

GROUND STATES OF SPIN- F BOSE-EINSTEIN CONDENSATES*TONGHUA TIAN[†], YONGYONG CAI[‡], XINMING WU[§], AND ZAIWEN WEN[¶]

Abstract. The computation of the ground states of spin- F Bose-Einstein condensates (BECs) can be formulated as an energy minimization problem with two quadratic constraints. We discretize the energy functional and constraints using the Fourier pseudospectral schemes and view the discretized problem as an optimization problem on manifold. Three different types of retractions to the manifold are designed. They enable us to apply various optimization methods on manifold to solve the problem. Specifically, an adaptive regularized Newton method is used together with a cascadic multigrid technique to accelerate the convergence. Our method is the first applicable algorithm for BECs with an arbitrary integer spin, including the complicated spin-3 BECs. Extensive numerical results on ground states of spin-1, spin-2, and spin-3 BECs with diverse interaction and optical lattice potential in one/two/three dimensions are reported to show the efficiency of our method and to demonstrate interesting physical phenomena.

Key words. Gross-Pitaevskii theory, spinor condensates, spin-2 ground state, spin-3 ground state, energy minimization

AMS subject classifications. 35Q55, 35A01, 81Q99

DOI. 10.1137/19M1271117

1. Introduction. Bose-Einstein condensate (BEC), first predicted by A. Einstein based on S. N. Bose's work, refers to the state of matter in which part of the bosons occupy the same quantum state at extremely low temperature. The earliest experimental observations of BECs were announced in 1995 [5, 14, 17] and have attracted numerous researchers into the study of condensates of dilute gases ever since [4, 15, 18, 24, 26, 28, 29]. While in early experiments the spin degrees of freedom are frozen due to the magnetic trapping, the experimental realizations of spin-1 and spin-2 condensates were achieved later by optical confinements [12, 20, 25, 30, 32] and revealed various exciting phenomena absent in single-component condensates.

Numerous theoretical studies of spinor condensates have been carried out after the experimental achievement [21, 23, 27, 31]. At zero temperature, a spin- F ($F = 1, 2, \dots$) BEC is described by a $2F + 1$ vector wave function $\Phi = (\phi_F, \dots, \phi_{-F})^T \in \mathcal{C}^{2F+1}$ and a generalized coupled Gross-Pitaevskii equation (GPE). Three important variants of it are the mass of the wave function, the magnetization, and the energy per particle. A fundamental problem in BECs is to find the condensate stationary states, which is obtained by minimizing the Gross-Pitaevskii (GP) energy functional

*Submitted to the journal's Computational Methods in Science and Engineering section June 27, 2019; accepted for publication (in revised form) May 4, 2020; published electronically July 30, 2020.
<https://doi.org/10.1137/19M1271117>

Funding: The work of the second author was supported by NSFC through grants 11771036 and 91630204. The work of the third author was supported by NSFC through grants 11971120 and 91730302 and by the Shanghai Science and Technology Commission through grant 17XD1400500. The work of the fourth author was supported by the NSFC through grants 11831002 and 1142110 and by the Beijing Academy of Artificial Intelligence.

[†]The School of Operations Research and Information Engineering, Cornell University, Ithaca, NY 14853 (tt543@cornell.edu).

[‡]School of Mathematical Sciences, Beijing Normal University, Beijing, 100875, China, and Beijing Computational Science Research Center, Beijing, 100193, China (yongyong.cai@bnu.edu.cn).

[§]Shanghai Key Laboratory for Contemporary Applied Mathematics, School of Mathematical Sciences, Fudan University, Shanghai, 200433, China (wuxinming@fudan.edu.cn).

[¶]Beijing International Center for Mathematical Research, Peking University, Beijing, 100871, China (wenzw@pku.edu.cn).

subject to the conservation of total mass and magnetization.

Different numerical methods have been proposed in the literature to compute the ground state of a spin-1 BEC [11, 7, 35, 36, 9]. Among them, a very popular method is the imaginary time method combined with a proper discretization scheme to evolve the resulted gradient flow equation under the normalization of the wave function [6, 8, 9, 11]. To apply the normalized gradient flow method (NGF) to compute the ground state of a spin- F BEC, $2F + 1$ projection constants have to be determined in the normalization step, while only two normalization conditions (i.e., the two constraints) are given. In the literature, this method is applied to compute the ground state of a spin-1 BEC through the introduction of a random variable [35, 36] or a third normalization condition [9]. Recently, a projection gradient method [11, 33] was proposed to compute ground states of spin-1 and spin-2 BECs, where a continuous normalized gradient flow was discretized by the Crank–Nicolson finite difference method with a proper and very special way to deal with the nonlinear terms. This scheme is proved to be mass- and magnetization-conservative and energy-diminishing in the discretized level. However, a fully nonlinear coupled system has to be solved at each time step.

Most of the existing numerical methods for computing the ground states of spinor BECs evolve from the gradient flow method, and thus converge at most linearly and/or require solving a large scale linear system per iteration, which leads to quite expensive computational cost. Most of them are specially designed for spin-1 or spin-2 BECs, but the spin-3 cases are rarely discussed. Meanwhile, over the last decade, some advanced optimization methods have been developed for solving minimization problems on matrix manifolds, such as the Riemannian Newton methods and trust-region methods [2, 1] with superlinear or quadratic convergence rate. Recently, Riemannian optimization has been introduced to compute the ground state of a single-component BEC problem [16]. The aim of this paper is to explore a new way to compute the ground states of spinor BECs, and to propose an efficient regularized Newton method for the general spin- F cases. We first discretize the energy functional and the constraints with the Fourier pseudospectral schemes and thus approximate the original infinite-dimensional problem by a finite-dimensional minimization problem, of which the feasible region can be proven to be a Riemannian manifold. We give the formulas of Riemannian gradient and Hessian on this manifold, and then aim to apply an adaptive regularized Newton method to solve the Riemannian optimization problem. To improve the efficiency and stability, we adopt the cascadic multigrid technique and use a Riemannian gradient method with the Barzilai–Borwein (BB) step size to compute initial points on each mesh. Three different retractions on the manifold are proposed for the implementation of Riemannian optimization algorithms. The first one is the classical projective retraction, and the second one comes from the normalized gradient flow [9]. The computation of them relies on finding a unique zero of a single-variable function, which can be done quite efficiently and accurately. The third retraction is proposed as an approximation of the first one, with a very brief closed-form formula. Extensive numerical experiments demonstrate that our approach can quickly compute an accurate approximation of the ground state and is more stable than the classical Riemannian trust-region method. The algorithm remains effective even for the complicated spin-3 BECs in 3D with an optical lattice potential, for which there exists no applicable algorithm before.

The rest of this paper is organized as follows. Specific problem statements of spin-1, spin-2, and spin-3 BECs are given in section 2. Discretizations of the energy functional and the constraints via the Fourier pseudospectral schemes are introduced

in section 3. In section 4, we give some preliminaries on Riemannian optimization, and investigate the manifold structure of the feasible region. In section 5, we present a modified version of the adaptive regularized Newton method for solving the discretized optimization problem. The three retractions are described in section 6, and detailed numerical results are reported in section 7 to illustrate the efficiency and accuracy of our algorithm. Finally, some conclusions are given in section 8.

2. Problem statement. The specific formulation of the minimization problem for computing the ground states of spin-1, spin-2, and spin-3 BECs is stated as follows:

Spin-1. For a spin-1 BEC, the GP energy functional for the spin-1 wave function is given by

(2.1)

$$E(\Phi(\cdot)) = \int_{\mathcal{R}^d} \left\{ \sum_{l=-1}^1 \left(\frac{1}{2} |\nabla \phi_l|^2 + (V(\mathbf{x}) - pl + ql^2) |\phi_l|^2 \right) + \frac{\beta_0}{2} |\Phi|^4 + \frac{\beta_1}{2} |\mathbf{F}|^2 \right\} d\mathbf{x},$$

where $\mathbf{x} = x$ in 1D, $\mathbf{x} = (x, y)^T$ in 2D, and $\mathbf{x} = (x, y, z)^T$ in 3D, $V(\mathbf{x})$ is the external confining potential, and p and q are the linear and quadratic Zeeman energy shifts, respectively. β_0 is the density-dependent interaction strength between the particles and β_1 is the spin-dependent interaction strength, $\mathbf{F} := \mathbf{F}(\Phi) = (F_x, F_y, F_z)^T \in \mathcal{R}^3$ is the spin vector given by

$$(2.2) \quad F_x = \Phi^* f_x \Phi, \quad F_y = \Phi^* f_y \Phi, \quad F_z = \Phi^* f_z \Phi,$$

where $\Phi^* = \bar{\Phi}^T$ is the conjugate transpose and f_α ($\alpha = x, y, z$) are the 3-by-3 spin-1 matrices

$$(2.3) \quad f_x = \frac{1}{\sqrt{2}} \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad f_y = \frac{i}{\sqrt{2}} \begin{pmatrix} 0 & -1 & 0 \\ 1 & 0 & -1 \\ 0 & 1 & 0 \end{pmatrix}, \quad f_z = \begin{pmatrix} 1 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 0 & -1 \end{pmatrix},$$

and $i = \sqrt{-1}$ is the imaginary unit. In detail, the components of spin vector \mathbf{F} can be written explicitly as $F_z = |\phi_1|^2 - |\phi_{-1}|^2$,

(2.4)

$$F_x = \frac{1}{\sqrt{2}} [\bar{\phi}_1 \phi_0 + \bar{\phi}_0 (\phi_1 + \phi_{-1}) + \bar{\phi}_{-1} \phi_0], \quad F_y = \frac{i}{\sqrt{2}} [-\bar{\phi}_1 \phi_0 + \bar{\phi}_0 (\phi_1 - \phi_{-1}) + \bar{\phi}_{-1} \phi_0].$$

Spin-2. For a spin-2 BEC, the GP energy is given by

(2.5)

$$E(\Phi(\cdot)) = \int_{\mathcal{R}^d} \left\{ \sum_{l=-2}^2 \left(\frac{1}{2} |\nabla \phi_l|^2 + (V(\mathbf{x}) - pl + ql^2) |\phi_l|^2 \right) + \frac{\beta_0}{2} |\Phi|^4 + \frac{\beta_1}{2} |\mathbf{F}|^2 + \frac{\beta_2}{2} |A_{00}|^2 \right\} d\mathbf{x},$$

where β_2 is the spin-singlet interaction strength and all the other parameters p, q, β_0, β_1 are the same as those in the spin-1 case, $\mathbf{F} := \mathbf{F}(\Phi) = (F_x, F_y, F_z)^T \in \mathcal{R}^3$ is the spin vector defined by (2.2), with f_α ($\alpha = x, y, z$) given by the 5-by-5 spin-2 matrices

$$(2.6) \quad f_x = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & \sqrt{\frac{3}{2}} & 0 & 0 \\ 0 & \sqrt{\frac{3}{2}} & 0 & \sqrt{\frac{3}{2}} & 0 \\ 0 & 0 & \sqrt{\frac{3}{2}} & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad f_y = i \begin{pmatrix} 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & -\sqrt{\frac{3}{2}} & 0 & 0 \\ 0 & \sqrt{\frac{3}{2}} & 0 & -\sqrt{\frac{3}{2}} & 0 \\ 0 & 0 & \sqrt{\frac{3}{2}} & 0 & -1 \\ 0 & 0 & 0 & 1 & 0 \end{pmatrix},$$

and

$$(2.7) \quad f_z = \text{diag}(2, 1, 0, -1, -2).$$

Therefore, the spin vector \mathbf{F} can be written explicitly as

$$(2.8) \quad \begin{aligned} F_x &= \bar{\phi}_2\phi_1 + \bar{\phi}_1\phi_2 + \bar{\phi}_{-1}\phi_{-2} + \bar{\phi}_{-2}\phi_{-1} + \frac{\sqrt{6}}{2}(\bar{\phi}_1\phi_0 + \bar{\phi}_0\phi_1 + \bar{\phi}_0\phi_{-1} + \bar{\phi}_{-1}\phi_0), \\ F_y &= i \left[\bar{\phi}_1\phi_2 - \bar{\phi}_2\phi_1 + \bar{\phi}_{-2}\phi_{-1} - \bar{\phi}_{-1}\phi_{-2} + \frac{\sqrt{6}}{2}(\bar{\phi}_0\phi_1 - \bar{\phi}_1\phi_0 + \bar{\phi}_{-1}\phi_0 - \bar{\phi}_0\phi_{-1}) \right], \\ F_z &= 2|\phi_2|^2 + |\phi_1|^2 - |\phi_{-1}|^2 - 2|\phi_{-2}|^2. \end{aligned}$$

Define the matrix

$$(2.9) \quad \mathbf{A} = \frac{1}{\sqrt{5}} \begin{pmatrix} 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 \end{pmatrix};$$

then $A_{00} := A_{00}(\Phi) = \Phi^T \mathbf{A} \Phi$ can be expressed as

$$(2.10) \quad A_{00} = \frac{1}{\sqrt{5}}(2\phi_2\phi_{-2} - 2\phi_1\phi_{-1} + \phi_0^2).$$

Spin-3. For a spin-3 BEC, the GP energy is given by

$$(2.11) \quad \begin{aligned} E(\Phi(\cdot)) &= \int_{\mathcal{R}^d} \left\{ \sum_{l=-3}^3 \left(\frac{1}{2}|\nabla\phi_l|^2 + (V(\mathbf{x}) - pl + ql^2)|\phi_l|^2 \right) + \frac{\beta_0}{2}|\Phi|^4 + \frac{\beta_1}{2}|\mathbf{F}|^2 + \frac{\beta_2}{2}|A_{00}|^2 \right. \\ &\quad \left. + \frac{\beta_3}{2} \sum_{l=-2}^2 |A_{2l}|^2 \right\} d\mathbf{x}, \end{aligned}$$

where β_3 is the spin-quintet interaction strength, and all the other parameters p , q , β_0 , β_1 , β_2 are the same as those in the spin-1 and spin-2 cases. $\mathbf{F} := \mathbf{F}(\Phi) = (F_x, F_y, F_z)^T \in \mathcal{R}^3$ is the spin vector defined by (2.2), with f_α ($\alpha = x, y, z$) given by the 7-by-7 spin-3 matrices

$$(2.12) \quad f_x = \begin{pmatrix} 0 & \sqrt{3/2} & 0 & 0 & 0 & 0 & 0 \\ \sqrt{3/2} & 0 & \sqrt{5/2} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{5/2} & 0 & \sqrt{3} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 & \sqrt{3} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3} & 0 & \sqrt{5/2} & 0 \\ 0 & 0 & 0 & 0 & \sqrt{5/2} & 0 & \sqrt{3/2} \\ 0 & 0 & 0 & 0 & 0 & \sqrt{3/2} & 0 \end{pmatrix},$$

$$(2.13) \quad f_y = i \begin{pmatrix} 0 & -\sqrt{3/2} & 0 & 0 & 0 & 0 & 0 \\ \sqrt{3/2} & 0 & -\sqrt{5/2} & 0 & 0 & 0 & 0 \\ 0 & \sqrt{5/2} & 0 & -\sqrt{3} & 0 & 0 & 0 \\ 0 & 0 & \sqrt{3} & 0 & -\sqrt{3} & 0 & 0 \\ 0 & 0 & 0 & \sqrt{3} & 0 & -\sqrt{5/2} & 0 \\ 0 & 0 & 0 & 0 & \sqrt{5/2} & 0 & -\sqrt{3/2} \\ 0 & 0 & 0 & 0 & 0 & \sqrt{3/2} & 0 \end{pmatrix},$$

and

$$(2.14) \quad f_z = \text{diag}(3, 2, 1, 0, -1, -2, -3).$$

The spin vector \mathbf{F} can be written explicitly as

$$\begin{aligned} F_+ &= F_x + iF_y = \sqrt{6}\psi_3\psi_2 + \sqrt{10}\psi_2\psi_1 + 2\sqrt{3}\psi_1\psi_0 + 2\sqrt{3}\psi_0\psi_{-1} \\ &\quad + \sqrt{10}\psi_{-1}\psi_{-2} + \sqrt{6}\psi_{-2}\psi_{-3}, \\ F_z &= 3|\psi_3|^2 + 2|\psi_2|^2 + |\psi_1|^2 - |\psi_{-1}|^2 - 2|\psi_{-2}|^2 - 3|\psi_{-3}|^2. \end{aligned}$$

Define the matrices

$$\mathbf{A} = \frac{1}{\sqrt{7}} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}, \quad \mathbf{A}_0 = \frac{1}{\sqrt{7}} \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{5}{2\sqrt{3}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \frac{-\sqrt{3}}{2} & 0 \\ 0 & 0 & 0 & \sqrt{\frac{2}{3}} & 0 & 0 & 0 \\ 0 & 0 & \frac{-\sqrt{3}}{2} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{5}{2\sqrt{3}} & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

and $\mathbf{A}_l = (a_{l,jk})_{7 \times 7}$ ($l = \pm 1, \pm 2$), where $a_{l,jk}$ is zero except for those $j+k = 8-l$. For the simplicity of notation, we denote $\vec{a}_l = (a_{l,1(7-l)}, a_{l,2(6-l)}, \dots, a_{l,(7-l)1})^T \in \mathcal{R}^{7-l}$ for $l = 1, 2$ and $\vec{a}_l = (a_{l,(1-l)7}, a_{l,(2-l)6}, \dots, a_{l,7(1-l)})^T \in \mathcal{R}^{7+l}$ for $l = -1, -2$ with

$$\begin{aligned} \vec{a}_{\pm 1} &= \frac{1}{\sqrt{7}} \left(\frac{5}{2\sqrt{3}}, -\frac{\sqrt{5}}{2}, \frac{1}{\sqrt{6}}, \frac{1}{\sqrt{6}}, -\frac{\sqrt{5}}{2}, \frac{5}{2\sqrt{3}} \right)^T, \\ \vec{a}_{\pm 2} &= \frac{1}{\sqrt{7}} \left(\sqrt{\frac{5}{6}}, -\sqrt{\frac{5}{3}}, \sqrt{2}, -\sqrt{\frac{5}{3}}, \sqrt{\frac{5}{6}} \right)^T. \end{aligned}$$

Then $A_{00} := A_{00}(\Psi) = \Psi^T \mathbf{A} \Psi$ and $A_{2l} := A_{2l}(\Psi) = \Psi^T \mathbf{A}_l \Psi$ can be expressed as

$$(2.15) \quad A_{00} = \frac{1}{\sqrt{7}} (2\psi_3\psi_{-3} - 2\psi_2\psi_{-2} + 2\psi_1\psi_{-1} - \psi_0^2),$$

$$(2.16) \quad A_{20} = \frac{1}{\sqrt{21}} (5\psi_3\psi_{-3} - 3\psi_1\psi_{-1} + \sqrt{2}\psi_0^2),$$

$$(2.17) \quad A_{2,\pm 1} = \frac{1}{\sqrt{21}} (5\psi_{\pm 3}\psi_{\mp 2} - \sqrt{15}\psi_{\pm 2}\psi_{\mp 1} + \sqrt{2}\psi_{\pm 1}\psi_0),$$

$$(2.18) \quad A_{2,\pm 2} = \frac{1}{\sqrt{21}} (\sqrt{10}\psi_{\pm 3}\psi_{\mp 1} - \sqrt{20}\psi_{\pm 2}\psi_0 + \sqrt{2}\psi_{\pm 1}^2).$$

For computing the ground state of a spin- F BEC, the energy functional $E(\Phi(\cdot))$ is usually subject to two constraints, i.e., the *mass* (or *normalization*),

$$(2.19) \quad \mathbf{N}(\Phi(\cdot)) := \|\Phi(\cdot)\|^2 = \int_{\mathbb{R}^d} \sum_{l=-F}^F |\phi_l(\mathbf{x})|^2 d\mathbf{x} = 1,$$

and the *magnetization* (with $M \in [-F, F]$),

$$(2.20) \quad \mathbf{M}(\Phi(\cdot)) := \int_{\mathbb{R}^d} \sum_{l=-F}^F l|\phi_l(\mathbf{x})|^2 d\mathbf{x} = M.$$

The ground state $\Phi_g(\mathbf{x})$ is obtained from the minimization of the energy functional subject to the conservation of total mass and magnetization:

Find $(\Phi_g \in S_M)$ such that

$$(2.21) \quad E_g := E(\Phi_g) = \min_{\Phi \in S_M} E(\Phi),$$

where the nonconvex set S_M is defined as

$$(2.22) \quad S_M = \left\{ \Phi = (\phi_F, \dots, \phi_{-F})^T \in \mathcal{C}^{2F+1} \mid \|\Phi\| = 1, \int_{\mathbb{R}^d} \sum_{l=-F}^F l|\phi_l(\mathbf{x})|^2 = M, E(\Phi) < \infty \right\}.$$

For $M = \pm F$ in the spin- F BEC, the constraints ensure only one component $\phi_{\pm F}$ is nonzero, and (2.21) reduces to the single-component BEC ground state problems which have been considered. Therefore, we will assume $|M| < F$ for the spin- F BEC ground states in the rest of the paper.

2.1. Notation. Given $X \in \mathcal{C}^{m \times n}$, denote \bar{X} , X^T , X^* , and $\Re(X)$ as the complex conjugate, the transpose, the complex conjugate transpose, and the real part of X , respectively. The trace of X , i.e., the sum of the diagonal elements of $X \in \mathcal{C}^{n \times n}$, is denoted by $\text{tr}(X)$. For a given vector $d \in \mathcal{C}^n$, the operator $\text{diag}(d)$ returns a square matrix in $\mathcal{C}^{n \times n}$ with the elements of d on the main diagonal, while $\text{diag}(X)$ gives a column vector in \mathcal{C}^n consisting of the main diagonal of X . The Euclidean inner product between two matrices $X, Y \in \mathcal{C}^{m \times n}$ is defined as $\langle X, Y \rangle := \sum_{jk} X_{jk} \bar{Y}_{jk} = \text{tr}(Y^* X)$.

3. Discretization schemes. In this section, we introduce discretization of the energy functional (2.11) and constraints (2.19)–(2.20) in the constrained minimization problem (2.21) for the spin-3 case. It is similar and much easier to deal with the spin-1 and spin-2 cases. Due to the external trapping potential, the ground state of (2.21) decays exponentially as $|\mathbf{x}| \rightarrow \infty$. Thus we can truncate the energy functional and constraints from the whole space \mathcal{R}^d to a bounded computational domain U which is chosen large enough such that the truncation error is negligible with periodic boundary conditions. Then we approximate spatial derivatives via the Fourier pseudospectral (FP) method and the integrals via the composite trapezoidal quadrature. For simplicity of notation, we only present the FP discretization in 1D. Extensions to 2D and 3D are straightforward for tensor grids, and details are omitted here for brevity.

For $d = 1$, we take a bounded interval $U = (a, b)$. Let $h = (b - a)/n$ be the spatial mesh size with n an even positive integer, and denote $x_j = a + jh$ for $j = 0, 1, \dots, n$. Let ϕ_{jl} be the numerical approximation of $\phi_l(x_j)$ for $j = 0, 1, \dots, n$ and $l = 3, \dots, -3$ satisfying $\phi_{0l} = \phi_{nl}$, and denote $X = (\sqrt{h}\phi_{jl}) \in \mathcal{C}^{n \times 7}$ ($j = 0, 1, \dots, n - 1$, $l = 3, \dots, -3$).

$$(3.1) \quad E(\Phi) = \sum_{j=0}^{n-1} \int_{x_j}^{x_{j+1}} \left\{ \sum_{l=-3}^3 \left(-\frac{1}{2} \bar{\phi}_l \partial_{xx} \phi_l + (V(x) - pl + ql^2) |\phi_l|^2 \right) \right. \\ \left. + \frac{\beta_0}{2} |\Phi|^4 + \frac{\beta_1}{2} |\mathbf{F}|^2 + \frac{\beta_2}{2} |A_{00}|^2 + \frac{\beta_3}{2} \sum_{l=-2}^2 |A_{2l}|^2 \right\} dx$$

$$(3.2) \quad \approx h \sum_{j=0}^{n-1} \left\{ \sum_{l=-3}^3 \left(-\frac{1}{2} \bar{\phi}_l(x_j) \partial_{xx}^f \phi_l \Big|_j + (V(x_j) - pl + ql^2) |\phi_l(x_j)|^2 \right) \right. \\ \left. + \frac{\beta_0}{2} |\Phi(x_j)|^4 + \frac{\beta_1}{2} |\mathbf{F}(x_j)|^2 + \frac{\beta_2}{2} |A_{00}(x_j)|^2 + \frac{\beta_3}{2} \sum_{l=-2}^2 |A_{2l}(x_j)|^2 \right\}$$

where the FP differential operator is given as

$$(3.3) \quad \partial_{xx}^f \phi \Big|_j = -\frac{1}{n} \sum_{p=-n/2}^{n/2-1} \lambda_p^2 \tilde{\phi}_{pl} e^{i \frac{2\pi j p}{n}},$$

with

$$(3.4) \quad \tilde{\phi}_{pl} = \sum_{j=0}^{n-1} \phi_l(x_j) e^{-i \frac{2\pi j p}{n}}, \quad \lambda_p = \frac{2\pi p}{b-a}, \quad p = -\frac{n}{2}, \dots, \frac{n}{2} - 1.$$

Introduce $\mathbf{V} = \text{diag}(V(x_0), \dots, V(x_{n-1}))$, $\mathbf{B} = \text{diag}(b)$ with $b = (b_3, b_2, \dots, b_{-3})^T$ ($b_l = -pl + ql^2$, $l = 3, \dots, -3$), $\Lambda = \text{diag}(\lambda_{-\frac{n}{2}}^2, \dots, \lambda_{\frac{n}{2}-1}^2)$, and $\mathbf{C} = (c_{jp}) \in \mathcal{C}^{n \times n}$ with entries $c_{jp} = e^{-i \frac{2\pi j p}{n}}$ for $j = 0, \dots, n-1$ and $p = -\frac{n}{2}, \dots, \frac{n}{2} - 1$. Plugging (3.3) and (3.4) into (3.2) and replacing $\phi_l(x_j)$ with ϕ_{jl} , we get the finite-dimensional approximation to the energy functional defined as

$$(3.5) \quad E_h(X) = \frac{1}{2} \text{tr}(X^* L X) + \text{tr}(X^* \mathbf{V} X) + \text{tr}(X \mathbf{B} X^*) \\ + \frac{\beta_0}{2h} \rho^T \rho + \frac{\beta_1}{2h} \sum_{\alpha=x,y,z} \mathbf{F}_\alpha^T \mathbf{F}_\alpha + \frac{\beta_2}{2h} \mathbf{A}_{00}^* \mathbf{A}_{00} + \frac{\beta_3}{2h} \sum_{l=-2}^2 \mathbf{A}_{2l}^* \mathbf{A}_{2l},$$

where $L = \mathbf{C}^* \Lambda \mathbf{C}$ is the matrix representation of the discrete negative Laplace operator and

$$(3.6) \quad \rho = \text{diag}(XX^*), \quad \mathbf{F}_\alpha = \text{diag}(X f_\alpha^T X^*), \quad \alpha = x, y, z,$$

$$(3.7) \quad \mathbf{A}_{00} = \text{diag}(X \mathbf{A} X^T), \quad \mathbf{A}_{2l} = \text{diag}(X \mathbf{A}_l X^T), \quad l = -2, \dots, 2,$$

are column vectors. In fact, the first term in (3.5) can be computed efficiently at cost $O(n \ln n)$ through the fast Fourier transform (FFT).

Similarly, letting $D = \text{diag}(3, \dots, -3)$, the constraints (2.19)–(2.20) can be truncated and discretized as

$$(3.8) \quad \mathbf{N}(\Phi(\cdot)) \approx h \sum_{j=0}^{n-1} \sum_{l=-3}^3 |\phi_{jl}|^2 = \text{tr}(X^* X) = 1,$$

$$(3.9) \quad \mathbf{M}(\Phi(\cdot)) \approx h \sum_{j=0}^{n-1} \sum_{l=-3}^3 l |\phi_{jl}|^2 = \text{tr}(X^* X D) = M,$$

which immediately implies that the set S_M can be discretized as

$$(3.10) \quad S_h = \{X \in \mathcal{C}^{n \times 7} \mid \text{tr}(X^* X) = 1, \text{tr}(X^* XD) = M, E_h(X) < \infty\}.$$

Hence, the original problem (2.21) with $d = 1$ can be approximated by the discretized minimization problem via the FP discretization:

$$(3.11) \quad E_{h,g} := E_h(X_g) = \min_{X \in S_h} E_h(X).$$

To solve the discrete minimization problem (3.11), it is often necessary to compute the gradient and Hessian matrix of the discrete energy $E_h(X)$. The second-order Taylor expansion of $E_h(X)$ can be expressed as

$$(3.12) \quad E_h(X + \Delta X) = E_h(X) + \Re \langle \nabla E_h, \Delta X \rangle + \frac{1}{2} \Re \langle (\nabla^2 E_h) \Delta X, \Delta X \rangle + \text{h.o.t.},$$

where h.o.t. is short for higher-order terms. By a simple calculation, we can get the gradient

$$(3.13) \quad \begin{aligned} \nabla E_h(X) = & LX + 2VX + 2XB \\ & + \frac{2\beta_0}{h} \text{diag}(\rho)X + \frac{2\beta_1}{h} \sum_{\alpha=x,y,z} \text{diag}(F_\alpha)Xf_\alpha^T \\ & + \frac{2\beta_2}{h} \text{diag}(A_{00})\bar{X}\mathbf{A} + \frac{2\beta_3}{h} \sum_{l=-2}^2 \text{diag}(A_{2l})\bar{X}\mathbf{A}_l \end{aligned}$$

and the Hessian-vector product

$$(3.14) \quad \begin{aligned} \nabla^2 E_h(X)[Z] = & LZ + 2VZ + 2ZB \\ & + \frac{2\beta_0}{h} (\text{diag}(\rho)Z + 2\text{diag}(\Re(ZX^*))X) \\ & + \frac{2\beta_1}{h} \sum_{\alpha=x,y,z} (\text{diag}(F_\alpha)Zf_\alpha^T + 2\text{diag}(\Re(Zf_\alpha^T X^*))Xf_\alpha^T) \\ & + \frac{2\beta_2}{h} (\text{diag}(A_{00})\bar{Z}\mathbf{A} + 2\text{diag}(Z\mathbf{A}X^T)\bar{X}\mathbf{A}) \\ & + \frac{2\beta_3}{h} \sum_{l=-2}^2 (\text{diag}(A_{2l})\bar{Z}\mathbf{A}_l + 2\text{diag}(Z\mathbf{A}_l X^T)\bar{X}\mathbf{A}_l). \end{aligned}$$

4. Manifold structure. In the ground state of a spin- F BEC, we have $M \leftrightarrow -M \iff \phi_l \leftrightarrow \phi_{-l}$. Thus we only discuss the cases where $M \geq 0$. Express X as $X = X_r + iX_i$, where $X_r, X_i \in \mathcal{R}^{n \times (2F+1)}$. Let $(X_r; X_i) = (u_F, u_{F-1}, \dots, u_0, \dots, u_{-F+1}, u_{-F}) \in \mathcal{R}^{2n \times (2F+1)}$, and let $u \in \mathcal{R}^N$ be the reconstructed column vector of this matrix, where $N = 2n(2F+1)$. Introduce

$$(4.1) \quad \Gamma = \text{diag}(FI_{2n}, \dots, I_{2n}, 0, -I_{2n}, \dots, -FI_{2n}) \in \mathcal{R}^{N \times N},$$

where I_{2n} denotes the identity matrix of size $2n$. Then the constraints can be discretized as

$$(4.2) \quad \mathcal{M} = \{u \in \mathcal{R}^N \mid u^T u = 1, u^T \Gamma u = M\},$$

in which $0 \leq M < F$. Defining $\tilde{E}(u) := E_h(X)$, our model problem can be formulated as

$$(4.3) \quad \min \quad \tilde{E}(u), \quad \text{s.t. } u \in \mathcal{M}.$$

This is a nonconvex optimization problem with constraints. Observe that \mathcal{M} is a level set of the function

$$(4.4) \quad \mathbf{G}(u) = \frac{1}{2}(u^T u - 1, u^T \Gamma u - M)^T.$$

When $M \notin \mathcal{Z}$, $\nabla \mathbf{G}(u) = (u, \Gamma u)^T$ has full rank at every point $u \in \mathcal{M}$; thus according to Proposition 3.3.3 in [2], \mathcal{M} is a closed embedded submanifold of \mathcal{R}^N of dimension $N - 2$. In the following discussion, we will let u be an arbitrary point on the manifold \mathcal{M} and assume f is a smooth real-valued function in a neighborhood of u .

Given a curve $\gamma(t) : \mathcal{R} \rightarrow \mathcal{M} \subset \mathcal{R}^N$ through u at $t = 0$, the associated tangent vector $\dot{\gamma}(0)$ is defined such that

$$(4.5) \quad \dot{\gamma}(0)f = \nabla f(u)^T \gamma'(0),$$

where $\gamma'(0)$ denotes its derivative in \mathcal{R}^N space [2]. Since \mathcal{M} is a level set of the constant-rank function \mathbf{G} , the tangent space $T_u \mathcal{M}$ reads

$$(4.6) \quad T_u \mathcal{M} = \ker(\nabla \mathbf{G}(u)) = \{\xi \in \mathcal{R}^N \mid u^T \xi = 0, u^T \Gamma \xi = 0\}.$$

We naturally define the inner product $\langle \cdot, \cdot \rangle_u$ and the norm $\|\cdot\|_u$ on $T_u \mathcal{M}$ as

$$(4.7) \quad \langle \xi, \zeta \rangle_u := \xi^T \zeta, \quad \|\xi\|_u := \sqrt{\xi^T \xi}, \quad \xi, \zeta \in T_u \mathcal{M}.$$

Under such a metric, the Riemannian gradient $\text{grad}f(u)$, defined as the unique element of $T_u \mathcal{M}$ satisfying

$$(4.8) \quad \langle \text{grad}f(u), \xi \rangle_u = \frac{d}{dt} f(u + t\xi)|_{t=0} \quad \forall \xi \in T_u \mathcal{M},$$

can be written as

$$(4.9) \quad \text{grad}f(u) = P_u \nabla f(u),$$

where P_u denotes the orthogonal projection from \mathcal{R}^N onto $T_u \mathcal{M}$. From (4.6) we can easily derive the formula of P_u .

LEMMA 4.1 (P_u). *For an arbitrary point $w \in \mathcal{R}^N$, the orthogonal projection of w onto $T_u \mathcal{M}$ reads*

$$(4.10) \quad P_u w = w - \frac{u^T \Gamma^2 u \cdot u^T w - M u^T \Gamma w}{u^T \Gamma^2 u - M^2} u + \frac{M u^T w - u^T \Gamma w}{u^T \Gamma^2 u - M^2} \Gamma u.$$

Proof. From (4.6) and the definition of P_u , we have

$$(4.11) \quad (T_u \mathcal{M})^\perp = \{\alpha u + \beta \Gamma u \mid \alpha, \beta \in \mathcal{R}\}$$

and $w - P_u w \in (T_u \mathcal{M})^\perp$; therefore there exists $\alpha_w, \beta_w \in \mathcal{R}$ such that

$$(4.12) \quad P_u w = w - \alpha_w u - \beta_w \Gamma u.$$

Noticing that $P_u w \in T_u \mathcal{M}$ implies

$$(4.13) \quad u^T (P_u w) = 0, \quad u^T \Gamma (P_u w) = 0,$$

we can obtain from (4.12) that

$$(4.14) \quad \alpha_w u^T u + \beta_w u^T \Gamma u = u^T w, \quad \alpha_w u^T \Gamma u + \beta_w u^T \Gamma^2 u = u^T \Gamma w.$$

In view of the fact that $u^T u = 1$ and $u^T \Gamma u = M$, (4.14) can be simplified as

$$(4.15) \quad \begin{pmatrix} 1 & M \\ M & u^T \Gamma^2 u \end{pmatrix} \begin{pmatrix} \alpha_w \\ \beta_w \end{pmatrix} = \begin{pmatrix} u^T w \\ u^T \Gamma w \end{pmatrix}.$$

It follows directly from *Cauchy–Schwarz inequality* that

$$(4.16) \quad u^T \Gamma^2 u = \sum_{l=-F}^F l^2 \|u_l\|_2^2 > \frac{(\sum_{l=-F}^F l \|u_l\|_2^2)^2}{\sum_{l=-F}^F \|u_l\|_2^2} = \frac{(u^T \Gamma u)^2}{u^T u} = M^2,$$

which ensures the linear system (4.15) has a unique solution:

$$(4.17) \quad \begin{pmatrix} \alpha_w \\ \beta_w \end{pmatrix} = \frac{1}{u^T \Gamma^2 u - M^2} \begin{pmatrix} u^T \Gamma^2 u & -M \\ -M & 1 \end{pmatrix} \begin{pmatrix} u^T w \\ u^T \Gamma w \end{pmatrix}.$$

Substituting (4.17) into (4.12) yields the formula (4.10). \square

Let $\mathfrak{X}(\mathcal{M})$ be the set of smooth vector fields on \mathcal{M} . The Riemannian Hessian $\text{Hess } f(u)$ is a linear mapping from $T_u \mathcal{M}$ into itself defined as

$$(4.18) \quad \text{Hess } f(u)[\Xi(u)] = (\tilde{\nabla}_\Xi \text{grad } f)(u), \quad \Xi \in \mathfrak{X}(\mathcal{M}),$$

where $\tilde{\nabla}$ denotes the Riemannian connection of \mathcal{M} . Since \mathcal{M} is a Riemannian submanifold of \mathbb{R}^N , according to [2] its Riemannian connection reads

$$(4.19) \quad (\tilde{\nabla}_\Xi \Pi)(u) = P_u(\nabla \Pi(u) \cdot \Xi(u)), \quad \Xi, \Pi \in \mathfrak{X}(\mathcal{M}).$$

Thus we have

$$(4.20) \quad \text{Hess } f(u)[\xi] = P_u(\nabla \text{grad } f(u) \cdot \xi), \quad \xi \in T_u \mathcal{M}.$$

The formula of $\text{Hess } f(u)$ is given in the following lemma.

LEMMA 4.2 ($\text{Hess } f(u)$). *Given a tangent vector $\xi \in T_u \mathcal{M}$, and letting g and H be the Euclidean gradient and Euclidean Hessians of f , respectively, we obtain*

$$(4.21) \quad \text{Hess } f(u)[\xi] = h_e - \alpha_g \xi - \beta_g \Gamma \xi - \frac{u^T \Gamma^2 u \cdot u^T h_e - M \beta_u}{\alpha_u} u + \frac{M u^T h_e - \beta_u}{\alpha_u} \Gamma u,$$

where

$$(4.22) \quad \alpha_g := \left(1 + \frac{M^2}{\alpha_u}\right) u^T g - \frac{M}{\alpha_u} u^T \Gamma g, \quad \beta_g := -\frac{M}{\alpha_u} u^T g + \frac{1}{\alpha_u} u^T \Gamma g,$$

and $h_e := H(u) \cdot \xi$, $\alpha_u := u^T \Gamma^2 u - M^2$, $\beta_u := u^T \Gamma h_e - \beta_g u^T \Gamma^2 \xi$.

Proof. Recalling Lemma 4.1 and (4.9), we get

$$(4.23) \quad \text{grad}f(u) = g(u) - \alpha_g u - \beta_g \Gamma u$$

and

$$\begin{aligned} \nabla \text{grad}f(u) \cdot \xi &= \nabla g(u) \cdot \xi - \nabla(\alpha_g u) \cdot \xi - \nabla(\beta_g \Gamma u) \cdot \xi \\ &= h_e - \alpha_g \xi - \beta_g \Gamma \xi - (\nabla \alpha_g^T \xi) u - (\nabla \beta_g^T \xi) \Gamma u. \end{aligned}$$

Since $(\nabla \alpha_g^T \xi) u, (\nabla \beta_g^T \xi) \Gamma u \in (T_u \mathcal{M})^\perp$, we have

$$(4.24) \quad \text{Hess}f(u)[\xi] = P_u(\nabla \text{grad}f(u) \cdot \xi) = P_u(h_e - \alpha_g \xi - \beta_g \Gamma \xi).$$

For $\xi \in T_u \mathcal{M}$, Lemma 4.1 and (4.24) lead to formula (4.21). \square

The first-order and second-order optimality conditions for optimization problems on Riemannian manifolds coincide with the conventional ones [34]. If u^* is a local solution of (4.3), we have $\text{grad}\tilde{E}(u^*) = 0$, and all the points u at which $\text{grad}\tilde{E}(u) = 0$ are called *stationary points* of \tilde{E} .

Line search optimization methods in \mathbb{R}^N are based on the update formula

$$(4.25) \quad u_{k+1} = u_k + t_k \eta_k,$$

where $\eta_k \in \mathbb{R}^N$ is the *search direction* and $t_k > 0$ is the *step size*. Correspondingly, when (4.25) is generalized to a manifold, η_k is selected as a tangent vector, and the line search procedure relies on the concept of *retraction*.

DEFINITION 4.3 (retraction). *A retraction on a manifold \mathcal{M} is a smooth mapping R from the tangent bundle $T\mathcal{M}$ onto \mathcal{M} with the following properties. Let R_u denote the restriction of R to $T_u\mathcal{M}$:*

- (i) $R_u(0_u) = u$, where 0_u denotes the zero element of $T_u\mathcal{M}$.
- (ii) With the canonical identification $T_{0_u}T_u\mathcal{M} \simeq T_u\mathcal{M}$, R_u satisfies

$$DR_u(0_u) = \text{id}_{T_u\mathcal{M}},$$

where $DR_u(0_u)$ denotes the directional derivative of R_u at 0_u and $\text{id}_{T_u\mathcal{M}}$ denotes the identity mapping on $T_u\mathcal{M}$.

The condition (ii) is referred to as the *local rigidity* condition. It ensures that for all $\xi \in T_u\mathcal{M}$, the curve $\gamma_\xi : t \rightarrow R_u(t\xi)$ satisfies $\dot{\gamma}_\xi(0) = \xi$ [2]. Conceptually, R_u is a mapping from $T_u\mathcal{M}$ to \mathcal{M} that preserves gradients at u and allows us to implement the curvilinear search procedure on manifold based on the update formula

$$(4.26) \quad u_{k+1} = R_{u_k}(t_k \eta_k).$$

Remark 4.1. When $M \in \mathcal{Z}$, \mathcal{M} is not a well-defined manifold. However, by restricting the feasible region to

$$\widetilde{\mathcal{M}} := \{u \in \mathcal{M} \mid \text{at least two components of } u \text{ are nonzero}\},$$

we can also define the above structures and the formulas still work. This modification does not change our numerical experiments.

5. A modified adaptive regularized Newton method. We aim to solve (4.3) with a modified version of the adaptive regularized Newton method (ARNT) developed in [22]. At the k th iteration, ARNT uses a second-order Taylor model with a penalization term to approximate the original objective function but keeps the constraint $u \in \mathcal{M}$, i.e.,

$$(5.1) \quad \min_{u \in \mathcal{M}} m_k(u) := \langle \nabla \tilde{E}(u_k), u - u_k \rangle + \frac{1}{2} \langle H_k(u - u_k), u - u_k \rangle + \frac{\sigma_k}{2} \|u - u_k\|_2^2,$$

where $\langle \cdot, \cdot \rangle$ denotes the dot product in \mathcal{R}^N , H_k is the Euclidean Hessian of \tilde{E} at u_k , and σ_k is the regularization parameter chosen carefully. The Riemannian trust-region method (RTR) [1] can be applied to (4.3) as well. The main difference is that RTR uses the standard trust-region framework with the Riemannian gradient and Hessian to construct the following subproblem on the Tangent space:

$$(5.2) \quad \min_{\xi \in T_{u_k} \mathcal{M}} \langle \text{grad } \tilde{E}(u_k), \xi \rangle + \frac{1}{2} \langle \text{Hess } \tilde{E}(u_k)[\xi], \xi \rangle, \quad \text{s.t. } \|\xi\|_{u_k} \leq \Delta_k,$$

where Δ_k is the trust-region radius. The detailed description of RTR is omitted for the sake of space.

The subproblem (5.1) is solved approximately by applying a modified conjugate gradient (CG) method to the linear system

$$(5.3) \quad \text{grad } m_k(u_k) + \text{Hess } m_k(u_k)[\xi] = 0.$$

The method terminates when either a certain accuracy is reached or negative curvature is detected. It outputs two vectors s_k and d_k , where s_k is the solution computed by the CG method and d_k represents the negative curvature information. The new search direction ξ_k is chosen as

$$(5.4) \quad \xi_k = \begin{cases} s_k + \tau_k d_k & \text{if } d_k \neq 0, \\ s_k & \text{if } d_k = 0, \end{cases} \quad \text{with } \tau_k := \frac{\langle d_k, \text{grad } m_k(u_k) \rangle_{u_k}}{\langle d_k, \text{Hess } m_k(u_k)[d_k] \rangle_{u_k}},$$

which is a descent direction (cf. [22, Lemma 7]).

After construction of ξ_k , a monotone Armijo-based curvilinear search is conducted to generate a trial point,

$$(5.5) \quad z_k = R_{u_k}(\alpha_0 \delta^\varsigma \xi_k),$$

where ς is the smallest integer satisfying

$$(5.6) \quad m_k(R_{u_k}(\alpha_0 \delta^\varsigma \xi_k)) \leq \rho \alpha_0 \delta^\varsigma \langle \text{grad } m_k(u_k), \xi_k \rangle_{u_k}$$

and $\rho, \delta \in (0, 1)$, $\alpha_0 \in (0, 1]$ are given constants.

In order to monitor the acceptance of the trial point z_k and adjust the regularization parameter σ_k , the above procedure is embedded in a trust-region framework where σ_k plays a similar role as the trust-region radius and is updated according to the ratio

$$(5.7) \quad \rho_k = \frac{\tilde{E}(z_k) - \tilde{E}(u_k)}{m_k(z_k)}.$$

ARNT may exhibit a certain instability when directly applied to solve (4.3). To improve its performance, we combine it with the cascadic multigrid method in [13].

Algorithm 1: A Modified Adaptive Regularized Newton Method.

Choose an initial mesh \mathcal{T}^0 and $u^{(0)}$. Set $j = 0$.

```

while  $j \leq m$  do
    >On each mesh, we first compute a point using RGBB.
    Input  $u_0 = u^{(j)}$ . Set  $k = 0, C_0 = \tilde{E}(u_0), Q_0 = 1$ .
    while stopping conditions not met do
        Compute  $\eta_k = -\text{grad } \tilde{E}(u_k)$ .
        Compute  $\gamma_k, C_k, Q_k$  and find the  $\varsigma$  satisfying (5.8).
        Set  $u_{k+1} \leftarrow R_{u_k}(\gamma_k \delta^\varsigma \eta_k)$ .
        Set  $k \leftarrow k + 1$ .
    >Then we switch to the Newton-type steps from the point obtained by RGBB.
    Input  $u_0 = u_k$ . Choose  $0 < \eta_1 \leq \eta_2 < 1, 0 < \gamma_0 < 1 < \gamma_1 \leq \gamma_2$  and an initial regularization parameter  $\sigma_0 > 0$ . Set  $k = 0$ .
    while stopping conditions not met do
        Compute a new trial point  $z_k$  according to (5.5) and (5.6).
        Compute the ratio via (5.7).
        if  $\rho_k \geq \eta_1$  then
            Set  $u_{k+1} = z_k$ .
            if  $\rho_k \geq \eta_2$  then choose  $\sigma_{k+1} \in (0, \gamma_0 \sigma_k]$ ,
            else choose  $\sigma_{k+1} \in [\gamma_0 \sigma_k, \gamma_1 \sigma_k]$ .
        else
            Set  $u_{k+1} = u_k$ .
            Choose  $\sigma_{k+1} \in [\gamma_1 \sigma_k, \gamma_2 \sigma_k]$ .
        Set  $k \leftarrow k + 1$ .
    Set  $u^{(j+1)} = u_k$ . Refine the mesh  $\mathcal{T}^j$  uniformly to obtain  $\mathcal{T}^{j+1}$ .
     $j \leftarrow j + 1$ .

```

In detail, we first solve (4.3) on the coarsest mesh, and then use the obtained solution as the initial guess of the problem on a finer mesh, and repeat until reaching the finest mesh.

The modified adaptive regularized Newton method (still referred to as ARNT in this paper) is presented in Algorithm 1. There are two inner loops on each mesh. In the first inner loop, we use the Riemannian gradient method with a BB step size (RGBB) in [22] to compute an initial point for ARNT. At the k th iteration, RGBB performs a nonmonotone Armijo-based curvilinear search along the steepest descent direction $\eta_k = -\text{grad } \tilde{E}(u_k)$. Given $\rho, \varrho, \delta \in (0, 1)$, it tries to find the smallest integer ς satisfying

$$(5.8) \quad \tilde{E}(R_{u_k}(\gamma_k \delta^\varsigma \eta_k)) \leq C_k + \rho \gamma_k \delta^\varsigma \langle \text{grad } \tilde{E}(u_k), \eta_k \rangle_{u_k},$$

where the initial step size γ_k is computed as $\gamma_k = |\langle s_{k-1}, v_{k-1} \rangle_{u_k}| / \langle v_{k-1}, v_{k-1} \rangle_{u_k}$, with $s_{k-1} = u_k - u_{k-1}$, $v_{k-1} = \text{grad } \tilde{E}(u_k) - \text{grad } \tilde{E}(u_{k-1})$. The value C_{k+1} is calculated via $C_{k+1} = (\varrho Q_k C_k + \tilde{E}(u_{k+1})) / Q_{k+1}$, with $C_0 = \tilde{E}(u_0)$, $Q_{k+1} = \varrho Q_k + 1$, and $Q_0 = 1$. Once certain inaccurate stopping conditions are met in RGBB, we switch to ARNT and start the second inner loop from the point obtained by RGBB. The detailed stopping conditions are described in subsection 7.1.

6. Retractions. The selection of retractions can affect the performance of Riemannian optimization algorithms. In this section, we try to find retractions of the form

$$(6.1) \quad R_u(\xi_u) := \psi(u + \xi_u), \quad u \in \mathcal{M}, \xi_u \in T_u \mathcal{M},$$

where ψ is some “projection” from a neighborhood of \mathcal{M} in \mathbb{R}^N to \mathcal{M} .

For $w = (w_F, w_{F-1}, \dots, w_{-F}) \in \mathbb{R}^N$ ($w_l \in \mathbb{R}^{2n}$, $N = 2n(2F+1)$, $l = F, \dots, -F$), we define two functions $f_1(w), f_2(w) \in \{F, F-1, \dots, -F\}$:

$$(6.2) \quad f_1(w) := \min\{l \mid w_l \neq 0\}, \quad f_2(w) := \max\{l \mid w_l \neq 0\} \quad \forall w \in \mathbb{R}^N.$$

Observe that at every point $u \in \mathcal{M}$, the constraints indicate

$$(6.3) \quad \sum_{l=-F}^F (l - M) \|u_l\|_2^2 = 0.$$

Thus when $M \notin \mathcal{Z}$, we have $f_1(u) < M, f_2(u) > M$, which is equivalent to

$$(6.4) \quad \sum_{l < M} \|u_l\|_2^2 > 0, \quad \sum_{l > M} \|u_l\|_2^2 > 0;$$

when $M \in \mathcal{Z}$, (6.4) also holds for $u \in \widetilde{\mathcal{M}}$. Define the open set Ω as

$$(6.5) \quad \Omega := \left\{ w \in \mathbb{R}^N \mid \sum_{l < M} \|w_l\|_2^2 > 0, \quad \sum_{l > M} \|w_l\|_2^2 > 0 \right\};$$

then $\Omega \supset \mathcal{M}$ for $M \notin \mathcal{Z}$ and $\Omega \supset \widetilde{\mathcal{M}}$ for $M \in \mathcal{Z}$. For simplicity of presentation, we will discuss three different retractions from Ω to \mathcal{M} ($M \notin \mathcal{Z}$), and all the results hold also for $\widetilde{\mathcal{M}}$ ($M \in \mathcal{Z}$).

6.1. Projective retraction. The most intuitive retraction ψ is given by the *projection operator* $\mathcal{P}_{\mathcal{M}}$, which is defined as

$$(6.6) \quad \mathcal{P}_{\mathcal{M}}(w) = \arg \min_{z \in \mathcal{M}} \frac{1}{2} \|z - w\|_2^2, \quad w \in \mathbb{R}^N.$$

According to [3], $\mathcal{P}_{\mathcal{M}}$ is a well-defined function (existence and uniqueness of the projection hold) in a neighborhood $\widetilde{\Omega} \subset \Omega$ of \mathcal{M} , and the mapping $R_u(\xi_u) := \mathcal{P}_{\mathcal{M}}(u + \xi_u)$ is a well-defined retraction on \mathcal{M} , called the *projective retraction* in this paper. The explicit formula is given as follows.

LEMMA 6.1. *For an arbitrary point $w \in \widetilde{\Omega}$, $\mathcal{P}_{\mathcal{M}}(w)$ reads*

$$(6.7) \quad \mathcal{P}_{\mathcal{M}}(w)_l = \begin{cases} 0 & \text{if } l < f_1(w) \text{ or } l > f_2(w), \\ \frac{w_l}{1-\mu-l\lambda} & \text{if } f_1(w) \leq l \leq f_2(w), \end{cases}$$

where

$$(6.8) \quad \lambda = r \sqrt{\sum_{l=f_1(w)}^{f_2(w)} \frac{\|w_l\|_2^2}{[1 - (l - M)r]^2}}, \quad \mu = 1 - (1 + Mr) \sqrt{\sum_{l=f_1(w)}^{f_2(w)} \frac{\|w_l\|_2^2}{[1 - (l - M)r]^2}},$$

and r is the unique zero of the function

$$(6.9) \quad h_1(t) = \sum_{l=f_1(w)}^{f_2(w)} \frac{(l-M)\|w_l\|_2^2}{[1-(l-M)t]^2}, \quad t \in \left(\frac{1}{f_1(w)-M}, \frac{1}{f_2(w)-M} \right).$$

Proof. Define the Lagrangian function of (6.6) as

$$(6.10) \quad L(z, \mu, \lambda) = \frac{1}{2}\|z-w\|_2^2 - \frac{\mu}{2}(\|z\|_2^2 - 1) - \frac{\lambda}{2} \left(\sum_{l=-F}^F l\|z_l\|_2^2 - M \right).$$

Let $z = \mathcal{P}_{\mathcal{M}}(w)$; then the condition $\nabla_l L(z, \mu, \lambda) = 0$ gives

$$(6.11) \quad (1 - \mu - l\lambda)z_l = w_l, \quad l = F, \dots, -F.$$

Since $\|z - w\|_2^2 = \|z\|_2^2 + \|w\|_2^2 - 2\sum_l w_l^T z_l = 1 + \|w\|_2^2 - 2\sum_l w_l^T z_l$ ($z \in \mathcal{M}$), the projection z maximizes $w_l^T z_l$, which leads to $1 - \mu - l\lambda > 0$ for $w_l \neq 0$. On the other hand, if $w_l = 0$, then $z_l = 0$; otherwise, substituting z_l with $-z_l$ yields a different projection of w , which contradicts the uniqueness.

Since $w_{f_1(w)}$ and $w_{f_2(w)}$ are nonzero, we have

$$(6.12) \quad 1 - \mu - f_1(w)\lambda > 0, \quad 1 - \mu - f_2(w)\lambda > 0,$$

which is equivalent to $1 - \mu - M\lambda > (f_1(w) - M)\lambda$, $1 - \mu - M\lambda > (f_2(w) - M)\lambda$, and

$$(6.13) \quad 1 - \mu - M\lambda > 0, \quad \frac{1}{f_1(w)-M} < \frac{\lambda}{1-\mu-M\lambda} < \frac{1}{f_2(w)-M}.$$

The inequalities in (6.12) indicate that $1 - \mu - l\lambda > 0$ for $l = f_1(w) + 1, \dots, f_2(w) - 1$, and from (6.11) and $z \in \mathcal{M}$ we have

$$(6.14) \quad \sum_{l=f_1(w)}^{f_2(w)} \frac{\|w_l\|_2^2}{(1-\mu-l\lambda)^2} = 1, \quad \sum_{l=f_1(w)}^{f_2(w)} \frac{l\|w_l\|_2^2}{(1-\mu-l\lambda)^2} = M.$$

In view of (6.13)–(6.14), denoting

$$(6.15) \quad s = (1 - \mu - M\lambda)^2, \quad r = \frac{\lambda}{1 - \mu - M\lambda},$$

recalling the definition of function $h_1(\cdot)$ in (6.9), we have $s > 0$, $\frac{1}{f_1(w)-M} < r < \frac{1}{f_2(w)-M}$, and

$$(6.16) \quad h_1(r) = 0, \quad s = \sum_{l=f_1(w)}^{f_2(w)} \frac{\|w_l\|_2^2}{[1-(l-M)r]^2}.$$

For any $t \in (\frac{1}{f_1(w)-M}, \frac{1}{f_2(w)-M})$, we have

$$(6.17) \quad h'_1(t) = \sum_{l=f_1(w)}^{f_2(w)} \frac{2(l-M)^2\|w_l\|_2^2}{[1-(l-M)t]^3} > 0.$$

In addition, noticing that

$$(6.18) \quad \lim_{t \rightarrow \frac{1}{f_1(w)-M}+0} h_1(t) = -\infty, \quad \lim_{t \rightarrow \frac{1}{f_2(w)-M}-0} h_1(t) = +\infty,$$

$h_1(\cdot)$ has exactly one zero in $(\frac{1}{f_1(w)-M}, \frac{1}{f_2(w)-M})$. Substituting (6.16) into (6.15), the formulas of λ and μ can be obtained accordingly. \square

We remark that (6.1) can be applied to any $w \in \Omega$. For spin-1 cases, the closed-form solution of (6.6) is computable.

LEMMA 6.2. *When $F = 1$, given any nonzero $w \in \Omega$, the optimal solution $z = (z_1; z_0; z_{-1})$ of (6.6) is as follows:*

(1) *If $M = 0$, then $z_0 = w_0/t$ and*

$$(6.19) \quad z_l = \frac{\|w_1\|_2 + \|w_{-1}\|_2}{2t\|w_l\|_2} w_l, \quad l = \pm 1,$$

with $t = \sqrt{\|w_0\|_2^2 + (\|w_1\|_2 + \|w_{-1}\|_2)^2/2}$.

(2) *If $M > 0$ and $w_0 = 0$, then $z_0 = 0$,*

$$(6.20) \quad z_l = \frac{\sqrt{1+lM}}{\sqrt{2}\|w_l\|_2} w_l, \quad l = \pm 1.$$

(3) *If $M > 0$ and $w_{-1} = 0$, then $z_{-1} = 0$,*

$$(6.21) \quad z_0 = \frac{\sqrt{1-M}}{\|w_0\|_2} w_0, \quad z_1 = \frac{\sqrt{M}}{\|w_1\|_2} w_1.$$

(4) *If $M > 0$ and $\|w_0\|, \|w_{-1}\| > 0$, then $z_l = w_l/(1 - \mu - l\lambda)$ ($l = \pm 1, 0$) with*

$$(6.22) \quad \mu = 1 - \frac{\|w_0\|_2}{\alpha}, \quad \lambda = \frac{\|w_1\|_2}{\sqrt{1+M-\alpha^2}} - \frac{\|w_0\|_2}{\alpha},$$

where

$$(6.23) \quad \alpha = \frac{\sqrt{1-M^2}\beta}{\sqrt{2M+(1+M)\beta^2}},$$

with β depending on $\|w_l\|$ as

$$(6.24) \quad \beta = \frac{\xi}{2} + S - \frac{1}{2}\sqrt{-4S^2 - 2p_0 - q_0/S}, \quad p_0 = \frac{-\xi^2 - 2\xi^2 + 2}{2}, \quad q_0 = -\xi(\xi^2 + 1),$$

$$(6.25) \quad \Delta_0 = (\xi^2 - \zeta^2 + 1)^2, \quad \Delta_1 = 2(\xi^2 - \zeta^2 + 1)^3 + 108\xi^2\zeta^2,$$

and

$$(6.26) \quad S = \begin{cases} \frac{1}{2}\sqrt{-\frac{2}{3}p_0 + \frac{1}{3}\left(Q + \frac{\Delta_0}{Q}\right)}, & \text{with } Q = \sqrt[3]{\frac{\Delta_1 + \sqrt{\Delta_1^2 - 4\Delta_0^3}}{2}}, \quad \xi^{2/3} + 1 \leq \eta^{2/3}, \\ \frac{1}{2}\sqrt{-\frac{2}{3}p_0 + \frac{2}{3}\sqrt{\Delta_0}\cos\left(\frac{1}{3}\arccos\left(\frac{\Delta_1}{2\sqrt{\Delta_0^3}}\right)\right)}, & \xi^{2/3} + 1 > \eta^{2/3}, \end{cases}$$

$$(6.27) \quad \xi = \frac{2\sqrt{M}\|w_0\|_2}{\sqrt{1+M}\|w_{-1}\|_2}, \quad \zeta = \frac{\sqrt{1-M}\|w_1\|_2}{\sqrt{1+M}\|w_{-1}\|_2}.$$

Proof. The first three cases are straightforward to verify. Here we only present the proof for case (4). If $M > 0$ and $\|w_0\|_2, \|w_{-1}\|_2 > 0$, then $z_l = w_l/(1 - \mu - l\lambda)$ ($l = 0, \pm 1$). Letting $\|z_0\|_2 = s$, we have from (6.11) that $1 - \mu = \|w_0\|_2/s$ and

$$(6.28) \quad \frac{\sqrt{2}\|w_1\|_2}{\sqrt{1+M-s^2}} + \frac{\sqrt{2}\|w_{-1}\|_2}{\sqrt{1-M-s^2}} = 2\frac{\|w_0\|_2}{s},$$

which implies

$$(6.29) \quad \frac{s}{\sqrt{2\sqrt{1+M-s^2}}} \cdot \frac{\|w_1\|_2}{\|w_0\|_2} + \frac{s}{\sqrt{2\sqrt{1-M-s^2}}} \cdot \frac{\|w_{-1}\|_2}{\|w_0\|_2} = 1.$$

There exists a unique solution $s \in (0, \sqrt{1-M})$ and the Lagrange multipliers can be identified. \square

6.2. Orthogonal retraction. Inspired by the projective retraction, we can consider ψ of the form

$$(6.30) \quad \psi(w)_l = \sigma_l w_l, \quad l = F, \dots, -F,$$

with undetermined positive coefficients $\sigma_F, \dots, \sigma_{-F}$. Besides the constraints

$$(6.31) \quad \sum_{l=-F}^F \|w_l\|_2^2 \sigma_l^2 = 1, \quad \sum_{l=-F}^F l \|w_l\|_2^2 \sigma_l^2 = M,$$

we have to introduce $2F - 1$ additional conditions to uniquely determine the $2F + 1$ coefficients.

In [9], Bao and Lim proposed the condition $\sigma_1 \sigma_{-1} = \sigma_0^2$ for spin-1 BECs. It can be generalized to

$$(6.32) \quad \begin{cases} \sigma_{l-1} \sigma_{l+1} = \sigma_l^2, & l = 1, 2, \dots, F-1, \\ \sigma_l \sigma_{-l} = \sigma_0^2, & l = 1, 2, \dots, F, \end{cases}$$

for spin- F cases. The mapping R characterized by (6.1) and (6.30)–(6.32) is called the *orthogonal retraction* in this paper.

LEMMA 6.3. *For an arbitrary point $w \in \Omega$, $\psi(w)$ defined by (6.30)–(6.32) reads*

$$(6.33) \quad \psi(w)_l = \sqrt{\frac{r^l}{\sum_{k=-F}^F \|w_k\|_2^2 r^k}} \cdot w_l, \quad l = F, \dots, -F,$$

where r is the unique positive zero of the polynomial

$$(6.34) \quad h_2(t) = \sum_{l=0}^{2F} (l - F - M) \|w_{l-F}\|_2^2 \cdot t^l.$$

Proof. Let $r = \sigma_1^2/\sigma_0^2 > 0$; then

$$(6.35) \quad \sigma_l = \sigma_0 r^{l/2}, \quad l = F, \dots, -F.$$

Substituting (6.35) into (6.31) yields

$$(6.36) \quad h_2(r) = 0, \quad \sigma_0^2 = \frac{1}{\sum_{l=-F}^F \|w_l\|_2^2 r^l}.$$

Introducing $h_3(t) = h_2(t)t^{-f_1(w)-F}$, we have

$$(6.37) \quad h_3(t) := \sum_{l=0}^{f_2(w)-f_1(w)} (l + f_1(w) - M) \|w_{l+f_1(w)}\|_2^2 \cdot t^l$$

and

$$(6.38) \quad h_3(0) = (f_1(w) - M) \|w_{f_1(w)}\|_2^2 < 0, \quad \lim_{t \rightarrow +\infty} h_3(t) = +\infty,$$

which implies the function $h_3(\cdot)$ has at least one positive zero.

Let \sum_1 and \sum_2 denote $\sum_{1 \leq l < M-f_1(w)}$ and $\sum_{M-f_1(w) < l \leq f_2(w)-f_1(w)}$, respectively. At a zero r_0 of h_3 , we have

$$\begin{aligned} & \sum_1 l(l + f_1(w) - M) \|w_{l+f_1(w)}\|_2^2 r_0^{l-1} \\ & \geq (M - f_1(w)) \sum_1 (l + f_1(w) - M) \|w_{l+f_1(w)}\|_2^2 r_0^{l-1} \\ & = \frac{M - f_1(w)}{r_0} \left[\sum_1 (l + f_1(w) - M) \|w_{l+f_1(w)}\|_2^2 r_0^l - h_3(r_0) \right] \\ & = (M - f_1(w)) \left[\frac{(M - f_1(w)) \|w_{f_1(w)}\|_2^2}{r_0} - \sum_2 (l + f_1(w) - M) \|w_{l+f_1(w)}\|_2^2 r_0^{l-1} \right], \end{aligned}$$

which leads to

$$\begin{aligned} h'_3(r_0) &= \sum_1 l(l + f_1(w) - M) \|w_{l+f_1(w)}\|_2^2 r_0^{l-1} + \sum_2 l(l + f_1(w) - M) \|w_{l+f_1(w)}\|_2^2 r_0^{l-1} \\ (6.39) \quad &\geq \frac{(M - f_1(w))^2 \|w_{f_1(w)}\|_2^2}{r_0} + \sum_2 (l + f_1(w) - M)^2 \|w_{l+f_1(w)}\|_2^2 r_0^{l-1} > 0. \end{aligned}$$

From (6.39) we can see that h_3 has exactly one positive zero, and h_2 has exactly one positive zero too. Substituting (6.36) into (6.35) leads to the formulas of the coefficients. \square

Notice that in spin-1 cases, h_2 degenerates to a quadratic polynomial, and the orthogonal retraction has a closed-form solution.

The well-definedness of the orthogonal retraction is guaranteed by the following theorem [2].

THEOREM 6.4. *Let \mathcal{M} be an embedded manifold of a vector space \mathcal{E} , and let \mathcal{N} be an abstract manifold such that $\dim(\mathcal{M}) + \dim(\mathcal{N}) = \dim(\mathcal{E})$. Assume that there is a diffeomorphism*

$$\varphi : \mathcal{M} \times \mathcal{N} \rightarrow \mathcal{E}_* : (u, v) \mapsto \varphi(u, v),$$

where \mathcal{E}_* is an open subset of \mathcal{E} , with a neural element e satisfying

$$\varphi(u, e) = u \quad \forall u \in \mathcal{M}.$$

Then the mapping

$$R_u(\xi_u) := \pi_1(\varphi^{-1}(u + \xi_u)),$$

where $\pi_1 : \mathcal{M} \times \mathcal{N} \rightarrow \mathcal{M}$ is the projection onto the first component, defines a retraction on \mathcal{M} .

LEMMA 6.5. *The orthogonal retraction is a well-defined retraction on \mathcal{M} .*

Proof. Take $\mathcal{N} = \mathcal{R}_+^2$, and \mathcal{N} is a manifold satisfying $\dim(\mathcal{M}) + \dim(\mathcal{N}) = \dim(\mathcal{R}^N)$. Define the mapping $\varphi : \mathcal{M} \times \mathcal{N} \rightarrow \Omega$ as

$$(6.40) \quad \varphi(u, v) := \left(\frac{1}{v_1 v_2^{l/2}} \cdot u_l \right)_{l=-F}^F \quad \forall (u, v) \in \mathcal{M} \times \mathcal{N}.$$

Lemma 6.3 shows that for any $w \in \Omega$ there exists a unique $u = \psi(w), v = (\sigma_0, r)^T$ such that $\varphi(u, v) = w$, and thus φ is a bijection. It is obvious to see that φ is smooth on $\mathcal{M} \times \mathcal{N}$, and $\varphi(u, \mathbf{1}) = u \ \forall u \in \mathcal{M}$.

From Lemma 6.3, we have

$$(6.41) \quad \varphi^{-1}(w) = \left(\left(\sigma_0 r^{l/2} w_l \right)_{l=-F}^F, (\sigma_0, r)^T \right) \quad \forall w \in \Omega,$$

where $\sigma_0 = \sqrt{\frac{1}{\sum_{l=-F}^F \|w_l\|_2^2 r^l}}$ and r is characterized by the equation

$$(6.42) \quad h_2(r) = \sum_{l=0}^{2F} (l - F - M) \|w_{l-F}\|_2^2 \cdot r^l = 0.$$

Since $h'_2(r) = (h_3(t) \cdot t^{f_1(w)+F})'|_{t=r} = h'_3(r) \cdot r^{f_1(w)+F} > 0$, it follows from the implicit function theorem that r , when considered as a function of w , is smooth. Then φ^{-1} is also a smooth function at every $w \in \Omega$, which makes φ a diffeomorphism. Thus the orthogonal retraction, given by

$$(6.43) \quad R_u(\xi_u) := \pi_1(\varphi^{-1}(u + \xi_u)),$$

is a retraction on \mathcal{M} . □

6.3. Closed-form retraction. In the projective retraction, the coefficients take the form

$$(6.44) \quad \sigma_l = \frac{1}{1 - \mu - l\lambda}.$$

When $w \rightarrow \mathcal{M}$, we have $\sigma_l \rightarrow 1$ and $\mu, \lambda \rightarrow 0$, and

$$(6.45) \quad \frac{1}{1 - \mu - l\lambda} = \sqrt{1 + 2\mu + 2l\lambda} + o(\mu + l\lambda).$$

Thus we can approximate the projective retraction by taking

$$(6.46) \quad \sigma_l \approx \sqrt{1 + 2\mu + 2l\lambda}, \quad l = F, \dots, -F.$$

As shown below, (6.46) has a closed-form formula, and the mapping R characterized by (6.1), (6.30), (6.31), and (6.46) is called the *closed-form retraction* in this paper. First, we introduce

$$(6.47) \quad S := \{w \in \mathcal{R}^N \mid w^T \Gamma^2 w - (M + l) w^T \Gamma w + Ml \cdot w^T w > 0, l = \pm F\}.$$

Apparently, S is an open set and $\mathcal{M} \subset S \subset \Omega$. We next discuss the computation of $\psi(w)$ for $w \in S$.

LEMMA 6.6. For an arbitrary point $w \in S$, $\psi(w)$ defined by (6.30), (6.31), and (6.46) reads

$$(6.48) \quad \psi(w)_l = \sqrt{\frac{w^T \Gamma^2 w - (M + l)w^T \Gamma w + Ml \cdot w^T w}{w^T w \cdot w^T \Gamma^2 w - (w^T \Gamma w)^2}} \cdot w_l, \quad l = F, \dots, -F.$$

Proof. Substituting (6.46) into (6.31) yields

$$(6.49) \quad \begin{cases} (1 + 2\mu) w^T w + 2\lambda w^T \Gamma w = 1, \\ (1 + 2\mu) w^T \Gamma w + 2\lambda w^T \Gamma^2 w = M, \end{cases}$$

and the solution is given by

$$(6.50) \quad \mu = \frac{1}{2} \cdot \frac{w^T \Gamma^2 w - Mw^T \Gamma w}{w^T w \cdot w^T \Gamma^2 w - (w^T \Gamma w)^2} - \frac{1}{2}, \quad \lambda = \frac{1}{2} \cdot \frac{-w^T \Gamma w + Mw^T w}{w^T w \cdot w^T \Gamma^2 w - (w^T \Gamma w)^2}.$$

The condition $w \in S$ ensures $1 + 2\mu + 2l\lambda > 0$ for $l = F, \dots, -F$. In view of the retraction (6.46), we obtain formula (6.48). \square

LEMMA 6.7. The closed-form retraction is a well-defined retraction.

Proof. Denote

$$(6.51) \quad \mathcal{N} = \{v = (v_1, v_2)^T \in \mathbb{R}^2 \mid 1 + 2v_1 + 2Fv_2 > 0, 1 + 2v_1 - 2Fv_2 > 0\}.$$

\mathcal{N} is an open subset of \mathbb{R}^2 and therefore a 2D manifold. Define the mapping $\varphi : \mathcal{M} \times \mathcal{N} \rightarrow \mathcal{R}^N$ as

$$(6.52) \quad \varphi(u, v) := \left(\frac{1}{\sqrt{1 + 2v_1 + 2lv_2}} \cdot u_l \right)_{l=-F}^F \quad \forall (u, v) \in \mathcal{M} \times \mathcal{N}.$$

For an arbitrary point $(u, v) \in \mathcal{M} \times \mathcal{N}$, let $w = \varphi(u, v)$.

- If $v_2 = 0$, then $w = \frac{1}{\sqrt{1+2v_1}} u$ ($u \in \mathcal{M}$), and

$$\begin{aligned} & w^T \Gamma^2 w - (M + l)w^T \Gamma w + Ml \cdot w^T w \\ &= \frac{1}{1 + 2v_1} [u^T \Gamma^2 u - (M + l)u^T \Gamma u + Ml \cdot u^T u] \\ &= \frac{1}{1 + 2v_1} (u^T \Gamma^2 u - M^2) > 0, \quad l = \pm F. \end{aligned}$$

- If $v_2 \neq 0$, then

$$\begin{aligned} & w^T \Gamma^2 w - (M + l)w^T \Gamma w + Ml \cdot w^T w \\ &= \sum_{k=-F}^F \frac{k^2 - (M + l)k + Ml}{1 + 2v_1 + 2kv_2} \|u_k\|_2^2 \\ &= \sum_{k=-F}^F \left[\frac{k}{2v_2} - \frac{M}{2v_2} + \frac{(1 + 2v_1 + 2lv_2)(M - k)}{2v_2(1 + 2v_1 + 2kv_2)} \right] \|u_k\|_2^2 \\ &= \frac{1}{2v_2} u^T \Gamma u - \frac{M}{2v_2} u^T u + \sum_{k=-F}^F \frac{(1 + 2v_1 + 2lv_2)(M - k)}{2v_2(1 + 2v_1 + 2kv_2)} \|u_k\|_2^2 \\ &= \frac{1 + 2v_1 + 2lv_2}{2} \sum_{k=-F}^F \frac{M - k}{v_2(1 + 2v_1 + 2kv_2)} \|u_k\|_2^2, \quad l = \pm F. \end{aligned}$$

Noticing that

$$(6.53) \quad \frac{M-k}{v_2(1+2v_1+2kv_2)} - \frac{M-k}{v_2(1+2v_1+2Mv_2)} = \frac{2(M-k)^2}{(1+2v_1+2kv_2)(1+2v_1+2Mv_2)} > 0,$$

we have

$$\begin{aligned} & w^T \Gamma^2 w - (M+l)w^T \Gamma w + Ml \cdot w^T w \\ & > \frac{1+2v_1+2lv_2}{2} \sum_{k=-F}^F \frac{M-k}{v_2(1+2v_1+2Mv_2)} \|u_k\|_2^2 \\ & = \frac{1+2v_1+2lv_2}{2v_2(1+2v_1+2Mv_2)} (Mu^T u - u^T \Gamma u) = 0, \quad l = \pm F. \end{aligned}$$

On the one hand, the above analysis shows $\varphi(\mathcal{M} \times \mathcal{N}) \subset S$; on the other hand, Lemma 6.6 indicates that for any $w \in S$, there exists a unique $u = \psi(w)$, $v = (\mu, \lambda)^T$ such that $\varphi(u, v) = w$. Hence φ is a bijection from $\mathcal{M} \times \mathcal{N}$ to S . It is straightforward to see that φ and φ^{-1} are both smooth functions, and $\varphi(u, \mathbf{0}) = u \ \forall u \in \mathcal{M}$. Thus from Theorem 6.4 we know that the closed-form retraction, given by

$$(6.54) \quad R_u(\xi_u) := \pi_1(\varphi^{-1}(u + \xi_u)),$$

is a retraction on \mathcal{M} . \square

7. Numerical experiments. In this section, we first compare the performance of Algorithm 1 with RGBB and RTR [1] by testing some BEC examples. RGBB and RTR are also expedited with the mesh refinement technique. We present numerical results of these algorithms under the three different retractions defined in section 6. A comparison between ARNT and NGF [9] is also presented. Then we apply Algorithm 1 to compute the ground states of spin-2 and spin-3 BECs with different parameters. All codes were written in MATLAB. All experiments were performed on a workstation with Intel Xeon E5-2680 v3 processors at 2.50GHz($\times 12$) and 128GB memory running CentOS 6.8 and MATLAB R2019b.

In the spin-1 BEC, the initial data is chosen as $\Phi_0(\mathbf{x}) = U\phi_0(\mathbf{x})$, where

$$(7.1) \quad \phi_0(\mathbf{x}) = \frac{1}{\pi^{d/4}} e^{-(x_1^2 + \dots + x_d^2)/2}, \quad \mathbf{x} \in \mathcal{R}^d,$$

$U = (\frac{\sqrt{1+3M}}{2}, \sqrt{\frac{1-M}{2}}, \frac{\sqrt{1-M}}{2})^T$ for the ferromagnetic interaction ($\beta_1 \leq 0$); $U = (\sqrt{\frac{1+M}{2}}, 0, \sqrt{\frac{1-M}{2}})^T$ for the antiferromagnetic interaction ($\beta_1 > 0$) [9].

In the spin-2 BEC, the initial data is chosen as $\Phi_0(\mathbf{x}) = U\phi_0(\mathbf{x})$, where

$$(7.2) \quad U = \left(\frac{m_1^4}{16}, \frac{m_1^3 m_2}{8}, \frac{\sqrt{6} m_1^2 m_2^2}{16}, \frac{m_1 m_2^3}{8}, \frac{m_2^4}{16} \right)^T$$

with $m_1 = \sqrt{2+M}$ and $m_2 = \sqrt{2-M}$ for the ferromagnetic interaction ($\beta_1 < 0$ and $\beta_2 > 20\beta_1$), and $U = (\frac{\sqrt{2+M}}{2}, 0, 0, 0, \frac{\sqrt{2-M}}{2})^T$ for the nematic interaction ($\beta_2 < 0$ and $\beta_2 < 20\beta_1$), $U = (\sqrt{\frac{M+1}{3}}, 0, 0, \sqrt{\frac{2-M}{3}}, 0)^T$ for the cyclic interaction ($\beta_1 > 0$ and $\beta_2 > 0$) [10]. In the spin-3 BEC, the initial data is chosen as $\Phi_0(\mathbf{x}) = U\phi_0(\mathbf{x})$, where $U \in \mathcal{R}^7$ is taken as the random vector. In all the examples, we take $p = q = 0$.

7.1. Performance of algorithms. In RGBB we used all of the default parameters. As for RTR, we added a rule $\|r_{j+1}\|_2 \leq \min\{0.1, 0.1\|r_0\|_2\}$ into the stopping criterion of the truncated CG method. All other default settings of RTR were used. For ARNT, we set $\eta_1 = 0.01, \eta_2 = 0.9, \gamma_0 = 0.2, \gamma_1 = 1, \gamma_2 = 10$, and $\sigma_k = \hat{\sigma}_k \|\text{grad}\tilde{E}(u_k)\|_2$, where $\hat{\sigma}_k$ is updated by the procedure in Algorithm 1 with $\hat{\sigma}_0 = 1$. Furthermore, when an estimation of the absolute value of the negative curvature, denoted by σ_{est} , is available at the k th subproblem, we can calculate

$$(7.3) \quad \sigma_{k+1}^{new} = \max\{\sigma_{k+1}, \sigma_{est} + \tilde{\gamma}\},$$

with some small $\tilde{\gamma} \geq 0$. Then the parameter σ_{k+1} is reset to σ_{k+1}^{new} .

On the finest mesh, ARNT, RGBB, and RTR terminate when either $\|\text{grad}\tilde{E}(u_k)\|_2 \leq 10^{-6}$ or the number of iterations reaches 10000, while on the coarse meshes they all terminate when $\|\text{grad}\tilde{E}(u_k)\|_2 \leq 10^{-5}$. In the implementation of ARNT, RGBB stops when either $\|\text{grad}\tilde{E}(u_k)\|_2 \leq 10^{-2}$ or the number of iterations reaches 2000. The maximum number of inner iterations in ARNT is chosen adaptively depending on $\|\text{grad}\tilde{E}(u_k)\|_2$. These parameters apply for all of the numerical experiments in this paper unless clarified otherwise.

In the subsequent tables, the columns “f,” “nrmG,” and “time” display the final objective function value, the final norm of the Riemannian gradient, and the total CPU time each algorithm spent to reach the stopping criterion. The column “iter” reports the number of iterations (the average numbers of inner iterations) on the finest mesh. The choice of retractions is shown in the column “retr,” where R1, R2, and R3 denote the projective retraction, the orthogonal retraction, and the closed-form retraction, respectively.

We present results of the following cases for spin-1, spin-2, and spin-3 BECs:

- Spin-1 BECs [9]:

2D antiferromagnetic case. $V(x, y) = \frac{1}{2}(x^2 + y^2) + 10[\sin^2(\frac{\pi x}{4}) + \sin^2(\frac{\pi y}{4})]$, $\beta_0 = 300, \beta_1 = 100, U = [-16, 16] \times [-16, 16], n = 2^9$.

3D ferromagnetic case. $V(x, y, z) = \frac{1}{2} \sum_{\alpha=x,y,z} (\alpha^2 + 200 \sin^2(\frac{\pi \alpha}{2}))$, $\beta_0 = 880, \beta_1 = -4.1, U = [-16, 16] \times [-16, 16] \times [-16, 16], n = 2^8$.

- Spin-2 BECs [19]:

2D antiferromagnetic case. $V(x, y) = \frac{1}{2}(x^2 + y^2) + 10 [\sin^2(\frac{\pi x}{2}) + \sin^2(\frac{\pi y}{2})]$, $\beta_0 = 243, \beta_1 = 12.1, \beta_2 = -13, U = [-8, 8] \times [-8, 8], n = 2^8$.

3D cyclic case. $V(x, y, z) = \frac{1}{2} \sum_{\alpha=x,y,z} (\alpha^2 + 200 \sin^2(\frac{\pi \alpha}{2}))$, $\beta_0 = 183.9, \beta_1 = 26.8, \beta_2 = 134.7, U = [-16, 16] \times [-16, 16] \times [-16, 16], n = 2^8$.

- Spin-3 BECs:

2D. $V(x, y) = \frac{1}{2}(x^2 + y^2) + 10 [\sin^2(\frac{\pi x}{2}) + \sin^2(\frac{\pi y}{2})]$, $\beta_0 = 100, \beta_1 = 1, \beta_2 = 10, \beta_3 = 1, U = [-8, 8] \times [-8, 8], n = 2^8$.

3D. $V(x, y, z) = \frac{1}{2} \sum_{\alpha=x,y,z} (\alpha^2 + 200 \sin^2(\frac{\pi \alpha}{2}))$, $\beta_0 = 100, \beta_1 = 1, \beta_2 = 10, \beta_3 = 1, U = [-8, 8] \times [-8, 8] \times [-8, 8], n = 2^7$.

In the 3D cases of spin-2 BECs, the stopping criterion on the finest mesh is set to be $\|\text{grad}\tilde{E}(u_k)\|_2 \leq 10^{-3}$ when $M = 0$. The detailed numerical results are reported in Tables 1–6. In most cases, all three algorithms converge to points with the same function values. For the spin-1 and spin-2 BECs, the choice of different retractions has small impact on the numerical performance, except for the spin-2 cases in 3D with $M = 0$, and the second-order algorithms ARNT and RTR exhibit higher stability than the first-order algorithm RGBB in response to the change of retractions. In the 3D case of spin-3 BEC (Table 6), RTR converges to a slightly larger function value than ARNT when $M = 1.5$ using retractions R1 and R2. Overall, taking both numerical

TABLE 1
Numerical results of spin-1 BECs in 2D.

retr	ARNT				RGBB				RTR			
	f	nrmG	iter	time	f	nrmG	iter	time	f	nrmG	iter	time
$M = 0.0$												
R1	15.1032	7.8e-07	4 (38)	16.9	15.1032	8.3e-07	239	20.5	15.1032	6.3e-07	20 (17)	30.4
R2	15.1032	7.8e-07	4 (38)	14.9	15.1032	6.3e-07	257	19.1	15.1032	6.1e-07	20 (17)	30.2
R3	15.1032	5.8e-07	4 (34)	14.1	15.1032	3.3e-07	255	17.7	15.1032	6.2e-07	20 (17)	30.9
$M = 0.2$												
R1	15.1411	7.6e-07	4 (45)	18.6	15.1411	9.3e-07	258	22.0	15.1411	2.6e-07	21 (26)	54.9
R2	15.1411	8.2e-07	4 (45)	20.1	15.1411	1.0e-06	254	20.9	15.1411	2.7e-07	21 (26)	54.6
R3	15.1411	6.3e-07	4 (45)	18.4	15.1411	7.9e-07	261	15.3	15.1411	2.7e-07	21 (26)	52.3
$M = 0.5$												
R1	15.3436	8.3e-07	4 (66)	24.2	15.3436	8.1e-07	431	37.5	15.3436	2.0e-07	21 (31)	62.0
R2	15.3436	3.5e-07	4 (65)	25.3	15.3436	9.1e-07	421	33.4	15.3436	2.1e-07	21 (31)	59.7
R3	15.3436	8.6e-07	4 (63)	25.1	15.3436	9.0e-07	429	30.8	15.3436	2.2e-07	21 (31)	64.6
$M = 0.9$												
R1	15.9621	5.3e-07	4 (56)	24.5	15.9621	8.8e-07	323	30.5	15.9621	3.9e-07	21 (26)	52.9
R2	15.9621	7.4e-07	4 (57)	24.6	15.9621	1.0e-06	279	22.5	15.9621	3.9e-07	21 (26)	55.6
R3	15.9621	4.0e-07	4 (61)	23.3	15.9621	7.6e-07	483	30.3	15.9621	3.3e-07	21 (26)	51.4

TABLE 2
Numerical results of spin-1 BECs in 3D.

retr	ARNT				RGBB				RTR			
	f	nrmG	iter	time	f	nrmG	iter	time	f	nrmG	iter	time
$M = 0.0$												
R1	55.4362	4.9e-07	4 (41)	1171.4	55.4362	8.0e-07	310	1844.3	55.4362	7.1e-07	17 (15)	1287.5
R2	55.4362	2.4e-07	4 (47)	1254.2	55.4362	9.5e-07	384	2914.5	55.4362	6.2e-07	17 (17)	1416.0
R3	55.4362	6.6e-07	4 (35)	1035.1	55.4362	7.0e-07	227	1013.8	55.4362	6.9e-07	17 (15)	1298.3
$M = 0.2$												
R1	55.4362	4.6e-07	4 (39)	1794.1	55.4363	9.6e-05	10000	43377.9	55.4362	5.5e-07	17 (16)	1330.1
R2	55.4362	3.5e-07	4 (38)	2471.0	55.4363	8.6e-05	10000	43039.6	55.4362	7.2e-07	17 (15)	1278.6
R3	55.4362	6.1e-07	4 (40)	1429.2	55.4362	1.2e-04	10000	42328.7	55.4362	6.1e-07	17 (17)	1412.1
$M = 0.5$												
R1	55.4362	7.6e-07	4 (36)	1188.3	55.4362	7.5e-07	341	1623.8	55.4362	7.0e-07	17 (15)	1291.2
R2	55.4362	5.2e-07	4 (34)	1228.7	55.4362	9.3e-07	666	6114.5	55.4362	7.2e-07	17 (16)	1376.0
R3	55.4362	5.7e-07	4 (30)	1083.9	55.4362	8.5e-07	225	1135.0	55.4362	6.9e-07	17 (15)	1283.3
$M = 0.9$												
R1	55.4362	2.5e-07	4 (41)	1316.8	55.4362	9.1e-07	631	5930.9	55.4362	7.4e-07	17 (15)	1295.0
R2	55.4362	7.6e-07	4 (36)	1194.7	55.4362	9.4e-07	200	1093.4	55.4362	7.3e-07	17 (15)	1274.7
R3	55.4362	3.3e-07	4 (41)	1313.0	55.4362	5.9e-07	335	1938.1	55.4362	7.1e-07	17 (15)	1287.0

stability and time cost into consideration, ARNT shows the best performance.

We next compare the performance of ARNT and NGF by testing on spin-1 BECs. For ARNT we apply the same parameters as above. For NGF, we set the step size Δt to be 10^{-3} , and we set the stopping criterion to be $\|\phi^k - \phi^{k-1}\|_\infty / \Delta t \leq 10^{-2}$ or $k \geq 10000$ on the coarse meshes and $\|\phi^k - \phi^{k-1}\|_\infty / \Delta t \leq 10^{-3}$ or $k \geq 10000$ on the finest mesh. In particular, the 1D cases with $V(x) = \frac{1}{2}x^2 + 25 \sin^2(\frac{\pi x}{4})$, $U = [-16, 16]$, $n = 2^{11}$, and following parameters are considered:

- Case I (ferromagnetic). $\beta_0 = 885$, $\beta_1 = -4.1$.
- Case II (antiferromagnetic). $\beta_0 = 241$, $\beta_1 = 7.5$.

TABLE 3
Numerical results of spin-2 BECs in 2D.

retr	ARNT				RGBB				RTR			
	f	nrmG	iter	time	f	nrmG	iter	time	f	nrmG	iter	time
<i>M = 0.0</i>												
R1	14.3386	5.9e-07	4 (32)	8.0	14.3386	9.0e-07	280	12.7	14.3386	9.3e-07	17 (18)	17.8
R2	14.3386	5.9e-07	4 (32)	7.6	14.3386	8.0e-07	238	9.8	14.3386	9.3e-07	17 (18)	16.4
R3	14.3386	5.9e-07	4 (32)	7.6	14.3386	9.0e-07	225	8.8	14.3386	9.3e-07	17 (18)	17.7
<i>M = 0.5</i>												
R1	14.3730	3.8e-07	4 (59)	15.2	14.3730	1.3e-07	346	16.0	14.3730	4.7e-07	18 (22)	23.5
R2	14.3730	3.8e-07	4 (59)	12.4	14.3730	9.9e-07	323	13.4	14.3730	4.7e-07	18 (22)	22.1
R3	14.3730	3.8e-07	4 (59)	12.4	14.3730	9.6e-08	523	35.7	14.3730	4.7e-07	18 (22)	22.0
<i>M = 1.5</i>												
R1	14.6754	4.3e-07	4 (62)	13.1	14.6754	9.5e-07	462	20.0	14.6754	2.8e-07	18 (28)	28.7
R2	14.6754	4.3e-07	4 (62)	12.4	14.6754	9.9e-07	519	30.7	14.6754	2.8e-07	18 (28)	26.0
R3	14.6754	4.3e-07	4 (62)	12.5	14.6754	8.5e-07	402	21.0	14.6754	2.8e-07	18 (28)	28.8

TABLE 4
Numerical results of spin-2 BECs in 3D.

retr	ARNT				RGBB				RTR			
	f	nrmG	iter	time	f	nrmG	iter	time	f	nrmG	iter	time
<i>M = 0.0</i>												
R1	45.6071	8.0e-04	3 (183)	38306.7	45.6071	9.8e-04	382	13294.7	45.6071	7.8e-04	27 (66)	15822.4
R2	45.6071	6.5e-04	5 (184)	29152.1	45.6071	9.9e-04	629	4625.7	45.6071	9.2e-04	24 (45)	8898.4
R3	45.6071	5.7e-04	4 (161)	31430.8	46.2917	9.3e-04	93	854.7	45.6071	9.0e-04	24 (45)	9618.9
<i>M = 0.5</i>												
R1	45.7403	7.5e-07	5 (75)	4090.1	45.7403	9.3e-07	1129	9028.3	45.7403	5.4e-07	19 (28)	4819.9
R2	45.7403	6.6e-07	5 (76)	4144.4	45.7403	9.2e-07	1050	7588.9	45.7403	4.7e-07	19 (28)	4750.9
R3	45.7403	5.4e-07	5 (80)	4310.3	45.7403	9.6e-07	1296	10671.6	45.7403	5.3e-07	19 (28)	4706.7
<i>M = 1.5</i>												
R1	46.8619	4.4e-07	4 (54)	2838.4	46.8619	9.8e-07	391	3048.2	46.8619	3.7e-07	19 (22)	3792.0
R2	46.8619	6.9e-07	4 (56)	2753.4	46.8619	8.9e-07	380	2970.2	46.8619	3.5e-07	19 (22)	3760.9
R3	46.8619	6.9e-07	4 (54)	2815.9	46.8619	9.3e-07	515	5197.9	46.8619	3.3e-07	19 (22)	3732.1

The results are reported in Table 7. In general, ARNT shows higher efficiency than NGF.

7.2. Application in spin-2 BEC. In this section, we apply the ARNT method with the projective retraction to compute the ground state of a spin-2 BEC in 1–3 dimensions and under different interactions. Specifically, the following cases are considered [19]:

- 1D, $V(x) = \frac{1}{2}x^2 + 25 \sin^2(\frac{\pi x}{4})$.
Case I (ferromagnetic). $\beta_0 = 130.6$, $\beta_1 = -25.4$, $\beta_2 = -125.3$, $U = [-8, 8]$, $n = 2^8$.
Case II (antiferromagnetic). $\beta_0 = 243$, $\beta_1 = 12.1$, $\beta_2 = -13$, $U = [-16, 16]$, $n = 2^9$.
Case III (cyclic). $\beta_0 = 183.9$, $\beta_1 = 26.8$, $\beta_2 = 134.7$, $U = [-16, 16]$, $n = 2^9$.
- 2D, $V(x, y) = \frac{1}{2}(x^2 + y^2) + 10 [\sin^2(\frac{\pi x}{2}) + \sin^2(\frac{\pi y}{2})]$, $U = [-8, 8] \times [-8, 8]$, $n = 2^8$.
Case I (ferromagnetic). $\beta_0 = 130.6$, $\beta_1 = -25.4$, $\beta_2 = -125.3$.

TABLE 5
Numerical results of spin-3 BECs in 2D.

retr	ARNT				RGBB				RTR			
	f	nrmG	iter	time	f	nrmG	iter	time	f	nrmG	iter	time
$M = 0.0$												
R1	11.8279	7.1e-07	4 (48)	21.7	11.8279	9.6e-07	548	47.7	11.8279	5.1e-07	28 (165)	349.7
R2	11.8279	6.6e-07	4 (58)	24.0	11.8279	6.6e-07	587	46.5	11.8279	5.1e-07	28 (165)	340.5
R3	11.8279	6.9e-07	4 (58)	21.0	11.8279	9.6e-07	484	30.5	11.8279	5.8e-07	32 (184)	442.3
$M = 0.5$												
R1	11.8334	5.9e-07	4 (73)	27.2	11.8334	9.8e-07	694	63.4	11.8334	1.8e-07	18 (79)	106.5
R2	11.8334	7.5e-07	4 (81)	27.7	11.8334	8.4e-07	482	41.1	11.8334	1.9e-07	18 (79)	110.5
R3	11.8334	4.9e-07	4 (75)	25.5	11.8334	7.5e-07	431	27.2	11.8334	1.9e-07	18 (42)	58.8
$M = 1.5$												
R1	11.8780	9.2e-07	5 (168)	50.3	11.8780	9.1e-07	3596	236.8	11.8780	3.0e-07	18 (49)	69.8
R2	11.8780	7.7e-07	6 (160)	58.5	11.8780	9.8e-07	3017	216.2	11.8780	2.6e-07	18 (63)	88.1
R3	11.8780	3.0e-07	5 (135)	41.0	11.8780	9.9e-07	737	52.4	11.8780	2.9e-07	18 (49)	67.0

TABLE 6
Numerical results of spin-3 BECs in 3D.

retr	ARNT				RGBB				RTR			
	f	nrmG	iter	time	f	nrmG	iter	time	f	nrmG	iter	time
$M = 0.0$												
R1	42.9752	5.4e-07	4 (97)	3690.4	42.9774	1.9e-04	10000	15016.8	42.9752	9.9e-07	17 (12)	415.4
R2	42.9752	5.0e-07	4 (42)	489.1	42.9772	3.1e-04	10000	15031.7	42.9752	1.0e-07	18 (31)	1038.8
R3	42.9752	9.3e-07	5 (157)	4466.9	42.9752	9.9e-07	729	1440.4	42.9752	8.0e-07	50 (64)	5599.4
$M = 0.5$												
R1	42.9822	8.7e-07	4 (115)	1042.6	42.9822	8.7e-07	810	1301.4	42.9822	6.9e-07	99 (146)	25560.9
R2	42.9822	8.8e-07	4 (115)	1033.5	42.9822	1.0e-06	1194	1791.7	42.9822	5.6e-07	97 (155)	26273.6
R3	42.9822	8.8e-07	4 (115)	1014.6	42.9822	8.1e-07	1181	2280.1	42.9822	6.8e-07	21 (40)	1493.2
$M = 1.5$												
R1	43.0399	2.5e-07	5 (153)	1820.1	43.0399	4.7e-04	10000	14885.7	43.0417	1.7e-07	22 (42)	1676.5
R2	43.0399	2.4e-07	5 (151)	1837.8	43.0400	3.0e-03	10000	15000.0	43.0417	1.6e-07	22 (37)	1478.6
R3	43.0399	2.8e-07	5 (150)	1877.2	43.0400	6.1e-04	10000	14733.2	43.0399	4.7e-07	314 (196)	106186.1

Case II (antiferromagnetic). $\beta_0 = 243$, $\beta_1 = 12.1$, $\beta_2 = -13$.

Case III (cyclic). $\beta_0 = 183.9$, $\beta_1 = 26.8$, $\beta_2 = 134.7$.

- 3D, $V(x, y, z) = \frac{1}{2}(x^2 + y^2 + z^2) + 100 \left[\sin^2\left(\frac{\pi x}{2}\right) + \sin^2\left(\frac{\pi y}{2}\right) + \sin^2\left(\frac{\pi z}{2}\right) \right]$

Case I (ferromagnetic). $\beta_0 = 130.6$, $\beta_1 = -25.4$, $\beta_2 = -125.3$, $U = [-8, 8] \times [-8, 8] \times [-8, 8], n = 2^7$.

Case II (antiferromagnetic). $\beta_0 = 243$, $\beta_1 = 12.1$, $\beta_2 = -13$, $U = [-16, 16] \times [-16, 16] \times [-16, 16], n = 2^8$.

Case III (cyclic). $\beta_0 = 183.9$, $\beta_1 = 26.8$, $\beta_2 = 134.7$, $U = [-16, 16] \times [-16, 16] \times [-16, 16], n = 2^8$.

In Case III of 3D BECs, when $M = 0$, the stopping criterion is set to $\|\text{grad} \tilde{E}(u_k)\|_2 \leq 10^{-3}$ on the finest mesh.

The ground state energies in the above cases are listed in Table 8. Under ferromagnetic interaction (Case I) the energy remains constant when M changes; under antiferromagnetic interaction and cyclic interaction (Cases II and III), the energy increases as M increases.

The wave functions of the ground states computed by ARNT are given in Figures 1–3. The peak function value under ferromagnetic interaction is lower than that under

TABLE 7
Comparison between ARNT and NGF.

M	ARNT			NGF		
	f	iter	time	f	iter	time
Case I						
0.0	47.69416804	4 (155)	2.1	47.69418167	5157	4.9
0.2	47.69416804	5 (190)	2.2	47.69419166	7457	6.3
0.5	47.69416804	5 (142)	1.8	47.69419353	9424	11.3
0.9	47.69416804	4 (153)	1.7	47.69419841	8935	12.1
Case II						
0.0	25.64789561	4 (69)	0.9	25.64791533	3077	3.6
0.2	25.65969259	6 (176)	2.3	25.65975843	10000	12.3
0.5	25.72898994	4 (151)	1.6	25.72901433	7242	8.9
0.9	25.93389752	4 (152)	1.6	25.93392689	5359	6.5

TABLE 8
Ground state energies of spin-2 BECs.

M	1D			2D			3D		
	Case I	Case II	Case III	Case I	Case II	Case III	Case I	Case II	Case III
0.0	10.3700	25.6185	22.6300	9.5754	14.3386	13.4946	39.0045	46.9770	45.6071
0.5	10.3700	25.7372	22.9404	9.5754	14.3730	13.5746	39.0045	47.0301	45.7403
1.5	10.3700	26.8415	25.4640	9.5754	14.6754	14.2734	39.0045	47.5117	46.8619

the other two types of interactions. By comparing Figures 1, 2, and 3, we can find a common property: when $M > 0$, in the ground states under nematic interaction, the components $\phi_1, \phi_0, \phi_{-1}$ are always zero-valued functions; and in the ground states under cyclic interaction, the components $\phi_1, \phi_0, \phi_{-2}$ are always zero-valued functions.

7.3. Application in spin-3 BECs. In this section, we apply the ARNT method with the projective retraction to compute the ground state of a spin-3 BEC in 1–3 dimensions under different interactions. In detail, the following cases are considered:

- 1D, $V(x) = \frac{1}{2}x^2 + 25 \sin^2(\frac{\pi x}{4})$, $U = [-8, 8]$, $n = 2^8$.
 - Case I. $\beta_0 = 100$, $\beta_1 = 1$, $\beta_2 = -10$, $\beta_3 = -1$.
 - Case II. $\beta_0 = 100$, $\beta_1 = 1$, $\beta_2 = 10$, $\beta_3 = 1$.
- 2D, $V(x, y) = \frac{1}{2}(x^2 + y^2) + 10 [\sin^2(\frac{\pi x}{2}) + \sin^2(\frac{\pi y}{2})]$, $U = [-8, 8] \times [-8, 8]$, $n = 2^8$.
 - Case I. $\beta_0 = 100$, $\beta_1 = 1$, $\beta_2 = -10$, $\beta_3 = -1$.
 - Case II. $\beta_0 = 100$, $\beta_1 = 1$, $\beta_2 = 10$, $\beta_3 = 1$.
- 3D, $V(x, y, z) = \frac{1}{2}(x^2 + y^2 + z^2) + 100 [\sin^2(\frac{\pi x}{2}) + \sin^2(\frac{\pi y}{2}) + \sin^2(\frac{\pi z}{2})]$, $U = [-8, 8] \times [-8, 8] \times [-8, 8]$, $n = 2^7$.
 - Case I. $\beta_0 = 100$, $\beta_1 = 1$, $\beta_2 = -10$, $\beta_3 = -1$.
 - Case II. $\beta_0 = 100$, $\beta_1 = 1$, $\beta_2 = 10$, $\beta_3 = 1$.

The ground state energies in the above cases are listed in Table 9. In each case, the energy increases as M increases. The wave functions of the ground states computed by ARNT are given in Figures 4–6. By comparing the figures, we can see that in Case I, when $M > 0$, the components $\phi_2, \phi_1, \phi_0, \phi_{-1}, \phi_{-2}$ are always close to zero; in Case II, the components $\phi_2, \phi_1, \phi_{-1}, \phi_{-2}$ are always close to zero (∞ -norm less than 10^{-6}).

8. Conclusions. The Fourier pseudospectral method was adopted to discretize the energy functional and constraints for computing the ground states of spin- F Bose–

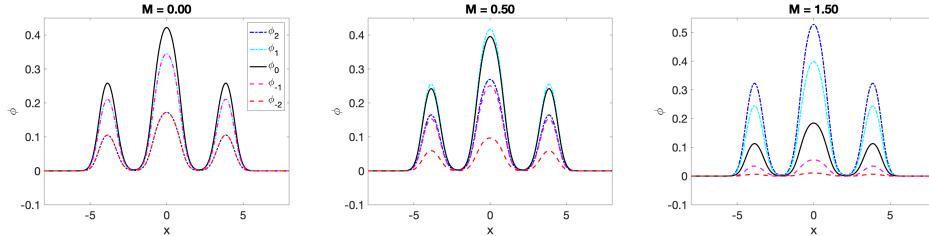


FIG. 1. Wave functions of the ground state, i.e., $\phi_2(x)$ (blue dash-dot line), $\phi_1(x)$ (light blue dash-dot line), $\phi_0(x)$ (black solid line), $\phi_{-1}(x)$ (purple dashed line), and $\phi_{-2}(x)$ (red dashed line) of a spin-2 BEC for Case I in 1D under different magnetizations $M = 0, 0.5, 1.5$. (Color available online.)

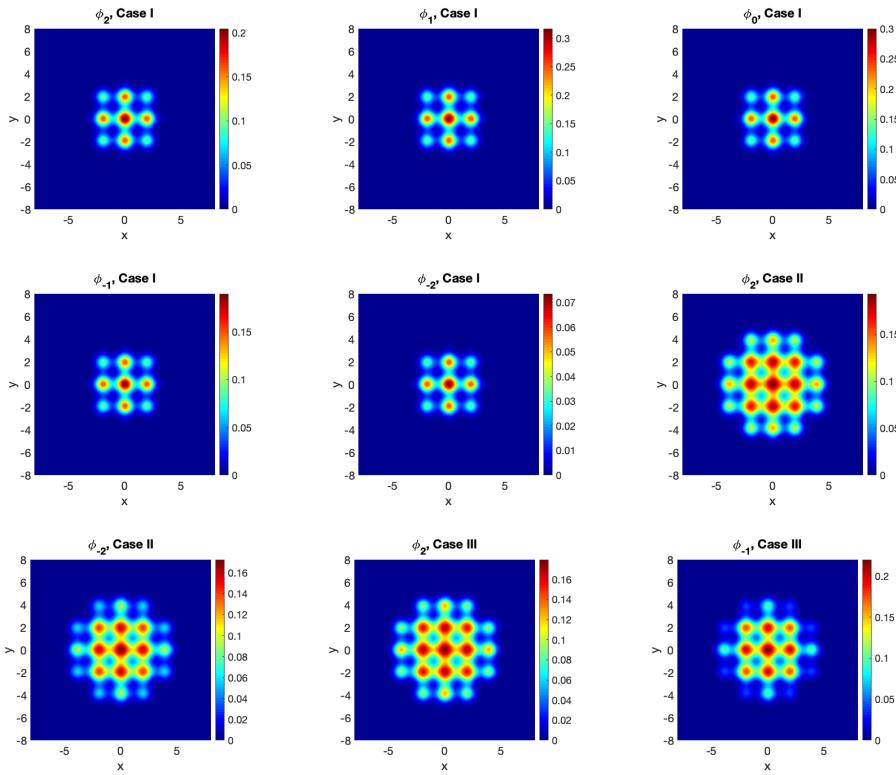


FIG. 2. Contour plots for the wave functions of the ground state, i.e., $\phi_2(x, y)$, $\phi_1(x, y)$, $\phi_0(x, y)$, $\phi_{-1}(x, y)$, $\phi_{-2}(x, y)$ of a spin-2 BEC in 2D with $M = 0.5$ under different interactions. In Case II, the components $\phi_1(x, y)$, $\phi_0(x, y)$, $\phi_{-1}(x, y)$ are zero. In Case III, the components $\phi_1(x, y)$, $\phi_0(x, y)$, $\phi_{-2}(x, y)$ are zero.

Einstein condensates (BECs). The original variational problem was reduced to a finite-dimensional Riemannian optimization problem. An adaptive regularized Newton method, combined with a Riemannian gradient method and a cascadic multigrid technique, was designed to solve the discretized problem. Three different retractions were proposed to implement the optimization algorithms on the manifold. Comparison with the Riemannian gradient method and trust-region method for different

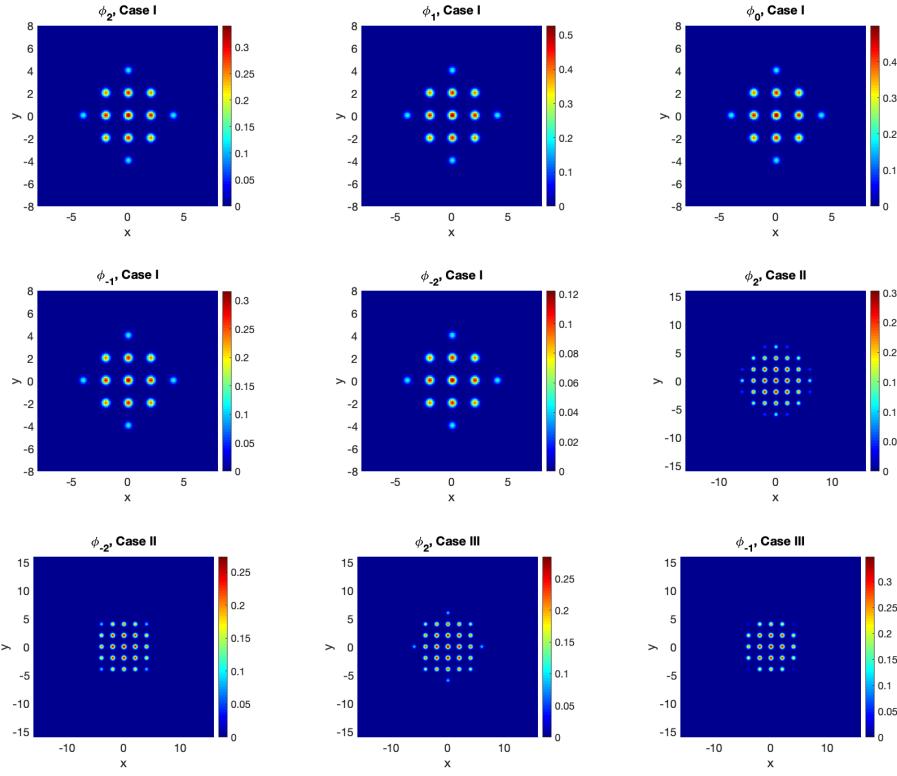


FIG. 3. Contour plots for the wave functions of the ground state, i.e., $\phi_2(x, y, 0)$, $\phi_1(x, y, 0)$, $\phi_0(x, y, 0)$, $\phi_{-1}(x, y, 0)$, $\phi_{-2}(x, y, 0)$ of a spin-2 BEC in 3D with $M = 0.5$ under different interactions. In Case II, the components $\phi_1(x, y, z)$, $\phi_0(x, y, z)$, $\phi_{-1}(x, y, z)$ are zero. In Case III, the components $\phi_1(x, y, z)$, $\phi_0(x, y, z)$, $\phi_{-2}(x, y, z)$ are zero.

TABLE 9
Ground state energies of spin-3 BECs.

M	1D		2D		3D	
	Case I	Case II	Case I	Case II	Case I	Case II
0.0	17.1091	17.2527	11.7811	11.8279	42.9028	42.9752
0.5	17.1289	17.2693	11.7877	11.8334	42.9115	42.9822
1.5	17.2905	17.4034	11.8415	11.8780	42.9825	43.0399

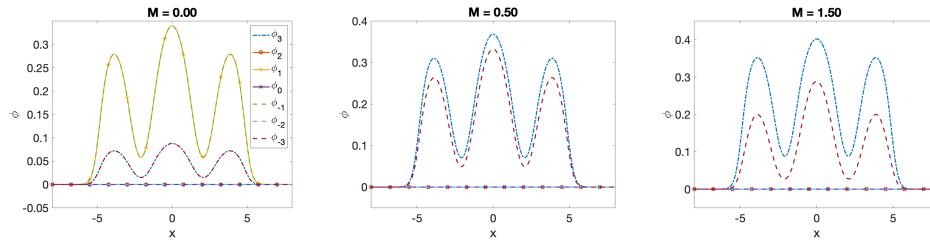


FIG. 4. Wave functions of the ground state, i.e., $\phi_3(x)$, $\phi_2(x)$, $\phi_1(x)$, $\phi_0(x)$, $\phi_{-1}(x)$, $\phi_{-2}(x)$ and $\phi_{-3}(x)$ of a spin-3 BEC for Case I in 1D under different magnetizations $M = 0, 0.5, 1.5$.

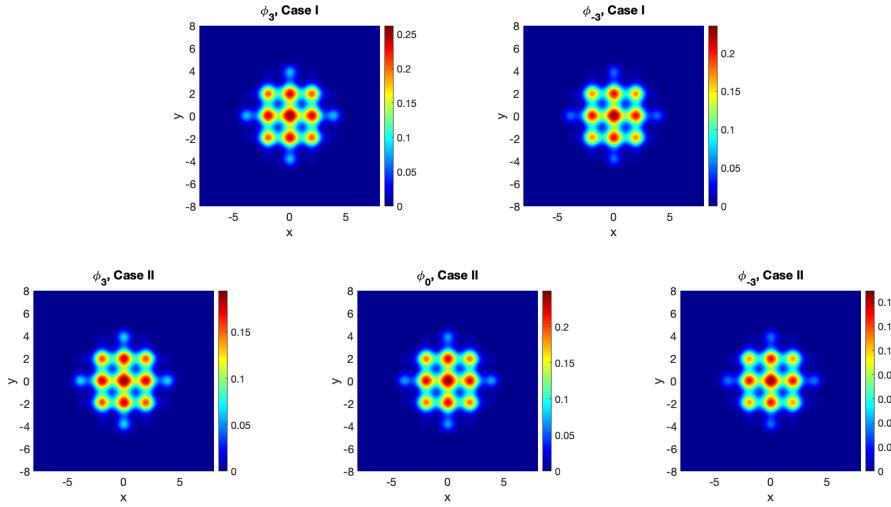


FIG. 5. Contour plots for the wave functions of the ground state, i.e., $\phi_3(x, y)$, $\phi_2(x, y)$, $\phi_1(x, y)$, $\phi_0(x, y)$, $\phi_{-1}(x, y)$, $\phi_{-2}(x, y)$, $\phi_{-3}(x, y)$ of a spin-3 BEC in 2D with $M = 0.5$ under different interactions. In Case I, the components $\phi_2(x, y)$, $\phi_1(x, y)$, $\phi_0(x, y)$, $\phi_{-1}(x, y)$, $\phi_{-2}(x, y)$ are close to zero. In Case II, the components $\phi_2(x, y)$, $\phi_1(x, y)$, $\phi_{-1}(x, y)$, $\phi_{-2}(x, y)$ are close to zero.

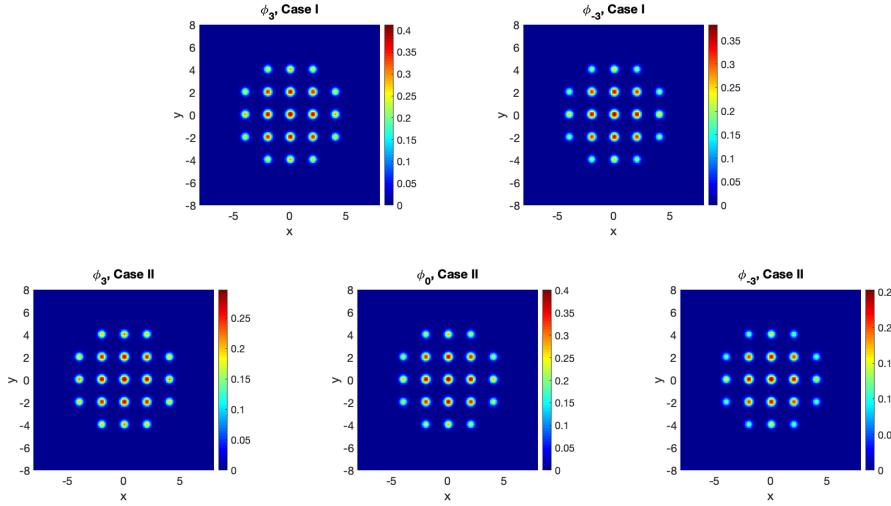


FIG. 6. Contour plots for the wave functions of the ground state, i.e., $\phi_3(x, y, 0)$, $\phi_2(x, y, 0)$, $\phi_1(x, y, 0)$, $\phi_0(x, y, 0)$, $\phi_{-1}(x, y, 0)$, $\phi_{-2}(x, y, 0)$, $\phi_{-3}(x, y, 0)$ of a spin-3 BEC in 3D with $M = 0.5$ under different interactions. In Case I, the components $\phi_2(x, y, z)$, $\phi_1(x, y, z)$, $\phi_0(x, y, z)$, $\phi_{-1}(x, y, z)$, $\phi_{-2}(x, y, z)$ are close to zero, In Case II, the components $\phi_2(x, y, z)$, $\phi_1(x, y, z)$, $\phi_{-1}(x, y, z)$, $\phi_{-2}(x, y, z)$ are close to zero.

retractions and parameters showed that our approach is more efficient and stable. Extensive numerical examples of spin-2 and spin-3 BECs in 1D, 2D, and 3D with optical lattice potential and various interaction demonstrated the robustness of our approach. The energy and wave functions of ground states are reported to reveal some interesting physical phenomena. Our method is the first one to explore spin-3 BECs

computationally. Although the spin-3 cases discussed in this paper are relatively simple, our algorithm is also applicable to cases with more diverse parameters.

Acknowledgments. The authors are grateful to Prof. Eric Cancès and two anonymous referees for their valuable comments and suggestions.

REFERENCES

- [1] P.-A. ABSIL, C. BAKER, AND K. GALLIVAN, *Trust-region methods on Riemannian manifolds*, Found. Comput. Math., 7 (2007), pp. 303–330.
- [2] P.-A. ABSIL, R. MAHONY, AND R. SEPULCHRE, *Optimization Algorithms on Matrix Manifolds*, Princeton University Press, Princeton, NJ, 2008.
- [3] P.-A. ABSIL AND J. MALICK, *Projection-like retractions on matrix manifolds*, SIAM J. Optim., 22 (2012), pp. 135–158, <https://doi.org/10.1137/100802529>.
- [4] J. ANDERSON, *Theory of the weakly interacting Bose gas*, Rev. Mod. Phys., 76 (2004), pp. 599–639.
- [5] M. ANDERSON, J. ENSHER, M. MATTEWS, C. WIEMAN, AND E. CORNELL, *Observation of Bose-Einstein condensation in a dilute atomic vapor*, Science, 269 (1995), pp. 198–201.
- [6] W. BAO, *Ground states and dynamics of multicomponent Bose-Einstein condensates*, Multiscale Model. Simul., 2 (2004), pp. 210–236, <https://doi.org/10.1137/030600209>.
- [7] W. BAO, I.-L. CHERN, AND Y. ZHANG, *Efficient numerical methods for computing ground states of spin-1 Bose-Einstein condensates based on their characterizations*, J. Comput. Phys., 253 (2013), pp. 189–208.
- [8] W. BAO AND Q. DU, *Computing the ground state solution of Bose-Einstein condensates by a normalized gradient flow*, SIAM J. Sci. Comput., 25 (2004), pp. 1674–1697, <https://doi.org/10.1137/S1064827503422956>.
- [9] W. BAO AND F. Y. LIM, *Computing ground states of spin-1 Bose-Einstein condensates by the normalized gradient flow*, SIAM J. Sci. Comput., 30 (2008), pp. 1925–1948, <https://doi.org/10.1137/070698488>.
- [10] W. BAO, Q. TANG, AND Y. YUAN, *Computing Ground States of Spin-2 Bose-Einstein Condensates by the Normalized Gradient Flow*, preprint.
- [11] W. BAO AND H. WANG, *A mass and magnetization conservative and energy-diminishing numerical method for computing ground state of spin-1 Bose-Einstein condensates*, SIAM J. Numer. Anal., 45 (2007), pp. 2177–2200, <https://doi.org/10.1137/070681624>.
- [12] M. BARRETT, J. SAUER, AND M. CHAPMAN, *All-optical formation of an atomic Bose-Einstein condensate*, Phys. Rev. Lett., 87 (2001), 010404.
- [13] F. BORNEMANN AND P. DEUFLHARD, *The cascadic multigrid method for elliptic problems*, Numer. Math., 75 (1996), pp. 135–152.
- [14] C. BRADLEY, C. SACKETT, J. TOLLETT, AND R. HULET, *Evidence of Bose-Einstein condensation in an atomic gas with attractive interactions*, Phys. Rev. Lett., 75 (1995), pp. 1687–1690.
- [15] F. DALFOVO, S. GIORGINI, L. PITAEVSKII, AND S. STRINGARI, *Theory of Bose-Einstein condensation in trapped gases*, Rev. Mod. Phys., 71 (1999), pp. 463–512.
- [16] I. DANAILA AND B. PROTAS, *Computation of ground states of the Gross-Pitaevskii functional via Riemannian optimization*, SIAM J. Sci. Comput., 39 (2017), pp. B1102–B1129, <https://doi.org/10.1137/17M1121974>.
- [17] K. DAVIS, M. MEWES, M. ANDREWS, N. VAN DRUTEN, D. DURFEE, D. KURN, AND W. KETTERLE, *Bose-Einstein condensation in a gas of sodium atoms*, Phys. Rev. Lett., 75 (1995), pp. 3969–3973.
- [18] A. FETTER, *Rotating trapped Bose-Einstein condensates*, Rev. Mod. Phys., 81 (2009), pp. 647–691.
- [19] S. GAUTAM AND S. K. ADHIKARI, *Analytic models for the density of a ground-state spinor condensate*, Phys. Rev. A, 92 (2015), 023616, <https://doi.org/10.1103/PhysRevA.92.023616>.
- [20] A. GÖRLITZ, T. L. GUSTAVSON, A. E. LEANHARDT, R. LÖW, A. P. CHIKKATUR, S. GUPTA, S. INOUYE, D. E. PRITCHARD, AND W. KETTERLE, *Sodium Bose-Einstein condensates in the $F = 2$ state in a large-volume optical trap*, Phys. Rev. Lett., 90 (2003), 090401.
- [21] T. HO, *Spinor Bose condensates in optical traps*, Phys. Rev. Lett., 81 (1998), pp. 742–745.
- [22] J. HU, A. MILZAREK, Z. WEN, AND Y. YUAN, *Adaptive quadratically regularized Newton method for Riemannian optimization*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 1181–1207, <https://doi.org/10.1137/17M1142478>.
- [23] C. LAW, H. PU, AND N. BIGELOW, *Quantum spins mixing in spinor Bose-Einstein condensates*,

- Phys. Rev. Lett., 81 (1998), pp. 5257–5261.
- [24] A. LEGGETT, *Bose-Einstein condensation in the alkali gases: Some fundamental concepts*, Rev. Mod. Phys., 73 (2001), pp. 307–356.
 - [25] H. MIESNER, D. STAMPER-KURN, J. STENGER, S. INOUYE, A. CHIKKATUR, AND W. KETTERLE, *Observation of metastable states in spinor Bose-Einstein condensates*, Phys. Rev. Lett., 82 (1999), pp. 2228–2231.
 - [26] O. MORSCH AND M. OBERTHALER, *Dynamics of Bose-Einstein condensates in optical lattices*, Rev. Mod. Phys., 78 (2006), pp. 179–215.
 - [27] T. OHMI AND K. MACHIDA, *Bose-Einstein condensation with internal degrees of freedom in alkali atom gases*, J. Phys. Soc. Japan, 67 (1998), pp. 1822–1825.
 - [28] R. OZERI, N. KATZ, J. STEINHAUER, AND N. DAVIDSON, *Colloquium: Bulk Bogoliubov excitations in a Bose-Einstein condensate*, Rev. Mod. Phys., 77 (2005), pp. 187–205.
 - [29] A. POSAZHENNIKOVA, *Colloquium: Weakly interacting, dilute Bose gases in 2D*, Rev. Mod. Phys., 78 (2006), pp. 1111–1134.
 - [30] D. STAMPER-KURN, M. ANDREWS, A. CHIKKATUR, S. INOUYE, H. MIESNER, J. STENGER, AND W. KETTERLE, *Optical confinement of a Bose-Einstein condensate*, Phys. Rev. Lett., 80 (1998), pp. 2027–2030.
 - [31] D. STAMPER-KURN AND W. KETTERLE, *Spinor condensates and light scattering from Bose-Einstein condensates*, in *Coherent Atomic Matter Waves*, Springer, 2001, pp. 139–217.
 - [32] J. STENGER, S. INOUYE, D. STAMPER-KURN, H. MIESNER, A. CHIKKATUR, AND W. KETTERLE, *Spin domains in ground state Bose-Einstein condensates*, Nature, 396 (1998), pp. 345–348.
 - [33] H. WANG, *A projection gradient method for computing ground state of spin-2 Bose-Einstein condensates*, J. Comput. Phys., 274 (2014), pp. 473–488.
 - [34] W. YANG, L.-H. ZHANG, AND R. SONG, *Optimality conditions for the nonlinear programming problems on Riemannian manifolds*, Pac. J. Optim., 10 (2014), pp. 415–434.
 - [35] W. ZHANG, S. YI, AND L. YOU, *Mean field ground state of a spin-1 condensate in a magnetic field*, New J. Phys., 5 (2003), pp. 77–89.
 - [36] W. ZHANG AND L. YOU, *An effective quasi-one-dimensional description of a spin-1 atomic condensate*, Phys. Rev. A, 71 (2005), 025603.