Numerische
Mathematik

The *J*-method for the Gross–Pitaevskii eigenvalue problem

Robert Altmann[1] · Patrick Henning[2,3] · Daniel Peterseim[1]

Abstract

This paper studies the *J*-method of [*E. Jarlebring, S. Kvaal, W. Michiels. SIAM J. Sci. Comput. 36-4:A1978-A2001, 2014*] for nonlinear eigenvector problems in a general Hilbert space framework. This is the basis for variational discretization techniques and a mesh-independent numerical analysis. A simple modification of the method mimics an energy-decreasing discrete gradient flow. In the case of the Gross–Pitaevskii eigenvalue problem, we prove global convergence towards an eigenfunction for a damped version of the *J*-method. More importantly, when the iterations are sufficiently close to an eigenfunction, the damping can be switched off and we recover a local linear convergence rate previously known from the discrete setting. This quantitative convergence analysis is closely connected to the *J*-method's unique feature of sensitivity with respect to spectral shifts. Contrary to classical gradient flows, this allows both the selective approximation of excited states as well as the amplification of convergence beyond linear rates in the spirit of the Rayleigh quotient iteration for linear eigenvalue problems. These advantageous convergence properties are demonstrated in a series of numerical experiments involving exponentially localized states under disorder potentials and vortex lattices in rotating traps.

✉ Patrick Henning
  patrick.henning@rub.de

  Robert Altmann
  robert.altmann@math.uni-augsburg.de

  Daniel Peterseim
  daniel.peterseim@math.uni-augsburg.de

[1] Department of Mathematics, University of Augsburg, Universitätsstr. 14, 86159 Augsburg, Germany

[2] Department of Mathematics, Ruhr-University Bochum, 44801 Bochum, Germany

[3] Department of Mathematics, KTH Royal Institute of Technology, SE-100 44 Stockholm, Sweden

1 Introduction

This paper studies eigenvalues of nonlinear differential operators $\mathcal{A}$ on a *real* Hilbert space V of the form

$$\mathcal{A}(v) = A(v, v),$$

where $A \colon V \times V \to V^*$ is a continuous, bounded mapping that is invariant under scaling of the first argument and real-linear in the second argument. A well-known example that can be cast in such a format is the Gross–Pitaevskii eigenvalue problem (GPEVP) in a bounded domain $\mathcal{D} \subset \mathbb{R}^d$ (with $d = 2, 3$). In the strong form the GPEVP reads

$$\left[-\Delta + W - \Omega \mathcal{L}_z + \kappa \, |u^*|^2 \right] u^* = \lambda^* u^*. \tag{1}$$

Here, $W \in L^\infty(\mathcal{D}, \mathbb{R})$ denotes a non-negative and space-dependent potential, $\Omega \in \mathbb{R}$ is the angular velocity, $\mathcal{L}_z = -\mathrm{i} \left(x \partial_y - y \partial_x \right)$ is the z-component of the angular momentum, and $\kappa \geq 0$ regulates the nonlinearity of the problem. This problem is related, e.g., to the modeling of Bose-Einstein condensates [17,26,30,45]. In this context, a solution u^* represents a stationary quantum state of the condensate, $|u^*|^2$ is the corresponding density, and λ^* the so-called chemical potential.

The numerical solution of the GPEVP (1) has been studied extensively in recent years. Popular methods for solving the nonlinear eigenvalue problem are for example the *Self Consistent Field Iteration* (SCF), cf. [19,23,24,29,52], which requires to solve a linearized eigenvalue problem in each iteration. Algorithms that belong to the SCF class are for instance the *Roothaan algorithm* [48] or the *optimal damping algorithm* proposed in [23]. Another important class of iterative methods is based on gradient flows for the energy functional associated with (1), where we mention the *Discrete Normalized Gradient Flow* (DNGF), cf. [11–13,15], which is based on an implicit Euler discretization of the L^2-gradient flow. Improvements of the approach by using conjugated gradients were proposed in [7]. Here we also mention the *Projected Sobolev Gradient Flows* (PSGFs), cf. [27,33,34,36,39,46,47,57], which form a subclass of the gradient flow methods. Sobolev gradients are the Riesz representants of the Fréchet derivative of the energy functional in a suitable Hilbert space. With this, PSGFs are based on computing such Sobolev gradients and then projecting them onto the tangent space of the normalization constraint. An improvement of PSGF by using Riemannian conjugate gradients was suggested in [28]. Another strategy to solve the GPEVP involves a direct minimization of the energy functional, cf. [14,18], which means that (1) is written as a nonlinear saddle point problem which is then solved by a Newton-type method. Global convergence results for solving the GPEVP are very rare in the literature. To the best of our knowledge, global convergence had been so far only

established for a damped PSGF suggested in [36], where corresponding analytical results can be found in [36,57].

The SCF and PSGF approaches are typically based on the simple linearization of the partial differential operator by replacing the density $|u^*|^2$ in (1) by the density of some given approximation of u^*. In the finite-dimensional case (after spatial discretization) these approaches can be interpreted as generalizations of the inverse iteration (inverse power method). Linear convergence is observed empirically which is in agreement with the convergence theory for linear matrix eigenvalue problems. However, in the presence of clustered eigenvalues, which are typically connected to interesting physical phenomena, linear convergence can be very slow and shifting (as in the shifted inverse iteration or the Rayleigh quotient iteration) is a well-established technique for the acceleration of matrix eigenvalue solvers. The problem is that the aforementioned schemes seem to be unable to achieve any speed-up by a sophisticated shift of the nonlinear operator. In [38], Jarlebring, Kvaal, and Michiels observed that this insensitivity towards spectral shifts is the result of an unsuitable choice of linearization. The authors propose a natural but conceptually different linearization using the Jacobian of the nonlinear operator $\mathcal{A}$. The resulting approach does not belong to any of the classes above and we will refer to it as the *J-method*.

This paper generalizes the *J*-method and its numerical analysis to an abstract Hilbert space setting. Hence, the resulting iteration scheme is based on a variational formulation which is mesh-independent (cf. [2] for a similar approach in electromagnetism). This then allows the application of any spatial discretization, including finite elements and spectral methods. As discovered in [38], the correct choice of the linearization in the *J*-method does not only lead to a competitive nonlinear eigenvalue solver, it also allows major theoretical progress such as a proof of local linear convergence based on a version of Ostrowski's theorem (see Sect. 3 for the corresponding result in the Hilbert space setting). The obtained linear convergence rate, which is essentially $|\lambda^* + \sigma|/|\mu + \sigma|$, depends on the spectral gap of the σ-shifted Jacobian, where $\mu + \sigma$ is the second-smallest eigenvalue of the σ-shifted Jacobian around u^*. Our results generalize the findings of [38] from the matrix case to the abstract setting. In the case of the GPEVP, which is discussed in Sect. 4, this marks the first quantified convergence result. Moreover, an adaptive choice of the shifts in the spirit of the Rayleigh quotient iteration amplifies convergence beyond the linear rate in representative numerical experiments, see Sect. 6. At the same time, the sensitivity to spectral shifts facilitates the computation of excited states.

As usual for nonlinear problems, the quantitative convergence results are of local nature in the sense that they require a sufficiently accurate initial approximation. In practical computations this may be a rather challenging task. A simple damping strategy inspired by an energy-dissipative gradient flow [36] resolves this problem. In the case of the GPEVP we prove global convergence of the method towards an eigenfunction with a guaranteed decrease of energy, cf. Sect. 5.

Altogether, the combination of the *J*-method with damping for globalization and shifting for acceleration provides a powerful methodology for the simulation and analysis of nonlinear PDE eigenvector problems such as the GPEVP.

2 Definition of the *J*-method in Hilbert spaces

2.1 Nonlinear eigenvector problem

We consider a nonlinear differential operator

$$\mathcal{A} \colon V \to V^*$$

that maps a *real* Hilbert space V into its dual space V^*. In particular, V is equipped with an $\mathbb{R}$-inner product and every $F \in V^*$ satisfies $\langle F, v \rangle_{V^*, V} \in \mathbb{R}$ for all $v \in V$. Note that V is still allowed to contain complex-valued functions. We assume that $\mathcal{A}$ has the form

$$\mathcal{A}(v) = A(v, v),$$

where $A \colon V \times V \to V^*$ is continuous, bounded, and (real-)linear in the second argument. Recall that $A(v, \cdot)$ being real-linear (cf. [16, Def. 3.50]) means that for all $v, w_1, w_2 \in V$ and $\alpha \in \mathbb{R}$ we have

$$A(v, w_1 + w_2) = A(v, w_1) + A(v, w_2) \quad \text{and} \quad A(v, \alpha w_1) = \alpha A(v, w_1).$$

Furthermore, we assume that $A(\cdot, \cdot)$ is sufficiently smooth on $V \backslash \{0\} \times V$ (in particular real-Fréchet differentiable in both arguments) and that it is real-scaling invariant in the first argument, i.e.,

$$A(u, v) = A(\alpha u, v)$$

for all $u, v \in V$ and all $\alpha \in \mathbb{R} \backslash \{0\}$.

For the weak formulation of the eigenvalue problem with a nonlinearity in the eigenfunction, we introduce another real Hilbert space H (the pivot space) with $\mathbb{R}$-inner product

$$(\cdot, \cdot)_H \colon H \times H \to \mathbb{R}$$

such that V, H, V^* forms a Gelfand triple [56, Ch. 23.4]. The corresponding H-norm $\|\cdot\| := \|\cdot\|_H$ will be used for the normalization condition of the eigenfunction. The corresponding embedding $\mathcal{I} \colon V \to V^*$ is defined via

$$\langle \mathcal{I}u, \cdot \rangle_{V^*, V} = (u, \cdot)_H$$

for $u \in V$. For brevity, we shall write $\langle \cdot, \cdot \rangle$ for the canonical duality pairing in V.

The goal of this paper is to solve the corresponding PDE eigenvalue problem: find an eigenfunction $u^* \in V$ with $\|u^*\| = 1$ and an eigenvalue $\lambda^* \in \mathbb{R}$ such that $\mathcal{A}(u^*) = \lambda^* \mathcal{I} u^*$. This is equivalent to the variational formulation

$$\langle \mathcal{A}(u^*), v \rangle = \langle \lambda^* \mathcal{I} u^*, v \rangle = \lambda^* (u^*, v)_H \quad \text{for all test functions } v \in V. \tag{2}$$

Throughout the paper, we denote by $\overline{\alpha}$ the complex conjugate of a complex number $\alpha \in \mathbb{C}$, by $\Re(v)$ the real part of v, and by $\Im(v)$ the imaginary part.

Example 1 We are particularly interested in the GPEVP (1). Here we consider $H = L^2(\mathcal{D}) := L^2(\mathcal{D}, \mathbb{C})$ as a real Hilbert space equipped with the inner product

$$(v, w)_H := \Re \left(\int_{\mathcal{D}} v \, \overline{w} \, dx \right)$$

and analogously $V = H_0^1(\mathcal{D}) := H_0^1(\mathcal{D}, \mathbb{C})$ as a real Hilbert space equipped with

$$(v, w)_V := \Re \left(\int_{\mathcal{D}} \nabla v \cdot \overline{\nabla w} \, dx \right).$$

The (real) dual space is denoted by $V^* = H^{-1}(\mathcal{D}) := H^{-1}(\mathcal{D}, \mathbb{C})$. The corresponding nonlinear operator reads

$$\langle A(u, v), w \rangle := \Re \left(\int_{\mathcal{D}} \nabla_R v \cdot \overline{\nabla_R w} + W_R \, v \, \overline{w} \, dx + \frac{\kappa}{\|u\|^2} \int_{\mathcal{D}} |u|^2 v \, \overline{w} \, dx \right). \quad (3)$$

Here, we have $W_R(x) = W(x) - \frac{1}{4} \Omega^2 |x|^2$ and the rotational gradient ∇_R is given by

$$\nabla_R v := \nabla v + i \frac{\Omega}{2} R^\top v$$

for the divergence-free vector field $R(x, y, z) := (y, -x, 0)$ if $d = 3$ and $R(x, y) := (y, -x)$ if $d = 2$. Note that the nonlinear term is multiplied by $\|u\|^{-2}$ in order to achieve the assumed scaling invariance of A in the first component.

2.2 Linearization and the *J*-operator

Recall that the real-Fréchet derivative of some $F\colon X \to Y$ in a point $u \in X$ is a bounded and $\mathbb{R}$-linear operator $F'(u)\colon X \to Y$ with the property

$$\lim_{h \in X \setminus \{0\}, \|h\|_X \to 0} \frac{\|F(u + h) - F(u) - F'(u)h\|_Y}{\|h\|_X} = 0.$$

We write $F'(u; h) := F'(u)h$ for the Fréchet derivative in u in direction of h. In the following, we speak about the Fréchet derivatives of operators, where we always mean the real-Fréchet derivative as stated above. Moreover, we neglect writing the supplement $h \in X \setminus \{0\}$ beneath the limit. For the operator A, we denote the partial Fréchet derivative with respect to the (nonlinear) first component by $\partial_1 A\colon V^3 \to V^*$. Because of the scaling invariance of A, i.e., $A(u, \cdot) = A(\alpha u, \cdot)$, we have $\partial_1 A(u, \cdot; u) = 0$, which can be seen by

$$\lim_{t \to 0} \frac{\|A(u + tu, \cdot) - A(u, \cdot) - 0\|_{\mathcal{L}(V, V^*)}}{t \|u\|_V} = \lim_{t \to 0} \frac{\|A(u, \cdot) - A(u, \cdot)\|_{\mathcal{L}(V, V^*)}}{t \|u\|_V} = 0.$$

The Fréchet derivative of $\mathcal{A}$ in $u \in V$ and in direction $w \in V$ can be expressed as

$$
\begin{aligned}
\mathcal{A}'(u; w) &= \lim_{t \to 0} \frac{A(u + tw, u + tw) - A(u, u)}{t} \\
&= \lim_{t \to 0} \frac{A(u + tw, u) - A(u, u)}{t} + \lim_{t \to 0} \frac{A(u + tw, u + tw) - A(u + tw, u)}{t} \\
&= \partial_1 A(u, u; w) + A(u, w).
\end{aligned}
$$

Due to $\partial_1 A(u, \cdot\,; u) = 0$, we conclude

$$
\mathcal{A}'(u; u) = \partial_1 A(u, u; u) + A(u, u) = A(u, u) = \mathcal{A}(u).
$$

For an element $u \in V$ we now define the operator $J(u) \colon V \to V^*$ by

$$
J(u) := \mathcal{A}'(u; \cdot\,). \tag{4}
$$

Hence, using $\langle J(u^*)u^*, v \rangle = \langle \mathcal{A}'(u^*, u^*), v \rangle = \langle \mathcal{A}(u^*), v \rangle$, we observe that the eigenvalue problem (2) can be rewritten as: find $u^* \in V$ with $\|u^*\| = 1$ and $\lambda^* \in \mathbb{R}$ such that

$$
\langle J(u^*)u^*, v \rangle = \lambda^*(u^*, v)_H \qquad \text{for all } v \in V. \tag{5}
$$

2.3 The shifted *J*-method

The shifted J-method is defined by applying the inverse power iteration to the reformulated eigenvalue problem (5) based on the linearization J and some spectral shift σ. For that, we consider a function $v \in V$ and $\sigma \in \mathbb{R}$ such that $J(v) + \sigma \mathcal{I} \colon V \to V^*$ is invertible. In practice, one may either choose σ large such that $J(v) + \sigma \mathcal{I}$ is coercive (see Sect. 5.1 for the example of the GPEVP) or close to the negative of the target eigenvalue (as it is done in the numerical experiments of Sect. 6).

For $F \in V^*$ we define $u = (J(v) + \sigma \mathcal{I})^{-1} F \in V$ as the unique solution of the variational problem

$$
\langle (J(v) + \sigma \mathcal{I})u, w \rangle = \langle F, w \rangle
$$

for all $w \in V$. Considering a fixed shift σ such that $-\sigma$ is no eigenvalue of $J(v)$, we define $\psi \colon V \to V$ by

$$
\psi(v) := (J(v) + \sigma \mathcal{I})^{-1} \mathcal{I} v.
$$

Including an additional normalization step finally leads to the operator $\phi \colon V \to V$,

$$
\phi(v) := \frac{\psi(v)}{\|\psi(v)\|}.
$$

Now consider a normalized eigenpair (u^*, λ^*) of the eigenvalue problem (2). Then, $\psi(u^*)$ satisfies due to $J(u^*)u^* = \mathcal{A}(u^*) = \lambda^* \mathcal{I} u^*$,

$$
\begin{aligned}
(\lambda^* + \sigma)\psi(u^*) &= (J(u^*) + \sigma\mathcal{I})^{-1} J(u^*)\, u^* + \sigma (J(u^*) + \sigma\mathcal{I})^{-1} \mathcal{I} u^* \\
&= (J(u^*) + \sigma\mathcal{I})^{-1}(J(u^*) + \sigma\mathcal{I})\, u^* = u^*.
\end{aligned}
$$

As a consequence, we observe that u^* is a fixed point of ϕ, i.e., $u^* = \phi(u^*)$. This motivates the following iteration scheme: given $u^0 \in V$ with $\|u^0\| = 1$ and a shift σ such that $J(u^0) + \sigma\mathcal{I}$ is invertible, compute iterates u^{n+1} for $n \geq 0$ by

$$
u^{n+1} = \phi(u^n) = \frac{\psi(u^n)}{\|\psi(u^n)\|} = \alpha_n \, (J(u^n) + \sigma\mathcal{I})^{-1} \mathcal{I} u^n \tag{6}
$$

with the normalization factor $\alpha_n = 1/\|(J(u^n) + \sigma\mathcal{I})^{-1}\mathcal{I} u^n\|$. For the matrix case, the convergence of this scheme was analyzed in [38]. We emphasize that this scheme is well-defined if the shift σ guarantees the invertibility of $J(u^n) + \sigma\mathcal{I}$. This property has to be checked within the specific application and will be proved for the GPEVP in Sect. 5.1.

2.4 The damped *J*-method

In many applications, eigenvalue problems (with a nonlinearity in the eigenfunction) can be equivalently formulated as a constrained energy minimization problem. In this case, the radius of convergence can be considerably enhanced by introducing a variable damping parameter τ_n. This damping parameter (or step size) is adaptively selected such that the energy is optimally minimized in each iteration. In [36], a geometric justification of this approach was given, by interpreting it as the discretization of a certain projected Sobolev gradient flow, where the inner product (with which respect the Sobolev gradient is computed) is based on a repeated linearization of the differential operator. Even though $\langle (J(v) + \sigma\mathcal{I})\cdot, \cdot \rangle$ typically does not define an inner product (and hence does not allow a geometric interpretation), it is still possible to generalize the *J*-method defined in (6) in the spirit of the results in [36]. In Sect. 5 below, we shall give a rigorous justification of this approach by proving global convergence of the damped *J*-method in the context of the GPEVP.

The damping strategy aims to find a (optimized) linear combination of u^n and $(J(u^n) + \sigma\mathcal{I})^{-1}\mathcal{I} u^n$. For brevity we introduce $J_\sigma(u) := J(u) + \sigma\mathcal{I}$ and assume for a moment that the shift σ is chosen such that $J_\sigma(u^n)$ remains coercive.

One step of the damped *J*-method then reads as follows: given $u^n \in V$ with $\|u^n\| = 1$ and a step size $\tau_n > 0$ compute

$$
\tilde{u}^{n+1} = (1 - \tau_n)u^n + \tau_n \gamma_n J_\sigma(u^n)^{-1}\mathcal{I} u^n, \qquad u^{n+1} = \frac{\tilde{u}^{n+1}}{\|\tilde{u}^{n+1}\|} \tag{7}
$$

with $\gamma_n^{-1} := (J_\sigma(u^n)^{-1}\mathcal{I}u^n, u^n)_H$. This choice provides the important H-orthogonality

$$(\tilde{u}^{n+1} - u^n, u^n)_H = 0. \tag{8}$$

Due to the normalization in (7), we recover the original J-method (6) for $\tau_n = 1$. Further, the H-orthogonality (8) implies that the mass (i.e., the H-norm) of $\tilde{u}^{n+1}$ is greater or equal to 1 (even strictly larger unless $u^{n+1} = u^n$). This can be seen from

$$1 = \|u^n\|^2 = (u^n, \tilde{u}^{n+1})_H \le \|\tilde{u}^{n+1}\|. \tag{9}$$

For the example of the GPEVP we will show in Sect. 5 that the iteration (7) is well-defined and that it converges to an eigenstate if the step size τ_n is sufficiently small.

To summarize, we have introduced a damped version of the J-method (7) (with typical choice $\sigma = 0$ or σ large) which intends to ensure global convergence and a shifted version (6) to accelerate the convergence (with typical choice σ close to $-\lambda^*$). Note, however, that we do not intend to use damping and shifting simultaneously as this seems hard to control numerically. Instead, we propose to apply damping for globalization of convergence and then switch to shifting if a certain accuracy is obtained. This practice then provides a powerful methodology as illustrated in Sect. 6.

3 Abstract local convergence of the shifted J-method

The sensitivity of the J-method to spectral shifts facilitates the numerical approximation of excited states. To see this, we transfer the local convergence result presented in [38] to the Hilbert space setting. This shows that the shifted J-method converges to an eigenfunction u^* of $\mathcal{A}$ if the starting function and the shift are sufficiently close to the eigenpair (u^*, λ^*) of interest. In this section, we consider the undamped version of the J-method, i.e., we consider the iteration $u^{n+1} = \phi(u^n)$, including an arbitrary shift $\sigma \in \mathbb{R}$ and assuming that $J_\sigma(u^n)$ is invertible.

In order to prove local convergence with a linear rate that depends on the shift, we make use of the following well-known proposition that is a version of the Ostrowski theorem [44] in Banach spaces.

Proposition 1 *Let $\phi \colon V \to V$ be a mapping on the Banach space V that is Fréchet-differentiable at a point $u^* \in V$ such that the Fréchet-derivative $\phi'(u^*) \colon V \to V$ is a bounded linear operator with spectral radius*

$$\rho^* := \rho(\phi'(u^*)) < 1.$$

Then there is an open neighborhood S of u^*, such that for all starting values $u^0 \in S$ we have that the fixed point iterations

$$u^{n+1} := \phi(u^n)$$

converge strongly in V to u^*, i.e., $\|u^n - u^*\|_V \to 0$ for $n \to \infty$. Furthermore, for every $\varepsilon > 0$ there exist a neighborhood S_ε of u^* and a (finite) constant $C_\varepsilon > 0$ such

that

$$\|u^n - u^*\|_V \le C_\varepsilon \, |\rho^* + \varepsilon|^n \, \|u^0 - u^*\|_V \quad \text{for all } u^0 \in S_\varepsilon \text{ and } n \ge 1.$$

Hence, ρ^* defines the asymptotic linear convergence rate of the fixed point iteration.

Proof The proof is based on the observation that under the assumptions of the lemma and for each $\varepsilon > 0$, there exists an induced operator norm $\|\cdot\|_\varepsilon$ such that $\|\phi'(u^*)\|_\varepsilon \le \rho^* + \varepsilon$, cf. [1, Lem. 4.3.7]. With this result at hand, we can exploit the norm equivalence together with the definition of Fréchet derivatives to conclude the desired local convergence rate. This simple argument is elaborated e.g. in [32,50] and generalizes the previous findings obtained in [40]. □

In the finite dimensional case and under some additional assumptions on the structure of $\phi'(u^*)$ it can be shown that $\|u^n - u^*\|_V \le C \, |\rho^*|^n \|u^0 - u^*\|_V$, cf. [42, Th. 10.1.3 and 10.1.4; NR 10.1-5].

In the spirit of Proposition 1 we need to study the spectrum of $\phi'(u^*)$ to conclude local convergence.

3.1 Derivative of ϕ

Since we interpret ψ and ϕ as operators from V to V, we have likewise for the first Fréchet derivative in $v \in V \setminus \{0\}$, $\psi'(v\,;\,\cdot)$, $\phi'(v\,;\,\cdot)\colon V \to V$. In order to apply Proposition 1, we need to compute the Fréchet derivative of ϕ in u^*. As a first step, we compute the derivative of the mapping $v \mapsto \|\psi(v)\|$, leading to

$$\langle \partial_v \|\psi(v)\|, w \rangle = \frac{(\psi'(v\,;\,w),\,\psi(v))_H}{\|\psi(v)\|} = (\psi'(v\,;\,w),\,\phi(v))_H$$

or, more compactly, $\partial_v \|\psi(v)\| = (\psi'(v),\,\phi(v))_H$. In order to compute $\phi'(v)$, we write $\psi'(v)$ in the form

$$\psi'(v) = \partial_v \left(\phi(v)\|\psi(v)\|\right) = \|\psi(v)\|\,\phi'(v) + (\psi'(v),\,\phi(v))_H \,\phi(v)$$

and conclude that

$$\phi'(v) = \frac{\psi'(v)}{\|\psi(v)\|} - \frac{1}{\|\psi(v)\|}(\psi'(v),\,\phi(v))_H \,\phi(v).$$

For $v = u^*$ we can exploit $\phi(u^*) = u^*$, $\|u^*\| = 1$, and $u^* = (\lambda^* + \sigma)\,\psi(u^*)$. This implies $\|\psi(u^*)\| = |\lambda^* + \sigma|^{-1}$ and thus,

$$\phi'(u^*) = |\lambda^* + \sigma| \left(\psi'(u^*) - (\psi'(u^*),\,u^*)_H \,u^*\right).$$

To compute $\psi'(u^*)$, in turn, we use the identity

$$\mathcal{I} = \partial_v(\mathcal{I}v) = \partial_v\left((J(v) + \sigma\mathcal{I})\psi(v)\right) = J'(v\,;\,\cdot)\,\psi(v) + (J(v) + \sigma\mathcal{I})\,\psi'(v).$$

Note that $J'(v) \colon V \times V \to V^*$ denotes here the Fréchet derivative of J in a fixed element $v \in V \setminus \{0\}$. We conclude

$$\psi'(u^*) = J_\sigma(u^*)^{-1}\mathcal{I} - J_\sigma(u^*)^{-1}J'(u^*; \cdot)\,\psi(u^*).$$

Next, we want to verify that $J'(u^*; \cdot)\psi(u^*) = 0$. For that, we consider the definition of J' which yields

$$
\begin{aligned}
J'(u^*; w)\,u^* &= \lim_{t \to 0} \frac{J(u^* + tw)\,u^* - J(u^*)\,u^*}{t} \\
&= \lim_{t \to 0} \frac{J(u^* + tw)(u^* + tw) - J(u^* + tw)(tw) - J(u^*)u^*}{t} \\
&= \lim_{t \to 0} \frac{\mathcal{A}(u^* + tw) - \mathcal{A}(u^*) - J(u^* + tw)(tw)}{t} \\
&= \lim_{t \to 0} \left(\mathcal{A}'(u^*)w - J(u^* + tw)w \right) \stackrel{(4)}{=} 0.
\end{aligned}
$$

Here we also used that in a neighborhood of u^* (which excludes the zero element due to $\|u^*\| = 1$) the operator J is continuous. In summary, we proved $J'(u^*; w)\,u^* = 0$ for all $w \in V$. Since u^* equals $\psi(u^*)$ up to a multiplicative constant, we conclude that $J'(u^*; \cdot)\,\psi(u^*) = 0$ and hence, $\psi'(u^*) = J_\sigma(u^*)^{-1}\mathcal{I}$. Plugging this into previous estimates, we directly obtain

$$\phi'(u^*) = |\lambda^* + \sigma| \left(J_\sigma(u^*)^{-1}\mathcal{I} - (J_\sigma(u^*)^{-1}\mathcal{I}, u^*)_H\, u^* \right). \tag{10}$$

3.2 Local convergence results

Since the J-method is of the form $u^{n+1} = \phi(u^n)$, we recall from Proposition 1 that the local convergence rate is strongly connected to the spectrum of $\phi'(u^*)$. Besides the modified expression in (10), the analysis of this spectrum also requires the following orthogonality result.

Lemma 1 *Let $(u, \lambda) \in V \times \mathbb{R}$ be an eigenpair of $J(u^*)$. Further, let $(w, \mu) \in V \times \mathbb{R}$ be an adjoint-eigenpair of $J(u^*)$ with $\mu \neq \lambda$, i.e.,*

$$\langle J(u^*)\,v, w \rangle = \mu\,(v, w)_H$$

for all $v \in V$. Then, it holds that $(u, w)_H = 0$.

Proof By definition we know that $\langle J(u^*)\,u, v \rangle = \lambda\,(u, v)_H$ for all $v \in V$. Thus, with the test function $v = w$ we have

$$\lambda\,(u, w)_H = \langle J(u^*)\,u, w \rangle = \mu\,(u, w)_H.$$

The assumption $\mu \neq \lambda$ then directly implies the H-orthogonality of primal and adjoint eigenfunctions. $\qquad\square$

We are prepared to study the spectrum of the linear operator $\phi'(u^*)\colon V \to V$ to make conclusions on the convergence rate using Proposition 1. The corresponding eigenvalue problem reads: find $z \in V\backslash\{0\}$ and $\mu \in \mathbb{R}$ so that

$$\langle \phi'(u^*)z, v \rangle = \mu\,(z, v)_H$$

for all $v \in V$. The subsequent lemma relates the spectra of $\phi'(u^*)$ and $J(u^*)$. Recall that $J(u^*)$ is a linear, bounded operator over $\mathbb{R}$. Thus, its (real) spectrum coincides with the (real) spectrum of the corresponding adjoint operator [37].

Lemma 2 *Assume that the (real) eigenvalues of $J(u^*)$ are countable and that there exists a basis of corresponding adjoint-eigenfunctions of V. Let the eigenvalue of interest λ^* be simple and $\sigma \in \mathbb{R}$ a shift such that $J_\sigma(u^*)$ is invertible (note that in particular we have $\sigma \neq -\lambda^*$). Then, the spectra of $J(u^*)$ and $\phi'(u^*)$ are connected in the following sense: $\mu \neq \lambda^*$ is an eigenvalue of $J(u^*)$ if and only if $|\lambda^*+\sigma|/(\mu+\sigma)$ is an eigenvalue of $\phi'(u^*)$ and the eigenvalue λ^* of $J(u^*)$ corresponds to the eigenvalue 0 of $\phi'(u^*)$.*

Proof Due to the assumption on the shift, $J_\sigma(u^*)$ is invertible and $J_\sigma(u^*)^{-1}\mathcal{I}$ has the eigenvalues $(\lambda_1+\sigma)^{-1}, (\lambda_2+\sigma)^{-1}, \ldots$, including $(\lambda^*+\sigma)^{-1}$ with eigenfunction u^*. Note that the eigenvalues are shifted but the primal- and adjoint-eigenfunctions are the same as for $J(u^*)$. We first consider the eigenvalue $(\lambda^*+\sigma)^{-1}$ of $J_\sigma(u^*)^{-1}\mathcal{I}$ with eigenfunction u^*. Here we note that

$$\phi'(u^*)\,u^* \stackrel{(10)}{=} |\lambda^*+\sigma|\left(J_\sigma(u^*)^{-1}\mathcal{I}u^* - (J_\sigma(u^*)^{-1}\mathcal{I}u^*, u^*)_H\,u^*\right) = 0,$$

i.e., 0 is the corresponding eigenvalue of $\phi'(u^*)$. Now let $(\mu+\sigma)^{-1}$ be an eigenvalue of $J_\sigma(u^*)^{-1}\mathcal{I}$ with $\mu \neq \lambda^*$. Since $J_\sigma(u^*)^{-1}\mathcal{I}$ is a linear, bounded operator over $\mathbb{R}$, its (real) spectrum coincides with the (real) spectrum of the corresponding adjoint operator (cf. [37]) such that there exists a corresponding adjoint-eigenfunction w. By definition this means that

$$\left(J_\sigma(u^*)^{-1}\mathcal{I}v, w\right)_H = \frac{1}{\mu+\sigma}\,(v, w)_H$$

for all $v \in V$. Using that w is also an adjoint-eigenfunction of $J(u^*)$ and the orthogonality of Lemma 1, we find that

$$(\phi'(u^*)\,v, w)_H = |\lambda^*+\sigma|\left[\left(J_\sigma(u^*)^{-1}\mathcal{I}v, w\right)_H - \left(J_\sigma(u^*)^{-1}\mathcal{I}v, u^*\right)_H(u^*, w)_H\right]$$

$$= \frac{|\lambda^*+\sigma|}{\mu+\sigma}\,(v, w)_H.$$

Thus, w is an adjoint-eigenfunction of $\phi'(u^*)$ to the eigenvalue $|\lambda^*+\sigma|/(\mu+\sigma)$.

For the reverse direction let (z, μ) be an eigenpair of $\phi'(u^*)$, i.e., for all $v \in V$ we have

$$\mu\,(z, v)_H = (\phi'(u^*)\,z, v)_H$$
$$= |\lambda^* + \sigma|\left[(J_\sigma(u^*)^{-1}\mathcal{I}z, v)_H - (J_\sigma(u^*)^{-1}\mathcal{I}z, u^*)_H\,(u^*, v)_H\right].$$

For $\mu = 0$ this directly leads to $z = \pm u^*$, since $\mu = 0$ yields the relation $J_\sigma(u^*)^{-1}\mathcal{I}z = (J_\sigma(u^*)^{-1}\mathcal{I}z, u^*)_H\,u^*$. An application of $J_\sigma(u^*)$ then implies that u^* and z coincide up to a multiplicative constant. Since both functions are normalized, we get $z = u^*$ or $z = -u^*$. In the case $\mu \neq 0$ we employ the adjoint-eigenfunctions $v_1, v_2, \ldots$ of $J(u^*)$ as test functions. This yields

$$\tfrac{\mu}{|\lambda^*+\sigma|}\,(z, v_j)_H = \left(J_\sigma(u^*)^{-1}\mathcal{I}z, v_j\right)_H - (J_\sigma(u^*)^{-1}\mathcal{I}z, u^*)_H\,(u^*, v_j)_H.$$

Note that the second term on the right-hand side vanishes due to Lemma 1. The definition of being an adjoint-eigenfunction then leads to

$$\tfrac{\mu}{|\lambda^*+\sigma|}\,(z, v_j)_H = \left(J_\sigma(u^*)^{-1}\mathcal{I}z, v_j\right)_H = \tfrac{1}{\lambda_j+\sigma}\,(z, v_j)_H$$

for $j = 1, 2, \ldots$. Obviously, $(z, v_j)_H$ cannot vanish for all test functions. This implies that $\mu = |\lambda^*+\sigma|/(\lambda_j+\sigma)$ for a certain index j. In other words, μ equals an eigenvalue of $J_\sigma(u^*)^{-1}\mathcal{I}$ up to the factor $|\lambda^* + \sigma|$. □

With the obtained knowledge about the spectrum of $\phi'(u^*)$ we can directly apply Proposition 1 to deduce an abstract local convergence result with a rate depending on the spectrum of $J(u^*)$ relative to the implemented shift as in the matrix case.

Theorem 1 (abstract local convergence) *Consider the assumptions of Lemma 2 and let $\sigma \in \mathbb{R}$ be a shift sufficiently close to $-\lambda^*$. By μ we denote the (real) eigenvalue of $J(u^*)$ which is closest to $-\sigma$ but different from λ^*. If the shift is selected such that*

$$\rho^* := \frac{|\lambda^* + \sigma|}{|\mu + \sigma|} < 1,$$

then the iterations of the J-method (6) are locally convergent to u^* in the V-norm. Furthermore, for every $\varepsilon > 0$ there exists a neighborhood S_ε of u^* and a constant $C_\varepsilon > 0$ such that

$$\|u^n - u^*\|_V \le C_\varepsilon\,|\rho^* + \varepsilon|^n\,\|u^0 - u^*\|_V \quad \text{for all } u^0 \in S_\varepsilon$$

and $n \ge 1$.

Below, we apply this abstract result to the GPEVP of Example 1.

4 Quantified local convergence of the shifted *J*-method for the GPEVP

As already mentioned in the introduction, we want to apply the damped J-method to the GPEVP (1), cf. Example 1. Thus, the task is to find an eigenfunction $u^* \in H_0^1(\mathcal{D})$

with $\|u^*\|^2_{L^2(\mathcal{D})} = 1$ and corresponding eigenvalue $\lambda^* \in \mathbb{R}$ such that

$$-\Delta u^* + W u^* - \Omega \mathcal{L}_z u^* + \kappa\, |u^*|^2 u^* = \lambda^* u^*.$$

Recall that the potential satisfies $W \in L^\infty(\mathcal{D}, \mathbb{R})$, whereas $\kappa \geq 0$ regulates the nonlinearity. Furthermore, $\mathcal{L}_z$ is the angular momentum operator with angular velocity $\Omega \in \mathbb{R}$. With these properties it is easily seen that the GPEVP can only have real eigenvalues. In the following we will also assume that

$$W(x) \geq \Omega^2 |x|^2. \tag{11}$$

This condition can be interpreted as that trapping frequencies are larger than the angular frequency. Physically speaking, this ensures that centrifugal forces do not become too strong compared to the strength of the trapping potential W.

4.1 The Gross–Pitaevskii energy

We consider homogeneous Dirichlet boundary conditions and the short-hand notation $L^2(\mathcal{D}) := L^2(\mathcal{D}, \mathbb{C})$, $H^1_0(\mathcal{D}) := H^1_0(\mathcal{D}, \mathbb{C})$, and $H^{-1}(\mathcal{D}) := H^{-1}(\mathcal{D}, \mathbb{C})$ for the function spaces over $\mathbb{R}$, cf. the details in Example 1. We set

$$V := H^1_0(\mathcal{D}), \qquad H := L^2(\mathcal{D}), \qquad V^* := H^{-1}(\mathcal{D}).$$

This means that $\| \cdot \|$ equals the $L^2(\mathcal{D})$-norm and $(\cdot, \cdot)_H := \Re(\cdot, \cdot)_{L^2(\mathcal{D})}$, where $(u, v)_{L^2(\mathcal{D})} := \int_{\mathcal{D}} u\, \bar{v}\, dx$. Similarly, we equip V with the inner product $(\cdot, \cdot)_V := \Re(\cdot, \cdot)_{H^1(\mathcal{D})}$, where $(u, v)_{H^1(\mathcal{D})} := \int_{\mathcal{D}} \nabla u \cdot \overline{\nabla v}\, dx$. As before, the weak formulation is given by $\mathcal{A}(u^*) = A(u^*, u^*) = \lambda^* \mathcal{I} u^*$, where the nonlinear operator A is defined in (3). The eigenvalue problem is equivalent to finding the critical points (on the manifold associated with the L^2-normalization constraint) of the energy functional $E: V \to \mathbb{R}$ given by

$$\begin{aligned} E(u) &:= \frac{1}{2} \int_{\mathcal{D}} |\nabla u|^2 + W|u|^2 - \Omega\, \bar{u}\, \mathcal{L}_z u + \frac{\kappa}{2}|u|^4\, dx \\ &= \frac{1}{2} \int_{\mathcal{D}} |\nabla_{\!R} u|^2 + W_{\!R}|u|^2 + \frac{\kappa}{2}|u|^4\, dx, \end{aligned} \tag{12}$$

cf. [27]. Recalling that $W_{\!R}(x) = W(x) - \frac{1}{4}\Omega^2 |x|^2$ and considering assumption (11), we see that the energy functional is bounded from below by a positive constant $c = c(\Omega, \mathcal{D})$, i.e.,

$$E(u) \geq c > 0 \qquad \text{for all } u \in V \text{ with } \|u\| = 1.$$

Hence, E is weakly lower semi-continuous and bounded from below, which yields the existence of a minimizer that is typically called a ground state.

If $\Omega = 0$, i.e., in the absence of a rotating potential, then the GPEVP has infinitely many eigenvalues $0 < \lambda_1^* < \lambda_2^* \leq \lambda_3^* \leq \cdots < \infty$, where the ground state eigenvalue λ_1^* is simple, cf. [21,36]. If $\Omega \neq 0$, the smallest eigenvalue is typically no longer simple [15]. This case refers to the physical phenomenon of a broken symmetry, where the ground state can have different shapes which differ in their number and location of vortices.

4.2 The *J*-operator for the GPEVP

In order to formulate the J-method for the GPEVP, we need to compute the linearization $J(u)$ according to (4). Hence, by calculating the Fréchet derivative of $\mathcal{A}(u) = A(u, u)$ using (3), we obtain

$$\langle J(u)v, w \rangle = \Re \langle \mathcal{J}(u)v, w \rangle, \tag{13}$$

where

$$\begin{aligned}
\langle \mathcal{J}(u)v, w \rangle := & \int_{\mathcal{D}} \nabla_{\mathrm{R}} v \cdot \overline{\nabla_{\mathrm{R}} w} + W_{\mathrm{R}} v \overline{w} \, \mathrm{d}x \\
& + \frac{\kappa}{\|u\|^2} \int_{\mathcal{D}} (u\overline{v} + 2v\overline{u}) \, u\overline{w} \, \mathrm{d}x \\
& - \frac{2\kappa \, \Re(u, v)_{L^2(\mathcal{D})}}{\|u\|^4} \int_{\mathcal{D}} |u|^2 u\overline{w} \, \mathrm{d}x.
\end{aligned} \tag{14}$$

Note that the operator $J(u)$ induces an $\mathbb{R}$-bilinearform $\langle J(u) \cdot, \cdot \rangle$, i.e., we still consider V as an $\mathbb{R}$-vector space and have bilinearity only for multiplicative constants in $\mathbb{R}$. Recall that we consider the real-Fréchet derivative, which also fits in the standard framework used in quantum mechanics. Moreover, the complex-Fréchet derivative does not exist for the present example.

We will show coercivity of $J(u)$ up to a shift in Lemma 3 below. Before that, it is worth to mention that the eigenvalue problem can be equivalently expressed in terms of $\mathcal{J}(u)$, which also yields the typical structure with standard inner products. We have the following result.

Proposition 2 *Consider the GPEVP with full operator $\mathcal{J}$ given by (14) and its real part J given by (13). Then $\lambda^* \in \mathbb{R}$ is an eigenvalue with L^2-normalized eigenfunction $u^* \in V$, i.e.,*

$$\langle J(u^*), v \rangle = \lambda^* \Re(u^*, v)_{L^2(\mathcal{D})} \quad \text{for all } v \in V$$

if and only if

$$\langle \mathcal{J}(u^*), v \rangle = \lambda^* (u^*, v)_{L^2(\mathcal{D})} \quad \text{for all } v \in V.$$

Proof If $(\lambda^*, u^*) \in \mathbb{R} \times V$ solves $\langle \mathcal{J}(u^*), v \rangle = \lambda^* (u^*, v)_{L^2(\mathcal{D})}$, then we can take the real part on both sides and obtain $\langle J(u^*), v \rangle = \lambda^* \Re(u^*, v)_{L^2(\mathcal{D})}$. Vice versa, if

$(\lambda^*, u^*) \in \mathbb{R} \times V$ solves $\Re\langle \mathcal{J}(u^*), v \rangle = \lambda^* \Re(u^*, v)_{L^2(\mathcal{D})}$, then we can use test functions of the form $\mathrm{i}v \in V$ to obtain $\Im\langle \mathcal{J}(u^*), v \rangle = \lambda^* \Im(u^*, v)_{L^2(\mathcal{D})}$. Multiplying the second equation with the complex number i and adding it to the first equation readily yields $\langle \mathcal{J}(u^*), v \rangle = \lambda^*(u^*, v)_{L^2(\mathcal{D})}$. $\qquad\square$

The proposition shows that we can either interpret the eigenvalue problem with real parts only (i.e., with J) or with the full operator $\mathcal{J}$. With analogous arguments, we can also prove that $J(u^*)^{-1}\mathcal{I} = \mathcal{J}(u^*)^{-1}\mathcal{I}^{\mathbb{C}}$, where

$$\mathcal{I}^{\mathbb{C}}v := (v, \,\cdot\,)_{L^2(\mathcal{D})}. \tag{15}$$

This justifies that we can interpret the iterations of the damped J-method (7) equivalently with J or $\mathcal{J}$. In particular, we have the following conclusion.

Conclusion 2 *Consider the GPEVP with full operator $\mathcal{J}$ given by* (14) *and its real part J defined in* (13). *Given $u^n \in V$ with $\|u^n\| = 1$ and a step size $\tau_n > 0$, we assume that $J_\sigma(u^n)^{-1}$ is invertible. Then the iterations of the J-method* (7) *can be equivalently characterized by the $\mathcal{J}$-iteration*

$$\tilde{u}^{n+1} = (1 - \tau_n)u^n + \tau_n \, \gamma_n^{\mathbb{C}} \, \mathcal{J}_\sigma(u^n)^{-1}\mathcal{I}^{\mathbb{C}}u^n, \qquad u^{n+1} = \frac{\tilde{u}^{n+1}}{\|\tilde{u}^{n+1}\|} \tag{16}$$

with $\mathcal{J}_\sigma(u) := \mathcal{J}(u) + \sigma\mathcal{I}^{\mathbb{C}}$ and $(\gamma_n^{\mathbb{C}})^{-1} := (\mathcal{J}_\sigma(u^n)^{-1}\mathcal{I}^{\mathbb{C}}u^n, u^n)_{L^2(\mathcal{D})}$. The assumed existence of $J_\sigma(u^n)^{-1}\mathcal{I}$ implies the existence (and uniqueness) of $\mathcal{J}_\sigma(u^n)^{-1}\mathcal{I}^{\mathbb{C}}$.

4.3 Local convergence

Finally, we apply Theorem 1 to the GPEVP.

Theorem 3 (quantified convergence for the GPEVP) *Consider the GPEVP as described in Example 1 and let u^n denote the iterates generated by the shifted J-method (without damping), i.e.,*

$$u^{n+1} = \phi(u^n) = \frac{J_\sigma(u^n)^{-1}\mathcal{I}u^n}{\|J_\sigma(u^n)^{-1}\mathcal{I}u^n\|}.$$

By $u^* \in V = H_0^1(\mathcal{D})$ we denote an L^2-normalized eigenfunction to (1) *with eigenvalue λ^*. Assume that λ^* is a simple eigenvalue of $J(u^*)$ and that the shift $\sigma \neq -\lambda^*$ is such that $J_\sigma(u^*)$ has a bounded inverse. Further, let the shift σ be sufficiently close to $-\lambda^*$ and let μ be the eigenvalue of $J_\sigma(u^*)$ so that $|\mu + \sigma|$ is minimal and*

$$\frac{|\lambda^* + \sigma|}{|\mu + \sigma|} < 1.$$

Then there exists a neighborhood $S \subset V$ of u^* such that for all starting values $u^0 \in S$ we have

$$\lim_{n \to \infty} \|u^n - u^*\|_V = 0.$$

Furthermore, for every $\varepsilon > 0$ there is a neighborhood $S_\varepsilon \subset V$ of u^* and a constant $C(\varepsilon) > 0$ such that for all $u^0 \in S_\varepsilon$ (and $n \geq 1$) it holds

$$\|u^n - u^*\|_V \ \leq \ C(\varepsilon) \left(\frac{|\lambda^* + \sigma|}{|\mu + \sigma|} + \varepsilon \right)^n \|u^0 - u^*\|_V.$$

Hence, if the shift σ is close to $-\lambda^*$ and u^n close to u^*, then we have locally a linear convergence with rate

$$\frac{|\lambda^* + \sigma|}{|\mu + \sigma|} + \varepsilon < 1.$$

Proof By definition of the problem (cf. Example 1) all eigenvalues are real. Furthermore, the spectrum is countable and does not have an accumulation point in $\mathbb{C}$. This is a direct consequence from the observation that the operator $J_\sigma(u^*)^{-1}\mathcal{I} \colon L^2(\mathcal{D}) \to L^2(\mathcal{D})$ is compact for all shifts $-\sigma$ that are not in the spectrum of $J(u^*)$ (i.e., we have compact resolvents). Hence, we can apply Theorem 1 which readily proves the result. □

In the finite dimensional case based on a finite difference discretization, a corresponding result was presented in [38]. Motivated by the above convergence result, a practical realization of the iterations can be based on the more natural formulation of the J-method given by (16). This is also the version for which we discuss the implementation in "Appendix B".

5 Global convergence of the damped *J*-method for the GPEVP

In this section we come back to the question of invertibility of the operator J (which hence also implies invertibility of $\mathcal{J}$). This will then lead to a globally convergent method.

5.1 Coercivity of the shifted *J*-operator

We first show that the operator J is – up to a shift – coercive.

Lemma 3 *Given $u \in V$ and assumption (11), the operator $J(u)$ corresponding to the Gross–Pitaevskii operator satisfies a Gårding inequality. More precisely, for any $\sigma \geq \frac{\kappa}{3}\|u\|_{L^4(\mathcal{D})}^4/\|u\|^4$ the bilinear form*

$$\langle (J(u) + \sigma\mathcal{I}) \cdot , \cdot \rangle \colon V \times V \to \mathbb{R}$$

is coercive and thus, the operator $J_\sigma(u) = J(u) + \sigma\mathcal{I}\colon V \to V^*$ is invertible.

Proof Consider $v \in V$. We start with considering the rotational gradient, for which we observe with Young's inequality that

$$\int_{\mathcal{D}} |\nabla_{\mathrm{R}} v|^2 \, dx \geq \frac{1}{2} \int_{\mathcal{D}} |\nabla v|^2 \, dx - \frac{3}{4}\Omega^2 \int_{\mathcal{D}} |x|^2 |v|^2 \, dx.$$

Hence, with $W_{\mathrm{R}} = W(x) - \frac{1}{4}\Omega^2 |x|^2$ and assumption (11) we have

$$\int_{\mathcal{D}} |\nabla_{\mathrm{R}} v|^2 + W_{\mathrm{R}} |v|^2 \, dx \geq \frac{1}{2} \int_{\mathcal{D}} |\nabla v|^2 \, dx. \tag{17}$$

This leads to

$$
\begin{aligned}
\langle J(u)v, v \rangle &= \int_{\mathcal{D}} |\nabla_{\mathrm{R}} v|^2 + W_{\mathrm{R}} |v|^2 \, dx + \frac{\kappa}{\|u\|^2} \int_{\mathcal{D}} |u|^2 |v|^2 \, dx \\
&\quad + \frac{2\kappa}{\|u\|^2} \int_{\mathcal{D}} |\Re(u\bar{v})|^2 \, dx - 2\kappa \frac{\Re(u,v)_{L^2(\mathcal{D})}}{\|u\|^4} \int_{\mathcal{D}} |u|^2 \Re(u\bar{v}) \, dx \\
&\overset{(17)}{\geq} \frac{1}{2} \|\nabla v\|^2 + \frac{\kappa}{\|u\|^2} \int_{\mathcal{D}} |u\bar{v}|^2 \, dx + \frac{2\kappa}{\|u\|^2} \int_{\mathcal{D}} |\Re(u\bar{v})|^2 \, dx \\
&\quad - 2\kappa \frac{\Re(u,v)_{L^2(\mathcal{D})}}{\|u\|^4} \int_{\mathcal{D}} |u|^2 \Re(u\bar{v}) \, dx \\
&\geq \frac{1}{2} \|\nabla v\|^2 + 3\frac{\kappa}{\|u\|^2} \int_{\mathcal{D}} |\Re(u\bar{v})|^2 \, dx - 2\kappa \frac{\Re(u,v)_{L^2(\mathcal{D})}}{\|u\|^4} \int_{\mathcal{D}} |u|^2 \Re(u\bar{v}) \, dx.
\end{aligned} \tag{18}
$$

To estimate the negative part, we apply once more Young's inequality with some parameter $\mu > 0$ and the Cauchy Schwarz inequality to get

$$
\begin{aligned}
2\,\Re(u,v)_{L^2(\mathcal{D})} \int_{\mathcal{D}} |u|^2 \Re(u\bar{v}) \, dx &\leq \frac{1}{\mu} |\Re(u,v)_{L^2(\mathcal{D})}|^2 + \mu \left(\int_{\mathcal{D}} |u|^2 \Re(u\bar{v}) \, dx \right)^2 \\
&\leq \frac{1}{\mu} \|u\|^2 \|v\|^2 + \mu \|u\|_{L^4(\mathcal{D})}^4 \int_{\mathcal{D}} |\Re(u\bar{v})|^2 \, dx.
\end{aligned}
$$

Using this estimate in (18) with the particular choice $\mu = 3\|u\|^2 / \|u\|_{L^4(\mathcal{D})}^4$, we obtain

$$
\begin{aligned}
\langle J(u)v, v \rangle &\geq \frac{1}{2} \|\nabla v\|^2 + \frac{3\kappa}{\|u\|^2} \int_{\mathcal{D}} |\Re(u\bar{v})|^2 \, dx - \frac{\kappa}{3} \frac{\|u\|_{L^4(\mathcal{D})}^4 \|v\|^2}{\|u\|^4} - \frac{3\kappa}{\|u\|^2} \int_{\mathcal{D}} |\Re(u\bar{v})|^2 \, dx \\
&= \frac{1}{2} \|\nabla v\|^2 - \frac{\kappa}{3} \frac{\|u\|_{L^4(\mathcal{D})}^4}{\|u\|^4} \|v\|^2.
\end{aligned}
$$

Thus, we conclude

$$\langle J(u)v, v \rangle \geq \frac{1}{2} \|v\|_V^2 - \sigma \|v\|^2 = \frac{1}{2} \|v\|_V^2 - \sigma \langle \mathcal{I}v, v \rangle.$$

$\square$

Recall the definition of the energy E in Sect. 4.1. Motivated by the previous lemma, we also define the shifted energy $E_\sigma(u) := E(u) + \frac{1}{2}\sigma \|u\|^2$ such that $E_0(u) = E(u)$. For $u \in V$ with normalization constraint $\|u\| = 1$, a sufficient shift in the sense of Lemma 3 is thus given by

$$\sigma := \frac{4}{3} E(u) \geq \frac{\kappa}{3} \|u\|_{L^4(\mathcal{D})}^4.$$

Note that for a normalized function u we can express the Rayleigh quotient in terms of the energy by

$$\lambda(u) := \langle \mathcal{A}(u), u \rangle = 2E(u) + \frac{\kappa}{2} \|u\|_{L^4(\mathcal{D})}^4.$$

In particular, this formula relates the eigenvalues with the energies of the eigenfunctions.

5.2 Feasibility of the *J*-method

For the feasibility of the damped J-method, we need to guarantee a priori that $J_\sigma(u^n)$ stays invertible throughout the iteration process. We fix the shift σ in the beginning of the iteration, e.g., by $\sigma := \frac{4}{3} E(u^0)$. Now, the aim is to show that the energy of the iterates does not increase such that $\sigma \geq \frac{4}{3} E(u^n)$ for all $n \geq 0$.

Lemma 4 *Recall the shifted energy $E_\sigma(u) := E(u) + \frac{1}{2}\sigma \|u\|^2$ with $E(u)$ being defined in (12). Further recall the definition of the J-method (7) with iteration steps $\tilde{u}^{n+1} = (1 - \tau)u^n + \tau \gamma_n J_\sigma(u^n)^{-1} \mathcal{I}u^n$. Then, for any step size $\tau \leq \frac{1}{2}$ we have the following guaranteed estimate for the difference of the shifted energies*

$$\begin{aligned} E_\sigma(u^n) - E_\sigma(\tilde{u}^{n+1}) \geq &-\frac{3\kappa}{2} \int_{\mathcal{D}} |\tilde{u}^{n+1} - u^n|^4 \, dx \\ &+ \left(\tfrac{1}{\tau} - \tfrac{1}{2}\right) \int_{\mathcal{D}} |\nabla_R(\tilde{u}^{n+1} - u^n)|^2 \\ &+ (\sigma + W_R)|(\tilde{u}^{n+1} - u^n)|^2 \, dx. \end{aligned} \tag{19}$$

Proof We start by establishing a couple of identities for different evaluations of J. Here we exploit the L^2-orthogonality (8), i.e., $(u^n, \tilde{u}^{n+1} - u^n) = 0$. Together with $\|u^n\| = 1$ this implies $(u^n, \tilde{u}^{n+1}) = 1$. Using these facts, we observe that

$$\langle J(u^n)\tilde{u}^{n+1}, \tilde{u}^{n+1} \rangle$$

$$= \int_{\mathcal{D}} |\nabla_R \tilde{u}^{n+1}|^2 + W_R |\tilde{u}^{n+1}|^2 \, dx + \kappa \int_{\mathcal{D}} |\tilde{u}^{n+1}|^2 \, |u^n|^2 \, dx$$

$$+ 2\kappa \int_{\mathcal{D}} \left(\Re(u^n \, \overline{\tilde{u}^{n+1}}) - |u^n|^2 \right) \Re(u^n \, \overline{\tilde{u}^{n+1}}) \, dx$$

$$= 2E(\tilde{u}^{n+1}) + \kappa \int_{\mathcal{D}} |\tilde{u}^{n+1}|^2 \left(|u^n|^2 - |\tilde{u}^{n+1}|^2 \right) dx$$

$$+ \frac{\kappa}{2} \int_{\mathcal{D}} |\tilde{u}^{n+1}|^4 \, dx$$

$$+ 2\kappa \int_{\mathcal{D}} \left(\Re(u^n \, \overline{\tilde{u}^{n+1}}) - |u^n|^2 \right) \Re(u^n \, \overline{\tilde{u}^{n+1}}) \, dx. \tag{20}$$

Next, using again the L^2-orthogonality together with the definition of $\tilde{u}^{n+1}$ in (7), we have

$$\frac{1}{\tau} \langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n \rangle$$

$$= -\langle J_\sigma(u^n) u^n, \tilde{u}^{n+1} - u^n \rangle + \gamma^n \langle J_\sigma(u^n) J_\sigma(u^n)^{-1} \mathcal{I} u^n, \tilde{u}^{n+1} - u^n \rangle$$

$$= -\langle J_\sigma(u^n) u^n, \tilde{u}^{n+1} - u^n \rangle + \gamma^n \langle u^n, \tilde{u}^{n+1} - u^n \rangle$$

$$= -\langle J_\sigma(u^n) u^n, \tilde{u}^{n+1} - u^n \rangle. \tag{21}$$

With this equality, we conclude

$$\langle J_\sigma(u^n) u^n, u^n \rangle = \langle J_\sigma(u^n) u^n, \tilde{u}^{n+1} \rangle - \langle J_\sigma(u^n) u^n, \tilde{u}^{n+1} - u^n \rangle$$

$$\overset{(21)}{=} \langle J_\sigma(u^n) u^n, \tilde{u}^{n+1} \rangle + \frac{1}{\tau} \langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n \rangle$$

$$= \langle J_\sigma(u^n) \tilde{u}^{n+1}, \tilde{u}^{n+1} \rangle - \langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), u^n \rangle$$

$$+ \left(\frac{1}{\tau} - 1 \right) \langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n \rangle. \tag{22}$$

Here we need to have a closer look at the term $\langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), u^n \rangle$. By the definition of $J(u^n)$ it is easily seen that

$$\langle J(u^n)(\tilde{u}^{n+1} - u^n), u^n \rangle - \langle J(u^n) u^n, \tilde{u}^{n+1} - u^n \rangle$$

$$= 2\kappa \int_{\mathcal{D}} |u^n|^2 \Re \left((\tilde{u}^{n+1} - u^n) \overline{u^n} \right) dx. \tag{23}$$

Plugging this into (22) yields for the shifted energy

$$2E_\sigma(u^n)$$

$$= \langle J_\sigma(u^n) u^n, u^n \rangle - \frac{\kappa}{2} \int_{\mathcal{D}} |u^n|^4 \, dx$$

$$\overset{(22)}{=} \langle J_\sigma(u^n) \tilde{u}^{n+1}, \tilde{u}^{n+1} \rangle - \langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), u^n \rangle$$

$$+ \left(\frac{1}{\tau} - 1 \right) \langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n \rangle - \frac{\kappa}{2} \int_{\mathcal{D}} |u^n|^4 \, dx$$

$$\overset{(21),(23)}{=} \langle J_\sigma(u^n)\tilde{u}^{n+1}, \tilde{u}^{n+1}\rangle + \tfrac{1}{\tau}\langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n\rangle$$

$$- \frac{\kappa}{2}\int_{\mathcal{D}} |u^n|^4\, \mathrm{d}x + (\tfrac{1}{\tau} - 1)\langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n\rangle$$

$$- 2\kappa \int_{\mathcal{D}} |u^n|^2\Re\left((\tilde{u}^{n+1} - u^n)\overline{u^n}\right)\, \mathrm{d}x$$

$$\overset{(20)}{=} 2E_\sigma(\tilde{u}^{n+1}) - \kappa \int_{\mathcal{D}} |\tilde{u}^{n+1}|^2\, (|\tilde{u}^{n+1}|^2 - |u^n|^2)\, \mathrm{d}x + \frac{\kappa}{2}\int_{\mathcal{D}} |\tilde{u}^{n+1}|^4\, \mathrm{d}x$$

$$- \frac{\kappa}{2}\int_{\mathcal{D}} |u^n|^4\, \mathrm{d}x - 2\kappa \int_{\mathcal{D}} |u^n|^2\Re\left((\tilde{u}^{n+1} - u^n)\overline{u^n}\right)\, \mathrm{d}x$$

$$+ 2\kappa \int_{\mathcal{D}} \left(\Re\left(u^n\, \overline{\tilde{u}^{n+1}}\right) - |u^n|^2\right)\Re(u^n\, \overline{\tilde{u}^{n+1}})\, \mathrm{d}x$$

$$+ (\tfrac{2}{\tau} - 1)\langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n\rangle.$$

Since $\Re z = \Re\overline{z}$ for all $z \in \mathbb{C}$, we conclude that

$$2E_\sigma(u^n) - 2E_\sigma(\tilde{u}^{n+1})$$

$$= -\kappa \int_{\mathcal{D}} |\tilde{u}^{n+1}|^2\, (|\tilde{u}^{n+1}|^2 - |u^n|^2)\, \mathrm{d}x + \frac{\kappa}{2}\int_{\mathcal{D}} |\tilde{u}^{n+1}|^4\, \mathrm{d}x - \frac{\kappa}{2}\int_{\mathcal{D}} |u^n|^4\, \mathrm{d}x$$

$$+ 2\kappa \int_{\mathcal{D}} \left(|u^n|^2 - \Re(\tilde{u}^{n+1}\overline{u^n})\right)^2\, \mathrm{d}x\ +\ (\tfrac{2}{\tau} - 1)\, \langle J_\sigma(u^n)(\tilde{u}^{n+1} - u^n), \tilde{u}^{n+1} - u^n\rangle. \tag{24}$$

Using that

$$\langle J(u^n)(\tilde{u}^{n+1} - u^n), (\tilde{u}^{n+1} - u^n)\rangle = \int_{\mathcal{D}} |\nabla_{\mathrm{R}}(\tilde{u}^{n+1} - u^n)|^2 + W_{\mathrm{R}}\, |\tilde{u}^{n+1} - u^n|^2\, \mathrm{d}x$$

$$+ \kappa \int_{\mathcal{D}} |\Re(u^n\, (\overline{\tilde{u}^{n+1} - u^n}))|^2 + |u^n|^2\, |\tilde{u}^{n+1} - u^n|^2\, \mathrm{d}x, \tag{25}$$

and $\Re(u^n(\overline{\tilde{u}^{n+1} - u^n})) = \Re(\tilde{u}^{n+1}\overline{u^n}) - |u^n|^2$ we see that

$$2E_\sigma(u^n) - 2E_\sigma(\tilde{u}^{n+1})$$

$$\overset{(24),(25)}{\geq} -\kappa \int_{\mathcal{D}} |\tilde{u}^{n+1}|^2(|\tilde{u}^{n+1}|^2 - |u^n|^2)\, \mathrm{d}x + \frac{\kappa}{2}\int_{\mathcal{D}} |\tilde{u}^{n+1}|^4\, \mathrm{d}x - \frac{\kappa}{2}\int_{\mathcal{D}} |u^n|^4\, \mathrm{d}x$$

$$+ (\tfrac{2}{\tau} - 1)\int_{\mathcal{D}} |\nabla_{\mathrm{R}}(\tilde{u}^{n+1} - u^n)|^2 + (W_{\mathrm{R}} + \sigma)\, |(\tilde{u}^{n+1} - u^n)|^2\, \mathrm{d}x$$

$$+ \kappa\, (\tfrac{2}{\tau} - 1)\int_{\mathcal{D}} |u^n|^2|\tilde{u}^{n+1} - u^n|^2\, \mathrm{d}x.$$

Finally, an application of the triangle and Young's inequality yields the estimate

$$-2 \int_{\mathcal{D}} |\tilde{u}^{n+1}|^2 (|\tilde{u}^{n+1}|^2 - |u^n|^2) \, dx - \int_{\mathcal{D}} |u^n|^4 \, dx + \int_{\mathcal{D}} |\tilde{u}^{n+1}|^4 \, dx$$

$$= - \int_{\mathcal{D}} |\tilde{u}^{n+1} + u^n|^2 |\tilde{u}^{n+1} - u^n|^2 \, dx$$

$$\geq - \int_{\mathcal{D}} 3 \, |\tilde{u}^{n+1} - u^n|^4 + 6 \, |u^n|^2 |\tilde{u}^{n+1} - u^n|^2 \, dx,$$

which then implies the desired inequality if $\tau \leq \frac{1}{2}$. $\qquad\square$

With this result we are now in the position to prove uniform boundedness of the energy of the iterates and even a reduction of the energy.

Theorem 4 (energy reduction) *Given $u^0 \in V$ with $\|u^0\| = 1$, we let $\sigma \geq \frac{4}{3} E(u^0)$ and $W(x) \geq \Omega^2 |x|^2$. Then, there exists a $\tau^* > 0$ (that only depends on u^0 and its energy) such that for all $0 < \tau_n \leq \tau^*$ the sequence obtained by the damped J-method (7) is well-posed and strictly energy diminishing for all n, i.e., it holds*

$$E(u^{n+1}) \leq E(u^n) \leq E(u^0),$$

where $E(u^{n+1}) < E(u^n)$ if u^n is not already a critical point of E and thus an eigenstate.

Proof We proceed inductively. From estimate (17) in the proof of Lemma 3 we know that $J_\sigma(u^0)$ is coercive with constant $1/2$, i.e., $\|\nabla v\|^2 = \|v\|_V^2 \leq 2 \langle J_\sigma(u^0) v, v \rangle$. Together with (21) and the L^2-orthogonality (8), this implies

$$\frac{1}{2\tau_0} \int_{\mathcal{D}} |\nabla \tilde{u}^1 - \nabla u^0|^2 \, dx$$

$$\leq - \langle J_\sigma(u^0) u^0, \tilde{u}^1 - u^0 \rangle$$

$$= \Re \left(\int_{\mathcal{D}} \nabla_R u^0 \cdot \overline{\nabla_R (u^0 - \tilde{u}^1)} + W_R u^0 \overline{(u^0 - \tilde{u}^1)} \, dx + \kappa \int_{\mathcal{D}} |u^0|^2 u^0 \overline{(u^0 - \tilde{u}^1)} \, dx \right)$$

$$\leq \|\nabla_R u^0\| \, \|\nabla_R (\tilde{u}^1 - u^0)\| + \|\sqrt{W_R} u^0\| \, \|\sqrt{W_R} (\tilde{u}^1 - u^0)\| + \kappa \|u^0\|_{L^6(\mathcal{D})}^3 \|\tilde{u}^1 - u^0\|$$

$$\leq C \sqrt{E(u^0)} \, \|\nabla(\tilde{u}^1 - u^0)\|.$$

Hence, we get

$$\|\nabla \tilde{u}^1 - \nabla u^0\|^2 = \int_{\mathcal{D}} |\nabla \tilde{u}^1 - \nabla u^0|^2 \, dx \leq 4 \, \tau_0^2 \, C^2 E(u^0) =: \tau_0^2 \, C_0.$$

Assume that $\tau_0 \leq \tau^* \leq 2$, where τ^* is selected sufficiently small compared to C_0. Then we have that $\|\nabla \tilde{u}^1 - \nabla u^0\| < 1$ and the Sobolev embedding of $L^4(\mathcal{D}) \hookrightarrow H_0^1(\mathcal{D})$ with constant C_S implies that

$$\int_{\mathcal{D}} |\tilde{u}^1 - u^0|^4 \, dx \leq C_S \int_{\mathcal{D}} |\nabla \tilde{u}^1 - \nabla u^0|^2 \, dx.$$

Consequently, we can use (19) and (17) to observe that the energy difference fulfills

$$
\begin{aligned}
E_\sigma(u^0) &- E_\sigma(\tilde{u}^1) \\
&\geq -\frac{3\kappa}{2} \int_{\mathcal{D}} |\tilde{u}^1 - u^0|^4 \, dx + \frac{1}{2} \left(\frac{1}{\tau_0} - \frac{1}{2} \right) \int_{\mathcal{D}} |\nabla(\tilde{u}^1 - u^0)|^2 \, dx \\
&\quad + \sigma \left(\frac{1}{\tau_0} - \frac{1}{2} \right) \int_{\mathcal{D}} |\tilde{u}^1 - u^0|^2 \, dx \\
&\geq \frac{1}{2} \left(\frac{1}{\tau_0} - \frac{1}{2} - 3\kappa \, C_S \right) \int_{\mathcal{D}} |\nabla(\tilde{u}^1 - u^0)|^2 \, dx \\
&\quad + \sigma \left(\frac{1}{\tau_0} - \frac{1}{2} \right) \int_{\mathcal{D}} |\tilde{u}^1 - u^0|^2 \, dx.
\end{aligned}
\tag{26}
$$

Hence, if

$$
\tau_0 \leq \tau^* < \min \left\{ \frac{2}{1 + 6\kappa C_S}, \; C_0^{-1/2}, \; \frac{1}{2} \right\},
$$

then we have

$$
E_\sigma(u^1) \leq E_\sigma(\tilde{u}^1) \leq E_\sigma(u^0),
$$

where we have used that the iterations increase the mass intermediately, i.e., $\|\tilde{u}^1\| \geq 1$ as shown in (9). Note that since u^0 and u^1 are normalized in $L^2(\mathcal{D})$, we can drop the shift, leading to $E(u^1) \leq E(u^0)$. Hence, we have $\sigma \geq \frac{4}{3} E(u_0) \geq \frac{4}{3} E(u_1)$ and Lemma 3 guarantees that $J_\sigma(u^1)$ is still coercive. Inductively, we can repeat the arguments for u^n with the same generic constant C_0 to show that for $\tau^n \leq \tau^*$ we have

$$
E_\sigma(u^{n+1}) \leq E_\sigma(u^n).
$$

Since the energy is diminished in every iteration, the coercivity of $\langle J_\sigma(u^n) \cdot, \cdot \rangle$ is maintained and all the iterations are well-defined. Finally, we note that because of (26) we have $E_\sigma(u^n) = E_\sigma(\tilde{u}^{n+1})$ if and only if $u^n = \tilde{u}^{n+1}$. However, this can only happen if

$$
J_\sigma(u^n) u^n = \gamma_n \mathcal{I} u^n,
$$

i.e., if u^n is already an eigenfunction with eigenvalue $\lambda = \gamma_n$.   □

It is interesting to note that the L^2-norm of $\tilde{u}^n$ cannot diverge. We see this in the following conclusion.

Conclusion 5 *In the setting of Theorem 4 it holds that* $\|\tilde{u}^n\| \to 1$ *for* $n \to \infty$.

Proof In the proof of Theorem 4 we have seen that

$$
E_\sigma(u^n) - E_\sigma(u^{n+1}) \geq \frac{1}{2} \left(\frac{1}{\tau^*} - \frac{1}{2} - 3\kappa \, C_S \right) \int_{\mathcal{D}} |\nabla(\tilde{u}^{n+1} - u^n)|^2 \, dx.
$$

Since $E_\sigma(u^n)$ is monotonically decreasing and bounded from below, we have $E_\sigma(u^n) - E_\sigma(u^{n+1}) \to 0$. This together with the Poincaré-Friedrichs inequality implies that

$$\|\tilde{u}^{n+1}\| \leq \|\tilde{u}^{n+1} - u^n\| + \|u^n\| \to 1$$

for $n \to \infty$. $\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\qquad\quad$ $\square$

5.3 Convergence and optimal damping

In this subsection we prove the convergence of the *J*-method for a suitable choice of damping parameters. We can make practical use of this result by selecting τ_n in each iteration step by the minimizer of a simple one-dimensional minimization problem.

Theorem 6 (global convergence) *Suppose that the assumptions of Theorem 4 are fulfilled. Additionally assume that τ_n is selected such that it does not degenerate, i.e., is uniformly bounded away from zero. Then there exists a limit energy $E^* := \lim_{n\to\infty} E(u^n)$ and, up to a subsequence, we have that the iterates u^n of the damped J-method converge strongly in $H^1(\mathcal{D})$ to a limit $u^* \in V$. The limit is an L^2-normalized eigenfunction with some eigenvalue $\lambda^* > 0$, i.e.,*

$$\mathcal{A}(u^*) = \lambda^* \mathcal{I} u^*$$

and we have $E(u^*) = E^*$. If u^* is the only eigenfunction on the energy level E^*, then we have convergence of the full sequence u^n.

Proof The proof is similar to the arguments presented in [36, Th. 4.9]. First, Theorem 4 guarantees the existence of the limit $E^* := \lim_{n\to\infty} E(u^n)$. Hence, u^n is uniformly bounded in V and we can extract a subsequence (still denoted by u^n) that converges weakly in $H^1(\mathcal{D})$ and strongly in $L^p(\mathcal{D})$ (for $p < 6$) to a limit $u^* \in V$ with $\|u^*\| = 1$. Since $J_\sigma(u^*)$ is a real-linear operator that depends continuously on the data and which induces the coercive bilinear form $\langle (J(u) + \sigma\mathcal{I}) \cdot, \cdot \rangle$, we have that

$$J_\sigma(u^*)^{-1}\mathcal{I}u^n \to J_\sigma(u^*)^{-1}\mathcal{I}u^* \quad \text{strongly in } H^1(\mathcal{D}).$$

Together with the strong convergence $u^n \to u^*$ in $L^4(\mathcal{D})$, we conclude that

$$J_\sigma(u^n)^{-1}\mathcal{I}u^n \to J_\sigma(u^*)^{-1}\mathcal{I}u^* \quad \text{strongly in } H^1(\mathcal{D}).$$

This shows that

$$(\gamma_n)^{-1} = (J_\sigma(u^n)^{-1}\mathcal{I}u^n, u^n) \xrightarrow{n\to\infty} (J_\sigma(u^*)^{-1}\mathcal{I}u^*, u^*) =: (\gamma^*)^{-1}.$$

Furthermore, we have seen in the proof of Theorem 4 (respectively Conclusion 5) that the strong energy reduction implies that for $n \to 0$

$$\|\tilde{u}^{n+1} - u^n\|_{H^1(\mathcal{D})} \to 0$$

and consequently we have with $\tilde{u}^{n+1} = (1 - \tau_n)u^n + \tau_n \gamma^n J_\sigma(u^n)^{-1}\mathcal{I}u^n$ and the boundedness of τ_n that

$$u^n = \gamma^n J_\sigma(u^n)^{-1}\mathcal{I}u^n - \tau_n^{-1}(\tilde{u}^{n+1} - u^n) \;\to\; \gamma^* J_\sigma(u^*)^{-1}\mathcal{I}u^*$$

strongly in $H^1(\mathcal{D})$. Since we already know that u^n converges weakly in $H^1(\mathcal{D})$ to u^*, we can now conclude that this is even a strong convergence and we have

$$J_\sigma(u^*)u^* = \gamma^*\mathcal{I}u^*.$$

This shows that u^* is an eigenfunction with eigenvalue γ^*. The strong convergence in $H^1(\mathcal{D})$ also implies convergence of the energies, i.e., it holds $E^* = \lim_{n\to\infty} E(u^n) = E(u^*)$. $\square$

For all sufficiently small τ_n, Theorem 4 proves the energy reduction and Theorem 6 global convergence. However, since we do not know a priori what a sufficiently small value for τ_n is, we can combine the damped J-method with a line search algorithm that optimizes τ_n in each iteration step such that the energy reduction is (quasi) optimal. Theorems 4 and 6 show that such an optimal τ_n exists and that it does not degenerate to zero. We stress that finding such a τ_n does not require any additional inversions, which makes the procedure very cheap, cf. "Appendix A" for details.

Conclusion 7 (*J-method with optimal damping*) *Consider a shift σ such that the assumptions of Theorem 4 are fulfilled. Given $u^n \in V$ with $\|u^n\| = 1$ the next iteration is obtained by selecting the optimal damping parameter with*

$$\tau_n := \underset{0 < \tau \le 2}{\arg\min}\; E\left(\frac{(1 - \tau)u^n + \tau\,\gamma_n\, J_\sigma(u^n)^{-1}\mathcal{I}u^n}{\|(1 - \tau)u^n + \tau\,\gamma_n\, J_\sigma(u^n)^{-1}\mathcal{I}u^n\|} \right)$$

and defining u^{n+1} as in (7). The approximations are energy diminishing and converge (up to a subsequence) strongly in V to an L^2-normalized eigenfunction of the GPEVP.

Finally, if there is no rotation, we can even achieve guaranteed global convergence to the ground state provided that the selected initial value is non-negative.

Proposition 3 (*global convergence to ground state*) *Assume the setting of Theorem 6. Furthermore, we consider that there is no rotation, i.e., $\Omega = 0$, and a non-negative and L^2-normalized starting value $u^0 \in V$, i.e., $u^0 \ge 0$. If $\tau_n \le 1$ and if the shift parameter $\sigma > 0$ is selected sufficiently large, then the iterates u^n of the damped J-method converge strongly in $H^1(\mathcal{D})$ to the unique (positive) ground state $u^* \ge 0$.*

Proof If $\Omega = 0$, then the problem can be fully formulated over $\mathbb{R}$ and admits a unique positive ground state $u^* \in V$, cf. [21]. The only other ground state is $-u^*$. Furthermore, all excited states (i.e., all other eigenfunctions) must necessarily change their sign on $\mathcal{D}$, cf. [36, Lem. 5.4]. Hence, if we can verify that the iterates of the damped J-method (7) preserve positivity, then Theorem 6 guarantees that the global H^1-limit must be the desired ground state u^*.

Recall the damped *J*-iteration from (7) together with (14), which (over $\mathbb{R}$) reduces to

$$\langle J_\sigma(u^n)v, w \rangle = \langle S_\sigma v, w \rangle + \langle Gv, w \rangle,$$

where S_σ is the linear self-adjoint operator given by

$$\langle S_\sigma v, w \rangle := (\nabla v, \nabla w)_{L^2(\mathcal{D})} + ((W + \sigma + 3\kappa |u^n|^2)\, v, w)_{L^2(\mathcal{D})}$$

and G characterizes the non-symmetric rank-1 remainder, i.e.,

$$\langle Gv, w \rangle := -(u^n, v)_{L^2(\mathcal{D})}\, (f, w)_{L^2(\mathcal{D})}, \qquad \text{where } f := 2\kappa\, |u^n|^2 u^n.$$

Analogously to the Sherman–Morrison formula for matrices [49], we can see that

$$(S_\sigma + G)^{-1} = S_\sigma^{-1} - \frac{S_\sigma^{-1} \circ G \circ S_\sigma^{-1}}{1 - (u^n, S_\sigma^{-1}\mathcal{I}f)_{L^2(\mathcal{D})}}.$$

Consequently we can write the effect of the inverse as

$$J_\sigma(u^n)^{-1}\mathcal{I}v = (S_\sigma + G)^{-1}\mathcal{I}v = S_\sigma^{-1}(\mathcal{I}v) - \frac{S_\sigma^{-1} \circ G \circ S_\sigma^{-1}(\mathcal{I}v)}{1 - (u^n, S_\sigma^{-1}\mathcal{I}f)_{L^2(\mathcal{D})}}.$$

Since $\sigma > 0$ and $W \geq 0$, S_σ is a self-adjoint elliptic operator and hence preserves positivity, i.e., we have $S_\sigma^{-1}\mathcal{I}v \geq 0$ if $v \geq 0$. This immediately follows by writing the action of S_σ^{-1} as an energy minimization problem. Starting (inductively) from a function $u^n \geq 0$ we conclude that

$$S_\sigma^{-1}(\mathcal{I}u^n) \geq 0 \quad \text{and} \quad -S_\sigma^{-1} \circ G \circ S_\sigma^{-1}(\mathcal{I}u^n) \geq 0.$$

If we can ensure that $1 - (u^n, S_\sigma^{-1}\mathcal{I}f)_{L^2(\mathcal{D})} > 0$, then we have $J_\sigma(u^n)^{-1}\mathcal{I}u^n \geq 0$. Let us hence consider $(u^n, S_\sigma^{-1}\mathcal{I}f)_{L^2(\mathcal{D})}$, for which we obtain

$$|(u^n, S_\sigma^{-1}\mathcal{I}f)_{L^2(\mathcal{D})}| \leq \|u^n\|\, \|S_\sigma^{-1}\|_{\mathcal{L}(V^*,V)}\|\mathcal{I}f\|_{V^*} \leq 2\kappa\, \|u^n\|^3_{L^6(\mathcal{D})}\, \|S_\sigma^{-1}\|_{\mathcal{L}(V^*,V)}.$$

Since S_σ is self-adjoint, we have

$$\|S_\sigma^{-1}\|_{\mathcal{L}(V^*,V)} = 1/\lambda_{\min}(S_\sigma) = 1/(\sigma + \lambda_{\min}(S_0)) \leq 1/\sigma,$$

where $\lambda_{\min}(S_0) > 0$ is the minimal eigenvalue of S_0. Consequently, if the shift is such that

$$2\kappa\, \|u^n\|^3_{L^6(\mathcal{D})} < \sigma,$$

then we have positivity of $J_\sigma(u^n)^{-1}\mathcal{I}u^n$. Note that by the energy reduction property, we can bound $\|u^n\|_{L^6(\mathcal{D})}^3$ uniformly for all n by a constant that only depends on the initial energy $E(u^0)$. Together with the obvious positivity of $(1 - \tau_n)u^n$ for $\tau_n \leq 1$, we conclude the existence of a sufficiently large shift σ so that $u^{n+1} \geq 0$ for all $n \geq 0$ and hence global convergence to the ground state. □

6 Numerical experiments

This section concerns the numerical performance of the proposed J-method enhanced by shifting and/or damping as outlined in Sects. 2.3 and 2.4 . As a general model, we seek critical points of the Gross–Pitaevskii energy (12) in a bounded domain $\mathcal{D} = (-L, L)^2$ for some parameter $L > 0$. The particular choice of L as well as the other physical parameters Ω, W, and κ will be specified separately in the various model problems below.

For the spatial discretization we always use bilinear finite elements on a Cartesian mesh of width $h = 2^{-8}L$. We will not investigate discretization errors with respect to the underlying PDE. For approximation properties of discrete eigenfunctions we refer to the analytical results presented in [20,21,25,35] and to [22,34,54] for a posteriori estimators and adaptivity. Our focus is the performance of the iterative eigenvalue solver promoted in this paper. As a measure of accuracy we will use the $L^2(\mathcal{D})$-norm of the residual $\mathcal{A}(u^n)u^n - \lambda^n\mathcal{I}u^n$ given an approximate finite element eigenpair (λ^n, u^n). We will stop the solver whenever the residual falls below the tolerance TOL $= 10^{-8}$.

For a better assessment of the performance of the J-method, we compare it with the projected a_z-Sobolev gradient flow introduced in [36]: given $u^0 \in V$ with $\|u^0\| = 1$, define for $n = 1, 2, \ldots$,

$$\hat{u}^{n+1} := A(u^n, \cdot)^{-1}\mathcal{I}u^n, \qquad \gamma_n^{-1} := \langle A(u^n, \hat{u}^{n+1}), \hat{u}^{n+1}\rangle,$$
$$u^{n+1} := \frac{(1 - \tau_n)u^n + \tau_n\gamma_n\hat{u}^{n+1}}{\|(1 - \tau_n)u^n + \tau_n\gamma_n\hat{u}^{n+1}\|}. \tag{27}$$

We will refer to this approach as the A-method (without shift). Note that every step of the A-method can be interpreted as an energy minimization problem such that the iteration is positivity preserving for the case without rotation. This guarantees global convergence to the ground state for every nonnegative starting value. The corresponding shifted version, which we will also consider in the experiments, considers $A_\sigma(u^n, \cdot) := A(u^n, \cdot) + \sigma\mathcal{I}$ in place of $A(u^n, \cdot)$ in the definition of $\hat{u}^{n+1}$ and a corresponding adjustment of the normalization factor γ_n. According to the numerical experiments of [36], this method is representative for the larger class of gradient flows in terms of accuracy-cost ratios. The cost per iteration step for both A- and J-method are proportional and of the same order. Tentatively, the A-method is cheaper by a fixed factor, since (due to the rank-1 matrix that appears in the J-version) an additional linear system per step has to be solved when the Sherman-Morrison formula is used, cf. "Appendix B". To what extent the computational overhead of the J-method can be reduced by suitable preconditioned iterative solvers is beyond the scope of the paper.

Notwithstanding the above, the experiments below clearly show that the *J*-method easily compensates its possible computational overhead per step by a notedly smaller iteration count.

6.1 Ground state in a harmonic potential

In the first model problem, we consider a harmonic trapping potential with trapping frequencies 1/2, i.e.,

$$W(x) = \tfrac{1}{2} |x|^2. \tag{28}$$

The angular momentum Ω is set to zero and the repulsion parameter to $\kappa = 1000$. The size of the domain is chosen as $L = 8$. This is larger than the Thomas-Fermi radius of the problem which can be estimated as $R^{\text{TF}} = \sqrt{2} \, (\kappa/\pi)^{1/4} \approx 5.97$, cf. [10]. We are interested in computing the ground state, i.e., the global minimizer u^{gs} of the energy (12). Note that, up to sign, u^{gs} is the unique eigenfunction that corresponds to the smallest eigenvalue λ^{gs} which is well-separated from the remaining spectrum. As an initial value for all variants of eigenvalue solvers we use the bi-quadratic bubble

$$u^0(x) = (1 - x_1^2/L^2)(1 - x_2^2/L^2), \tag{29}$$

interpolated in the finite element space and normalized in $L^2(\mathcal{D})$. For this simple model problem, there are certainly more sophisticated initial guesses such as the Thomas-Fermi approximation. Our uneducated initial guess marks an additional challenge. As ground state energy we computed $E^{\text{gs}} := E(u^{\text{gs}}) \leq 6.019$. The corresponding eigenvalue approximation is $\lambda^{\text{gs}} \approx 17.93$. Clearly, the accuracy of these numbers is limited by the choice of the discretization parameter h. Mesh adaptivity as used in [34] or a higher-order method would certainly help to improve on these numbers.

Figure 1 shows the evolution of the residuals during the iteration of several variants of the *J* and *A*-method. Unless specified differently, the general *J*- and *A*-methods refer to a combination of damping and shifting. More precisely, we use damping for globalization of convergence until the residual falls below 10^{-3}. Then we freeze the time step $\tau = 1$ and switch to a Rayleigh shift strategy to possibly accelerate convergence, i.e., in each step we choose the shift $\sigma = -\lambda^n$ to be the current eigenvalue approximation. According to our numerical experience the coexistence of damping and shifting is hard to control. The transition from damping to shifting is clearly seen in the convergence plot of Fig. 1. We observe linear but fairly slow convergence in the damped phase. As soon as we switch to shifting, the convergence is beyond linear. The *A*-method performs similarly in the damping phase but diverges as soon as the shift is turned on. It is this phenomenon already observed in [38] in a less extreme characteristic (see also the second model problem below) that motivated the derivation of a shift-sensitive *J*-method. Note that the improved rate can be explained by its connection to Newton's method, cf. [38, Sect. 3.4] . Convergence proofs with higher rate, however, are only given in the discrete setting for linear and particular nonlinear eigenvalue problems [41,43].
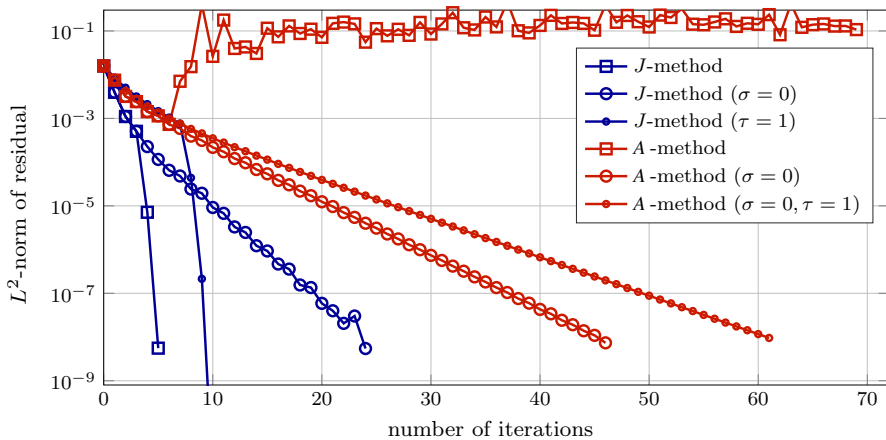
Fig. 1 Computation of the ground state for a harmonic potential, cf. Sect. 6.1 for details. The figure shows the $L^2(\mathcal{D})$-norms of the residuals (logarithmic scale) vs. the iteration count for several methods indicated in the legend

In Fig. 1 we also show results for the variants of the J- and A-methods where either τ, σ, or both are fixed. Their performance is in between the aforementioned combined approaches.

6.2 Exponentially localized ground state in a disorder potential

In the second model problem the non-negative external potential W reflects a high degree of disorder and the repulsion parameter κ is small. In this situation, the low-energy eigenstates essentially localize in the sense of an exponential decay of their moduli.

The numerical approximation of localized Schrödinger eigenstates in the linear case, i.e., for $\kappa = 0$, has recently caused a large interest in the fields of computational physics and scientific computing [8,9,31,51,55]. In particular, the results of [4] provide a mathematical justification of the observed localization. In the nonlinear case the phenomenon is still observable but locality deteriorates with increasing interaction [3,5,6]. Here, we choose $\kappa = 1$ which leads to a fairly localized ground state as it can be seen in Fig. 2 (right). Its computation turns out to be much more challenging than in the case of a harmonic trapping potential in the sense that convergence rates are slower and iteration counts larger. This is related to a possible clustering of the lowermost eigenvalues. In particular, we expect $\lambda_1/\lambda_2 \approx 1$ in this example such that shifting can provide a considerable speed up. Due to the small repulsion parameter κ, however, we expect a significant gap within the first few eigenvalues as in the linear case.

We have tested the same solvers as in the previous subsection. The J-method involving shifting performs best by far. Surprisingly, the variant without adaptive time step in the damping phase even performed better. This is no contradiction with the theory as we are showing residuals rather than energies (Fig. 3). Moreover, a locally
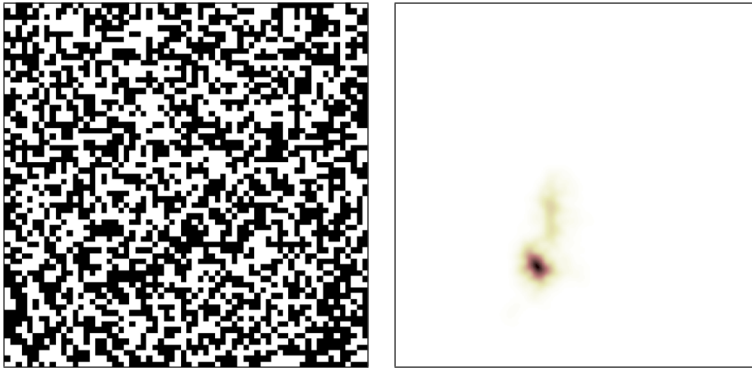
Fig. 2 Exponentially localized ground state in a disorder potential, cf. Sect. 6.2. Left: Disorder potential (random i.i.d. checkerboard) taking values 0 (white) and $(2\varepsilon L)^{-2}$ (black) for parameters $L = 8$, $\varepsilon = 2^{-6}$. Right: Corresponding ground state density for $\Omega = 0$, $\kappa = 1$
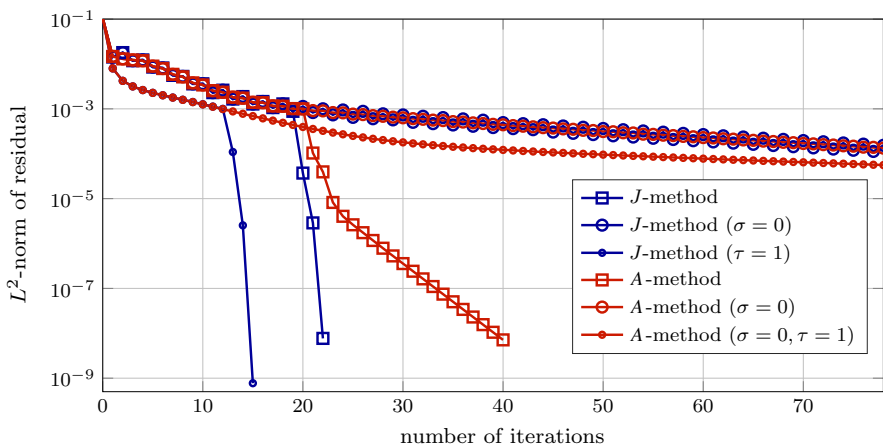
Fig. 3 Computation of the exponentially localized ground state in a disorder potential, cf. Sect. 6.2 for details. The figure shows the $L^2(\mathcal{D})$-norms of the residuals (logarithmic scale) vs. the iteration count for several variants of *J*- and *A*-method

optimal energy decrease does not necessarily lead to better global performance. Still the difference is not too big. As a general recommendation from our numerical experience we would favor to use an adaptive time step because it was more robust. Another difference with regard to the harmonic potential is that, this time, the *A*-method reacts upon shifting in a positive way. For a few steps the convergence is indeed accelerated. However, thereafter the method turns back to a linear regime of convergence which cannot compete with the shifted *J*-method. Nevertheless, the *A*-method does not fail completely as seen in the previous example. This may be caused by the small value of κ, compared with the considered potential.

6.3 Vortex lattices in a fast rotating trap

We close the numerical illustration of the J-method with a qualitative study of vortex lattice states in the present of fast rotating potentials. We choose a harmonic potential W as in (28) and set $\Omega = 0.85$, $\kappa = 1000$, and the size of the computational domain to $L = 10$.

We have computed four different eigenfunctions using the J-method. This was only possible using the shift-sensitivity of the J-method. We shall briefly describe the computational parameters. We use the bi-quadratic bubble (29) as the initial value for all computations. To compute the (tentative) ground state u_1 (see Fig. 4, upper left) we used the combined strategy as before. However, we switched from damping to shifting only once the residual falls below 10^{-6}. Switching earlier led to states of higher energy. E.g., switching at a tolerance of 10^{-3} lead to the eigenfunction u_2 depicted in the upper right of Fig. 4. It is interesting to observe that while $E_1 := E(u_1) < E(u_2) := E_2$ the corresponding eigenvalues are ordered the other way around.

Two further excited states are found by limiting the adaptive shift to the interval $[15.0, 15.6]$ for u_3 (see lower left of Fig. 4) and to the interval $[15.2, 15.45]$ for the state u_4 that does not show any vortices (see lower right of Fig. 4). In both cases we used the lower end of the interval as the shift in the damping phase and we switched to adaptive shifting at residual tolerance 10^{-4} for u_3 and 10^{-3} for u_4. Note that u_1 seems to be the global energy minimizer of the (discretized) problem but the exited states u_2, u_3, u_4 do not necessarily represent the next higher energy levels 2 to 4 but some levels of higher energy.

From this rather complicated derivation one can see that it is by no means trivial to compute these excited states. We shall also say that it is not always easy to control the shifting. If one shifts too early in the sense that the approximation is not yet close to the target eigenfunction (e.g. in terms of number of vortices) the procedure may fail completely. Despite this difficulty which is intrinsic to the nonlinear eigenvalue problem at hand, the J-method along with shifting and damping enables the selective approximation of excited states as well as the amplification of convergence beyond linear rates in the spirit of the Rayleigh quotient iteration even in this challenging regime of vortex pattern formation.

7 Conclusion

In this paper we have generalized the J-method proposed in [38] to the abstract Hilbert space setting. This gives rise to a variational formulation that is straightforwardly accessible by Galerkin-type discretizations, e.g., based on finite elements. Moreover, we have transferred the proof of local convergence of the J-method from the discrete setting (cf. [38]) to the abstract setting and recovered a quantitative convergence rate that depends on the spectral shift. Since this fast convergence is indeed a local feature, we have proposed a damped J-method. For the GPEVP, the damping step can be seen as a discretization of a generalized gradient flow and guarantees reduction of the energy associated to the Gross–Pitaevskii operator. This energy reduction is the key to the global convergence of the damped method.
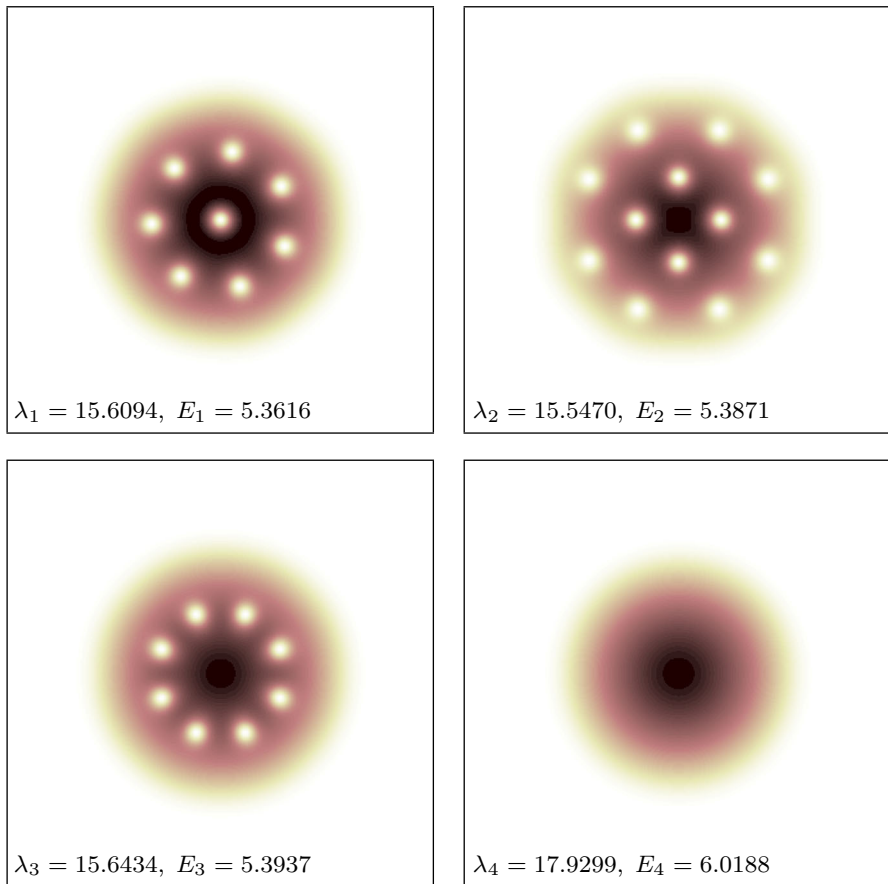
Fig. 4 Computation of vortex lattices in a fast rotating trap at different energy levels, cf. Sect. 6.3. The parameters are $L = 10$, $\Omega = 0.85$, $\kappa = 1000$. The upper left figure depicts the density of the tentative ground state. The other three figures show densities corresponding to excited states

We have proposed a combined strategy of damping and shifting, depending on the residual error. The damping part guides the iterates to a sufficiently small neighborhood of an eigenfunction. Therein, the shifting significantly improves the linear rate of convergence. With a Rayleigh-type shifting strategy remarkable speed-ups beyond linear convergence are observed. In numerical experiments we have demonstrated the excellent performance of the arising method and its suitability for both the computation of ground states and the selective computation of excited states. We believe that the proposed strategy can be also an efficient tool for treating other types of eigenvalue problems with nonlinearities, in particular those that can be rephrased as finding the critical points of constraint energy minimization problems.

A Energy-diminishing step size control

We consider the damped J-method (7) in the case of the Gross–Pitaevskii equation. In order to implement an efficient step size control, we consider the function

$$f(\tau) := E\left(\frac{(1-\tau)u^n + \tau\,\gamma_n\,J_\sigma(u^n)^{-1}\mathcal{I}u^n}{\|(1-\tau)u^n + \tau\,\gamma_n\,J_\sigma(u^n)^{-1}\mathcal{I}u^n\|}\right)$$

that we want to minimize for $\tau \in (0, 2)$. Based on u^n, we compute

$$w^n := \gamma_n\,J_\sigma(u^n)^{-1}\mathcal{I}u^n.$$

Note that this implies

$$\int_{\mathcal{D}} |(1-\tau)u^n + \tau w^n|^4\,\mathrm{d}x = (1-\tau)^4 \int_{\mathcal{D}} |u^n|^4\,\mathrm{d}x + 4(1-\tau)^3\tau \int_{\mathcal{D}} \Re(u^n\overline{w^n})\,|u^n|^2\,\mathrm{d}x$$

$$+ (1-\tau)^2\tau^2 \int_{\mathcal{D}} 2|w^n|^2|u^n|^2 + 4\,|\Re(u^n\overline{w^n})|^2\,\mathrm{d}x$$

$$+ 4\,\tau^3(1-\tau) \int_{\mathcal{D}} \Re(u^n\overline{w^n})\,|w^n|^2\,\mathrm{d}x + \int_{\mathcal{D}} \tau^4|w^n|^4\,\mathrm{d}x.$$

With this, we precompute various terms, which are given by

$$\alpha_0 := \int_{\mathcal{D}} |\nabla_{\!\mathrm{R}} u^n|^2 + W_{\mathrm{R}}|u^n|^2\,\mathrm{d}x, \quad \alpha_1 := 2\int_{\mathcal{D}} \Re\left(\nabla_{\!\mathrm{R}} u^n \cdot \overline{\nabla_{\!\mathrm{R}} w^n} + W_{\mathrm{R}}\,u^n\,\overline{w^n}\right)\,\mathrm{d}x,$$

$$\alpha_2 := \int_{\mathcal{D}} |\nabla_{\!\mathrm{R}} w^n|^2 + W_{\mathrm{R}}|w^n|^2\,\mathrm{d}x$$

as well as

$$\beta_0 := \frac{\kappa}{2} \int_{\mathcal{D}} |u^n|^4\,\mathrm{d}x, \quad \beta_1 := 2\kappa \int_{\mathcal{D}} \Re(u^n\overline{w^n})\,|u^n|^2\,\mathrm{d}x,$$

$$\beta_2 := \kappa \int_{\mathcal{D}} |w^n|^2|u^n|^2 + 2|\Re(u^n\overline{w^n})|^2\,\mathrm{d}x,$$

$$\beta_3 := 2\kappa \int_{\mathcal{D}} \Re(u^n\overline{w^n})\,|w^n|^2\,\mathrm{d}x, \quad \beta_4 := \frac{\kappa}{2} \int_{\mathcal{D}} |w^n|^4\,\mathrm{d}x$$

$$\zeta_0 := \int_{\mathcal{D}} |u^n|^2\,\mathrm{d}x, \quad \zeta_1 := 2\int_{\mathcal{D}} \Re(u^n\,w^n)\,\mathrm{d}x, \quad \zeta_2 := \int_{\mathcal{D}} |w^n|^2\,\mathrm{d}x.$$

Note that the terms α_i, β_i, and ζ_i have to be computed only once per time step (e.g. within one loop over the grid elements). Finally, we define the function

$$s^n(\tau) := \left(\sum_{i,j\geq 0:\ i+j=2} (1-\tau)^i\, \tau_n^j\, \zeta_j \right)^{-1/2}$$

and observe that $f(\tau)$ is given by

$$f(\tau) = \frac{1}{2}\left(\sum_{i,j\geq 0:\ i+j=2} |s^n(\tau)|^2 (1-\tau)^i\, \tau^j\, \alpha_j + \sum_{i,j\geq 0:\ i+j=4} |s^n(\tau)|^4 (1-\tau)^i\, \tau^j\, \beta_j \right).$$

After precomputing α_i, β_i, and ζ_i, the function $f(\tau)$ can be cheaply evaluated. The minimization step, i.e., $\tau_n := \arg\min\{f(\tau)|\ \tau \in (0, 2)\}$ can be easily implemented using, e.g., a golden section search. Note that the energy of u^{n+1} is now given by $f(\tau_n)$.

B Matrix representation of the $\mathcal{J}$-operator

The $\mathcal{J}$-operator in case of the GPEVP was derived in Sect. 4.2 and we consider the iteration given by (16). For the implementation we need to discuss the handling of the nonlinear terms. Using the identity $(u\bar{v} + 2v\bar{u})u\,\overline{w} = 2\Re(u\bar{v})\,u\,\overline{w} + |u|^2 v\,\overline{w}$, we end up with the integrals

$$I_1 := \int_{\mathcal{D}} \Re(u\bar{v})\,u\,\overline{w}\,\mathrm{d}x, \qquad I_2 := \int_{\mathcal{D}} |u|^2 v\,\overline{w}\,\mathrm{d}x, \qquad I_3 := \Re(u, v)_{L^2(\mathcal{D})}\int_{\mathcal{D}} |u|^2 u\,\overline{w}\,\mathrm{d}x.$$

Note that we only need to consider real test functions w and decompose u, v into its real and imaginary part, i.e., $u = u_R + iu_I$, $v = v_R + iv_I$. This is also done in the finite element discretization, i.e., we work with real vectors of double dimension. For the first integral we note that

$$\int_{\mathcal{D}} \Re(u\bar{v})\,u\,\overline{w}\,\mathrm{d}x = \int_{\mathcal{D}} \left(u_R v_R + u_I v_I \right) \cdot \left(u_R w + iu_I w \right)\mathrm{d}x$$

$$= \int_{\mathcal{D}} \left(v_R u_R^2 + v_I u_R u_I + iv_R u_R u_I + iv_I u_I^2 \right)w\,\mathrm{d}x.$$

Introducing M_v as the mass matrix weighted by v, this leads to the matrix representation

$$\begin{bmatrix} M_{u_R u_R} & M_{u_R u_I} \\ M_{u_R u_I} & M_{u_I u_I} \end{bmatrix}.$$

For the second term we have $|u|^2 v\,\overline{w} = |u|^2\bigl(v_R w + \mathrm{i}v_I w\bigr)$. Thus, the corresponding finite element matrix is block diagonal and reads $\mathrm{diag}(M_{|u|^2}, M_{|u|^2})$. Finally, we have

$$
\begin{aligned}
I_3 &= \Re(u, v)_{L^2(\mathcal{D})} \int_{\mathcal{D}} |u|^2 u\,\overline{w}\,\mathrm{d}x \\
&= \bigl[(u_R, v_R)_{L^2(\mathcal{D})} + (u_I, v_I)_{L^2(\mathcal{D})}\bigr] \int_{\mathcal{D}} |u|^2\bigl(u_R w + \mathrm{i}u_I w\bigr)\,\mathrm{d}x.
\end{aligned}
$$

A simple rearrangement shows that this corresponds to the rank-one matrix

$$
\begin{bmatrix} M u_R u_R^T M_{|u|^2} & M u_R u_I^T M_{|u|^2} \\ M u_I u_R^T M_{|u|^2} & M u_I u_I^T M_{|u|^2} \end{bmatrix} = \begin{bmatrix} M u_R \\ M u_I \end{bmatrix} \begin{bmatrix} M_{|u|^2} u_R \\ M_{|u|^2} u_I \end{bmatrix}^T.
$$

In total, a finite element discretization of the J-method calls for a solution of a linear system which is decomposed of several sparse matrices and the latter rank-1 update. This can be easily inverted using the Sherman-Morrison formula, cf. [49,53].

References

1. Abraham, R., Marsden, J.E., Ratiu, T.: Manifolds, tensor analysis, and applications, Applied Mathematical Sciences, vol. 75, 2nd edn. Springer-Verlag, New York (1988). https://doi.org/10.1007/978-1-4612-1029-0
2. Altmann, R., Froidevaux, M.: PDE eigenvalue iterations with applications in two-dimensional photonic crystals. ESAIM Math. Model. Numer. Anal. (M2AN) **54**(5), 1751–1776 (2020)
3. Altmann, R., Henning, P., Peterseim, D.: Localization and delocalization of ground states of Bose-Einstein condensates under disorder (2020). arXiv:2006.00773
4. Altmann, R., Henning, P., Peterseim, D.: Quantitative Anderson localization of Schrödinger eigenstates under disorder potentials. Math. Models Methods Appl. Sci. **30**(5), 917–955 (2020). https://doi.org/10.1142/S0218202520500190
5. Altmann, R., Peterseim, D.: Localized computation of eigenstates of random Schrödinger operators. SIAM J. Sci. Comput. **41**, B1211–B1227 (2019)
6. Altmann, R., Peterseim, D., Varga, D.: Localization studies for ground states of the Gross–Pitaevskii equation. PAMM **18**(1), e201800343 (2018). https://doi.org/10.1002/pamm.201800343
7. Antoine, X., Levitt, A., Tang, Q.: Efficient spectral computation of the stationary states of rotating Bose-Einstein condensates by preconditioned nonlinear conjugate gradient methods. J. Comput. Phys. **343**, 92–109 (2017). https://doi.org/10.1016/j.jcp.2017.04.040
8. Arnold, D.N., David, G., Filoche, M., Jerison, D., Mayboroda, S.: Computing spectra without solving eigenvalue problems. SIAM J. Sci. Comput. **41**(1), B69–B92 (2019). https://doi.org/10.1137/17M1156721
9. Arnold, D.N., David, G., Jerison, D., Mayboroda, S., Filoche, M.: Effective confining potential of quantum states in disordered media. Phys. Rev. Lett. **116**, 056602 (2016). https://doi.org/10.1103/PhysRevLett.116.056602
10. Bao, W.: Mathematical models and numerical methods for Bose-Einstein condensation. In: Proceedings of the International Congress for Mathematicians 2014, (2014)
11. Bao, W., Chern, I.L., Lim, F.Y.: Efficient and spectrally accurate numerical methods for computing ground and first excited states in Bose-Einstein condensates. J. Comput. Phys. **219**(2), 836–854 (2006). https://doi.org/10.1016/j.jcp.2006.04.019
12. Bao, W., Du, Q.: Computing the ground state solution of Bose-Einstein condensates by a normalized gradient flow. SIAM J. Sci. Comput. **25**(5), 1674–1697 (2004). https://doi.org/10.1137/S1064827503422956

13. Bao, W., Shen, J.: A generalized-Laguerre-Hermite pseudospectral method for computing symmetric and central vortex states in Bose-Einstein condensates. J. Comput. Phys. **227**(23), 9778–9793 (2008). https://doi.org/10.1016/j.jcp.2008.07.017
14. Bao, W., Tang, W.: Ground-state solution of Bose-Einstein condensate by directly minimizing the energy functional. J. Comput. Phys. **187**(1), 230–254 (2003). https://doi.org/10.1016/S0021-9991(03)00097-4
15. Bao, W., Wang, H., Markowich, P.A.: Ground, symmetric and central vortex states in rotating Bose-Einstein condensates. Commun. Math. Sci. **3**(1), 57–88 (2005). http://projecteuclid.org/euclid.cms/1111095641
16. Borzì, A., Ciaramella, G., Sprengel, M.: Formulation and numerical solution of quantum control problems. Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA (2017). https://doi.org/10.1137/1.9781611974843.ch1
17. Bose, S.N.: Plancks Gesetz und Lichtquantenhypothese. Zeitschrift für Physik **26**(1), 178–181 (1924)
18. Caliari, M., Ostermann, A., Rainer, S., Thalhammer, M.: A minimisation approach for computing the ground state of Gross–Pitaevskii systems. J. Comput. Phys. **228**(2), 349–360 (2009). https://doi.org/10.1016/j.jcp.2008.09.018
19. Cancès, E.: SCF algorithms for HF electronic calculations. In: Mathematical models and methods for ab initio quantum chemistry, Lecture Notes in Chem., vol. 74, pp. 17–43. Springer, Berlin (2000). https://doi.org/10.1007/978-3-642-57237-1_2
20. Cancès, E., Chakir, R., He, L., Maday, Y.: Two-grid methods for a class of nonlinear elliptic eigenvalue problems. IMA J. Numer. Anal. **38**(2), 605–645 (2018). https://doi.org/10.1093/imanum/drw053
21. Cancès, E., Chakir, R., Maday, Y.: Numerical analysis of nonlinear eigenvalue problems. J. Sci. Comput. **45**(1–3), 90–117 (2010). https://doi.org/10.1007/s10915-010-9358-1
22. Cancès, E., Dusson, G., Maday, Y., Stamm, B., Vohralík, M.: A perturbation-method-based a posteriori estimator for the planewave discretization of nonlinear Schrödinger equations. C. R. Math. Acad. Sci. Paris **352**(11), 941–946 (2014). https://doi.org/10.1016/j.crma.2014.09.014
23. Cancès, E., Le Bris, C.: Can we outperform the DIIS approach for electronic structure calculations? Int. J. Quantum Chem. **79**(2), 82–90 (2000)
24. Cancès, E., Le Bris, C.: On the convergence of SCF algorithms for the Hartree–Fock equations. M2AN Math. Model. Numer. Anal. **34**(4), 749–774 (2000). https://doi.org/10.1051/m2an:2000102
25. Chen, H., Gong, X., Zhou, A.: Numerical approximations of a nonlinear eigenvalue problem and applications to a density functional model. Math. Methods Appl. Sci. **33**(14), 1723–1742 (2010). https://doi.org/10.1002/mma.1292
26. Dalfovo, F., Giorgini, S., Pitaevskii, L.P., Stringari, S.: Theory of Bose-Einstein condensation in trapped gases. Rev. Mod. Phys. **71**(3), 463–512 (1999)
27. Danaila, I., Kazemi, P.: A new Sobolev gradient method for direct minimization of the Gross–Pitaevskii energy with rotation. SIAM J. Sci. Comput. **32**(5), 2447–2467 (2010). https://doi.org/10.1137/100782115
28. Danaila, I., Protas, B.: Computation of ground states of the Gross–Pitaevskii functional via Riemannian optimization. SIAM J. Sci. Comput. **39**(6), B1102–B1129 (2017). https://doi.org/10.1137/17M1121974
29. Dion, C.M., Cancès, E.: Ground state of the time-independent Gross–Pitaevskii equation. Comput. Phys. Commun. **177**(10), 787–798 (2007). https://doi.org/10.1016/j.cpc.2007.04.007
30. Einstein, A.: Quantentheorie des einatomigen idealen Gases, pp. 261–267. Sitzber. Kgl. Preuss. Akad. Wiss. (1924)
31. Filoche, M., Mayboroda, S.: Universal mechanism for Anderson and weak localization. Proc. Natl. Acad. Sci. USA **109**(37), 14761–14766 (2012). https://doi.org/10.1073/pnas.1120432109
32. Finta, B.: About Ostrowski's iterative method in infinite-dimensional spaces. Comput. Math. Appl. **38**(9–10), 193–196 (1999). https://doi.org/10.1016/S0898-1221(99)00274-6
33. García-Ripoll, J.J., Pérez-García, V.M.: Optimizing Schrödinger functionals using Sobolev gradients: applications to quantum mechanics and nonlinear optics. SIAM J. Sci. Comput. **23**(4), 1316–1334 (2001). https://doi.org/10.1137/S1064827500377721
34. Heid, P., Stamm, B., Wihler, T.P.: Gradient flow finite element discretizations with energy-based adaptivity for the Gross–Pitaevskii equation. J. Comput. Phys. **436**, 110165 (2021). https://doi.org/10.1016/j.jcp.2021.110165

35. Henning, P., Målqvist, A., Peterseim, D.: Two-level discretization techniques for ground state computations of Bose-Einstein condensates. SIAM J. Numer. Anal. **52**(4), 1525–1550 (2014). https://doi.org/10.1137/130921520

36. Henning, P., Peterseim, D.: Sobolev gradient flow for the Gross–Pitaevskii eigenvalue problem: global convergence and computational efficiency. SIAM J. Numer. Anal. **58**(3), 1744–1772 (2020). https://doi.org/10.1137/18M1230463

37. Huhtanen, M., Ruotsalainen, S.: Real linear operator theory and its applications. Integral Equ. Oper. Theory **69**(1), 113–132 (2011). https://doi.org/10.1007/s00020-010-1825-4

38. Jarlebring, E., Kvaal, S., Michiels, W.: An inverse iteration method for eigenvalue problems with eigenvector nonlinearities. SIAM J. Sci. Comput. **36**(4), A1978–A2001 (2014). https://doi.org/10.1137/S1064827500366124

39. Kazemi, P., Eckart, M.: Minimizing the Gross–Pitaevskii energy functional with the Sobolev gradient—analytical and numerical results. Int. J. Comput. Methods **7**(3), 453–475 (2010). https://doi.org/10.1142/S0219876210002301

40. Kitchen, J.W.: Concerning the convergence of iterates to fixed points. Studia Math. **27**, 247–249 (1966). https://doi.org/10.4064/sm-27-3-247-249

41. Mehrmann, V., Voss, H.: Nonlinear eigenvalue problems: a challenge for modern eigenvalue methods. GAMM-Mitteilungen **27**(2), 121–152 (2004). https://doi.org/10.1002/gamm.201490007

42. Ortega, J.M., Rheinboldt, W.C.: Iterative Solution of Nonlinear Equations in Several Variables. Academic Press, New York-London (1970)

43. Osborne, M.R.: A new method for the solution of eigenvalue problems. Comput. J. **7**(3), 228–232 (1964). https://doi.org/10.1093/comjnl/7.3.228

44. Ostrowski, A.M.: Solution of equations and systems of equations. Pure and Applied Mathematics, vol. 9, 2nd edn. Academic Press, New York-London (1966)

45. Pitaevskii, L.P., Stringari, S.: Bose-Einstein Condensation. Oxford University Press, Oxford (2003)

46. Raza, N., Sial, S., Butt, A.R.: Numerical approximation of time evolution related to Ginzburg-Landau functionals using weighted Sobolev gradients. Comput. Math. Appl. **67**(1), 210–216 (2014). https://doi.org/10.1016/j.camwa.2013.11.006

47. Raza, N., Sial, S., Siddiqi, S.S., Lookman, T.: Energy minimization related to the nonlinear Schrödinger equation. J. Comput. Phys. **228**(7), 2572–2577 (2009). https://doi.org/10.1016/j.jcp.2008.12.016

48. Roothaan, C.C.J.: New developments in molecular orbital theory. Rev. Mod. Phys. **23**, 69–89 (1951). https://doi.org/10.1103/RevModPhys.23.69

49. Sherman, J., Morrison, W.J.: Adjustment of an inverse matrix corresponding to a change in one element of a given matrix. Ann. Math. Statist. **21**(1), 124–127 (1950). https://doi.org/10.1214/aoms/1177729893

50. Shih, M.H.: A further generalization of the Ostrowski theorem in Banach spaces. Proc. Japan Acad. Ser. A Math. Sci. **57**(3), 168–170 (1981). http://projecteuclid.org/euclid.pja/1195516492

51. Steinerberger, S.: Localization of quantum states and landscape functions. Proc. Am. Math. Soc. **145**(7), 2895–2907 (2017). https://doi.org/10.1090/proc/13343

52. Upadhyaya, P., Jarlebring, E., Rubensson, E.H.: A density matrix approach to the convergence of the self-consistent field iteration. Numer. Algebra Control Optim. **11**(1), 99–115 (2021). https://doi.org/10.3934/naco.2020018

53. Woodbury, M.A.: Inverting modified matrices. Statistical Research Group, Memo. Rep. no. 42. Princeton University, Princeton, N. J. (1950)

54. Xie, H., Xie, M.: Computable error estimates for ground state solution of Bose-Einstein condensates (2016). ArXiv e-print arXiv:1604.05228

55. Xie, H., Zhang, L., Owhadi, H.: Fast eigenpairs computation with operator adapted wavelets and hierarchical subspace correction. SIAM J. Numer. Anal. **57**(6), 2519–2550 (2019)

56. Zeidler, E.: Nonlinear functional analysis and its applications. II/A. Springer-Verlag, New York (1990). https://doi.org/10.1007/978-1-4612-0985-0

57. Zhang, Z.: Exponential convergence of Sobolev gradient descent for a class of nonlinear eigenproblems (2019). ArXiv e-print arXiv:1912.02135