambda^{h_k}, u^{h_k}) \in X_{h_k}$.
2. *Set* $\theta_{k+1} = 1$.
3. *Solve the linearized equation derived from Newton iteration in the finite element space* $X_{h_{k+1}}$: *Find* $(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}) \in X_{h_{k+1}}$ *such that for any* $(\mu, v_{h_{k+1}}) \in X_{h_{k+1}}$, *there holds*

$$\langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}), (\mu, v_{h_{k+1}}) \rangle = -\langle \mathcal{G}(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle$$
$$+ \langle \mathcal{G}'(\lambda^{h_k}, u^{h_k})(\lambda^{h_k}, u^{h_k}), (\mu, v_{h_{k+1}}) \rangle.$$

   The above equation can be solved by the mixed finite element method in the following form: Find $(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}) \in \mathbb{R} \times V_{h_{k+1}}$ *such that for any* $(\mu, v_{h_{k+1}}) \in \mathbb{R} \times V_{h_{k+1}}$, *there holds*

$$\begin{cases} a(\lambda^{h_k}, u^{h_k}; \widehat{u}^{h_{k+1}}, v_{h_{k+1}}) + b(u^{h_k}; v_{h_{k+1}}, \widehat{\lambda}^{h_{k+1}}) & = \langle \mathcal{F}'_u(\lambda^{h_k}, u^{h_k})u^{h_k}, v_{h_{k+1}} \rangle \\ & \quad -\langle \mathcal{F}(\lambda^{h_k}, u^{h_k}), v_{h_{k+1}} \rangle - \lambda^{h_k}(u^{h_k}, v_{h_{k+1}}), \\ b(u^{h_k}; \widehat{u}^{h_{k+1}}, \mu) & = -\mu/2 - \mu(u^{h_k}, u^{h_k})/2. \end{cases}$$

4. *Set*

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) = (1 - \theta_{k+1})(\lambda^{h_k}, u^{h_k}) + \theta_{k+1}(\widehat{\lambda}^{h_{k+1}}, \widehat{u}^{h_{k+1}}).$$

5. *If* $Resi(\lambda^{h_{k+1}}, u^{h_{k+1}}) \leq Resi(\lambda^{h_k}, u^{h_k})$, *stop. Else set* $\theta_{k+1} = \theta_{k+1}/2$ *and goto Step 4.*

In order to simplify the notation, we use the following symbol to denote the above solving steps:

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) = Mixing\_Iteration(\lambda^{h_k}, u^{h_k}, \theta_{k+1}, X_{h_{k+1}}).$$

Algorithm 53 *Multigrid method based on mixing scheme*

1. *Solve the following nonlinear eigenvalue problem directly in the initial space* V_{h_1}: *Find* $(\lambda^{h_1}, u^{h_1}) \in \mathbb{R} \times V_{h_1}$ *such that* $\|u^{h_1}\|_0 = 1$ *and*

$$(\mathcal{A}\nabla u^{h_1}, \nabla v_{h_1}) + (V u^{h_1}, v_{h_1}) + (f((u^{h_1})^2)u^{h_1}, v_{h_1}) = \lambda^{h_1}(u^{h_1}, v_{h_1}), \quad \forall v_{h_1} \in V_{h_1}.$$

2. *For* $k = 1, \cdots, n-1$, *obtain a new approximate solution* $(\lambda^{h_{k+1}}, u^{h_{k+1}}) \in X_{h_{k+1}}$ *through*

$$(\lambda^{h_{k+1}}, u^{h_{k+1}}) = Mixing\_Iteration(\lambda^{h_k}, u^{h_k}, \theta_{k+1}, X_{h_{k+1}}),$$

   End for.
3. *Finally, we obtain the approximate solution* $(\lambda^{h_n}, u^{h_n}) \in X_{h_n}$.

6 Numerical results

In this section, we propose some numerical experiments to support our theoretical conclusions and demonstrate the efficiency of the presented multigrid methods. For the linearized boundary value problems derived by Newton iteration, we used the V-cycle multigrid method to obtain the numerical solutions. The multigrid method includes two pre-smoothing steps and two post-smoothing steps. The adopted smoother for the pre-smoothing and post-smoothing steps is the distributive Gauss-Seidel (DGS) iteration [5].

Fig. 1 The initial mesh for Example 1
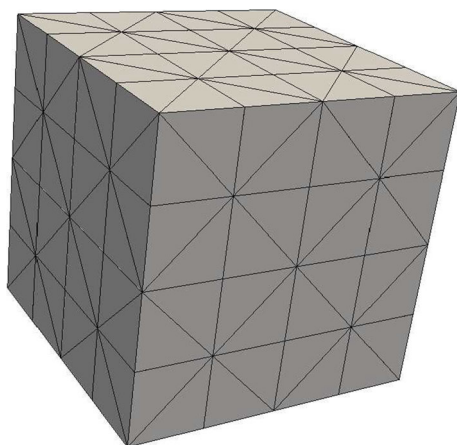
6.1 Example 1

In the first example, we use Algorithm 42 to solve the Gross-Pitaevskii equation: Find (λ, u) such that

$$
\begin{cases}
-\triangle u + V u + \zeta |u|^2 u = \lambda u, & \text{in } \Omega, \\
\qquad\qquad\qquad\quad u = 0, & \text{on } \partial\Omega, \\
\qquad\quad \int_\Omega u^2 d\Omega = 1,
\end{cases}
\tag{59}
$$

where $\Omega = [0, 1]^3$, $V = x_1^2 + 2x_2^2 + 4x_3^2$ and $\zeta = 1$.

The quadratic finite element space was adopted in this example. The sequence of meshes was produced through a one-time uniform refinement. Thus, the refinement index β between two consecutive meshes equals 2. The initial mesh is depicted in Fig. 1.

Because the exact solution of (59) is unknown, we select an adequately accurate approximation on a sufficiently fine mesh as the exact solution ($\lambda = 34.819449$). The error estimates of the approximate solutions derived by Algorithm 42 are presented in Fig. 2. The results show that Algorithm 42 is able to derive the approximate solutions with the optimal error estimates. In order to intuitively illustrate the efficiency of Algorithm 42, the computational time of Algorithm 42 is presented in Table 1. Besides, the computational time of the direct finite element method (solve the nonlinear eigenvalue problem directly in the final finite element space) is also presented. We can see that the linear computational complexity of Algorithm 42 can be obtained with the refinement of mesh. In addition, Algorithm 42 has a great advantage over the direct finite element method.

6.2 Example 2

In the second example, we use Algorithm 53 to solve the Gross-Pitaevskii equation (59) in $\Omega = [0, 1]^3$, where $W = x_1^2 + x_2^2 + x_3^2 + sin^2(2\pi x_1) + sin^2(2\pi x_2) + sin^2(2\pi x_3)$ and $\zeta = 100$.

In this example, due to the strong nonlinearity, Algorithm 42 does not converge. Thus, we used Algorithm 53 to solve the nonlinear eigenvalue problem and the convergent results are obtained.
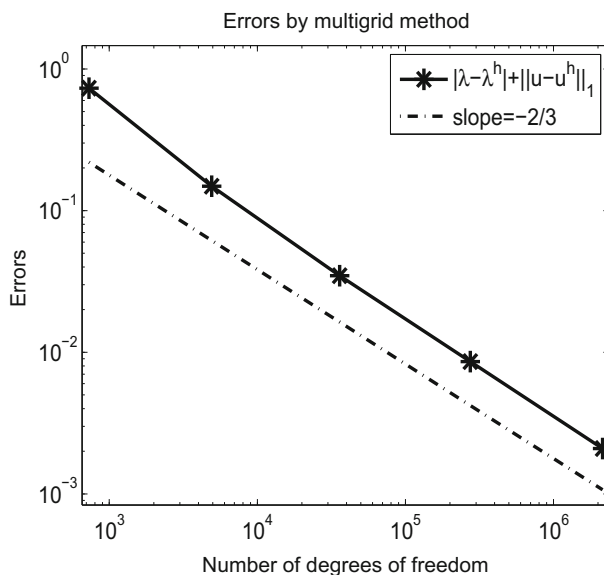
Fig. 2 The errors of approximate solutions derived by Algorithm 42 for Example 1

Table 1 The CPU time of Algorithm 42 and the direct finite element method for Example 1

Mesh level	Number of degrees of freedom	Time of Algorithm 42	Time of direct FEM
1	729	0.1643	0.1643
2	4913	0.2642	1.1434
3	35,937	1.1011	12.2292
4	274,625	8.3957	151.3532
5	2,146,689	66.9435	3947.5318
6	16,974,593	540.2690	–

The quadratic finite element space was adopted in this example. The sequence of meshes was produced through a one-time uniform refinement. Thus, the refinement index β between two consecutive meshes equals 2. The initial mesh is depicted in Fig. 3.

Because the exact solution of (59) is unknown, we select an adequately accurate approximation on a sufficiently fine mesh as the exact solution ($\lambda = 205.112532$). The error estimates of the approximate solutions derived by Algorithm 53 are shown in Fig. 4. The results show that Algorithm 53 is able to derive the approximate solutions with the optimal error estimates. In order to intuitively illustrate the efficiency of Algorithm 53, the computational time of Algorithm 53 and the direct finite element method are presented in Table 2. We can see that the linear computational complexity can be obtained with the refinement of mesh. In addition, Algorithm 53 has a great advantage over the direct finite element method.

In Table 3, we also present the value of θ_k in each finite element space X_{h_k}. The residuals corresponding to different θ_k are also presented. Based on the adaptive strategy, we can see that the value of the residual monotonically decreases with the refinement of mesh, and the results are consistent with Theorem 6.

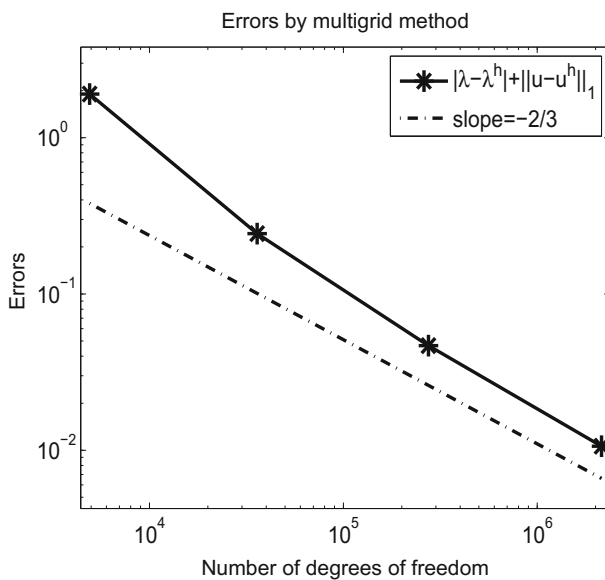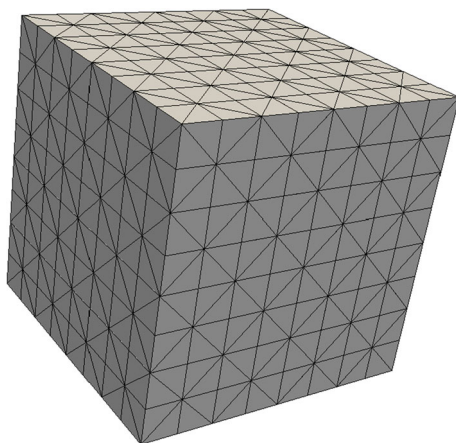**Fig. 3** The initial mesh for Example 2

Fig. 4 The errors of approximate solutions derived by Algorithm 53 for Example 2

Table 2 The CPU time of Algorithm 53 and the direct finite element method for Example 2

Mesh level	Number of degrees of freedom	Time of Algorithm 53	Time of direct FEM
1	4913	2.1944	2.1944
2	35,937	3.0003	24.5324
3	274,625	10.7252	289.6608
4	2,146,689	80.1640	7560.2502
5	16,974,593	676.3930	–

Table 3 The value of θ_k in Algorithm 53 for Example 2

Mesh level	The value of θ_k	The value of $Resi$
1	0.5	17.8814
2	0.5	8.9953
3	0.5	4.4768
4	0.5	2.2358
5	0.5	1.1355

7 Concluding remarks

In this study, we designed a novel multigrid method to solve the nonlinear eigenvalue problem on the basis of the Newton iteration. The novel scheme transforms the nonlinear eigenvalue problem into a series of linearized boundary value problems defined in a sequence of product finite element spaces. Because of avoiding solving large-scale nonlinear eigenvalue problems directly, the overall solving efficiency is significantly improved. Besides, the optimal error estimate and linear complexity can be derived simultaneously. In addition, an improved multigrid method coupled with the mixing scheme is also introduced. The improved scheme can make the iteration scheme converge for more complicated models. More importantly, a convergence result is derived which is missing in the existing literature on the mixing scheme for nonlinear eigenvalue problems.

Declarations

Conflict of interest The authors declare that they have no conflicts of interest/competing interests.

Code availability The custom codes generated during the current study are available from the corresponding author on reasonable request.

References

1. Adams, R.A.: Sobolev Spaces. Academic Press, New York (1975)
2. Anderson, D.G.: Iterative procedures for nonlinear integral equations. J. Assoc. Comput. Mach. **12**, 547–560 (1965)
3. Bao, G., Hu, G., Liu, D.: An h-adaptive finite element solver for the calculations of the electronic structures. J. Comput. Phys. **231**(14), 4967–4979 (2012)
4. Bao, W., Du, Q.: Computing the ground state solution of Bose–Einstein condensates by a normalized gradient flow. SIAM J. Sci. Comput. **25**, 1674–1697 (2004)
5. Benzi, M., Golub, G.H., Liesen, J.: Numerical solution of saddle point problems. Acta Numer **14**, 1–137 (2005)
6. Brandt, A.: Multi-level adaptive solutions to boundary-value problems. Math. Comput. **31**, 333–390 (1977)
7. Brezzi, F., Fortin, F.: Mixed and Hybrid Finite Element Methods. Springer, New York (1991)
8. Brenner, S., Scott, L.: The Mathematical Theory of Finite Element Methods. Springer, New York (1994)

9. Cai, Y., Zhang, L., Bai, Z., Li, R.: On an eigenvector-dependent nonlinear eigenvalue problem. SIAM J. Matrix Anal. Appl. **39**(3), 1360–1382 (2018)

10. Cancès, E., Chakir, R., Maday, Y.: Numerical analysis of nonlinear eigenvalue problems. J. Sci. Comput. **45**(1–3), 90–117 (2010)

11. Cancès, E., Chakir, R., Maday, Y.: Numerical analysis of the planewave discretization of some orbital-free and Kohn-Sham models. ESAIM Math. Model. Numer. Anal. **46**, 341–388 (2012)

12. Chen, H., He, L., Zhou, A.: Finite element approximations of nonlinear eigenvalue problems in quantum physics. Comput. Methods Appl. Mech. Engrg. **200**(21), 1846–1865 (2011)

13. Chen, H., Liu, F., Zhou, A.: A two-scale higher-order finite element discretization for Schrödinger equation. J. Comput. Math. **27**, 315–337 (2009)

14. Chen, H., Xie, H., Xu, F.: A full multigrid method for eigenvalue problems. J. Comput. Phys. **322**, 747–759 (2016)

15. Ciarlet, P.G., Lions, J.L. (eds.): Finite Element Methods, Handbook of Numerical Analysis, vol. II. North Holland, Amsterdam (1991)

16. Fang, H., Saad, Y.: Two classes of multisecant methods for nonlinear acceleration. Numer. Linear Algebra Appl. **16**(3), 197–221 (2009)

17. Fedorenko, R.P.: A relaxation method for solving elliptic difference equations. USSR Comput. Math. Math. Phys. **1**(4), 1092–1096 (1961)

18. Hackbusch, W.: Multi-grid Methods and Applications. Springer, Berlin (1985)

19. Harrison, R., Moroz, I., Tod, K.P.: A numerical study of the Schrödinger–Newton equations. Nonlinearity **16**, 101–122 (2003)

20. Hu, G., Xie, H., Xu, F.: A multilevel correction adaptive finite element method for Kohn–Sham equation. J. Comput. Phys. **355**, 436–449 (2018)

21. Jia, S., Xie, H., Xie, M., Xu, F.: A full multigrid method for nonlinear eigenvalue problems. Sci China Math **59**, 2037–2048 (2016)

22. Kohn, W., Sham, L.: Self-consistent equations including exchange and correlation effects. Phys. Rev. A **140**, 1133–1138 (1965)

23. Lieb, E.H.: Thomas-Fermi and related theories of atoms and molecules. Rev. Mod. Phys. **53**, 603–641 (1981)

24. Lin, L., Yang, C.: Elliptic preconditioner for accelerating the self-consistent field iteration in Kohn-Sham density functional theory. SIAM J. Sci. Comput. **35**(5), S277–S298 (2013)

25. Lott, P.A., Walker, H.F., Woodward, C.S., Yang, U.M.: An accelerated Picard method for nonlinear systems related to variably saturated flow. Adv. Water Resour. **38**, 92–101 (2012)

26. Motamarri, P., Iyer, M., Knap, J., Gavini, V.: Higher-order adaptive finite-element methods for orbital-free density functional theory. J. Comput. Phys. **231**, 6596–6621 (2012)

27. Pollock, S., Rebholz, L., Xiao, M.: Anderson-accelerated convergence of picard iterations for incompressible Navier-Stokes equations. SIAM J. Numer. Anal. **57**(2), 615–637 (2019)

28. Stasiak, P., Matsen, M.W.: Efficiency of pseudo-spectral algorithms with anderson mixing for the SCFT of periodic block-copolymer phases. Eur. Phys. J. E **34**(110), 1–9 (2011)

29. Shaidurov, V.V.: Multigrid Methods for Finite Element. Kluwer Academic Publics, Netherlands (1995)

30. Toth, A., Kelley, C.T.: Convergence analysis for Anderson acceleration. SIAM J. Numer. Anal. **53**(2), 805–819 (2015)

31. Xu, J.: Iterative methods by space decomposition and subspace correction. SIAM Rev. **34**, 581–613 (1992)

32. Xu, J., Zhou, A.: A two-grid discretization scheme for eigenvalue problems. Math. Comp. **70**, 17–25 (2001)

33. Yserentant, H.: On the regularity of the electronic Schrödinger equation in Hilbert spaces of mixed derivatives. Numer. Math. **98**(4), 731–759 (2004)

34. Zhou, A.: An analysis of finite-dimensional approximations for the ground state solution of Bose-Einstein condensates. Nonlinearity **17**, 541–550 (2004)

35. Zhou, A.: Finite dimensional approximations for the electronic ground state solution of a molecular system. Math. Methods Appl. Sci. **30**, 429–447 (2007)