Constrained high-index saddle dynamics for the solution landscape with equality constraints

Jianyuan Yin · Zhen Huang · Lei Zhang

Abstract We propose a constrained high-index saddle dynamics (CHiSD) method to search for index-k saddle points of an energy functional subject to equality constraints. With Riemannian manifold tools, the CHiSD is derived in a minimax framework, and its linear stability at an index-k saddle point is proved. To ensure the manifold property, the CHiSD is numerically implemented using retractions and vector transport. Then we present a numerical approach by combining CHiSD with downward and upward search algorithms to construct the solution landscape in the presence of equality constraints. We apply the Thomson problem and the Bose–Einstein condensation as numerical examples to demonstrate the efficiency of the proposed method.

1 Introduction

The energy landscape, which maps all possible configurations of a system to their corresponding energy [44], has been successfully applied to various scientific issues,

J. Yin
School of Mathematical Sciences, Peking University, Beijing 100871, China
E-mail: yinjy@pku.edu.cn

Z. Huang
School of Mathematical Sciences, Peking University, Beijing 100871, China
E-mail: hertz@pku.edu.cn

L. Zhang
Beijing International Center for Mathematical Research, Center for Quantitative Biology, Peking University, Beijing 100871, China
E-mail: zhangl@math.pku.edu.cn

arXiv:2011.13173v3 [math.NA] 11 Apr 2022

such as particle clusters [33, 35], protein folding [25, 32], and soft matter [12, 20]. The stationary point, at which the gradient vanishes on the energy landscape, plays an important role to determine physical or chemical properties of the system. The stability of a stationary point is determined by its Hessian matrix. For instance, a stationary point is a minimizer if all eigenvalues of its Hessian are positive. The stationary points with both positive and negative eigenvalues are called saddle points, which can be further classified by the Morse index. The Morse index is equal to the number of negative eigenvalues of the Hessian matrix at each stationary point. In particular, the index-1 saddle point, which has one and only one negative eigenvalue, is referred to as the transition state connecting two minima on the energy landscape.

Searching for saddle points on a complicated energy landscape has attracted plenty of attention during the past decades [52, 16, 54]. Compared to finding stable minima, computation of saddle points is much more challenging due to their unstable nature. Extensive numerical algorithms have been developed to compute index-1 saddle points, including path-finding methods [22, 15] and surface-walking methods [39]. In particular, the surface-walking method starts from one initial state on the energy landscape and searches for index-1 saddle points based on local derivative information without a priori knowledge of the final state. Examples of the surface-walking methods include the dimer-type methods [21, 51, 53], the gentlest ascent dynamics [17, 40], and the activation-relaxation technique [9, 31]. Moreover, the surface-walking methods can be generalized to search for high-index (i.e. index greater than one) saddle points. Quapp and Bofill developed a generalized gentlest ascent dynamics to locate high-index saddle points on the energy landscape with the calculation of the Hessian matrix [40]. The minimax method based on the local minimax theorem was proposed to find multiple high-index saddle points with a priori knowledge of low-index saddle points [26, 27, 28]. We refer to some excellent reviews for more information [16, 54]. Recently, Yin et al. proposed a high-index saddle dynamics (HiSD) and developed a high-index optimization-based shrinking dimer method for finding index-k saddle points on the energy landscape [49]. Later, a generalized high-index saddle dynamics (GHiSD) was developed to compute any-index saddle points of dynamical systems [48].

In many practical applications, the challenge of searching for saddle points is further increased by nonlinear equality constraints on the state variables. For example, in the Oseen–Frank theory for nematic liquid crystals, the director $\mathbf{n}(\boldsymbol{x})$ is a vector field subject to a unit-length constraint almost everywhere that describes the average orientation of liquid crystal molecules at the position $\boldsymbol{x}$ [19, 45]. In the Kohn–Sham density functional theory, the electron orbitals Ψ are supposed to satisfy the orthonormality constraint [30]. A simple approach to deal with nonlinear equality constraints is reparametrization via unconstrained variables, but this is often cumbersome and computationally inefficient. Therefore, a number of numerical methods have been proposed to compute index-1 saddle points in the presence of equality constraints. As examples of path-finding methods, the constrained string method [14] and the geodesic nudged elastic band method [7] are able to find constrained minimum energy path and index-1 saddle points. Alternatively, the constrained shrinking dimer dynamics [50] was developed to locate index-1 saddle points on a constrained energy function using the projected Hessian. Müller et al. identified multiple transitions of Skyrmions in magnetic systems using a surface-walking method, where each magnetic vector is restricted to

a unit length [37]. The physical space, which is the direct product of N spheres, is naturally embedded into the Euclidean space, and the theory of Riemannian manifolds is considered to derive the Hessian. The unphysical degrees of freedom in the embedding space are removed to calculate the true eigenvectors required in this saddle-point searching method. However, how to numerically compute the constrained high-index (index> 1) saddle points is still unclear at present.

In this article, we present a constrained high-index saddle dynamics (CHiSD) to search for high-index saddle points subject to equality constraints. The CHiSD is a constrained version of the HiSD [49] and derived with Riemannian gradients and Hessians. The retraction operator and vector transport are introduced to implement the numerical algorithm of CHiSD. With the CHiSD algorithm, we are able to construct the solution landscape on a constrained manifold. The solution landscape is a pathway map consisting of all stationary points and their connections, which not only provides an efficient approach to find multiple stationary points, including both minima and saddle points, without tuning initial guesses, but also shows the relationships between different stationary points [47].

The rest of this article is organized as follows. In Sect. 2, we first introduce the constrained saddle points and Riemannian manifold tools. After briefly reviewing the HiSD for index-k saddle points, the CHiSD is derived with Riemannian gradient and Hessian in a similar manner, and the linear stability of the index-k saddle point is proved. In Sect. 3, retractions and vector transport are applied to maintain the manifold property in numerical schemes, and then the CHiSD algorithm is presented. Furthermore, the solution landscape in equality-constrained cases can be constructed with the downward and upward search algorithms based on the CHiSD algorithm. In Sect. 4, we show two numerical examples to demonstrate the efficiency of the CHiSD algorithm. First, we construct the solution landscape of the Thomson problem to identify all possible stationary points in the cases of the particle number $N = 5, 7$ and 9. Second, as a nonlinear elliptic eigenvalue problem, the excited states of the Bose–Einstein condensates (BEC) are calculated with a combination of the CHiSD and the upward search algorithm. Some conclusions and discussions are presented in Sect. 5.

2 Constrained high-index saddle dynamics

2.1 Constrained saddle points

Given a twice Fréchet differentiable energy functional $E(\boldsymbol{x})$ defined on a d-dimensional real Hilbert space $\mathcal{H}$ with an inner product $\langle \cdot, \cdot \rangle$ and the norm $\| \cdot \|$, we let $\nabla E(\cdot) : \mathcal{H} \to \mathcal{H}$ denotes the Riesz map applied to the Fréchet derivative, and $\nabla^2 E(\boldsymbol{x}) \in \mathcal{L}(\mathcal{H})$ denotes the Hessian. We consider the functional $E(\boldsymbol{x})$ for $\boldsymbol{x} \in \mathcal{H}$ subject to m equality constraints,

$$\boldsymbol{c}(\boldsymbol{x}) = (c_1(\boldsymbol{x}), \cdots, c_m(\boldsymbol{x})) = (0, \cdots, 0). \tag{1}$$

where each $c_p : \mathcal{H} \to \mathbb{R}$ is a smooth function. Let $\mathbf{A}(\boldsymbol{x})$ denote $(\nabla c_1(\boldsymbol{x}), \cdots, \nabla c_m(\boldsymbol{x}))$, and it is always assumed that $\nabla c_1(\boldsymbol{x})$, $\cdots$, $\nabla c_m(\boldsymbol{x})$ are linearly independent for each $\boldsymbol{x}$ subject to (1), which is often referred to as the linear independence constraint qualification (LICQ) in optimization theory [38]. From the regular level set

theorem [43], the feasible set consisting of all feasible points,

$$\mathcal{M} = \{\boldsymbol{x} \in \mathcal{H} : \boldsymbol{c}(\boldsymbol{x}) = (0, \cdots, 0)\}, \tag{2}$$

is a $(d-m)$-dimensional smooth Riemannian manifold with induced metric. For $\boldsymbol{x} \in \mathcal{H}$, the normal space of the isosurface of $\boldsymbol{c}(\boldsymbol{x})$ is defined as $N(\boldsymbol{x}) = \text{span}\{\mathbf{A}(\boldsymbol{x})\}$, and the tangent space is defined as its orthogonal complement $T(\boldsymbol{x}) = N(\boldsymbol{x})^{\perp} = \{\boldsymbol{v} \in \mathcal{H} : \mathbf{A}(\boldsymbol{x})^{\top}\boldsymbol{v} = \mathbf{0}\}$. For $\mathbf{A} = (\boldsymbol{a}_1, \cdots, \boldsymbol{a}_p)$ and $\mathbf{B} = (\boldsymbol{b}_1, \cdots, \boldsymbol{b}_q)$ with columns in $\mathcal{H}$, $\mathbf{A}^{\top}\mathbf{B}$ denotes a $p \times q$ matrix whose (i,j)-entry is $\langle \boldsymbol{b}_j, \boldsymbol{a}_i \rangle$, and $\mathbf{A}^{\top}$ denotes the corresponding linear operator for simplicity. The orthogonal projection operators $\mathbf{P}$ on these spaces have the forms of,

$$\mathbf{P}_{N(\boldsymbol{x})} = \mathbf{A}(\boldsymbol{x}) \left(\mathbf{A}(\boldsymbol{x})^{\top}\mathbf{A}(\boldsymbol{x})\right)^{-1} \mathbf{A}(\boldsymbol{x})^{\top}, \qquad \mathbf{P}_{T(\boldsymbol{x})} = \mathbf{I} - \mathbf{P}_{N(\boldsymbol{x})}, \tag{3}$$

where $\mathbf{A}^{\top}\mathbf{A}$ is positive definite for each $\boldsymbol{x} \in \mathcal{M}$ from LICQ, and remains positive definite in a neighbourhood $U(\boldsymbol{x}) \subset \mathcal{H}$.

Since the functional E is constrained on the Riemannian manifold $\mathcal{M}$, the Riemannian gradient and Hessian should be considered in replacement of $\nabla E(\boldsymbol{x})$ and $\nabla^2 E(\boldsymbol{x})$ in unconstrained cases. For $\boldsymbol{x} \in \mathcal{M}$, the Riemannian gradient is defined as,

$$\text{grad}\, E(\boldsymbol{x}) = \mathbf{P}_{T(\boldsymbol{x})} \nabla E(\boldsymbol{x}), \tag{4}$$

and the Riemannian Hessian $\text{Hess}\, E(\boldsymbol{x}) : \boldsymbol{\eta} \in T(\boldsymbol{x}) \to T(\boldsymbol{x}) \subset \mathcal{H}$ is defined as,

$$\begin{aligned}
\text{Hess}\, E(\boldsymbol{x})[\boldsymbol{\eta}] &= \mathbf{P}_{T(\boldsymbol{x})} \left(\partial_{\boldsymbol{\eta}} \text{grad}\, E(\boldsymbol{x})\right) \\
&= \mathbf{P}_{T(\boldsymbol{x})} \left(\nabla^2 E(\boldsymbol{x})\boldsymbol{\eta} - \nabla^2 \boldsymbol{c}(\boldsymbol{x})\boldsymbol{\eta} \left(\mathbf{A}(\boldsymbol{x})^{\top}\mathbf{A}(\boldsymbol{x})\right)^{-1} \mathbf{A}(\boldsymbol{x})^{\top}\nabla E(\boldsymbol{x})\right),
\end{aligned} \tag{5}$$

where $\nabla^2 \boldsymbol{c}(\boldsymbol{x})\boldsymbol{\eta} = \left(\nabla^2 c_1(\boldsymbol{x})\boldsymbol{\eta}, \cdots, \nabla^2 c_m(\boldsymbol{x})\boldsymbol{\eta}\right)$ [1]. The Riemannian Hessian (5) is a symmetric operator on $T(\boldsymbol{x})$, and can be extended as a symmetric operator on $\mathcal{H}$,

$$\widetilde{\text{Hess}}\, E(\boldsymbol{x})[\boldsymbol{\eta}] := \text{Hess}\, E(\boldsymbol{x})[\mathbf{P}_{T(\boldsymbol{x})}\boldsymbol{\eta}], \tag{6}$$

to avoid the problem of definition domains. The Riemannian gradient (4) and Hessian (5) can be naturally extended to a neighbourhood $U(\boldsymbol{x}) \subset \mathcal{H}$ of each $\boldsymbol{x} \in \mathcal{M}$ with the same expressions, and can be thought of as the Riemannian gradient and Hessian on some isosurface of $\boldsymbol{c}(\boldsymbol{x})$.

A point $\hat{\boldsymbol{x}} \in \mathcal{M}$ is called a *stationary point* (or *critical point*) of the functional E subject to equality constraints (1) if $\text{grad}\, E(\hat{\boldsymbol{x}}) = \mathbf{0}$, and a stationary point $\hat{\boldsymbol{x}}$ is said to be *nondegenerate* if $\text{Hess}\, E(\hat{\boldsymbol{x}})$ has a bounded inverse on $T(\hat{\boldsymbol{x}})$. A stationary point $\hat{\boldsymbol{x}} \in \mathcal{M}$ is a *(constrained) saddle point*, if $\hat{\boldsymbol{x}}$ is not a local extremum on $\mathcal{M}$. According to the Morse theory, the *(Morse) index* of a stationary point $\hat{\boldsymbol{x}}$ is defined as the maximal dimension of a subspace $\mathcal{K} \subseteq T(\hat{\boldsymbol{x}})$ on which $\text{Hess}\, E(\hat{\boldsymbol{x}})$ is negative definite [36]. For an index-k saddle point (a k-saddle) $\hat{\boldsymbol{x}}$, the Riemannian Hessian $\text{Hess}\, E(\hat{\boldsymbol{x}})$ has $(d-m)$ eigenvalues $\hat{\lambda}_1 \leqslant \cdots \hat{\lambda}_k < 0 \leqslant \hat{\lambda}_{k+1} \leqslant \cdots \leqslant \hat{\lambda}_{d-m}$ with corresponding orthonormal eigenvectors $\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_{d-m}$. A nondegenerate index-$k$ saddle point $\hat{\boldsymbol{x}}$ is a local maximum on a k-dimensional submanifold $\mathcal{M}_-(\hat{\boldsymbol{x}})$ of $\mathcal{M}$, and a local minimum on a $(d-m-k)$-dimensional submanifold $\mathcal{M}_+(\hat{\boldsymbol{x}})$, and the tangent spaces at $\hat{\boldsymbol{x}}$ of the two submanifolds $\mathcal{M}_-(\hat{\boldsymbol{x}})$ and $\mathcal{M}_+(\hat{\boldsymbol{x}})$ are respectively $\text{span}\{\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_k\}$ and $\text{span}\{\hat{\boldsymbol{v}}_{k+1}, \cdots, \hat{\boldsymbol{v}}_{d-m}\}$. This minimax structure inspires us

to develop numerical methods for searching for saddle points with a certain index on the manifold.

Another frequently-used approach to studying stationary points in a constrained problem is the Lagrangian function. The Lagrangian function $L_E(\boldsymbol{x}, \boldsymbol{\xi})$ of the energy E with the equality constraints (1) is,

$$L_E(\boldsymbol{x}, \boldsymbol{\xi}) = E(\boldsymbol{x}) - \boldsymbol{c}(\boldsymbol{x})\boldsymbol{\xi}, \tag{7}$$

where $\boldsymbol{\xi} \in \mathbb{R}^m$ is the Lagrangian multiplier. A stationary point $\hat{\boldsymbol{x}} \in \mathcal{M}$ can be equivalently defined as where the first-order Karush–Kuhn–Tucker (KKT) condition,

$$\nabla_{\boldsymbol{x}} L_E(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\xi}}) = \nabla E(\hat{\boldsymbol{x}}) - \mathbf{A}(\hat{\boldsymbol{x}})\hat{\boldsymbol{\xi}} = \boldsymbol{0}, \tag{8}$$

holds for some multiplier $\hat{\boldsymbol{\xi}} \in \mathbb{R}^m$ [38]. From the KKT condition (8) and LICQ, the multiplier can be calculated as $\hat{\boldsymbol{\xi}} = \left(\mathbf{A}(\hat{\boldsymbol{x}})^\top \mathbf{A}(\hat{\boldsymbol{x}})\right)^{-1} \mathbf{A}(\hat{\boldsymbol{x}})^\top \nabla E(\hat{\boldsymbol{x}})$, and consequently, the Lagrangian Hessian at a stationary point $\hat{\boldsymbol{x}}$ is

$$\nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L_E(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\xi}}) = \nabla^2 E(\hat{\boldsymbol{x}}) - \nabla^2 \boldsymbol{c}(\hat{\boldsymbol{x}}) \left(\mathbf{A}(\hat{\boldsymbol{x}})^\top \mathbf{A}(\hat{\boldsymbol{x}})\right)^{-1} \mathbf{A}(\hat{\boldsymbol{x}})^\top \nabla E(\hat{\boldsymbol{x}}). \tag{9}$$

Since only the vectors in the tangent space $T(\hat{\boldsymbol{x}})$ are physical directions, a projected Hessian $\mathbf{P}_{T(\hat{\boldsymbol{x}})} \nabla_{\boldsymbol{x}\boldsymbol{x}}^2 L_E(\hat{\boldsymbol{x}}, \hat{\boldsymbol{\xi}}) \mathbf{P}_{T(\hat{\boldsymbol{x}})}$ is considered to determine the second-order properties of the stationary point, which accords with (5) in the tangent space.

We take the unit sphere constraint $c(\boldsymbol{x}) = (\boldsymbol{x}^\top \boldsymbol{x} - 1)/2$ in the Euclidean space $\mathbb{R}^d$ as a simple example, where the feasible set is customarily denoted as $S^{d-1} = \left\{\boldsymbol{x} \in \mathbb{R}^d : \|\boldsymbol{x}\|_2^2 = 1\right\}$. For $\boldsymbol{x} \neq \boldsymbol{0}$ and $\boldsymbol{\eta} \in T(\boldsymbol{x}) = \text{span}\{\boldsymbol{x}\}^\perp$, we have,

$$\begin{aligned} \mathbf{P}_{N(\boldsymbol{x})} &= \boldsymbol{x}(\boldsymbol{x}^\top \boldsymbol{x})^{-1} \boldsymbol{x}^\top, \quad \text{grad}\, E(\boldsymbol{x}) = \left(\mathbf{I} - \boldsymbol{x}(\boldsymbol{x}^\top \boldsymbol{x})^{-1} \boldsymbol{x}^\top\right) \nabla E(\boldsymbol{x}), \\ \text{Hess}\, E(\boldsymbol{x})[\boldsymbol{\eta}] &= \left(\mathbf{I} - \boldsymbol{x}(\boldsymbol{x}^\top \boldsymbol{x})^{-1} \boldsymbol{x}^\top\right) \nabla^2 E(\boldsymbol{x})\boldsymbol{\eta} - (\boldsymbol{x}^\top \boldsymbol{x})^{-1} \boldsymbol{x}^\top \nabla E(\boldsymbol{x})\boldsymbol{\eta}. \end{aligned} \tag{10}$$

2.2 Review of the HiSD method

The HiSD method aims to search for a high-index saddle point in an unconstrained case [49]. The HiSD for a k-saddle is given by

$$\begin{cases} \dot{\boldsymbol{x}} = -\left(\mathbf{I} - \sum_{i=1}^{k} 2\boldsymbol{v}_i \boldsymbol{v}_i^\top\right) \nabla E(\boldsymbol{x}), \\ \dot{\boldsymbol{v}}_i = -\left(\mathbf{I} - \boldsymbol{v}_i \boldsymbol{v}_i^\top - \sum_{j=1}^{i-1} 2\boldsymbol{v}_j \boldsymbol{v}_j^\top\right) \nabla^2 E(\boldsymbol{x})\boldsymbol{v}_i, \quad i = 1, \cdots, k, \end{cases} \tag{11}$$

which involves a position variable $\boldsymbol{x}$ and k directional variables $\boldsymbol{v}_i$ with an initial condition,

$$\boldsymbol{x} = \boldsymbol{x}^{(0)} \in \mathcal{H}, \quad \boldsymbol{v}_i = \boldsymbol{v}_i^{(0)} \in \mathcal{H}, \quad \text{s.t.} \left\langle \boldsymbol{v}_j^{(0)}, \boldsymbol{v}_i^{(0)} \right\rangle = \delta_{ij}, \quad i, j = 1, \cdots, k. \tag{12}$$

With straightforward calculations, for the HiSD (11) with the initial condition (12), the vectors $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$ always satisfy the orthonormal condition $\langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = \delta_{ij}$. The dynamics for $\boldsymbol{x}$ is actually a transformed gradient flow,

$$\dot{\boldsymbol{x}} = \mathbf{P}_{\mathcal{V}} \nabla E(\boldsymbol{x}) - (\mathbf{I} - \mathbf{P}_{\mathcal{V}}) \nabla E(\boldsymbol{x}), \tag{13}$$

where $\mathbf{P}_{\mathcal{V}} \nabla E(\boldsymbol{x})$ is the gradient ascent direction on $\mathcal{V} = \mathrm{span}\{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\}$ and $(\mathbf{I} - \mathbf{P}_{\mathcal{V}}) \nabla E(\boldsymbol{x})$ is the gradient descent direction on $\mathcal{V}^{\perp}$. Since a nondegenerate k-saddle is a local maximum along k orthogonal directions and a local minimum along other orthogonal directions, this dynamics can find a k-saddle with proper $\{\boldsymbol{v}_i : i = 1, \cdots, k\}$. The dynamics for $\boldsymbol{v}_i$ renews the subspace $\mathcal{V}$ in (13) by finding the eigenvectors corresponding to the smallest k eigenvalues of the Hessian $\nabla^2 E(\boldsymbol{x})$ at the current position $\boldsymbol{x}$. The eigenvector $\boldsymbol{v}_i$ corresponding to the i-th smallest eigenvalue of the Hessian $\nabla^2 E(\boldsymbol{x})$ can be obtained by solving a constrained optimization problem of the Rayleigh quotient,

$$\min_{\boldsymbol{v}_i \in \mathcal{H}} \quad \langle \nabla^2 E(\boldsymbol{x}) \boldsymbol{v}_i, \boldsymbol{v}_i \rangle \quad \text{s.t.} \langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = \delta_{ij}, \quad j = 1, \cdots, i, \tag{14}$$

with the knowledge of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{i-1}$, and the $\boldsymbol{v}_i$ dynamics in (11) solves the constrained optimization problem (14) using gradient flow. Then the subspace $\mathcal{V}$ in (13) is spanned by the vectors $\{\boldsymbol{v}_i : i = 1, \cdots, k\}$. In practice, the Hessian in (11) is often approximated by dimers with a length of $2l$ [21],

$$\nabla^2 E(\boldsymbol{x}) \boldsymbol{v}_i \approx (\nabla E(\boldsymbol{x} + l\boldsymbol{v}_i) - \nabla E(\boldsymbol{x} - l\boldsymbol{v}_i)) / 2l, \tag{15}$$

and the HiSD with dimer approximations and dimer shrinkage $\dot{l} = -l$ is referred to as the high-index optimization-based shrinking dimer method [49]. In numerical implementation, the HiSD can be simply discretized with an explicit Euler scheme.

2.3 Formulation of CHiSD

In order to search for a nondegenerate k-saddle $\hat{\boldsymbol{x}}$ on the manifold $\mathcal{M}$, the HiSD method (11) is supposed to be generalized to the equality-constrained case (1). With the minimax structure of saddle points on manifolds, we set the dynamics for a k-saddle $\hat{\boldsymbol{x}}$ as a transformed gradient flow similarly,

$$\dot{\boldsymbol{x}} = -(\mathbf{I} - 2\mathbf{P}_{\mathcal{V}}) \operatorname{grad} E(\boldsymbol{x}), \tag{16}$$

which is gradient ascent on the subspace $\mathcal{V}$ and gradient descent on its orthogonal complement $\mathcal{V}^{\perp}$. The k-dimensional subspace $\mathcal{V} \subseteq T(\boldsymbol{x})$ is spanned by k orthonormal directions $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k$, and the orthogonal projection operator $\mathbf{P}_{\mathcal{V}}$ possesses a simple form of $\sum_{i=1}^{k} \boldsymbol{v}_i \boldsymbol{v}_i^{\top}$. At a k-saddle $\hat{\boldsymbol{x}}$, the k directions $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k \in T(\hat{\boldsymbol{x}})$ should be the eigenvectors of the smallest k eigenvalues of the Riemannian Hessian $\operatorname{Hess} E(\hat{\boldsymbol{x}})$ (5). Therefore, the direction $\boldsymbol{v}_i$ at the current position $\boldsymbol{x}$ is approximated by the i-th eigenvector of $\operatorname{Hess} E(\boldsymbol{x})$, which can be obtained by a constrained optimization problem with the knowledge of $\boldsymbol{v}_1, \cdots, \boldsymbol{v}_{i-1}$,

$$\min_{\boldsymbol{v}_i \in T(\boldsymbol{x})} \langle \operatorname{Hess} E(\boldsymbol{x})[\boldsymbol{v}_i], \boldsymbol{v}_i \rangle, \quad \text{s.t.} \quad \langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = \delta_{ij}, \quad j = 1, \cdots, i. \tag{17}$$

Equivalently, we deal with another constrained optimization problem,

$$\min_{\boldsymbol{v}_i \in \mathcal{H}} \left\langle \widetilde{\text{Hess}}\, E(\boldsymbol{x})[\boldsymbol{v}_i], \boldsymbol{v}_i \right\rangle,$$
$$\text{s.t.} \quad \mathbf{A}(\boldsymbol{x})^\top \boldsymbol{v}_i = \mathbf{0}, \quad \langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = \delta_{ij}, \quad j = 1, \cdots, i, \tag{18}$$

with the operator $\widetilde{\text{Hess}}\, E(\boldsymbol{x}) : \mathcal{H} \to T(\boldsymbol{x}) \subset \mathcal{H}$ in (6). The Lagrangian function of (18) is,

$$L_i(\boldsymbol{v}_i, \boldsymbol{\xi}_i, \boldsymbol{\mu}_i) = \left\langle \widetilde{\text{Hess}}\, E(\boldsymbol{x})[\boldsymbol{v}_i], \boldsymbol{v}_i \right\rangle - \sum_{j=1}^{i} \xi_{ij} \left(\langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle - \delta_{ij} \right) - \boldsymbol{v}_i^\top \mathbf{A}(\boldsymbol{x}) \boldsymbol{\mu}_i, \tag{19}$$

where $\boldsymbol{\xi}_i \in \mathbb{R}^i$ and $\boldsymbol{\mu}_i \in \mathbb{R}^m$ are Lagrangian multipliers, and the dynamics of $\boldsymbol{v}_i$ is,

$$\dot{\boldsymbol{v}}_i = -\frac{1}{2} \nabla_{\boldsymbol{v}_i} L_i(\boldsymbol{v}_i, \boldsymbol{\xi}, \boldsymbol{\mu}) = -\widetilde{\text{Hess}}\, E(\boldsymbol{x})[\boldsymbol{v}_i] + \xi_i \boldsymbol{v}_i + \sum_{j=1}^{i-1} \frac{\xi_{ij}}{2} \boldsymbol{v}_j + \frac{1}{2} \mathbf{A}(\boldsymbol{x}) \boldsymbol{\mu}_i. \tag{20}$$

The dynamics (16) and (20) should maintain the manifold property,

$$\boldsymbol{x} \in \mathcal{M}, \quad \boldsymbol{v}_1, \cdots, \boldsymbol{v}_k \in T(\boldsymbol{x}), \tag{21}$$

as well as the orthonormal condition,

$$\langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = \delta_{ij}, \quad i, j = 1, \cdots, k. \tag{22}$$

Therefore, the Lagrangian multipliers $\boldsymbol{\xi}_i, \boldsymbol{\mu}_i$ should satisfy,

$$\frac{\mathrm{d}}{\mathrm{d}t} (\mathbf{A}(\boldsymbol{x})^\top \boldsymbol{v}_i) = \mathbf{A}(\boldsymbol{x})^\top \dot{\boldsymbol{v}}_i + \left(\nabla^2 \boldsymbol{c}(\boldsymbol{x}) \dot{\boldsymbol{x}} \right)^\top \boldsymbol{v}_i = \mathbf{0},$$
$$\frac{\mathrm{d}}{\mathrm{d}t} \langle \boldsymbol{v}_j, \boldsymbol{v}_i \rangle = \langle \dot{\boldsymbol{v}}_j, \boldsymbol{v}_i \rangle + \langle \boldsymbol{v}_j, \dot{\boldsymbol{v}}_i \rangle = 0, \tag{23}$$

and are obtained as,

$$\xi_{ii} = \left\langle \widetilde{\text{Hess}}\, E(\boldsymbol{x})[\boldsymbol{v}_i], \boldsymbol{v}_i \right\rangle, \quad \xi_{ij} = 4 \left\langle \widetilde{\text{Hess}}\, E(\boldsymbol{x})[\boldsymbol{v}_i], \boldsymbol{v}_j \right\rangle,$$
$$\boldsymbol{\mu}_i = -2 \left(\mathbf{A}(\boldsymbol{x})^\top \mathbf{A}(\boldsymbol{x}) \right)^{-1} \left(\nabla^2 \boldsymbol{c}(\boldsymbol{x}) \dot{\boldsymbol{x}} \right)^\top \boldsymbol{v}_i. \tag{24}$$

As a result, we obtain the CHiSD for a k-saddle (k-CHiSD) as,

$$\begin{cases} \dot{\boldsymbol{x}} = -\left(\mathbf{I} - \sum_{i=1}^{k} 2\boldsymbol{v}_i \boldsymbol{v}_i^\top \right) \text{grad}\, E(\boldsymbol{x}), \\ \dot{\boldsymbol{v}}_i = -\left(\mathbf{I} - \boldsymbol{v}_i \boldsymbol{v}_i^\top - \sum_{j=1}^{i-1} 2\boldsymbol{v}_j \boldsymbol{v}_j^\top \right) \widetilde{\text{Hess}}\, E(\boldsymbol{x})[\boldsymbol{v}_i] \\ \qquad - \mathbf{A}(\boldsymbol{x}) \left(\mathbf{A}(\boldsymbol{x})^\top \mathbf{A}(\boldsymbol{x}) \right)^{-1} \left(\nabla^2 \boldsymbol{c}(\boldsymbol{x}) \dot{\boldsymbol{x}} \right)^\top \boldsymbol{v}_i, \quad i = 1, \cdots, k, \end{cases} \tag{25}$$

with an initial condition at $t = 0$,

$$\boldsymbol{x} = \boldsymbol{x}^{(0)} \in \mathcal{M}, \quad \boldsymbol{v}_i = \boldsymbol{v}_i^{(0)} \in T(\boldsymbol{x}^{(0)}), \quad \left\langle \boldsymbol{v}_j^{(0)}, \boldsymbol{v}_i^{(0)} \right\rangle = \delta_{ij}, \quad i, j = 1, \cdots, k. \tag{26}$$

Remark 1 With straightforward calculations, it can be verified that the CHiSD (25) with the initial condition (26) satisfies (21) and (22) for $t > 0$.

Remark 2 The second term in the dynamics of $\boldsymbol{v}_i$ in (25), which is derived from the constraints $\boldsymbol{v}_i \in T(\boldsymbol{x})$, corresponds to the parallel translation induced by the Riemannian connection. For a smooth curve $\boldsymbol{x}(t)$ on the manifold $\mathcal{M}$, the tangent space $T(\boldsymbol{x})$ changes accordingly, and the dynamics

$$\dot{\boldsymbol{v}} = -\mathbf{A}(\boldsymbol{x})\left(\mathbf{A}(\boldsymbol{x})^\top \mathbf{A}(\boldsymbol{x})\right)^{-1}\left(\nabla^2 \boldsymbol{c}(\boldsymbol{x})\dot{\boldsymbol{x}}\right)^\top \boldsymbol{v}, \tag{27}$$

transports the tangent vector $\boldsymbol{v}(0) \in T(\boldsymbol{x}(0))$ to $\boldsymbol{v}(t) \in T(\boldsymbol{x}(t))$ along the curve $\boldsymbol{x}(t)$ on the manifold $\mathcal{M}$ in a parallel way, i.e. $\frac{\mathrm{D}}{\mathrm{d}t}\boldsymbol{v}(t) := \mathbf{P}_{T(\boldsymbol{x}(t))}\frac{\mathrm{d}}{\mathrm{d}t}\boldsymbol{v}(t) \equiv \boldsymbol{0}$.

For the unit sphere S^{d-1}, the CHiSD can be simplified as,

$$\begin{cases} \dot{\boldsymbol{x}} = -\left(\mathbf{I} - \boldsymbol{x}\boldsymbol{x}^\top - \sum_{i=1}^k 2\boldsymbol{v}_i\boldsymbol{v}_i^\top\right)\nabla E(\boldsymbol{x}), \\ \dot{\boldsymbol{v}}_i = -\left(\mathbf{I} - \boldsymbol{x}\boldsymbol{x}^\top - \boldsymbol{v}_i\boldsymbol{v}_i^\top - \sum_{j=1}^{i-1} 2\boldsymbol{v}_j\boldsymbol{v}_j^\top\right)\nabla^2 E(\boldsymbol{x})\boldsymbol{v}_i - \boldsymbol{x}\boldsymbol{v}_i^\top\nabla E(\boldsymbol{x}), \quad i = 1, \cdots, k. \end{cases} \tag{28}$$

The sphere constraint is a particularly special case because the Hessian of the constraint $\nabla^2 c(\boldsymbol{x})$ is a scalar multiple of the identity. Therefore, $(\mathbf{I} - \boldsymbol{x}\boldsymbol{x}^\top)\nabla^2 E(\boldsymbol{x})$ and the Riemannian Hessian $\mathrm{Hess}(\boldsymbol{x})$ share the same eigenvectors in $T(\boldsymbol{x})$ with a translation in eigenvalues, and the Riemannian Hessian in (17) can be replaced by $\nabla^2 E(\boldsymbol{x})$, which is not valid for general constraints.

2.4 Linear stability

Although CHiSD (25) with the initial condition (26) satisfies the property (21) and (22), a small perturbation may easily deviate $\boldsymbol{x}$ away from the manifold $\mathcal{M}$. To achieve better stability, a modified term with $\mu > 0$ is attached to the dynamics of $\boldsymbol{x}$ to reinforce the equality constraints (1), leading to a modified CHiSD for a k-saddle,

$$\begin{cases} \dot{\boldsymbol{x}} = -\left(\mathbf{I} - \sum_{i=1}^k 2\boldsymbol{v}_i\boldsymbol{v}_i^\top\right)\mathrm{grad}\,E(\boldsymbol{x}) - \mu\sum_{l=1}^m c_l(\boldsymbol{x})\nabla c_l(\boldsymbol{x}), \\ \dot{\boldsymbol{v}}_i = -\left(\mathbf{I} - \boldsymbol{v}_i\boldsymbol{v}_i^\top - \sum_{j=1}^{i-1} 2\boldsymbol{v}_j\boldsymbol{v}_j^\top\right)\widetilde{\mathrm{Hess}}\,E(\boldsymbol{x})[\boldsymbol{v}_i] \\ \qquad - \mathbf{A}(\boldsymbol{x})\left(\mathbf{A}(\boldsymbol{x})^\top\mathbf{A}(\boldsymbol{x})\right)^{-1}\left(\nabla^2\boldsymbol{c}(\boldsymbol{x})\dot{\boldsymbol{x}}\right)^\top\boldsymbol{v}_i, \quad i = 1, \cdots, k. \end{cases} \tag{29}$$

The following theorem shows the linear stability of k-saddles in (29).

Theorem 1 *Assume that $E(\boldsymbol{x})$ is a $\mathcal{C}^3$ functional, $\boldsymbol{x}^* \in \mathcal{M}$, $\{\boldsymbol{v}_i^*\}_{i=1}^k \subset T(\boldsymbol{x}^*)$ satisfies $\|\boldsymbol{v}_i^*\| = 1$, $\mu > 0$, and $\mathrm{Hess}\,E(\boldsymbol{x}^*)$ is nondegenerate, whose eigenvalues are $\lambda_1^* < \cdots < \lambda_k^* \leqslant \lambda_{k+1}^* \leqslant \cdots \leqslant \lambda_{d-m}^*$. Then $(\boldsymbol{x}^*, \boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_k^*)$ is a linearly stable steady state of (29), if and only if $\boldsymbol{x}^*$ is a k-saddle and $\mathrm{Hess}\,E(\boldsymbol{x}^*)[\boldsymbol{v}_i^*] = \lambda_i^*\boldsymbol{v}_i^*$ for $i = 1, \cdots, k$.*

Proof We consider the Jacobian operator of the dynamics (29),

$$\mathbb{J} = \frac{\partial(\dot{\boldsymbol{x}}, \dot{\boldsymbol{v}}_1, \dot{\boldsymbol{v}}_2, \cdots, \dot{\boldsymbol{v}}_k)}{\partial(\boldsymbol{x}, \boldsymbol{v}_1, \boldsymbol{v}_2, \cdots, \boldsymbol{v}_k)} = \begin{pmatrix} \mathbb{J}_{\boldsymbol{x}} & \mathbb{J}_{\boldsymbol{x}1} & \mathbb{J}_{\boldsymbol{x}2} & \cdots & \mathbb{J}_{\boldsymbol{x}k} \\ \star & \mathbb{J}_1 & \mathbb{O} & \cdots & \mathbb{O} \\ \star & \star & \mathbb{J}_2 & \cdots & \mathbb{O} \\ \vdots & \vdots & \vdots & & \vdots \\ \star & \star & \star & \cdots & \mathbb{J}_k \end{pmatrix}, \tag{30}$$

whose blocks have the following expressions,

$$\mathbb{J}_{\boldsymbol{x}} = \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{x}} = -\left(\mathbf{I} - \sum_{i=1}^{k} 2\boldsymbol{v}_i \boldsymbol{v}_i^\top \right) \left(\mathbf{H}(\boldsymbol{x}) - \mathbf{A}(\boldsymbol{x}) \left(\mathbf{A}(\boldsymbol{x})^\top \mathbf{A}(\boldsymbol{x}) \right)^{-1} \right.$$

$$\left. \left(\nabla^2 \boldsymbol{c}(\boldsymbol{x}) \operatorname{grad} E(\boldsymbol{x}) \right)^\top \right) - \mu \sum_{l=1}^{m} \left(c_l(\boldsymbol{x}) \nabla^2 c_l(\boldsymbol{x}) + \nabla c_l(\boldsymbol{x}) \nabla c_l(\boldsymbol{x})^\top \right),$$

$$\mathbb{J}_{\boldsymbol{x}i} = \frac{\partial \dot{\boldsymbol{x}}}{\partial \boldsymbol{v}_i} = 2\boldsymbol{v}_i^\top \operatorname{grad} E(\boldsymbol{x})\mathbf{I} + 2\boldsymbol{v}_i \operatorname{grad} E(\boldsymbol{x})^\top,$$

$$\mathbb{J}_i = \frac{\partial \dot{\boldsymbol{v}}_i}{\partial \boldsymbol{v}_i} = -\left(\mathbf{I} - \sum_{j=1}^{i} 2\boldsymbol{v}_j \boldsymbol{v}_j^\top \right) \widetilde{\operatorname{Hess}}\, E(\boldsymbol{x}) + \left\langle \boldsymbol{v}_i, \widetilde{\operatorname{Hess}}\, E(\boldsymbol{x})[\boldsymbol{v}_i] \right\rangle \mathbf{I}$$

$$- \mathbf{A}(\boldsymbol{x}) \left(\mathbf{A}(\boldsymbol{x})^\top \mathbf{A}(\boldsymbol{x}) \right)^{-1} \left(\nabla^2 \boldsymbol{c}(\boldsymbol{x})\dot{\boldsymbol{x}} \right)^\top.$$

Here, $\mathbf{H}(\boldsymbol{x})$ is an asymmetric extension of the Riemannian Hessian (5) defined as

$$\mathbf{H}(\boldsymbol{x}) = \mathbf{P}_{T(\boldsymbol{x})} \left(\nabla^2 E(\boldsymbol{x}) - \nabla^2 \boldsymbol{c}(\boldsymbol{x}) \left(\mathbf{A}(\boldsymbol{x})^\top \mathbf{A}(\boldsymbol{x}) \right)^{-1} \mathbf{A}(\boldsymbol{x})^\top \nabla E(\boldsymbol{x}) \right). \tag{31}$$

In the following, $\mathbb{J}(\boldsymbol{x}^*, \boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_k^*)$ is denoted as $\mathbb{J}^*$ and the blocks of $\mathbb{J}^*$ are denoted as $\mathbb{J}^*$ with corresponding subscripts.

"$\Leftarrow$": Supposing that $\boldsymbol{x}^*$ is a k-saddle and $\operatorname{Hess} E(\boldsymbol{x}^*)[\boldsymbol{v}_i^*] = \lambda_i^* \boldsymbol{v}_i^*$ for $i = 1, \cdots, k$, we have $\operatorname{grad} E(\boldsymbol{x}^*) = \boldsymbol{0}$ and $\lambda_k^* < 0 < \lambda_{k+1}^*$. With simple calculations, $(\boldsymbol{x}^*, \boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_k^*)$ is a steady state of (29). Note that $\mathbb{J}_{\boldsymbol{x}i}$ is null if $\operatorname{grad} E(\boldsymbol{x}) = \boldsymbol{0}$, so the Jacobian (30) is block lower triangular, whose eigenvalues are determined by the diagonal blocks,

$$\mathbb{J}_{\boldsymbol{x}}^* = -\left(\mathbf{I} - \sum_{i=1}^{k} 2\boldsymbol{v}_i^* \boldsymbol{v}_i^{*\top} \right) \mathbf{H}(\boldsymbol{x}^*) - \mu \sum_{l=1}^{m} \nabla c_l(\boldsymbol{x}^*) \nabla c_l(\boldsymbol{x}^*)^\top,$$

$$\mathbb{J}_i^* = -\left(\mathbf{I} - \sum_{j=1}^{i} 2\boldsymbol{v}_j^* \boldsymbol{v}_j^{*\top} \right) \widetilde{\operatorname{Hess}}\, E(\boldsymbol{x}^*) + \lambda_i^* \mathbf{I}.$$

For $\operatorname{Hess} E(\boldsymbol{x}^*)$, we denote $\boldsymbol{v}_i^*$ $(i = k+1, \cdots, d-m)$ as the eigenvector of its eigenvalue λ_i^* so that $\{\boldsymbol{v}_i^*\}_{i=1}^{d-m}$ is an orthonormal basis of $T(\boldsymbol{x}^*)$. The equations

$$\mathbb{J}_{\boldsymbol{x}}^* \boldsymbol{v}_j^* = -\left(\mathbf{I} - \sum_{i=1}^{k} 2\boldsymbol{v}_i^* \boldsymbol{v}_i^{*\top} \right) \operatorname{Hess} E(\boldsymbol{x}^*)[\boldsymbol{v}_j^*] = \begin{cases} \lambda_j^* \boldsymbol{v}_j^*, & 1 \leqslant j \leqslant k, \\ -\lambda_j^* \boldsymbol{v}_j^*, & k < j \leqslant d-m, \end{cases} \tag{32}$$

indicate that $\mathbb{J}_x^*$ has eigenvalues of $\lambda_1^*, \cdots, \lambda_k^*, -\lambda_{k+1}^*, \cdots, -\lambda_{d-m}^*$, and the equations

$$\nabla c_j(\boldsymbol{x}^*)^\top \mathbb{J}_x^* = -\mu \|\nabla c_j(\boldsymbol{x}^*)\|^2 \nabla c_j(\boldsymbol{x}^*)^\top, \quad 1 \leqslant j \leqslant m. \tag{33}$$

indicate that $\mathbb{J}_x^*$ has m eigenvalues of $-\mu\|\nabla c_1(\boldsymbol{x}^*)\|^2, \cdots, -\mu\|\nabla c_m(\boldsymbol{x}^*)\|^2$, which are different from the previous ones due to μ. Therefore, all the eigenvalues of $\mathbb{J}_x^*$ are negative. For $\mathbb{J}_i^*$, we have

$$\mathbb{J}_i^* \boldsymbol{v}_j^* = \begin{cases} (\lambda_i^* + \lambda_j^*)\boldsymbol{v}_j^*, & 1 \leqslant j \leqslant i, \\ (\lambda_i^* - \lambda_j^*)\boldsymbol{v}_j^*, & i < j \leqslant d-m, \end{cases} \tag{34}$$

indicating eigenvalues of $\lambda_i^* + \lambda_1^*, \cdots, \lambda_i^* + \lambda_i^*, \lambda_i^* - \lambda_{i+1}^*, \cdots, \lambda_i^* - \lambda_{d-m}^*$, and

$$\mathbb{J}_i^* \nabla c_j(\boldsymbol{x}^*) = \lambda_i^* \nabla c_j(\boldsymbol{x}^*), \quad 1 \leqslant j \leqslant m, \tag{35}$$

indicating an eigenvalues λ_i^* with multiplicity m. Therefore, all the eigenvalues of $\mathbb{J}_i^*$ are negative, and the steady state $(\boldsymbol{x}^*, \boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_k^*)$ is linearly stable.

"$\Rightarrow$": Supposing that $(\boldsymbol{x}^*, \boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_k^*)$ is a linearly stable steady state, we have $\dot{\boldsymbol{x}} = \boldsymbol{0}$ and $\dot{\boldsymbol{v}}_i = \boldsymbol{0}$, indicating

$$\left(\mathbf{I} - \sum_{j=1}^{i-1} 2\boldsymbol{v}_j^*\boldsymbol{v}_j^{*\top} \right) \widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*)[\boldsymbol{v}_i^*] = \mu_i^* \boldsymbol{v}_i^*, \tag{36}$$

where $\mu_i^* = \left\langle \widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*)[\boldsymbol{v}_i^*], \boldsymbol{v}_i^* \right\rangle$. We now show by induction that for $i = 1, \cdots, k$,

$$\widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*)[\boldsymbol{v}_i^*] = \mu_i^* \boldsymbol{v}_i^* \neq \boldsymbol{0}, \quad \langle \boldsymbol{v}_j^*, \boldsymbol{v}_i^* \rangle = \delta_{ij}, \quad j = 1, \cdots, i-1. \tag{37}$$

The $i = 1$ case is obtained from (36) directly, and $\boldsymbol{v}_1^* \in T(\boldsymbol{x}^*)$ indicates $\mu_1 \neq 0$ due to the nondegeneracy. Assumed that (37) holds for $1 \leqslant i < l$, by taking $i = l$ in (36) we have,

$$\left(\widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*) - \sum_{j=1}^{l-1} 2\mu_j^*\boldsymbol{v}_j^*\boldsymbol{v}_j^{*\top} \right) \boldsymbol{v}_l^* = \mu_l^* \boldsymbol{v}_l^*, \tag{38}$$

from the symmetry of $\widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*)$. Since $\boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_{l-1}^*$ are eigenvectors of $\widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*)$ according to the inductive assumption, $\widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*)$ and $\widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*) - \sum\limits_{j=1}^{l-1} 2\mu_j^*\boldsymbol{v}_j^*\boldsymbol{v}_j^{*\top}$ share the same eigenvectors, so $\boldsymbol{v}_l^* \in T(\boldsymbol{x}^*)$ is also an eigenvector of $\widetilde{\mathrm{Hess}}\, E(\boldsymbol{x}^*)$ with an eigenvalue $\mu_l^* \neq 0$. Furthermore, $\sum\limits_{j=1}^{l-1} \mu_j^*\langle \boldsymbol{v}_l^*, \boldsymbol{v}_j^* \rangle \boldsymbol{v}_j^* = \boldsymbol{0}$ leads to $\langle \boldsymbol{v}_l^*, \boldsymbol{v}_j^* \rangle = 0$, which completes the induction of (37). Consequently, from $\dot{\boldsymbol{x}} = \boldsymbol{0}$ and $\boldsymbol{x}^* \in \mathcal{M}$, we have $\mathrm{grad}\, E(\boldsymbol{x}^*) = \boldsymbol{0}$ and the Jacobian (30) is block lower triangular.

Finally, we show that the index of the stationary point $\boldsymbol{x}^*$ is k, and $\mu_i^* = \lambda_i^*$ for $i = 1, \cdots, k$. Since $\mathrm{Hess}\, E(\boldsymbol{x}^*)$ is a symmetric operator on $T(\boldsymbol{x}^*)$ with some eigenpairs $\{(\mu_i^*, \boldsymbol{v}_i^*)\}_{i=1}^k$, we denote $\boldsymbol{v}_i^*$ $(i = k+1, \cdots, d-m)$ as the eigenvector of its eigenvalue μ_i^* such that $\{\boldsymbol{v}_i^*\}_{i=1}^{d-m}$ is an orthonormal basis of $T(\boldsymbol{x}^*)$. Similarly to (32) and (33), the eigenvalues of $\mathbb{J}_x^*$ are

$$\mu_1^*, \cdots, \mu_k^*; \quad -\mu_{k+1}^*, \cdots, -\mu_{d-m}^*; \quad -\mu\|\nabla c_1(\boldsymbol{x}^*)\|^2, \cdots, -\mu\|\nabla c_m(\boldsymbol{x}^*)\|^2; \tag{39}$$

which have negative real parts from the linear stability. Therefore, $\{\mu_i^*\}_{i=1}^k$ are negative and $\{\mu_i^*\}_{i=k+1}^{d-m}$ are positive, so $\boldsymbol{x}^*$ is a k-saddle. Similarly to (34) and (35), the eigenvalues of $\mathbb{J}_i^*$,

$$\mu_i^* + \mu_1^*, \cdots, \mu_i^* + \mu_i^*; \quad \mu_i^* - \mu_{i+1}^*, \cdots, \mu_i^* - \mu_{d-m}^*; \quad \mu_i^*(\text{with multiplicity } m); \quad (40)$$

have negative real parts as well, which indicates $\mu_i^* < \mu_{i+1}^*$ for $i = 1, \cdots, k$. Comparing the negative eigenvalues of $\operatorname{Hess} E(\boldsymbol{x}^*)$, we have $\mu_i^* = \lambda_i^*$, which completes our proof.

Remark 3 The CHiSD (25) actually searches for k-saddles on the isosurface of $\boldsymbol{c}(\boldsymbol{x})$, while the initial condition (26) makes the dynamics search for saddle points on the manifold $\mathcal{M}$. The additional term in (29) is applied to pull the dynamics towards the manifold in case of perturbation in order to obtain linear stability.

From Remark 1, the dynamics (29) and the dynamics (25) with the initial condition (26) have the same orbits. Therefore, a linearly stable steady state $(\boldsymbol{x}^*, \boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_k^*)$ of (29) is also asymptotically stable for (25) in the following sense. As long as the initial condition (26) is sufficiently close to $(\boldsymbol{x}^*, \boldsymbol{v}_1^*, \cdots, \boldsymbol{v}_k^*)$, the dynamics (25) will converge to this steady state as well.

3 Numerical implementation

3.1 Retractions and vector transport

The numerical scheme of CHiSD (25) is supposed to maintain the manifold property (21) in each iteration step. Since a simple explicit Euler scheme can easily push $\boldsymbol{x}$ away from the manifold, we introduce the tools of retractions and vector transport in manifold optimization to discretize CHiSD (25), and we refer to [1] for more detailed information.

Strictly speaking, a tangent vector $\boldsymbol{\xi_x}$ to the manifold $\mathcal{M}$ at a point $\boldsymbol{x} \in \mathcal{M}$ is a mapping from the set of germs of functions $\mathfrak{F_x}(\mathcal{M})$ to $\mathbb{R}$. Each tangent vector $\boldsymbol{\xi_x}$ can be represented by a curve γ on $\mathcal{M}$ with $\gamma(0) = \boldsymbol{x}$ satisfying $\boldsymbol{\xi_x} f = \frac{\mathrm{d}}{\mathrm{d}t} f(\gamma(t))\big|_{t=0}$ for $\forall f \in \mathfrak{F_x}(\mathcal{M})$. The tangent space at $\boldsymbol{x} \in \mathcal{M}$, denoted by $T_{\boldsymbol{x}}\mathcal{M}$, can now be canonically identified with $T(\boldsymbol{x})$ via the one-to-one correspondence $\boldsymbol{\xi_x} \in T_{\boldsymbol{x}}\mathcal{M} \mapsto \gamma'(0) = \lim_{\tau \to 0} \frac{\gamma(\tau) - \gamma(0)}{\tau} \in T(\boldsymbol{x})$. Therefore, for $\boldsymbol{x} \in \mathcal{M}$, we treat the mapping $\boldsymbol{\xi_x} \in T_{\boldsymbol{x}}\mathcal{M}$ and the tangent vector object $\boldsymbol{\xi} \in T(\boldsymbol{x})$ at $\boldsymbol{x}$ equally throughout the article. The tangent bundle $T\mathcal{M} := \bigcup_{\boldsymbol{x} \in \mathcal{M}} T_{\boldsymbol{x}}\mathcal{M}$ denotes the set of all tangent vectors to $\mathcal{M}$.

The notion of moving $\boldsymbol{x}$ on the manifold in the direction of a tangent vector is generalized by a retraction mapping. A retraction R is a smooth mapping from $T\mathcal{M}$ to $\mathcal{M}$, and the retraction of R to $T_{\boldsymbol{x}}\mathcal{M}$, denoted by $R_{\boldsymbol{x}}$, satisfies $R_{\boldsymbol{x}}(\boldsymbol{0_x}) = \boldsymbol{x}$ and $\frac{\mathrm{d}}{\mathrm{d}t} R_{\boldsymbol{x}}(t\boldsymbol{\eta_x})\big|_{t=0} = \boldsymbol{\eta_x}$ for $\forall \boldsymbol{\eta_x} \in T_{\boldsymbol{x}}\mathcal{M}$, where $\boldsymbol{0_x}$ is the zero element in $T_{\boldsymbol{x}}\mathcal{M}$. As a natural retraction on the Riemannian manifold in the geometric sense, the exponential mapping at $\boldsymbol{x} \in \mathcal{M}$, denoted by $\operatorname{Exp}_{\boldsymbol{x}}$, maps $\boldsymbol{\eta} \in T_{\boldsymbol{x}}\mathcal{M}$ to $\gamma(1; \boldsymbol{x}, \boldsymbol{\eta}) \in \mathcal{M}$, where $\gamma(t; \boldsymbol{x}, \boldsymbol{\eta})$ is the unique geodesic such that $\gamma(0) = \boldsymbol{x}$ and $\dot{\gamma}(0) = \boldsymbol{\eta}$. For the unit sphere S^{d-1}, the exponential mapping has a computable form,

$$\operatorname{Exp}_{\boldsymbol{x}} \boldsymbol{\eta} = (\cos \|\boldsymbol{\eta}\|)\boldsymbol{x} + \frac{\sin \|\boldsymbol{\eta}\|}{\|\boldsymbol{\eta}\|} \boldsymbol{\eta}, \quad (41)$$

while the exponential mapping of a general Riemannian manifold poses significant numerical challenges to calculating cheaply. Therefore, other computable retractions are often applied as approximations of the exponential mapping. It should be pointed out that for the same manifold, we can have different choices of retractions, which might lead to different computational results. For the unit sphere S^{d-1}, another clear form of the retraction is

$$R_{\boldsymbol{x}}(\boldsymbol{\eta}) = \frac{\boldsymbol{x}+\boldsymbol{\eta}}{\|\boldsymbol{x}+\boldsymbol{\eta}\|} = \mathrm{Exp}_{\boldsymbol{x}}\left(\frac{\arctan\|\boldsymbol{\eta}\|}{\|\boldsymbol{\eta}\|}\boldsymbol{\eta}\right). \tag{42}$$

Next we consider enforcing the directions $\boldsymbol{v}_i$ on the tangent space. The notion of vector transport $\mathcal{T}$ specifies how to transport a tangent vector $\boldsymbol{\xi}_{\boldsymbol{x}} \in T_{\boldsymbol{x}}\mathcal{M}$ as $\boldsymbol{x}$ moves on the Riemannian manifold. A vector transport on a manifold $\mathcal{M}$ is a smooth mapping from the Whitney sum $T\mathcal{M} \oplus T\mathcal{M} := \{(\boldsymbol{\eta}_{\boldsymbol{x}}, \boldsymbol{\xi}_{\boldsymbol{x}}) : \boldsymbol{\eta}_{\boldsymbol{x}}, \boldsymbol{\xi}_{\boldsymbol{x}} \in T_{\boldsymbol{x}}\mathcal{M}, \boldsymbol{x} \in \mathcal{M}\}$ to $T\mathcal{M}$ satisfying following properties:

1. $\mathcal{T}_{\boldsymbol{\eta}_{\boldsymbol{x}}}\boldsymbol{\xi}_{\boldsymbol{x}} \in T_{R_{\boldsymbol{x}}(\boldsymbol{\eta}_{\boldsymbol{x}})}\mathcal{M}$, where R is the retraction associated with $\mathcal{T}$;
2. $\mathcal{T}_{\boldsymbol{0}_{\boldsymbol{x}}}\boldsymbol{\xi}_{\boldsymbol{x}} = \boldsymbol{\xi}_{\boldsymbol{x}}$ for $\forall \boldsymbol{\xi}_{\boldsymbol{x}} \in T_{\boldsymbol{x}}\mathcal{M}$;
3. $\mathcal{T}_{\boldsymbol{\eta}_{\boldsymbol{x}}}(a\boldsymbol{\xi}_{\boldsymbol{x}} + b\boldsymbol{\zeta}_{\boldsymbol{x}}) = a\mathcal{T}_{\boldsymbol{\eta}_{\boldsymbol{x}}}\boldsymbol{\xi}_{\boldsymbol{x}} + b\mathcal{T}_{\boldsymbol{\eta}_{\boldsymbol{x}}}\boldsymbol{\zeta}_{\boldsymbol{x}}$ for $\forall \boldsymbol{\xi}_{\boldsymbol{x}}, \boldsymbol{\zeta}_{\boldsymbol{x}} \in T_{\boldsymbol{x}}\mathcal{M}$ and $\forall a, b \in \mathbb{R}$.

The parallel translation (27) is a natural vector transport along a curve $\boldsymbol{x}(t)$ on the manifold, where the parallel translation of orthonormal vectors remains orthonormal. For the unit sphere S^{d-1}, as $\boldsymbol{x} \in S^{d-1}$ moves to $\mathrm{Exp}_{\boldsymbol{x}}\boldsymbol{\eta}$ along the geodesics, the parallel translation of a vector $\boldsymbol{\xi} \in T_{\boldsymbol{x}}\mathcal{M}$ has a closed form,

$$\mathcal{T}_{\boldsymbol{\eta}}\boldsymbol{\xi} = \boldsymbol{\xi} + \frac{\cos\|\boldsymbol{\eta}\|-1}{\|\boldsymbol{\eta}\|^2}\langle\boldsymbol{\eta},\boldsymbol{\xi}\rangle\boldsymbol{\eta} - \frac{\sin\|\boldsymbol{\eta}\|}{\|\boldsymbol{\eta}\|}\langle\boldsymbol{\eta},\boldsymbol{\xi}\rangle\boldsymbol{x} \in T_{\mathrm{Exp}_{\boldsymbol{x}}\boldsymbol{\eta}}\mathcal{M}, \tag{43}$$

where the exponential mapping (41) is the associated retraction. Parallel translation is not the only way to achieve vector transport, and is often difficult to calculate numerically for general Riemannian manifolds as well. Alternatively, there is considerable flexibility in how to choose the vector transport, and two typical approaches to generating computationally tractable vector transport according to the retraction operator R are

$$\mathcal{T}_{\boldsymbol{\eta}}\boldsymbol{\xi} = \left.\frac{\mathrm{d}}{\mathrm{d}t}R_{\boldsymbol{x}}(\boldsymbol{\eta}+t\boldsymbol{\xi})\right|_{t=0}, \text{ and } \mathcal{T}_{\boldsymbol{\eta}}\boldsymbol{\xi} = \mathbf{P}_{T(R_{\boldsymbol{x}}(\eta))}\boldsymbol{\xi}. \tag{44}$$

Note that vector transport generally may not maintain the orthonormality of vectors. For the unit sphere S^{d-1} with the retraction (42), the two approaches yield

$$\mathcal{T}_{\boldsymbol{\eta}}\boldsymbol{\xi} = \frac{1}{\|\boldsymbol{x}+\boldsymbol{\eta}\|}\boldsymbol{\xi} - \frac{\langle\boldsymbol{\xi},\boldsymbol{x}+\boldsymbol{\eta}\rangle}{\|\boldsymbol{x}+\boldsymbol{\eta}\|^3}(\boldsymbol{x}+\boldsymbol{\eta}), \text{ and } \mathcal{T}_{\boldsymbol{\eta}}\boldsymbol{\xi} = \boldsymbol{\xi} - \frac{\langle\boldsymbol{\xi},\boldsymbol{x}+\boldsymbol{\eta}\rangle}{\|\boldsymbol{x}+\boldsymbol{\eta}\|^2}(\boldsymbol{x}+\boldsymbol{\eta}). \tag{45}$$

3.2 Numerical algorithms of CHiSD

With the help of vector transport $\mathcal{T}$ and the retraction R associated with $\mathcal{T}$, we are able to implement CHiSD (25) numerically with the initial condition (12). Specifically, at the $(n+1)$-th iteration step, we aim to calculate $\boldsymbol{x}^{(n+1)}$ and $\boldsymbol{v}_i^{(n+1)}$ using $\boldsymbol{x}^{(n)}$ and $\boldsymbol{v}_i^{(n)}$ in the previous step. The retraction $R_{\boldsymbol{x}}$ provides a practical way

to pull the points from $\boldsymbol{x} + T(\boldsymbol{x})$ back onto the manifold $\mathcal{M}$, so we can implement an explicit scheme with a retraction as

$$\boldsymbol{x}^{(n+1)} = R_{\boldsymbol{x}^{(n)}}\left(\alpha^{(n)}\boldsymbol{g}^{(n)}\right), \quad \boldsymbol{g}^{(n)} = -\left(\mathbf{I} - \sum_{i=1}^{k} 2\boldsymbol{v}_i^{(n)}\boldsymbol{v}_i^{(n)\top}\right)\operatorname{grad} E\left(\boldsymbol{x}^{(n)}\right),$$
(46)

to calculate $\boldsymbol{x}^{(n+1)}$. Note that $\boldsymbol{v}_1^{(n)}, \cdots, \boldsymbol{v}_k^{(n)}$ are orthonormal vectors in $T(\boldsymbol{x}^{(n)})$ which approximate the corresponding eigenvectors of Hess $E(\boldsymbol{x}^{(n)})$, and $\alpha^{(n)}$ is the step size to be determined. Therefore, $\boldsymbol{g}^{(n)}$ lies in the tangent space $T(\boldsymbol{x}^{(n)})$, and this iteration scheme ensures that $\boldsymbol{x}^{(n)}$ always satisfies the equality constraints (1).

The dynamics of $\boldsymbol{v}_i$ consists of two terms: the first term solves the eigenvalue problem (18) using gradient flow, while the second one corresponds to the parallel translation as discussed in Remark 2. Accordingly, we deal with the two terms successively. We first transport the previous vectors $\{\boldsymbol{v}_i^{(n)}\}$ to $\{\check{\boldsymbol{v}}_i^{(n)}\}$ in the tangent space $T(\boldsymbol{x}^{(n+1)})$ using the vector transport $\mathcal{T}_{\alpha^{(n)}\boldsymbol{g}^{(n)}}$ at $\boldsymbol{x}^{(n)}$, and then solve the eigenvalue problem (18). Although an exact solution to the eigenvalue problem works well, we recommend to find a rough solution to reduce the computational costs, and $\boldsymbol{v}_i$ will converge as $\boldsymbol{x}$ approaches a k-saddle. Since $\boldsymbol{x}^{(n+1)}$ is close to $\boldsymbol{x}^{(n)}$, the transported vector $\check{\boldsymbol{v}}_i^{(n)}$ provides a good initial guess of the corresponding eigenvector, so we simply apply one-step gradient flow for each eigenvector. Practically, the Riemannian Hessian Hess $E(\boldsymbol{x})$ is often expensive to calculate, so a dimer approximation $\boldsymbol{u}_i^{(n)}$ is applied to approximate $\widetilde{\operatorname{Hess}} E(\boldsymbol{x}^{(n+1)})[\check{\boldsymbol{v}}_i^{(n)}]$ with a small $l > 0$ [21]. Finally, $\left\{\boldsymbol{v}_i^{(n+1)}\right\}$ is calculated from $\left\{\tilde{\boldsymbol{v}}_i^{(n+1)}\right\}$ using a normalized orthogonalization function orth$(\cdot)$ which can be realized by a Gram–Schmidt procedure. Consequently, $\{\boldsymbol{v}_i^{(n+1)}\}$ can be calculated with the following procedure in each iteration step,

$$\begin{cases} \check{\boldsymbol{v}}_i^{(n)} = \mathcal{T}_{\alpha^{(n)}\boldsymbol{g}^{(n)}}\boldsymbol{v}_i^{(n)}, \\[2mm] \boldsymbol{u}_i^{(n)} = \mathbf{P}_{T(\boldsymbol{x}^{(n+1)})}\dfrac{\operatorname{grad} E(\boldsymbol{x}^{(n+1)} + l\check{\boldsymbol{v}}_i^{(n)}) - \operatorname{grad} E(\boldsymbol{x}^{(n+1)} - l\check{\boldsymbol{v}}_i^{(n)})}{2l}, \\[3mm] \boldsymbol{d}_i^{(n)} = -\boldsymbol{u}_i^{(n)} + \left\langle\boldsymbol{u}_i^{(n)}, \check{\boldsymbol{v}}_i^{(n)}\right\rangle\check{\boldsymbol{v}}_i^{(n)} + \sum_{j=1}^{i-1} 2\left\langle\boldsymbol{u}_i^{(n)}, \check{\boldsymbol{v}}_j^{(n)}\right\rangle\check{\boldsymbol{v}}_j^{(n)}, \\[3mm] \tilde{\boldsymbol{v}}_i^{(n+1)} = \check{\boldsymbol{v}}_i^{(n)} + \beta^{(n)}\boldsymbol{d}_i^{(n)}, \\[2mm] \left[\boldsymbol{v}_1^{(n+1)}, \cdots, \boldsymbol{v}_k^{(n+1)}\right] = \operatorname{orth}\left(\left[\tilde{\boldsymbol{v}}_1^{(n+1)}, \cdots, \tilde{\boldsymbol{v}}_k^{(n+1)}\right]\right). \end{cases}$$
(47)

The iteration (46)–(47) is terminated if $\|\operatorname{grad} E(\boldsymbol{x}^{(n)})\|$ is smaller than the tolerance.

Remark 4 The above iteration (47) for $\boldsymbol{v}_i$ can be regarded as a one-step application of the power method for $\mathbf{I} - \beta^{(n)}\widetilde{\operatorname{Hess}} E(\boldsymbol{x}^{(n+1)})$ on the tangent space $T(\boldsymbol{x}^{(n+1)})$.

3.3 Construction of the solution landscape

In this subsection, we introduce a systematic numerical procedure to construct the solution landscape of an energy functional subject to equality constraints.

This procedure consists of a downward search algorithm and an upward search algorithm, which is a generalization of the pathway map to constrained cases [47].

First, we present a toy model to illustrate our basic idea. Consider an energy functional constrained on S^2,

$$E(x_1, x_2, x_3) = (x_1^2 - 1)^2 + x_2^2 + 2x_3^2,$$
$$\text{s.t.} \quad c(x_1, x_2, x_3) = x_1^2 + x_2^2 + x_3^2 - 1 = 0, \tag{48}$$

whose energy landscape is shown in Fig. 1(a). Two minima $C_1(1, 0, 0)$ and $C_2(-1, 0, 0)$ are connected by two 1-saddles $B_1(0, 1, 0)$ and $B_2(0, -1, 0)$. We aim to find these minima and 1-saddles down from an index-2 critical point $A(0, 0, 1)$, and construct the solution landscape as shown in Fig. 1(b), which depicts how lower-index saddle points are connected to the higher-index ones. The arrows in Fig. 1(b) are realized using downward search algorithms based on the CHiSD.
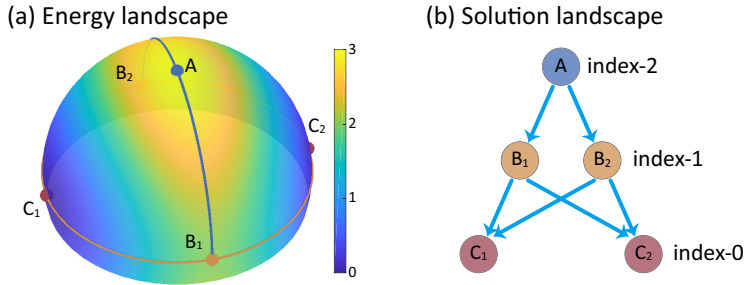
Fig. 1 Illustration of (a) an energy landscape of the function (48) and (b) the solution landscape starting from the maximum, A, down to minima C_1 and C_2. Two saddle points B_1 and B_2 are connected to A. The indices labeled in (b) are those according to the Morse definition.

Starting from a parent state (high-index saddle point), the downward search is the core procedure to search for the stationary points with lower indices that are connected to this parent state. Given a k-saddle $\hat{\boldsymbol{x}} \in \mathcal{M}$, we let $\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_k \in T(\hat{\boldsymbol{x}})$ denote the k orthonormal eigenvectors of Hess $E(\hat{\boldsymbol{x}})$ with negative eigenvalues $\hat{\lambda}_1 \leqslant \cdots \leqslant \hat{\lambda}_k < 0$. First we slightly perturb $\hat{\boldsymbol{x}}$ along an unstable direction $\hat{\boldsymbol{v}}_j$ chosen from the unstable directions $\{\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_k\}$ with $0 < \varepsilon \ll 1$. Then an m-CHiSD ($m < k$) is started from the point $R_{\hat{\boldsymbol{x}}}(\pm \varepsilon \hat{\boldsymbol{v}}_j)$ and the m initial directions $\{\boldsymbol{v}_i(0)\}_{i=1}^m$ are chosen from the unstable directions excluding $\hat{\boldsymbol{v}}_j$. A typical choice of the initial condition is $(R_{\hat{\boldsymbol{x}}}(\pm \varepsilon \hat{\boldsymbol{v}}_{m+1}), \hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_m)$. Once a new m-saddle is found, we implement another downward search from this newly-found saddle recursively until minima are found. The downward search algorithm is presented in detail as Algorithm 1. Note that each element $(\boldsymbol{x}, m, \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\})$ in the queue $\mathcal{A}$ represents to find m-saddles from the k-saddle $\boldsymbol{x}$.

With the downward search, we can systematically search for multiple stationary points from a given parent state in a controlled procedure. However, if the parent state is unknown, or in the case that multiple parent states exist, we need a numerical procedure to find a parent state. Therefore, an upward search algorithm is required to find a higher-index saddle from a low-index stationary point (usually a minimum). Fortunately, the CHiSD also embeds a mechanism to search

---

Algorithm 1 Downward search

Input: A $\hat{k}$-saddle $\hat{\boldsymbol{x}}$, $\varepsilon > 0$.
1: Calculate the $\hat{k}$ eigenvectors $\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_{\hat{k}} \in T(\hat{\boldsymbol{x}})$ of Hess $E(\hat{\boldsymbol{x}})$;
2: Set the queue $\mathcal{A} = \{(\hat{\boldsymbol{x}}, \hat{k} - 1, \{\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_{\hat{k}}\})\}$, the solution set $\mathcal{S} = \{\hat{\boldsymbol{x}}\}$, and the relation
    set $\mathcal{R} = \varnothing$;
3: **while** $\mathcal{A}$ is not empty **do**
4:     Pop $(\boldsymbol{x}, m, \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\})$ from $\mathcal{A}$;
5:     Push $(\boldsymbol{x}, m - 1, \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_k\})$ into $\mathcal{A}$ if $m \geqslant 1$;
6:     **for** $j = 1 : k$ **do**
7:         Determine the initial directions: $\{\boldsymbol{v}_i : i = 1, \cdots, m + 1, i \neq \min(j, m + 1)\}$;
8:         **if** m-CHiSD from $R_{\boldsymbol{x}} (\pm \varepsilon \boldsymbol{v}_j)$ converges to $(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{v}}_1, \cdots, \widetilde{\boldsymbol{v}}_m)$ **then**
9:             $\mathcal{R} \leftarrow \mathcal{R} \cup \{(\boldsymbol{x}, \widetilde{\boldsymbol{x}})\}$;
10:            **if** $\widetilde{\boldsymbol{x}} \notin \mathcal{S}$ **then**
11:                $\mathcal{S} \leftarrow \mathcal{S} \cup \{\widetilde{\boldsymbol{x}}\}$;
12:                Push $(\widetilde{\boldsymbol{x}}, m - 1, \{\widetilde{\boldsymbol{v}}_1, \cdots, \widetilde{\boldsymbol{v}}_m\})$ into $\mathcal{A}$ if $m \geqslant 1$;
Output: The solution set $\mathcal{S}$ and the relation set $\mathcal{R}$.

---

upward. Given a k-saddle $\hat{\boldsymbol{x}} \in \mathcal{M}$, more eigenvectors of Hess $E(\hat{\boldsymbol{x}})$ with smallest eigenvalues are calculated as $\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_K \in T(\boldsymbol{x})$, where K $(K > k)$ is the highest index of the saddle point to search for. Different from the downward search, we choose a direction $\hat{\boldsymbol{v}}_j$ from the stable directions $\{\hat{\boldsymbol{v}}_{k+1}, \cdots, \hat{\boldsymbol{v}}_K\}$, and slightly perturb $\hat{\boldsymbol{x}}$ along $\hat{\boldsymbol{v}}_j$. An m-CHiSD $(m > k)$ procedure is then started from $R_{\hat{\boldsymbol{x}}} (\pm \varepsilon \hat{\boldsymbol{v}}_j)$, while the m initial directions $\{\boldsymbol{v}_i(0)\}_{i=1}^m$ need to include $\hat{\boldsymbol{v}}_j$. Then this procedure is implemented to newly-found saddles recursively until K-saddles are found or no higher-index saddles can be found. A typical choice of the initial condition is $(R_{\hat{\boldsymbol{x}}} (\pm \varepsilon \hat{\boldsymbol{v}}_m), \hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_m)$. Besides the difference in initial conditions, the upward search only aims to find parent states of the solution landscape, while the downward search presents the relations between stationary points and gives a complete picture of the solution landscape by exhaustively searching for multiple stationary points. Algorithm 2 presents the upward search algorithm in detail with the typical choice.

Algorithm 2 Upward search

Input: A $\hat{k}$-saddle $\hat{\boldsymbol{x}}$, $\varepsilon > 0$, the highest index K.
1: Calculate the $\hat{k}$ eigenvectors $\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_{\hat{K}} \in T(\hat{\boldsymbol{x}})$ of Hess $E(\hat{\boldsymbol{x}})$;
2: Set the stack $\mathcal{A} = \{(\hat{\boldsymbol{x}}, \hat{k} + 1, \{\hat{\boldsymbol{v}}_1, \cdots, \hat{\boldsymbol{v}}_K\})\}$ and the solution set $\mathcal{S} = \{\hat{\boldsymbol{x}}\}$,
3: **while** $\mathcal{A}$ is not empty **do**
4:     Pop $(\boldsymbol{x}, m, \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K\})$ from $\mathcal{A}$;
5:     Push $(\boldsymbol{x}, m + 1, \{\boldsymbol{v}_1, \cdots, \boldsymbol{v}_K\})$ into $\mathcal{A}$ if $m < K$;
6:     **if** m-CHiSD from $(R_{\boldsymbol{x}}(\pm \varepsilon \boldsymbol{v}_m), \boldsymbol{v}_1, \cdots, \boldsymbol{v}_m)$ converges to $(\widetilde{\boldsymbol{x}}, \widetilde{\boldsymbol{v}}_1, \cdots, \widetilde{\boldsymbol{v}}_m)$ **then**
7:         **if** $\widetilde{\boldsymbol{x}} \notin \mathcal{S}$ **then**
8:             $\mathcal{S} \leftarrow \mathcal{S} \cup \{\widetilde{\boldsymbol{x}}\}$ and calculate more eigenvectors $\widetilde{\boldsymbol{v}}_{m+1}, \cdots, \widetilde{\boldsymbol{v}}_{\widetilde{K}}$ of Hess $E(\widetilde{\boldsymbol{x}})$;
9:             Push $(\widetilde{\boldsymbol{x}}, m + 1, \{\widetilde{\boldsymbol{v}}_1, \cdots, \widetilde{\boldsymbol{v}}_K\})$ into $\mathcal{A}$ if $m < K$;
Output: The solution set $\mathcal{S}$.

The computation cost of downward and upward search mainly depends on the CHiSD method, since each process can implement the CHiSD search individually, the algorithm can be naturally parallelized to speed it up. In practice, the solution landscape is achieved by a combination of downward search and upward search,

and we can navigate up and down systematically on the energy landscape to construct a complete solution landscape.

4 Numerical examples

4.1 Thomson problem

The Thomson problem considers the minimal-energy configuration of N classical charged particles confined to a sphere which interact with each other via a Coulomb potential $f(r) = r^{-1}$, and was originally proposed as a representation of the atomic structure [42]. The Thomson problem has attached much attention as a special case of the 7th problem in Steven Smale's eighteen problems for the 21st century [41]. Many numerical attempts have been made to find local and global minima [2, 18], transition states [33,51] and high-index saddle points [34] to fully understand the energy landscape of the Thomson problem.

In the Thomson problem, the coordinate of the i-th particle $\boldsymbol{x}_i = (x_i, y_i, z_i) \in \mathbb{R}^3$ is constrained on the unit sphere S^2, and the energy function of a configuration $(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N) \in \mathbb{R}^{3 \times N}$ is,

$$E(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N) = \sum_{i<j} f\left(\left\|\boldsymbol{x}_i - \boldsymbol{x}_j\right\|\right). \tag{49}$$

Because of the rotation symmetry, we specify the first particle to be the north pole and the second particle to lie on the yz-plane [51]. In other words, the state variables are constrained on a manifold,

$$\mathcal{M} = \left\{(\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N) \in \mathbb{R}^{3 \times N} : \boldsymbol{x}_1 = (0, 0, 1), x_2 = 0, \|\boldsymbol{x}_i\|_2 = 1\right\}, \tag{50}$$

so that the Riemannian Hessian at a stationary point has no zero eigenvalues generally. For $\boldsymbol{x} = (\boldsymbol{x}_1, \cdots, \boldsymbol{x}_N) \in \mathcal{M}$ and $\boldsymbol{\eta} = (\mathbf{0}, (0, \eta_2), \boldsymbol{\eta}_3, \cdots, \boldsymbol{\eta}_N) \in T_{\boldsymbol{x}}\mathcal{M}$ where $\eta_2 \in T_{(y_2, z_2)}S^1$ and $\boldsymbol{\eta}_i \in T_{\boldsymbol{x}_i}S^2$ for $i \geqslant 3$, the retraction operator is,

$$R_{\boldsymbol{x}}\boldsymbol{\eta} = \left(\boldsymbol{x}_1, (x_2, R^{S^1}_{(y_2, z_2)}\eta_2), R^{S^2}_{\boldsymbol{x}_3}\boldsymbol{\eta}_3, \cdots, R^{S^2}_{\boldsymbol{x}_N}\boldsymbol{\eta}_N\right), \tag{51}$$

where $R^{S^{d-1}}$ is the retraction operator on the unit sphere S^{d-1} chosen as the exponential mapping (41). For $\boldsymbol{\xi} = (\mathbf{0}, (0, \xi_2), \boldsymbol{\xi}_3, \cdots, \boldsymbol{\xi}_N) \in T_{\boldsymbol{x}}\mathcal{M}$, the vector transport from $\boldsymbol{x}$ to $R_{\boldsymbol{x}}\boldsymbol{\eta}$ is

$$\mathcal{T}_{\boldsymbol{\eta}}\boldsymbol{\xi} = \left(\mathbf{0}, (0, \mathcal{T}^{S^1}_{\eta_2}\xi_2), \mathcal{T}^{S^2}_{\boldsymbol{\eta}_3}\boldsymbol{\xi}_3, \cdots, \mathcal{T}^{S^2}_{\boldsymbol{\eta}_N}\boldsymbol{\xi}_N\right), \tag{52}$$

where $\mathcal{T}^{S^{d-1}}$ is the vector transport for the unit sphere S^{d-1} with the form of (43). Note that if N is large, the constraint amount m will be about as large as N. Nevertheless, the computational costs of the retraction operator (51) and the vector transport (52) will be $\mathcal{O}(N)$ because of the product structure of (50).

We aim to present multiple stationary points of the Thomson problem in order to offer a full description for the energy landscape without random initial guesses. A trivial stationary point of the Thomson problem is a planar polygon configuration (PP) that N particles are evenly spaced on a great circle of S^2, and is an
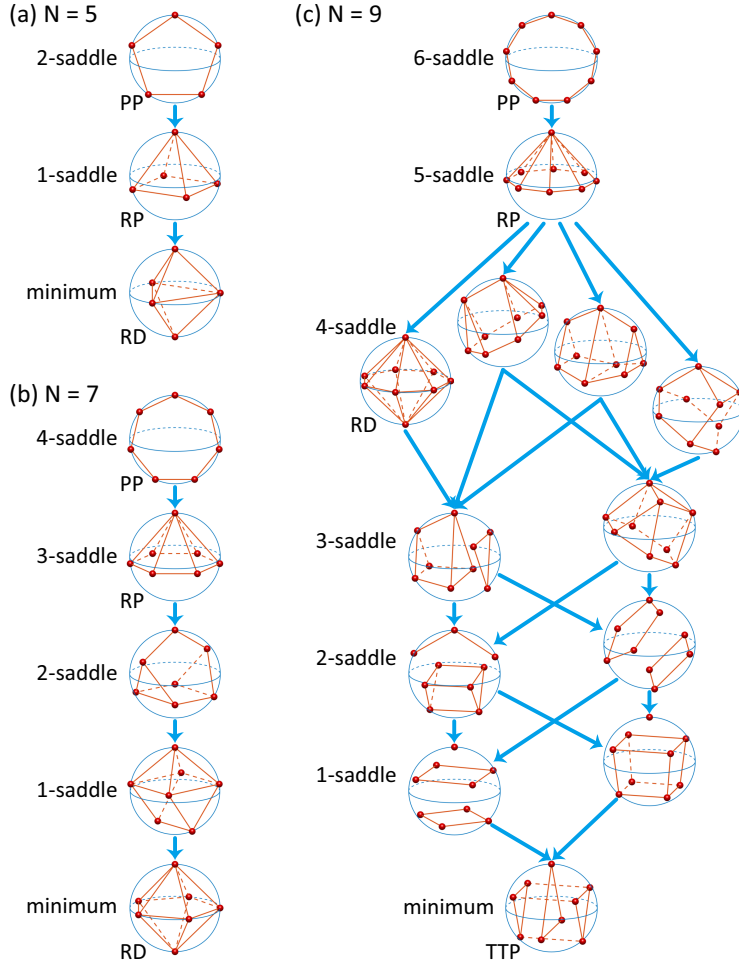
Fig. 2 The solution landscape of the Thomson problem with particle numbers (a) $N = 5$. (b) 7, and (c) 9. The configurations of stationary points are presented in the solution landscape. Each red ball represents a particle on the sphere, and the red line segments are drawn to show the relative positions. The height of each configuration approximately corresponds to its relative energy. The index of each stationary point is labelled on the left side, and some important stationary points are further labelled with their configurations. Each arrow from a higher-index stationary point to a lower-index stationary point corresponds to a CHiSD pathway by the downward search.

$(N-3)$-saddle for $N \geqslant 3$. This trivial stationary point presents a natural parent state for downward search which is implemented with fixed step sizes $\alpha^{(n)} = 10^{-4}$ and $\beta^{(n)} = 10^{-3}$. In this low-dimensional example, we adopt the canonical inner product and the Hessian is evaluated exactly. The computational results for some small particle numbers $N = 5, 7$ and 9 are presented in Fig. 2, where the stationary points with the same configuration are only shown once. From the trivial $(N-3)$-saddle, an $(N-4)$-saddle of a regular pyramid configuration (RP) is first located during the downward search, where $(N-1)$ particles evenly spaced

on a latitude ring. The regular dipyramid configuration (RD) with two antipodal particles and $(N-2)$ other particles evenly spaced on the equator is the minimal-energy configuration for $N = 5, 7$, and a 4-saddle configuration for $N = 9$, while the minimal-energy configuration for $N = 9$ is a triaugmented triangular prism (TTP).

4.2 Bose–Einstein condensation

The experimental realization of BEC in vapors of magnetically-trapped alkali atoms that occupy a single quantum state has attracted great interest in the atomic physics community [3,13], The properties of BEC at an ultra-low temperature are well described by the macroscopic wave function $\psi(\boldsymbol{x}, t)$, whose evolution is governed by the nonlinear Schrödinger equation (Gross–Pitaevskii equation),

$$\mathrm{i} \frac{\partial}{\partial t} \psi(\boldsymbol{x}, t) = \left(-\frac{1}{2} \nabla^2 + V(\boldsymbol{x}) + \beta |\psi(\boldsymbol{x}, t)|^2 \right) \psi(\boldsymbol{x}, t), \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{53}$$

where $V(\boldsymbol{x})$ is a real-valued trapping potential on $\mathbb{R}^d$ and β is the dimensionless interaction coefficient [4]. Two important invariants of (53) are the normalization of the wave function,

$$\int_{\mathbb{R}^d} |\psi(\boldsymbol{x}, t)|^2 \mathrm{d}\boldsymbol{x} = 1, \tag{54}$$

and the energy per particle,

$$E(\psi(\cdot, t)) = \int_{\mathbb{R}^d} \left[\frac{1}{2} |\nabla \psi(\boldsymbol{x}, t)|^2 + V(\boldsymbol{x}) |\psi(\boldsymbol{x}, t)|^2 + \frac{\beta}{2} |\psi(\boldsymbol{x}, t)|^4 \right] \mathrm{d}\boldsymbol{x}. \tag{55}$$

To find a stationary solution to (53), substituting $\psi(\boldsymbol{x}, t) = \mathrm{e}^{-\mathrm{i}\mu t} \phi(\boldsymbol{x})$ into (53) gives

$$\mu \, \phi(\boldsymbol{x}) = \left(-\frac{1}{2} \nabla^2 + V(\boldsymbol{x}) + \beta |\phi(\boldsymbol{x})|^2 \right) \phi(\boldsymbol{x}), \quad \boldsymbol{x} \in \mathbb{R}^d, \tag{56}$$

where μ is the chemical potential and $\phi(\boldsymbol{x})$ satisfies a unit complex sphere constraint,

$$\phi \in \mathcal{M} = \left\{ \varphi \in L^2(\mathbb{R}^d, \mathbb{C}) : E(\varphi) < \infty, \int_{\mathbb{R}^d} |\varphi(\boldsymbol{x})|^2 \mathrm{d}\boldsymbol{x} = 1 \right\}. \tag{57}$$

The eigenfunction to the nonlinear elliptic eigenvalue problem (56) is the stationary point of the energy

$$E(\phi) = \int_{\mathbb{R}^d} \left[\frac{1}{2} |\nabla \phi(\boldsymbol{x})|^2 + V(\boldsymbol{x}) |\phi(\boldsymbol{x})|^2 + \frac{\beta}{2} |\phi(\boldsymbol{x})|^4 \right] \mathrm{d}\boldsymbol{x}. \tag{58}$$

with the sphere constraint (57), and the chemical potential μ can be calculated as,

$$\mu = \int_{\mathbb{R}^d} \left[\frac{1}{2} |\nabla \phi(\boldsymbol{x})|^2 + V(\boldsymbol{x}) |\phi(\boldsymbol{x})|^2 + \beta |\phi(\boldsymbol{x})|^4 \right] \mathrm{d}\boldsymbol{x} = E(\phi) + \int_{\mathbb{R}^d} \frac{\beta}{2} |\phi(\boldsymbol{x})|^4 \mathrm{d}\boldsymbol{x}. \tag{59}$$

The ground state of BEC is defined as the eigenfunction of (56) subject to the sphere constraint (57) with the lowest energy (58). Any eigenfunction of (56)

subject to the sphere constraint (57) with a higher energy is usually called excited states in the physics literature.

We consider a two-dimensional BEC system with a repulsive interaction parameter $\beta = 300$, where the condensates are tightly confined in the other dimension. The potential $V(\boldsymbol{x}) : \mathbb{R}^2 \to \mathbb{R}$ is simply chosen as a radial-symmetric harmonic oscillator $V(\boldsymbol{x}) = \frac{1}{2}|\boldsymbol{x}|^2$. The existence, uniqueness up to a phase factor, and smoothness of the ground state have been well studied in previous researches [4,29], and there are many efficient numerical methods for computing the ground state [4,5,46]. Furthermore, there have been some attempts to find several excited states of BEC as well. By choosing proper initial guesses such as odd functions, a few methods for finding ground states can also be applied to compute some excited states with certain symmetry of BEC [5,46]. Some excited states of BEC have also been obtained with Newton's methods [23,24] and deflated continuation algorithms [11,8,10], while how to systematically compute the excited states of BEC remains a huge challenge.

The energy (58) implicates some invariance, which leads to zero eigenvalues of Hessians at stationary points. For a stationary solution $\phi \in \mathcal{M}$, both $\mathrm{e}^{\mathrm{i}\vartheta}\phi(\boldsymbol{x})$ and $\phi(R_\vartheta \boldsymbol{x})$ are also stationary solutions with the same index and energy for $\forall \vartheta \in \mathbb{R}$, where $R_\vartheta = \begin{pmatrix} \cos\vartheta & -\sin\vartheta \\ \sin\vartheta & \cos\vartheta \end{pmatrix}$ is the rotation around the origin. This invariance accounts for two zero eigenvalues for Hessians of general stationary solutions. As a special case, the ground state and central vortex states can be expressed as $\phi(\boldsymbol{x}) = \mathrm{e}^{\mathrm{i}m\theta}\varphi_m(r)$ using the polar coordinate (r, θ) and the winding number $m \in \mathbb{Z}$ of the central vortex, so the Hessians at these stationary solutions have only one zero eigenvalue [4,6].

In the numerical computation, the wave function ϕ is truncated into a bounded domain $D = [-M, M]^2$ with homogeneous Dirichlet boundary conditions,

$$E(\phi) = \int_D \left[\frac{1}{2}|\nabla\phi(\boldsymbol{x})|^2 + V(\boldsymbol{x})|\phi(\boldsymbol{x})|^2 + \frac{\beta}{2}|\phi(\boldsymbol{x})|^4 \right] \mathrm{d}\boldsymbol{x},$$
$$\phi(\boldsymbol{x}) = 0, \quad \boldsymbol{x} \in \partial D, \tag{60}$$

because the stationary states decay to zero exponentially fast in the far field from the effect of the trapping potential [4]. We discretize the wave function $\phi(\boldsymbol{x}) \in L^2(D)$ with $M = 8$ using finite difference methods with $N = 128$ nodes along each dimension. It should be noted that the complex-valued function space $L^2(D)$ is a real Hilbert space with the real inner product,

$$\langle \phi, \psi \rangle = \int_D \frac{\phi(\boldsymbol{x})\bar{\psi}(\boldsymbol{x}) + \bar{\phi}(\boldsymbol{x})\psi(\boldsymbol{x})}{2} \mathrm{d}\boldsymbol{x}, \tag{61}$$

and the gradient of (55), which is often referred to as the Wirtinger derivative, is

$$\nabla E(\phi) = 2\frac{\delta E}{\delta\bar{\phi}} = -\nabla^2\phi + 2V(\boldsymbol{x})\phi + 2\beta\phi|\phi|^2. \tag{62}$$

The retraction operator and the corresponding vector transport for this unit sphere are chosen as (41) and (43).

From the ground state, we implement the upward search with fixed step sizes $\alpha^{(n)} = 10^{-6}$ and $\beta^{(n)} = 10^{-3}$ to search for excited states, which is slightly different from Algorithm 2 because of the zero eigenvalues of Hessians at stationary
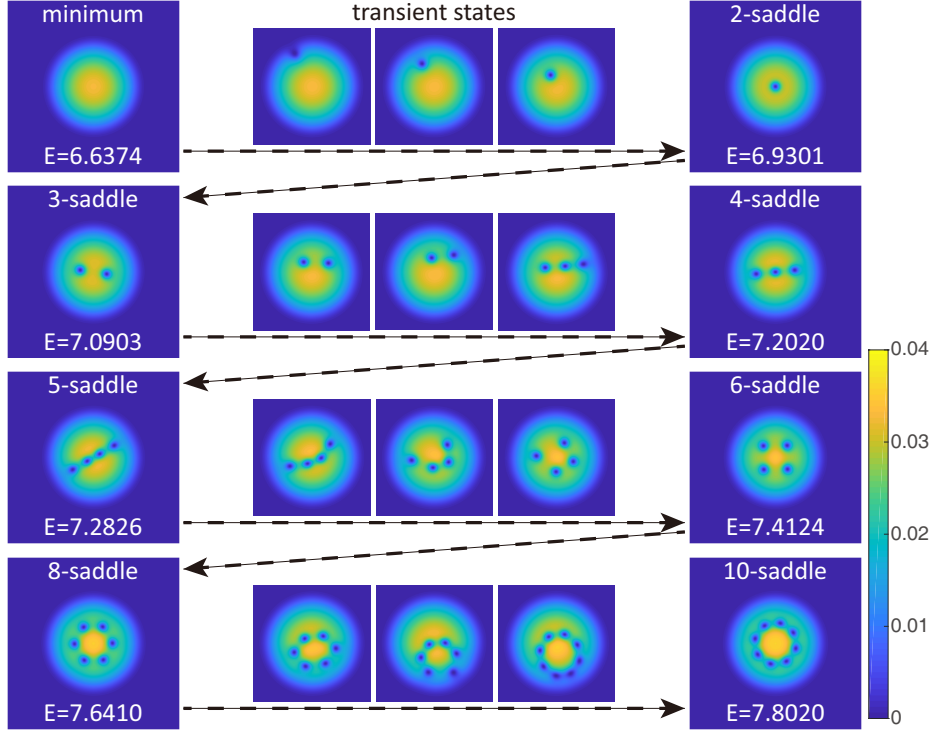
Fig. 3 An upward pathway sequence from the ground state to a 10-saddle of BEC. Each dashed arrow represents an upward search to an excited state, and three transient states on some upward dynamical pathways are presented. We show the probability density $|\phi|^2$ for each state, and present the energy of each stationary solution.

solutions. For each stationary solution ϕ_i, we let $\eta_{i,1}, \cdots, \eta_{i,m} \in T(\phi_i)$ denote m orthonormal eigenvectors of Hess $E(\hat{x})$ corresponding to the smallest m eigenvalues $\lambda_{i,1} \leqslant \cdots \leqslant \lambda_{i,m}$. Since the Hessian at the ground state ϕ_0 has only one zero eigenvalue $\lambda_{0,1} = 0$, the 1-CHiSD from the ground state cannot climb out of the basin. Alternatively, we treat this degenerate minimum ϕ_0 as a 1-saddle by regarding the zero eigenvector $\eta_{i,1}$ as an unstable direction, and a 2-CHiSD from $\left(R_{\phi_0}(\varepsilon\eta_{0,2}), \eta_{0,1}, \eta_{0,2} \right)$ finally converges to a 2-saddle ϕ_2. This 2-CHiSD upward search involves a quantized vortex of winding number $+1$ moving from the domain edge to the center, and the 2-saddle ϕ_2 is a central vortex state with only one zero eigenvalue as shown in Fig. 3. Here, the jump of Morse index comes from the degeneracy of this BEC system, that is, zero eigenvalues of the Riemannian Hessians at the critical points. Similarly, we implement upward search from ϕ_2 with a 4-CHiSD from $\left(R_{\phi_2}(\varepsilon\eta_{2,4}), \eta_{2,1}, \cdots, \eta_{2,4} \right)$ and find a degenerate 3-saddle ϕ_3. The 3-saddle ϕ_3 has two vortices with winding numbers $+1$ and -1 respectively, and its Hessian has two zero eigenvalues $\lambda_{3,4} = \lambda_{3,5} = 0$. Therefore, the upward search from ϕ_3 is implemented with a 6-CHiSD from $\left(R_{\phi_3}(\varepsilon\eta_{3,6}), \eta_{3,1}, \cdots, \eta_{3,6} \right)$ to find an excited state with a higher index, which turns out to be a 4-saddle. This 6-CHiSD upward search involves a new quantized vortex, and the three vortices of the 4-saddle are aligned. This upward search procedure can be repeated to newly-found

excited states, and Fig. 3 shows a sequence of seven upward-search attempts from the ground state eventually to a 10-saddle. It should be noted that the upward search $(k + 3)$-CHiSD from a k-saddle finally converges to a $(k + 1)$-saddle for $k = 3, 4, 5$, and a $(k + 2)$-saddle for $k = 6, 8$.

5 Conclusions and discussions

In this article, we proposed a CHiSD method for searching for index-k saddle points subject to general equality constraints. Applying the Riemannian gradients and Hessians, we derived the dynamical system with a transformed gradient flow as the formulation of CHiSD. The linear stability of CHiSD at the index-k constrained saddle points is proved. In the numerical implementation of CHiSD, the retraction operator and vector transport are introduced to discretize the dynamics. Combined with the CHiSD method, one can construct the solution landscape on a constrained manifold using downward/upward search algorithms.

We presented two numerical examples as applications of the CHiSD method. For the Thomson problem, we constructed the solution landscape from a planar configuration. Our results can be regarded as an improvement of the saddle point amounts and the kinetic transition networks [33, 34]. Although only cases with 5, 7 and 9 particles are considered here, our method can be applied to study more particles in a straightforward way. The other application is to search for the excited states of BEC, which correspond to a nonlinear eigenvalue problem. The upward search algorithm is slightly varied in this case due to the zero eigenvalues of Hessians at stationary points. We calculated the excited states using the upward search algorithm, which can further applied to the 10-saddle, as it is conjectured that the nonlinear eigenvalue problem (56) admits infinitely many eigenfunctions which are linearly independent [6]. Furthermore, in a rotating BEC system, vortices are energetically favored above a critical rotational frequency and this system can admit multiple stable/metastable states [4, 6]. We will systematically explore the solution landscape of rotating BEC and identify probable mechanisms of excitation in a subsequent work.

There are naturally some issues worthy of further investigations in order to improve numerical efficiency. Since the computational costs mainly depend on calculations of the eigenvectors of Hessians, one may consider more efficient methods for computing the eigenvectors. For instance, if the problem is ill-conditioned, the power iteration method (47) may present poor approximations of eigenvectors, leading to a failure of convergence. Thus, other efficient numerical algorithms will be needed to deal with such difficulties. Moreover, the fixed step size is used in current version of the CHiSD method for simplicity, which can be adopted by adaptive step sizes to accelerate the convergent rate. Furthermore, at the end of CHiSD iterations, we can apply the second-order methods to accelerate the convergence. Another important task is to apply the CHiSD method to explore more practical applications, in which the framework of CHiSD will be adapted to the numerical schemes originally designed for the gradient dynamics.

Conflict of interest

The authors declare that they have no conflict of interest.

References

1. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press, Princeton, NJ (2008)
2. Altschuler, E.L., Williams, T.J., Ratner, E.R., Dowla, F., Wooten, F.: Method of constrained global optimization. Phys. Rev. Lett. **72**(17), 2671–2674 (1994). DOI 10.1103/PhysRevLett.72.2671
3. Anderson, M.H., Ensher, J.R., Matthews, M.R., Wieman, C.E., Cornell, E.A.: Observation of Bose–Einstein condensation in a dilute atomic vapor. Science **269**(5221), 198–201 (1995). DOI 10.1126/science.269.5221.198
4. Bao, W., Cai, Y.: Mathematical theory and numerical methods for Bose–Einstein condensation. Kinet. Relat. Models **6**(1), 1–135 (2013). DOI 10.3934/krm.2013.6.1
5. 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). DOI 10.1016/j.jcp.2006.04.019
6. Bao, W., Wang, H., Markowich, P.A.: Ground, symmetric and central vortex states in rotating Bose–Einstein condensates. Comm. Math. Sci. **3**(1), 57–88 (2005). URL `https://projecteuclid.org/euclid.cms/1111095641`
7. Bessarab, P.F., Uzdin, V.M., Jónsson, H.: Method for finding mechanism and activation energy of magnetic transitions, applied to skyrmion and antivortex annihilation. Comput. Phys. Commun. **196**, 335–347 (2015). DOI 10.1016/j.cpc.2015.07.001
8. Boullé, N., Charalampidis, E.G., Farrell, P.E., Kevrekidis, P.G.: Deflation-based identification of nonlinear excitations of the three-dimensional Gross-Pitaevskii equation. Phys. Rev. A **102**, 053307 (2020). DOI 10.1103/PhysRevA.102.053307
9. Cancès, E., Legoll, F., Marinica, M.C., Minoukadeh, K., Willaime, F.: Some improvements of the activation-relaxation technique method for finding transition pathways on potential energy surfaces. J. Chem. Phys. **130**(11), 114711 (2009). DOI 10.1063/1.3088532
10. Charalampidis, E.G., Boullé, N., Farrell, P.E., Kevrekidis, P.G.: Bifurcation analysis of stationary solutions of two-dimensional coupled Gross–Pitaevskii equations using deflated continuation. Commun. Nonlinear Sci. Numer. Simul. **87**, 105255 (2020). DOI 10.1016/j.cnsns.2020.105255
11. Charalampidis, E.G., Kevrekidis, P.G., Farrell, P.E.: Computing stationary solutions of the two-dimensional Gross–Pitaevskii equation with deflated continuation. Commun. Nonlinear Sci. Numer. Simul. **54**, 482–499 (2018). DOI 10.1016/j.cnsns.2017.05.024
12. Cheng, X., Lin, L., E, W., Zhang, P., Shi, A.C.: Nucleation of ordered phases in block copolymers. Phys. Rev. Lett. **104**(14), 148301 (2010). DOI 10.1103/PhysRevLett.104.148301
13. Davis, K.B., Mewes, M.O., Andrews, M.R., van Druten, N.J., Durfee, D.S., Kurn, D.M., Ketterle, W.: Bose–Einstein condensation in a gas of sodium atoms. Phys. Rev. Lett. **75**(22), 3969–3973 (1995). DOI 10.1103/PhysRevLett.75.3969
14. Du, Q., Zhang, L.: A constrained string method and its numerical analysis. Commun. Math. Sci. **7**(4), 1039–1051 (2009). DOI 10.4310/CMS.2009.v7.n4.a11
15. E, W., Ren, W., Vanden-Eijnden, E.: String method for the study of rare events. Phys. Rev. B **66**(5), 052301 (2002). DOI 10.1103/PhysRevB.66.052301
16. E, W., Vanden-Eijnden, E.: Transition-path theory and path-finding algorithms for the study of rare events. Annu. Rev. Phys. Chem. **61**(1), 391–420 (2010). DOI 10.1146/annurev.physchem.040808.090412
17. E, W., Zhou, X.: The gentlest ascent dynamics. Nonlinearity **24**(6), 1831–1842 (2011). DOI 10.1088/0951-7715/24/6/008

18. Erber, T., Hockney, G.M.: Equilibrium configurations of N equal charges on a sphere. J. Phys. A: Math. Gen. **24**(23), L1369–L1377 (1991). DOI 10.1088/0305-4470/24/23/008

19. Frank, F.C.: I. Liquid crystals. On the theory of liquid crystals. Discuss. Faraday Soc. **25**, 19–28 (1958). DOI 10.1039/DF9582500019

20. Han, Y.C., Hu, Y.C., Zhang, P.W., Zhang, L.: Transition pathways between defect patterns in confined nematic liquid crystals. J. Comput. Phys. **396**, 1–11 (2019). DOI 10.1016/j.jcp.2019.06.028

21. Henkelman, G., Jónsson, H.: A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. J. Chem. Phys. **111**(15), 7010–7022 (1999). DOI 10.1063/1.480097

22. Henkelman, G., Jónsson, H.: Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. J. Chem. Phys. **113**(22), 9978–9985 (2000). DOI 10.1063/1.1323224

23. Law, K.J.H., Kevrekidis, P.G., Tuckerman, L.S.: Stable vortex–bright-soliton structures in two-component Bose–Einstein condensates. Phys. Rev. Lett. **105**, 160405 (2010). DOI 10.1103/PhysRevLett.105.160405

24. Law, K.J.H., Neely, T.W., Kevrekidis, P.G., Anderson, B.P., Bradley, A.S., Carretero-González, R.: Dynamic and energetic stabilization of persistent currents in Bose–Einstein condensates. Phys. Rev. A **89**, 053606 (2014). DOI 10.1103/PhysRevA.89.053606

25. Leeson, D.T., Gai, F., Rodriguez, H.M., Gregoret, L.M., Dyer, R.B.: Protein folding and unfolding on a complex energy landscape. Proc. Natl. Acad. Sci. U.S.A. **97**(6), 2527–2532 (2000). DOI doi.org/10.1073/pnas.040580397

26. Li, Y., Zhou, J.: A minimax method for finding multiple critical points and its applications to semilinear PDEs. SIAM J. Sci. Comput. **23**(3), 840–865 (2001). DOI 10.1137/S1064827599365641

27. Li, Z., Ji, B., Zhou, J.: A local minimax method using virtual geometric objects: Part I—for finding saddles. J. Sci. Comput. **78**(1), 202–225 (2019). DOI 10.1007/s10915-018-0774-y

28. Li, Z., Zhou, J.: A local minimax method using virtual geometric objects: Part II—for finding equality constrained saddles. J. Sci. Comput. **78**(1), 226–245 (2019). DOI 10.1007/s10915-018-0775-x

29. Lieb, E.H., Seiringer, R., Yngvason, J.: Bosons in a trap: A rigorous derivation of the Gross-Pitaevskii energy functional. Phys. Rev. A **61**, 043602 (2000). DOI 10.1103/PhysRevA.61.043602

30. Lin, L., Lu, J., Ying, L.: Numerical methods for Kohn–Sham density functional theory. Acta Numerica **28**, 405–539 (2019). DOI 10.1017/S0962492919000047

31. Machado-Charry, E., Béland, L.K., Caliste, D., Genovese, L., Deutsch, T., Mousseau, N., Pochet, P.: Optimized energy landscape exploration using the *ab initio* based activation-relaxation technique. J. Chem. Phys. **135**(3), 034102 (2011). DOI 10.1063/1.3609924

32. Mallamace, F., Corsaro, C., Mallamace, D., Vasi, S., Vasi, C., Baglioni, P., Buldyrev, S.V., Chen, S.H., Stanley, H.E.: Energy landscape in protein folding and unfolding. Proc. Natl. Acad. Sci. U.S.A. **113**(12), 3159–3163 (2016). DOI 10.1073/pnas.1524864113

33. Mehta, D., Chen, J., Chen, D.Z., Kusumaatmaja, H., Wales, D.J.: Kinetic transition networks for the Thomson problem and Smale's seventh problem. Phys. Rev. Lett. **117**(2), 028301 (2016). DOI 10.1103/PhysRevLett.117.028301

34. Mehta, D., Chen, T., Morgan, J.W.R., Wales, D.J.: Exploring the potential energy landscape of the Thomson problem via Newton homotopies. J. Chem. Phys. **142**(19), 194113 (2015). DOI 10.1063/1.4921163

35. Meng, G., Arkus, N., Brenner, M.P., Manoharan, V.N.: The free-energy landscape of clusters of attractive hard spheres. Science **327**(5965), 560–563 (2010). DOI 10.1126/science.1181263

36. Milnor, J.: Morse Theory. Princeton University Press, Princeton, NJ (1963)

37. Müller, G.P., Bessarab, P.F., Vlasov, S.M., Lux, F., Kiselev, N.S., Blügel, S., Uzdin, V.M., Jónsson, H.: Duplication, collapse, and escape of magnetic skyrmions revealed using a systematic saddle point search method. Phys. Rev. Lett. **121**(19), 197202 (2018). DOI 10.1103/PhysRevLett.121.197202

38. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer, New York (2006). DOI 10.1007/978-0-387-40065-5

39. Olsen, R.A., Kroes, G.J., Henkelman, G., Arnaldsson, A., Jónsson, H.: Comparison of methods for finding saddle points without knowledge of the final states. J. Chem. Phys. **121**(20), 9776–9792 (2004). DOI 10.1063/1.1809574

40. Quapp, W., Bofill, J.M.: Locating saddle points of any index on potential energy surfaces by the generalized gentlest ascent dynamics. Theor. Chem. Acc. **133**(8), 1510 (2014). DOI 10.1007/s00214-014-1510-9
41. Smale, S.: Mathematical problems for the next century. Math. Intell. **20**(2), 7–15 (1998). DOI 10.1007/BF03025291
42. Thomson, J.J.: XXIV. On the structure of the atom: an investigation of the stability and periods of oscillation of a number of corpuscles arranged at equal intervals around the circumference of a circle; with application of the results to the theory of atomic structure. London, Edinburgh, Dublin Phil. Mag. J. Sci. **7**(39), 237–265 (1904). DOI 10.1080/14786440409463107
43. Tu, L.W.: An Introduction to Manifolds. Springer, New York (2011). DOI https://doi.org/10.1007/978-1-4419-7400-6
44. Wales, D.J.: Energy Landscapes. Cambridge University Press, Cambridge, England (2003)
45. Wang, W., Zhang, L., Zhang, P.: Modelling and computation of liquid crystals. Acta Numerica **30**, 765–851 (2021). DOI 10.1017/S0962492921000088
46. Wu, X., Wen, Z., Bao, W.: A regularized Newton method for computing ground states of Bose–Einstein condensates. J. Sci. Comput. **73**, 303–329 (2017). DOI 10.1007/s10915-017-0412-0
47. Yin, J., Wang, Y., Chen, J.Z.Y., Zhang, P., Zhang, L.: Construction of a pathway map on a complicated energy landscape. Phys. Rev. Lett. **124**(9), 090601 (2020). DOI 10.1103/PhysRevLett.124.090601
48. Yin, J., Yu, B., Zhang, L.: Searching the solution landscape by generalized high-index saddle dynamics. Sci. China Math. **64**, 1801–1816 (2021). DOI 10.1007/s11425-020-1737-1
49. Yin, J., Zhang, L., Zhang, P.: High-index optimization-based shrinking dimer method for finding high-index saddle points. SIAM J. Sci. Comput. **41**(6), A3576–A3595 (2019). DOI 10.1137/19M1253356
50. Zhang, J., Du, Q.: Constrained shrinking dimer dynamics for saddle point search with constraints. J. Comput. Phys. **231**(14), 4745–4758 (2012). DOI 10.1016/j.jcp.2012.03.006
51. Zhang, J., Du, Q.: Shrinking dimer dynamics and its applications to saddle point search. SIAM J. Numer. Anal. **50**(4), 1899–1921 (2012). DOI 10.1137/110843149
52. Zhang, L., Chen, L.Q., Du, Q.: Mathematical and numerical aspects of a phase-field approach to critical nuclei morphology in solids. J. Sci. Comput. **37**(1), 89–102 (2008). DOI 10.1007/s10915-008-9207-7
53. Zhang, L., Du, Q., Zheng, Z.: Optimization-based shrinking dimer method for finding transition states. SIAM J. Sci. Comput. **38**(1), A528–A544 (2016). DOI 10.1137/140972676
54. Zhang, L., Ren, W., Samanta, A., Du, Q.: Recent developments in computational modelling of nucleation in phase transformations. NPJ Comput. Mater. **2**(1), 1–9 (2016). DOI 10.1038/npjcompumats.2016.3