J Sci Comput (2018) 75:1–25
https://doi.org/10.1007/s10915-017-0520-x

Computing the p-Spectral Radii of Uniform Hypergraphs with Applications

Jingya Chang[1,2] · Weiyang Ding[2] · Liqun Qi[2] · Hong Yan[3]

Received: 14 March 2017 / Revised: 21 July 2017 / Accepted: 25 July 2017 / Published online: 4 August 2017

Abstract The p-spectral radius of a uniform hypergraph covers many important concepts, such as Lagrangian and spectral radius of the hypergraph, and is crucial for solving spectral extremal problems of hypergraphs. In this paper, we establish a spherically constrained maximization model and propose a first-order conjugate gradient algorithm to compute the p-spectral radius of a uniform hypergraph (CSRH). By the semialgebraic nature of the adjacency tensor of a uniform hypergraph, CSRH is globally convergent and obtains the global maximizer with a high probability. When computing the spectral radius of the adjacency tensor of a uniform hypergraph, CSRH outperforms existing approaches. Furthermore, CSRH is competent to calculate the p-spectral radius of a hypergraph with millions of vertices and to approximate the Lagrangian of a hypergraph. Finally, we show that the CSRH method is

J. Chang work was partially supported by the National Natural Science Foundation of China (Grant Nos. 11401539 and 11571178). W. Ding work was partially supported by the Hong Kong Research Grant Council (Grant No. C1007-15G). L. Qi work was partially supported by the Hong Kong Research Grant Council (Grant Nos. PolyU 501913, 15302114, 15300715, 15301716 and C1007-15G). H. Yan work was supported in part by the Hong Kong Research Grants Council (Grant No. C1007-15G).

✉ Liqun Qi
maqilq@polyu.edu.hk

Jingya Chang
jychang@zzu.edu.cn

Weiyang Ding
weiyang.ding@gmail.com

Hong Yan
h.yan@cityu.edu.hk

1 School of Mathematics and Statistics, Zhengzhou University, Zhengzhou 450001, China

2 Department of Applied Mathematics, The Hong Kong Polytechnic University, Hung Hom, Kowloon, Hong Kong

3 Department of Electronic Engineering, City University of Hong Kong, Hung Hom, Kowloon, Hong Kong

capable of ranking real-world data set based on solutions generated by the p-spectral radius model.

Keywords Eigenvalue · Hypergraph · Large scale tensor · Network analysis · Pagerank · p-spectral radius

Mathematics Subject Classification 05C65 · 15A18 · 15A69 · 65F15 · 65K05 · 90C35 · 90C53

1 Introduction

With the emergence of big data in various fields of our social life, it becomes significant and challenging to analyze massive data and extract valuable information from them. Hypergraph, as an extension of graph, provides an efficient way to represent complex relationships among objects in scientific computation, such as chemistry [32,36], computer science [23,29,55], and image processing [5,11,18]. The spectral hypergraph theory has been studied widely [13,28,31,41,52,64,66], which reveals combinatorial and geometric structures of hypergraphs. Moreover, spectral hypergraph approaches are useful tools to address issues in the real world. Spectral hypergraph partitioning and spectral hypergraph clustering have broad applications in network analysis [42,58], image segmentation [17], multi-label classification [60], machine learning [67], and data analysis [2,39]. Hypergraph spectral hashing techniques highly contribute to problems of similarity search and retrieval of social images [40,68].

In this paper, we focus on the computation of p-spectral radii of uniform hypergraphs. The p-spectral radius of a hypergraph was introduced in [31] and linked with extremal hypergraph problems. Extremal graph theory, as a branch of graph theory, is one of the most attractive and best studied area in combinatorics. It was started in 1941 when Turán [62] introduced the Turán graph and Turán theorem. Naturally, the Turán-type questions were extended from graph to hypergraph in [63]. Although the Turán-type problems are adequately studied for graphs, the formulations are much more complex for hypergraphs. In [47], Nikiforov proved the spectral Turán-type inequality which generalized the Turán theorem. In [31], the p-spectral version of Nikiforov's inequality and the p-spectral version of a hypergraph were given, and it was shown that this result can be employed in solving 'degenerate' Turán-type problems. Furthermore, it was proved that the edge extremal problems are asymptotically equivalent to the extremal p-spectral radius problems in [48].

The p-spectral radius of a hypergraph covers not only the number of edges in extremal problems, but also the notions, such as Lagrangian, and the spectral radius of a hypergraph [41]. When $p = 1$, the p-spectral radius of a hypergraph is just its Lagrangian. The Lagrangians of graph and hypergraph were proposed in [43] to prove the Turán's theorem for graphs. The Lagrangians of hypergraphs were used to disprove the conjecture of Erdös [19,21] and to find non-jumping numbers for hypergraphs [22,53,54]. The Turán density of a hypergraph is an asymptotic solution to the (non-degenerate) Turán problem, and the Lagrangian of a hypergraph is associated with problems of determining Turán densities of hypergraphs [6,30,44,59]. When $p = 2$, the p-spectral radius of a uniform hypergraph is the largest Z-eigenvalue [56] of the corresponding adjacency tensor. When p is even and equals the order of the hypergraph, the p-spectral radius becomes the largest H-eigenvalue of the related adjacency tensor. Therefore, the p-spectral radius is connected with the (adjacency) spectral radius of a hypergraph [26,38,41]. Additionally, Kang et al. provided solutions to

several p-spectral radius related extremal problems in [28]. Nikiforov in [49] did a comprehensive study and obtained many theoretical conclusions about p-spectral radius.

Apart from the application to the extremal hypergraph theory, the p-spectral radius model can be employed for quantifying the importance of objects or centrality in networks. Evaluating the significance or popularity of objects is an important problem in data mining. It can be used to determine the importance of web pages [15,35,51], forecast customer behaviours [37], and retrieve images [27], etc. In the p-spectral radius model, entries of the vector associated with the p-spectral radius of a hypergraph are called the p-optimal weighting and represent the significance of its corresponding vertices. The ranking result varies when p changes. We will explain the meanings of different ranking results and show the numerical performance of our algorithm in sorting real world data in Sect. 6.

Several methods for computing tensor eigenvalues are related to the calculation of p-spectral radii of hypergraphs. Algorithms for tensor eigenvalues, such as the shifted symmetric higher-order power method (SS-HOPM) [33], the generalized eigenproblem adaptive power (GEAP) method [34], an extension of Collatz's method (NQZ) [45], and the CEST method, can be employed to compute the p-spectral radius when p equals 2 or when p equals the order of an even-uniform hypergraph. When p is even, the p-spectral radius problem is equivalent to the generalized tensor eigenvalue problem [9,16]. Therefore, methods for computing the generalized tensor eigenvalue, such as the polynomial optimization related algorithm for finding all real eigenvalues of a symmetric tensor developed by Cui et al. [14], and the homotopy approach for all eigenpairs of general real or complex tensors proposed by Chen et al. [10] can be used to compute even p-spectral radius of small scale hypergraphs. However, the problem of computing p-spectral radii of arbitrary hypergraph is still open. This is the main motivation of our paper.

To solve the p-spectral radius problem, we introduce a spherically constrained maximization model, which is equivalent to the original problem. Then we use an effective conjugate gradient method to determine an ascent direction for the constrained optimization model. Next, we employ the Cayley transform to project the ascent direction onto the unit sphere. It is proved that there exists a positive parameter in the curvilinear line search such that the Wolfe conditions hold. Based on these strategies, we propose a numerical method for Computing p-Spectral Radii of Hypergraphs (CSRH) with $p > 1$. When $p = 1$, the CSRH method is able to approximate the 1-spectral radius (Lagrangian) of a hypergraph. Furthermore, we prove that the CSRH algorithm is convergent and the sequence generated by this algorithm converges to the global optimization point with high probability. Numerical experiments show that CSRH is preponderant when compared to existing methods for computing Z-eigenvalues and H-eigenvalues of adjacency tensors. Moreover, CSRH is capable of calculating p-spectral radii of hypergraphs with millions of vertices effectively. In addition, we find that the significance of hypergraph vertices is related to the order of elements of the p-optimal weighting. Therefore, we apply the CSRH method to network analysis by ranking the vertices of the corresponding hypergraph. We show the ranking results of vertices in a small weighted hypergraph as an example. Moreover, we have successfully ranked 10305 authors based on their publication information by establishing a hypergraph model and using CSRH to solve the corresponding p-spectral radius problem. We sort the authors by individual and group entries respectively. The results of our ranking can be reasonably explained and are consistent with existing consequences [46].

The paper is organized as follows. In Sect. 2, we introduce mathematical notions. The computational issues about p-spectral radius are addressed in Sect. 3, where our new method CSRH is explained. In Sect. 4, we analyze the convergent property of CSRH. Our numerical experiments are discussed in Sect. 5. In Sect.6, we show the application of CSRH to network

analysis. The ranking results of a toy example and a large scale real-world problem are presented. Finally, we draw conclusions in Sect. 7.

2 Preliminary

In this section, we introduce useful notions and important results on hypergraphs and tensors. Let $\mathbb{R}^{[r,n]}$ be the rth order n-dimensional real-valued tensor space, i.e.,

$$\mathbb{R}^{[r,n]} \equiv \mathbb{R}^{\overbrace{n \times n \times \cdots \times n}^{r\text{-times}}}.$$

A tensor $\mathcal{T} = (t_{i_1 \cdots i_r}) \in \mathbb{R}^{[r,n]}$ is considered symmetric if $t_{i_1 \cdots i_r}$ is invariant under any permutation of indices [12]. Two operations involving $\mathcal{T}$ and any vector $\mathbf{x} \in \mathbb{R}^n$ are defined as

$$\mathcal{T}\mathbf{x}^r \equiv \sum_{i_1=1}^{n} \cdots \sum_{i_r=1}^{n} t_{i_1 \cdots i_r} x_{i_1} \cdots x_{i_r}$$

and

$$(\mathcal{T}\mathbf{x}^{r-1})_i \equiv \sum_{i_2=1}^{n} \cdots \sum_{i_r=1}^{n} t_{i i_2 \cdots i_r} x_{i_2} \cdots x_{i_r}, \quad \text{for} \quad i = 1, \ldots, n.$$

Here $\mathcal{T}\mathbf{x}^r \in \mathbb{R}$ and $\mathcal{T}\mathbf{x}^{r-1} \in \mathbb{R}^n$ are a scalar and a vector respectively, and $\mathcal{T}\mathbf{x}^r = \mathbf{x}^\top(\mathcal{T}\mathbf{x}^{r-1})$.

If there exists a real number λ and a nonzero real vector $\mathbf{x}$ satisfying

$$\mathcal{T}\mathbf{x}^{r-1} = \lambda \mathbf{x}^{[r-1]}, \tag{2.1}$$

then λ is called an H-eigenvalue of $\mathcal{T}$ with $\mathbf{x}$ being the associated H-eigenvector [56,57]. Additionally, $\mathbf{x}^{[r-1]} \in \mathbb{R}^n$ is a vector, of which the ith element is x_i^{r-1}. When a real vector $\mathbf{x}$ and a real number λ satisfy the following equations

$$\begin{cases} \mathcal{T}\mathbf{x}^{r-1} = \lambda \mathbf{x} \\ \mathbf{x}^\top \mathbf{x} = 1, \end{cases}$$

λ is called a Z-eigenvalue of $\mathcal{T}$ and $\mathbf{x}$ is the corresponding Z-eigenvector [56].

Definition 2.1 (Hypergraph) A hypergraph is defined as $G = (V, E)$, where $V = \{1, 2, \ldots, n\}$ is the vertex set and $E = \{e_1, e_2, \ldots, e_m\} \subseteq 2^V$ (the powerset of V) is the edge set. We call G an r-uniform hypergraph when $|e_p| = r \geq 2$ for $p = 1, \ldots, m$ and $e_i \neq e_j$ in case $i \neq j$.

If each edge of a hypergraph is linked with a positive number $s(e)$, then this hyperpragh is called a weighted hypergraph and $s(e)$ is the weight associated with the edge e. An ordinary hypergraph can be regarded as a weighted hypergraph with the weight of each edge being 1.

In the rest of this paper, an r-uniform hypergraph is abbreviated as an r-graph for convenience and hence hypergraph G refers to an r-graph. The degree of a vertex $i \in V$ is given by $d(i) = \text{sum}\{s(e) : i \in e, e \in E\}$. The *weight polynomial* of G [61] is defined as

$$w(G, \mathbf{x}) = \sum_{e=\{i_1, \ldots, i_r\} \in E} s(e)\, x_{i_1} \cdots x_{i_r}, \tag{2.2}$$

in which $\mathbf{x}$ is a vector in $\mathbb{R}^n$, $e = \{i_1, \ldots, i_r\}$ is an edge of G and $s(e)$ is the weight of e.

Definition 2.2 (p-spectral radius [28,31,49]) When $p \geq 1$, the p-spectral radius of G, denoted by $\lambda^{(p)}(G)$, is defined as

$$\lambda^{(p)}(G) = r! \max_{\|\mathbf{x}\|_p=1} w(G, \mathbf{x}) \tag{2.3}$$

where $\|\mathbf{x}\|_p = (\sum_{k=1}^{n} |\mathbf{x}_k|)^{\frac{1}{p}}$, and we call any vector $\mathbf{x}$ solving (2.3) a p-optimal weighting of G [7].

When $p = 1$, the p-spectral radius of an ordinary hypergraph G coincides with its Lagrangian $\lambda_L(G)$ [20,61], which is defined as

$$\lambda_L(G) = \begin{cases} \max\ w(G, \mathbf{x}) \\ \text{s.t.}\ \ \sum_{i=1}^{r} x_i = 1, \\ \qquad x_i \geq 0, \quad \text{for} \quad i = 1, \ldots, r. \end{cases} \tag{2.4}$$

The vector $\mathbf{x}$ related to the Lagrangian of G is named the optimal legal weighting [7,61].

Definition 2.3 (Adjacency tensor) The adjacency tensor $\mathcal{A}$ of a weighted r-graph G is defined as an rth order symmetric tensor with its elements being

$$a_{i_1 \cdots i_r} = \begin{cases} \dfrac{s(e)}{(r-1)!} & \text{if } \{i_1, \ldots, i_r\} \in E, \\ 0 & \text{otherwise.} \end{cases}$$

We can observe from (2.3) that the 2-spectral radius is exactly the product of $(r-1)!$ and the largest Z-eigenvalue of the adjacency tensor $\mathcal{A}$. When r is even, the r-spectral radius is $(r-1)!$ times the largest H-eigenvalue of $\mathcal{A}$ [56].

To the best of our knowledge, there is no formula or algorithm designed for computing general p-spectral radius of a hypergraph. However, research on p-spectral radius of hypergraphs with certain structures or for certain p values has made some progress.

Theorem 2.1 ([49]) *Let r-graph G be a β-star with m edges .*

a. If $p > r-1$, then $\lambda^{(p)}(G) = r! r^{-\frac{r}{p}} m^{(1-\frac{r-1}{p})}$.
b. If $p < r-1$, then $\lambda^{(p)}(G) = r! r^{-\frac{r}{p}}$.
c. If $p = r-1$, then $\lambda^{(p)}(G) = (r-1)! r^{-\frac{1}{r-1}}$.

Proposition 2.1 ([7]) *If G is a complete r-graph with n vertices, then the Lagrangian of G is*

$$\lambda_L(G) = \binom{n}{r} \frac{1}{n^r}. \tag{2.5}$$

A multiset is an extension of the ordinary set, such that the objects or elements in the multiset are repeatable. If the edge set E of a hypergraph G is a set of multisets, then G is called a multi-hypergraph [52]. Naturally, the p-spectral radius problem can be extended from hypergraph to muli-hypergraph. The algorithm and theoretical analysis in the following part of this paper are also applicable to p-spectral radius problems of multi-hypergraphs. In the rest of this paper, the symbol $\|\cdot\|$ refers to ℓ_2 norm.

3 Computation of the p-Spectral Radius of a Hypergraph

We transform the p-spectral radius in (2.3) into a spherically constraint optimization problem and propose an iterative algorithm to solve it.

3.1 Spherically Constraint Form for $\lambda^{(p)}(G)$

The p-spectral radius of G in (2.3) can be reformulated as

$$\lambda^{(p)}(G) = \max_{\|\mathbf{x}\|_p=1} (r-1)!\mathcal{A}\mathbf{x}^r \tag{3.1}$$

where $\mathcal{A}$ is the adjacency tensor of G. In an unconstrained format, the maximization problem (3.1) is equivalent to

$$\lambda^{(p)}(G) = \max_{\mathbf{x}\neq 0}(r-1)!\frac{\mathcal{A}\mathbf{x}^r}{\|\mathbf{x}\|_p^r}. \tag{3.2}$$

In order to restrict the search region and keep the vector $\mathbf{x}$ away from zero, we add a spherically constraint on $\lambda^{(p)}(G)$ in (3.2). Due to the zero-order homogeneous property of $\mathcal{A}\mathbf{x}^r/\|\mathbf{x}\|_p^r$, we can obtain $\lambda^{(p)}(G)$ by solving the following problem

$$\begin{cases} \max f(\mathbf{x}) = (r-1)!\frac{\mathcal{A}\mathbf{x}^r}{\|\mathbf{x}\|_p^r} \\ \text{s.t. } \|\mathbf{x}\|_2 = 1. \end{cases} \tag{3.3}$$

When $p > 1$, the objective function $f(\mathbf{x})$ is differentiable for any nonzero $\mathbf{x}$ and the gradient of $f(\mathbf{x})$ is

$$\nabla f(\mathbf{x}) = \frac{r!}{\|\mathbf{x}\|_p^r}\left(\mathcal{A}\mathbf{x}^{r-1} - \mathcal{A}\mathbf{x}^r\|\mathbf{x}\|_p^{-p}\mathbf{x}^{\langle p-1\rangle}\right), \tag{3.4}$$

where $\mathbf{x}^{\langle p-1\rangle}$ represents a vector whose ith element is $(\mathbf{x}^{\langle p-1\rangle})_i = |x_i|^{p-1}\text{sgn}(x_i)$. Since $f(\mathbf{x})$ is zero-order homogeneous, we have

$$\mathbf{x}^\top \nabla f(\mathbf{x}) = 0 \tag{3.5}$$

for any $0 \neq \mathbf{x} \in \mathbb{R}^n$.

Based on the spherically constrained form in (3.3), we have the following proposition, which provides a way to approximate the p-spectral radius of a hypergraph when it cannot be computed directly.

Proposition 3.1 *Let p_ϑ be a sequence such that*

$$\lim_{\vartheta\to\infty} p_\vartheta = p_*, \tag{3.6}$$

where each $p_\vartheta > 1$. Then

$$\lim_{\vartheta\to\infty} \lambda^{(p_\vartheta)}(G) = \lambda^{(p_*)}(G). \tag{3.7}$$

Proof We restrict the domain of $\mathbf{x}$ on a unit sphere, which is denoted as $\mathbb{S}^{n-1} \equiv \{\mathbf{x} \in \mathbb{R}^n : \mathbf{x}^\top\mathbf{x} = 1\}$. Renaming the function in (3.3) as

$$\hat{f}(\mathbf{x}, p) = (r-1)!\frac{\mathcal{A}\mathbf{x}^r}{\|\mathbf{x}\|_p^r} \qquad (\mathbf{x}, p) \in \mathbb{S}^{n-1} \times (1, +\infty),$$

we have

$$\lambda^{(p)}(G) = \max_{\mathbf{x}\in\mathbb{S}^{n-1}} \hat{f}(\mathbf{x}, p).$$

Here $\hat{f}(\mathbf{x}, p)$ is continuous. Let $\{\mathbf{x}_\vartheta^*\}$ be an infinite sequence on the compact space $\mathbb{S}^{n-1}$, such that

$$\hat{f}(\mathbf{x}_\vartheta^*, p_\vartheta) = \lambda^{(p_\vartheta)}(G). \tag{3.8}$$

If there are more than one point satisfying Eq. (3.8), we randomly choose one of them to be $\mathbf{x}_\vartheta^*$. Suppose $\{\mathbf{x}_\vartheta^*\}$ is a convergent sequence without loss of generality. Since the sequence is bounded, there exists a point $\mathbf{x}_0^* \in \mathbb{S}^{n-1}$ satisfying

$$\lim_{\vartheta \to \infty} \mathbf{x}_\vartheta^* = \mathbf{x}_0^*. \tag{3.9}$$

For any $\tilde{\mathbf{x}} \in \mathbb{S}^{n-1}$, we have

$$\hat{f}(\tilde{\mathbf{x}}, p_\vartheta) \leq \hat{f}(\mathbf{x}_\vartheta^*, p_\vartheta) \tag{3.10}$$

from (3.8), which indicates that

$$\lim_{\vartheta \to \infty} \hat{f}(\tilde{\mathbf{x}}, p_\vartheta) \leq \lim_{\vartheta \to \infty} \hat{f}(\mathbf{x}_\vartheta^*, p_\vartheta).$$

Then we obtain

$$\hat{f}(\tilde{\mathbf{x}}, p_*) \leq \hat{f}(\mathbf{x}_0^*, p_*) \tag{3.11}$$

based on (3.6) and (3.9). Therefore, we have $\hat{f}(\mathbf{x}_0^*, p_*) = \max_{\mathbf{x} \in \mathbb{S}^{n-1}} \hat{f}(\mathbf{x}, p_*) = \lambda^{(p_*)}(G)$. Since

$$\hat{f}(\mathbf{x}_0^*, p_*) = \lim_{\vartheta \to \infty} \hat{f}(\mathbf{x}_\vartheta^*, p_\vartheta) = \lim_{\vartheta \to \infty} \lambda^{(p_\vartheta)}(G),$$

conclusion (3.7) is then obtained. □

3.2 The CSRH Algorithm

We employ an iterative algorithm to solve (3.3).

Suppose that the current iterate is a unit vector $\mathbf{x}_k$. Our task is to find a new iterate $\mathbf{x}_{k+1}$, which satisfies the following two conditions:

1. $\mathbf{x}_{k+1}$ is on the unit sphere;
2. $\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ is an ascent direction, i.e.,

$$\mathbf{d}_k^\top \nabla f(\mathbf{x}_k) > 0. \tag{3.12}$$

In Fig. 1, the current iterate $\mathbf{x}_k$ is on the unit sphere and we can see that $\mathbf{x}_{k+1}$ is a unit vector if and only if the vector $\mathbf{x}_{k+1} + \mathbf{x}_k$ and the vector $\mathbf{d}_k = \mathbf{x}_{k+1} - \mathbf{x}_k$ are perpendicular to each other, i.e.

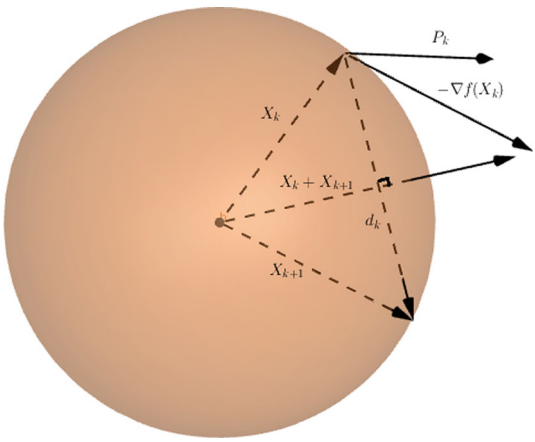

Fig. 1 Illustration of the new iterate on the unit sphere

$$(\mathbf{x}_{k+1} + \mathbf{x}_k)^\top \mathbf{d}_k = 0. \tag{3.13}$$

Let W_k be a skew-symmetric matrix, i.e., $W_k = -W_k^\top$. Then we have

$$(\mathbf{x}_k + \mathbf{x}_{k+1})^\top W_k(\mathbf{x}_{k+1} + \mathbf{x}_k) = -(\mathbf{x}_{k+1} + \mathbf{x}_k)^\top W_k(\mathbf{x}_{k+1} + \mathbf{x}_k) = 0.$$

Therefore, Eq. (3.13) is feasible and the first condition of $\mathbf{x}_{k+1}$ holds when

$$\mathbf{d}_k = W_k(\mathbf{x}_k + \mathbf{x}_{k+1}). \tag{3.14}$$

Furthermore, based on the optimization techniques, it is possible to find an ascent direction $\mathbf{p}_k$ such that

$$\mathbf{p}_k^\top \nabla f(\mathbf{x}_k) > 0. \tag{3.15}$$

Motivated by (3.15) and (3.5) , we construct $\mathbf{d}_k$ as a combination of $\mathbf{x}_k$ and $\mathbf{p}_k$, i.e.,

$$\mathbf{d}_k = a\mathbf{x}_k + b\mathbf{p}_k, \tag{3.16}$$

and obtain

$$\mathbf{d}_k^\top \nabla f(\mathbf{x}_k) = a\mathbf{x}_k^\top \nabla f(\mathbf{x}_k) + b\mathbf{p}_k^\top \nabla f(\mathbf{x}_k) = b\mathbf{p}_k^\top \nabla f(\mathbf{x}_k). \tag{3.17}$$

Therefore, if $b > 0$ in (3.16), $\mathbf{d}_k$ is an ascent direction with $\mathbf{d}_k^\top \nabla f(\mathbf{x}_k) > 0$ from (3.15) and (3.17).

The previous analysis shows that the two conditions of $\mathbf{x}_{k+1}$ are valid when $\mathbf{d}_k$ satisfies (3.14) and (3.16) for $b > 0$. Therefore, we construct the skew-symmetric matrix W_k based on $\mathbf{x}_k$ and $\mathbf{p}_k$. Let

$$W_k = \frac{1}{2}\alpha(\mathbf{p}_k\mathbf{x}_k^\top - \mathbf{x}_k\mathbf{p}_k^\top) \in \mathbb{R}^{n\times n} \tag{3.18}$$

where α is a positive parameter. Combining (3.14) and (3.18), we have

$$b = \frac{1}{2}\alpha\mathbf{x}_k^\top(\mathbf{x}_k + \mathbf{x}_{k+1}) = \frac{1}{2}\alpha(1 + \mathbf{x}_k^T\mathbf{x}_{k+1}) \geq 0$$

in (3.16). However if $b = 0$, i.e., $\mathbf{x}_{k+1} = -\mathbf{x}_k$, there is a contradiction when we substitute $\mathbf{x}_{k+1}$ by $-\mathbf{x}_k$ in (3.14). Hence, we have $b > 0$ and Eqs. (3.14) and (3.16) hold, which means the two conditions of $\mathbf{x}_{k+1}$ are satisfied when W_k is the matrix in (3.18) with $\mathbf{p}_k$ being an ascent direction.

Lemma 3.1 *The new iterate $\mathbf{x}_{k+1}$ can be expressed as*

$$\mathbf{x}_{k+1}(\alpha) = \frac{[(2 - \alpha\mathbf{x}_k^\top\mathbf{p}_k)^2 - \|\alpha\mathbf{p}_k\|^2]\mathbf{x}_k + 4\alpha\mathbf{p}_k}{4 + \|\alpha\mathbf{p}_k\|^2 - (\alpha\mathbf{x}_k^\top\mathbf{p}_k)^2}, \tag{3.19}$$

from (3.14) *and* (3.18)*. Furthermore, we have*

$$\|\mathbf{x}_{k+1}(\alpha) - \mathbf{x}_k\| = 2\left(\frac{\|\alpha\mathbf{p}_k\|^2 - (\alpha\mathbf{x}_k^\top\mathbf{p}_k)^2}{4 + \|\alpha\mathbf{p}_k\|^2 - (\alpha\mathbf{x}_k^\top\mathbf{p}_k)^2}\right)^{\frac{1}{2}}. \tag{3.20}$$

Proof From (3.14), we obtain $\mathbf{x}_{k+1} = Q\mathbf{x}_k$, where

$$Q = (I - W_k)^{-1}(I + W_k).$$

That is, the orthogonal transform is exactly the Cayley transform. The proof is then similar to Lemma 3.2 in [8]. □

For the new point $\mathbf{x}_{k+1}$ in (3.19), a crucial step is to find an ascent direction $\mathbf{p}_k$ to guarantee the ascent property in (3.15). Since problems related to hypergraphs and tensors are often large and time-consuming for computation, we employ the nonlinear conjugate gradient method, which is proposed for large-scale nonlinear optimization problems, to determine a suitable $\mathbf{p}_k$. The nonlinear conjugate gradient method does not need the Hessian matrices of the objective function and is usually faster than the steepest descent method. In [24,25], a nonlinear conjugate gradient method called CG_DESCENT was given and it was proved that the CG_DESCENT possesses a good descent property. We adopt the construction of parameter β_k in CG_DESCENT and obtain the ascent direction $\mathbf{p}_k$ as

$$\mathbf{p}_k = \nabla f(\mathbf{x}_k) + \beta_{k-1}\mathbf{d}_{k-1}. \tag{3.21}$$

The scalar β_{k-1} above is defined as $\beta_{k-1} = \max(0, \tilde{\beta}_{k-1})$, where

$$\tilde{\beta}_{k-1} = \begin{cases} \left(\tau \mathbf{d}_{k-1} \frac{\|\mathbf{y}_{k-1}\|^2}{\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}} - \mathbf{y}_{k-1}\right)^\top \frac{\nabla f(\mathbf{x}_k)}{\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}} & \text{if } |\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}| \geq \epsilon \|\mathbf{d}_{k-1}\| \|\mathbf{y}_{k-1}\| \\ 0 & \text{otherwise,} \end{cases} \tag{3.22}$$

$\mathbf{y}_{k-1} = \nabla f(\mathbf{x}_k) - \nabla f(\mathbf{x}_{k-1})$, and parameters $\tau > 1/4$ and $\epsilon > 0$. The initial direction is chosen as $\mathbf{p}_0 = \nabla f(\mathbf{x}_0)$. The direction $\mathbf{p}_k$ in (3.21) is proved to satisfy the ascent property in the following Lemma.

Lemma 3.2 *The search direction $\mathbf{p}_k$ generated by (3.21) satisfies the sufficient ascent condition, i.e.*

$$\mathbf{p}_k^\top \nabla f(\mathbf{x}_k) \geq \left(1 - \frac{1}{4\tau}\right) \|\nabla f(\mathbf{x}_k)\|^2, \tag{3.23}$$

and there exists a constant $M_0 > 1$ such that

$$\|\mathbf{p}_k\| \leq M_0 \|\nabla f(\mathbf{x}_k)\|. \tag{3.24}$$

Proof When $\beta_{k-1} = 0$, it is easy to show that the two inequalities hold. For $\beta_{k-1} \neq 0$, we have

$$\begin{aligned} \mathbf{p}_k^\top \nabla f(\mathbf{x}_k) &= \|\nabla f(\mathbf{x}_k)\|^2 + \tau \frac{\mathbf{d}_{k-1}^\top \nabla f(\mathbf{x}_k)}{\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}} \frac{\|\mathbf{y}_{k-1}\|^2}{\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}} \mathbf{d}_{k-1}^\top \nabla f(\mathbf{x}_k) - \frac{\mathbf{y}_{k-1}^\top \nabla f(\mathbf{x}_k)}{\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}} \mathbf{d}_{k-1}^\top \nabla f(\mathbf{x}_k) \\ &= \frac{1}{4\tau} \|\nabla f(\mathbf{x}_k)\|^2 - \frac{\mathbf{d}_{k-1}^\top \nabla f(\mathbf{x}_k)}{\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}} \mathbf{y}_{k-1}^\top \nabla f(\mathbf{x}_k) + \tau \frac{(\mathbf{d}_{k-1}^\top \nabla f(\mathbf{x}_k))^2}{(\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1})^2} \|\mathbf{y}_{k-1}\|^2 \\ &\quad + \left(1 - \frac{1}{4\tau}\right) \|\nabla f(\mathbf{x}_k)\|^2 \geq \left(1 - \frac{1}{4\tau}\right) \|\nabla f(\mathbf{x}_k)\|^2. \end{aligned}$$

Since

$$\|\mathbf{d}_{k-1} \cdot \mathbf{y}_{k-1}^\top\| = \|\mathbf{d}_{k-1}\| \cdot \|\mathbf{y}_{k-1}\| \quad \text{and} \quad \|\mathbf{d}_{k-1} \cdot \mathbf{d}_{k-1}^\top\| = \|\mathbf{d}_{k-1}\|^2,$$

we obtain

$$\begin{aligned} \|\beta_{k-1}\mathbf{d}_{k-1}\| &\leq \left\| \frac{\tau \|\mathbf{y}_{k-1}\|^2 \cdot \mathbf{d}_{k-1} \cdot \mathbf{d}_{k-1}^\top - \mathbf{d}_{k-1}^\top \mathbf{y}_{k-1} \cdot \mathbf{d}_{k-1} \cdot \mathbf{y}_{k-1}^\top}{(\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1})^2} \right\| \cdot \|\nabla f(\mathbf{x}_k)\| \\ &\leq \left[\frac{\|\mathbf{d}_{k-1}\| \|\mathbf{y}_{k-1}\|}{|\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1}|} + \frac{\tau \|\mathbf{y}_{k-1}\|^2 \|\mathbf{d}_{k-1}\|^2}{(\mathbf{d}_{k-1}^\top \mathbf{y}_{k-1})^2} \right] \cdot \|\nabla f(\mathbf{x}_k)\| \\ &\leq \left[\frac{1}{\epsilon} + \frac{\tau}{\epsilon^2} \right] \|\nabla f(\mathbf{x}_k)\|. \end{aligned}$$

Then we deduce that

$$\|\mathbf{p}_k\| \leq \|\nabla f(\mathbf{x}_k)\| + \|\beta_{k-1}\mathbf{d}_{k-1}\| \leq \left[1 + \frac{1}{\epsilon} + \frac{\tau}{\epsilon^2}\right]\|\nabla f(\mathbf{x}_k)\|.$$

Inequality (3.24) is valid when $M_0 = 1 + \frac{1}{\epsilon} + \frac{\tau}{\epsilon^2}$.

In the curvilinear line search, the parameter α in (3.19) is chosen to ensure that the Wolfe conditions hold. We provide the details in the next subsection.

3.3 Feasibility of Wolfe Conditions

In this section, we prove that there exists a step length α_k satisfying the Wolfe conditions for the curvilinear search in (3.19) in each iteration. First, we compute the derivative of α, which plays an important role in line search.

Lemma 3.3 *Let $f'(\alpha)$ be the derivative of $f(\mathbf{x}_{k+1}(\alpha))$ at point α. Then we have*

$$\alpha f'(\alpha) = -\nabla f(\mathbf{x}_{k+1}(\alpha))^\top \mathbf{x}_k. \tag{3.25}$$

Proof Equation (3.19) means that

$$[4 + \alpha^2\|\mathbf{p}_k\|^2 - \alpha^2(\mathbf{x}_k^\top \mathbf{p}_k)^2]\mathbf{x}_{k+1}(\alpha) = [(2 - \alpha\mathbf{x}_k^\top \mathbf{p}_k)^2 - \alpha^2\|\mathbf{p}_k\|^2]\mathbf{x}_k + 4\alpha\mathbf{p}_k.$$

Then we take derivative with respect to α as follows

$$\begin{aligned} &2\alpha(\|\mathbf{p}_k\|^2 - (\mathbf{x}_k^\top \mathbf{p}_k)^2)\mathbf{x}_{k+1}(\alpha) + [4 + \alpha^2\|\mathbf{p}_k\|^2 - \alpha^2(\mathbf{x}_k^\top \mathbf{p}_k)^2]\mathbf{x}'_{k+1}(\alpha) \\ &= [-4\mathbf{x}_k^\top \mathbf{p}_k + 2\alpha(\mathbf{x}_k^\top \mathbf{p}_k)^2 - 2\alpha\|\mathbf{p}_k\|^2]\mathbf{x}_k + 4\mathbf{p}_k. \end{aligned} \tag{3.26}$$

By multiplying both sides of (3.26) by α, we obtain

$$\alpha\mathbf{x}'_{k+1}(\alpha) = \frac{-2\alpha^2(\|\mathbf{p}_k\|^2 - (\mathbf{x}_k^\top \mathbf{p}_k)^2)}{4 + \alpha^2\|\mathbf{p}_k\|^2 - \alpha^2(\mathbf{x}_k^\top \mathbf{p}_k)^2}\mathbf{x}_{k+1}(\alpha) + \mathbf{x}_{k+1}(\alpha) - \mathbf{x}_k \tag{3.27}$$

from (3.19). Since $\nabla f(\mathbf{x}_{k+1}(\alpha))^\top \mathbf{x}_{k+1}(\alpha) = 0$, from (3.27) we obtain

$$\alpha f'(\alpha) = \alpha\nabla f(\mathbf{x}_{k+1}(\alpha))^\top \mathbf{x}'_{k+1}(\alpha) = -\nabla f(\mathbf{x}_{k+1}(\alpha))^\top \mathbf{x}_k.$$

□

Since $f(\mathbf{x})$ is twice continuously differentiable in the compact set $\mathbb{S}^{n-1}$, we can find a constant M such that

$$|f(\mathbf{x})| \leq M, \quad \|\nabla f(\mathbf{x})\| \leq M, \quad \text{and} \quad \|\nabla^2 f(\mathbf{x})\| \leq M. \tag{3.28}$$

For a given optimization algorithm which enjoys a good ascent or descent property, it is proved in [50, Lemma 3.1] that step lengths that satisfy the Wolfe conditions exist for a monotonous line search. In the following theorem, we show that Wolfe conditions are practicable for the curvilinear line search in our algorithm.

Theorem 3.1 *If $0 < c_1 < c_2 < 1$, there exists $\alpha_k > 0$ satisfying*

$$f\big(\mathbf{x}_{k+1}(\alpha_k)\big) \geq f(\mathbf{x}_k) + c_1\alpha_k\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k, \tag{3.29}$$

$$\nabla f(\mathbf{x}(\alpha_k))^\top \mathbf{p}_k \leq c_2\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k. \tag{3.30}$$

Proof Let $\mathbf{x}(\alpha) = \mathbf{x}_{k+1}(\alpha)$ and $f(\alpha) = f(\mathbf{x}_{k+1}(\alpha))$. From (3.19), we have $\mathbf{x}'_{k+1}(0) = -\mathbf{x}_k^\top \mathbf{p}_k \mathbf{x}_k + \mathbf{p}_k$, and

$$\begin{aligned} f'(0) &= \left.\frac{\mathrm{d} f(\mathbf{x}_{k+1}(\alpha))}{\mathrm{d}\alpha}\right|_{\alpha=0} = \nabla f(\mathbf{x}_{k+1}(0))^\top \mathbf{x}'_{k+1}(0) \\ &= \nabla f(\mathbf{x}_k)^\top \left(-\mathbf{x}_k^\top \mathbf{p}_k \mathbf{x}_k + \mathbf{p}_k\right) = \nabla f(\mathbf{x}_k)^\top \mathbf{p}_k. \end{aligned}$$

Denote a linear function $l(\alpha) = f(\mathbf{x}_k) + c_1 \alpha \nabla f(\mathbf{x}_k)^\top \mathbf{p}_k$. Then $f(0) = l(0) = f(\mathbf{x}_k)$ and $f'(0) > l'(0) > 0$ due to $0 < c_1 < 1$ and $\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k > 0$ in (3.23). Since $f(\alpha)$ is bounded above, the graph of $f(\alpha)$ must intersect with the line $l(\alpha)$ at least once when $\alpha > 0$. Suppose $\bar{\alpha}$ is the smallest intersection point, we obtain

$$f(\mathbf{x}_{k+1}(\bar{\alpha})) = f(\mathbf{x}_k) + c_1 \bar{\alpha} \nabla f(\mathbf{x}_k)^\top \mathbf{p}_k. \tag{3.31}$$

By the mean value theorem, we can find $\rho \in (0, \bar{\alpha})$ satisfying

$$\begin{aligned} f(\mathbf{x}_{k+1}(\bar{\alpha})) - f(\mathbf{x}_k) &= \bar{\alpha} f'(\rho) \\ \text{[By (3.25]} \qquad &= -\frac{\bar{\alpha}}{\rho} \nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{x}_k. \end{aligned} \tag{3.32}$$

On the other hand, from (3.5) and (3.19) we have

$$\begin{aligned} \nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{x}_{k+1}(\rho) &= \frac{[(2 - \rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2] \nabla f(\mathbf{x}_{k+1})(\rho)^\top \mathbf{x}_k}{4 + \|\rho \mathbf{p}_k\|^2 - (\rho \mathbf{x}_k^\top \mathbf{p}_k)^2} \\ &\quad + \frac{4\rho \nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{p}_k}{4 + \|\rho \mathbf{p}_k\|^2 - (\rho \mathbf{x}_k^\top \mathbf{p}_k)^2} \\ &= 0. \end{aligned}$$

Then we have

$$-[(2 - \rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2] \nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{x}_k = 4\rho \nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{p}_k \tag{3.33}$$

Combining (3.32) and (3.33), we have

$$[(2 - \rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2][f(\mathbf{x}_{k+1}(\bar{\alpha})) - f(\mathbf{x}_k)] = 4\bar{\alpha} \nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{p}_k. \tag{3.34}$$

Further, from (3.31) we obtain

$$\begin{aligned} &[(2 - \rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2][f(\mathbf{x}_{k+1}(\bar{\alpha})) - f(\mathbf{x}_k)] \\ &= [(2 - \rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2] c_1 \bar{\alpha} \nabla f(\mathbf{x}_k)^\top \mathbf{p}_k. \end{aligned} \tag{3.35}$$

Combing (3.34) and (3.35), we have

$$4 \nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{p}_k = [(2 - \rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2] c_1 \nabla f(\mathbf{x}_k)^\top \mathbf{p}_k$$

Since

$$\begin{aligned} \mathbf{x}_k^\top \mathbf{p}_k &= \mathbf{x}_k^\top (\nabla f(\mathbf{x}_k) + \beta_{k-1} \mathbf{d}_{k-1}) \\ &= \beta_{k-1} \mathbf{x}_k^\top (\mathbf{x}_k - \mathbf{x}_{k-1}) \\ &= \beta_{k-1} (1 - \mathbf{x}_k^\top \mathbf{x}_{k-1}) \\ &\geq 0 \end{aligned}$$

Algorithm CSRH Computing p-spectral radius of a hypergraph

1: For a uniform hypergraph G, $p > 1$, choose parameters $0 < c_1 < c_2 < 1$, $\tau > 1/4$, $\epsilon > 0$, an initial unit point $\mathbf{x}_0$, and $k \leftarrow 0$. Calculate $\mathbf{p}_0 = \nabla f(\mathbf{x}_0)$.
2: **while** the sequence of iterates does not converge **do**
3: Use interpolation method to find α_k such that (3.29) and(3.30) hold.
4: Update the new iterate $\mathbf{x}_{k+1} = \mathbf{x}_{k+1}(\alpha_k)$ by (3.19).
5: Compute $\mathbf{d}_k$, $\nabla f(\mathbf{x}_{k+1})$, β_k, and $\mathbf{p}_{k+1}$ by (3.21).
6: $k \leftarrow k + 1$.
7: **end while**

and $|\mathbf{x}_k^\top \mathbf{p}_k| \le \|\mathbf{p}_k\|$, we have

$$\begin{aligned}(2 - \rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2 &= 4 - 4\rho \mathbf{x}_k^\top \mathbf{p}_k + (\rho \mathbf{x}_k^\top \mathbf{p}_k)^2 - \|\rho \mathbf{p}_k\|^2 \\ &\le 4 - 4\rho \mathbf{x}_k^\top \mathbf{p}_k \\ &\le 4.\end{aligned}$$

Since $\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k \ge 0$,

$$\nabla f(\mathbf{x}_{k+1}(\rho))^\top \mathbf{p}_k \le c_1 \nabla f(\mathbf{x}_k)^\top \mathbf{p}_k. \tag{3.36}$$

Since $c_2 > c_1$, inequality (3.30) holds when $\alpha_k = \rho$. Also from the condition $\rho \in (0, \bar{\alpha})$, we have $f(\alpha_k) > l(\alpha_k)$ and (3.29) is obtained.

The algorithm CSRH for computing the p-spectral radius of a hypergraph is summarized above. First, we transform the original model of $\lambda^{(p)}(G)$ into an equivalent constrained optimization problem on the unit sphere (3.3). To solve the constrained model, we compute the ascent direction $\mathbf{p}_k$ from (3.4), (3.22) and (3.21), and choose a proper α_k so that the next iterate gained via (3.19) satisfies the Wolfe conditions (3.29) and (3.30). A fast computation method for calculating $\mathcal{A}\mathbf{x}^r$ and $\mathcal{A}\mathbf{x}^{r-1}$ was proposed in [8], which improves the efficiency of products of adjacency tensor and vector. We also adopt this technique in our algorithm.

4 Convergence Analysis

In this section we prove that the CSRH algorithm converges to a stationary point of $f(\mathbf{x})$ and touches the exact p-spectral radius with a high probability. Our CSRH algorithm terminates finitely when there exits a constant c such that $\nabla f(\mathbf{x}_c) = 0$. The following convergence analysis is for the case in which the sequence $\{\mathbf{x}_k\}$ is infinite and $\nabla f(\mathbf{x}_k)$ is always a nonzero vector.

4.1 Convergence Results

The following theorem shows that CSRH algorithm is convergent.

Theorem 4.1 *Suppose the sequence* $\{\mathbf{x}_k\}$ *is generated by the algorithm CSRH from any* $\mathbf{x}_0 \in \mathbb{S}^n$. *Then we have*

$$\lim_{k\to\infty} \|\nabla f(\mathbf{x}_k)\| = 0.$$

Proof The proof is divided into two steps. First, we show that the Zoutendijk condition holds, i.e.,

$$\sum_{k=0}^{\infty} \cos^2 \varphi_k \|\nabla f(\mathbf{x}_k)\|^2 < \infty. \tag{4.1}$$

Here φ_k is the angle between $\nabla f(\mathbf{x}_k)$ and $\mathbf{p}_k$, which is denoted as

$$\varphi_k \equiv \arccos \frac{\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k}{\|\nabla f(\mathbf{x}_k)\|\|\mathbf{p}_k\|}.$$

Since $\nabla^2 f(\mathbf{x})$ is bounded, we have $\nabla f(\mathbf{x})$ is Lipschitz continuous on $\mathbb{S}^{n-1}$, i.e.,

$$\|\nabla f(\mathbf{x}_1) - \nabla f(\mathbf{x}_2)\| \le L\|\mathbf{x}_1 - \mathbf{x}_2\| \qquad \forall\, \mathbf{x}_1, \mathbf{x}_2 \in \mathbb{S}^{n-1} \tag{4.2}$$

for a constant $L > 0$. From (3.18), we have

$$\|W\| = \left\| \frac{\alpha_k}{2}(\mathbf{x}_k\mathbf{p}_k^\top - \mathbf{p}_k\mathbf{x}_k^\top) \right\| \le \frac{\alpha_k}{2}\left(\|\mathbf{x}_k\mathbf{p}_k^\top\| + \|\mathbf{p}_k\mathbf{x}_k^\top\| \right) \le \alpha_k\|\mathbf{p}_k\|.$$

Hence from (3.14)

$$\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \le \|W_k\|(\|\mathbf{x}_{k+1}\| + \|\mathbf{x}_k\|) \le 2\alpha_k\|\mathbf{p}_k\|. \tag{4.3}$$

From (4.2) and (4.3), we have

$$(\nabla f(\mathbf{x}_k) - \nabla f(\mathbf{x}_{k+1}))^\top \mathbf{p}_k \le L\|\mathbf{x}_{k+1} - \mathbf{x}_k\|\|\mathbf{p}_k\| \le 2L\alpha_k\|\mathbf{p}_k\|^2.$$

From (3.30), we obtain

$$(\nabla f(\mathbf{x}_{k+1}) - \nabla f(\mathbf{x}_k))^\top \mathbf{p}_k \le (c_2 - 1)\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k. \tag{4.4}$$

By using the above two relations, we can derive the inequality

$$(1 - c_2)\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k \le 2L\alpha_k\|\mathbf{p}_k\|^2,$$

which implies

$$\alpha_k \ge \frac{1 - c_2}{2L} \frac{\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k}{\|\mathbf{p}_k\|^2}. \tag{4.5}$$

Then from (3.29), we obtain

$$f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k) \ge \frac{c_1(1 - c_2)}{2L} \frac{(\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k)^2}{\|\mathbf{p}_k\|^2} = \frac{c_1(1 - c_2)}{2L} \cos^2\varphi_k \|\nabla f(\mathbf{x}_k)\|^2,$$

which derives the following inequality

$$f(\mathbf{x}_{k+1}) - f(\mathbf{x}_0) = \sum_{i=0}^{k} f(\mathbf{x}_{i+1}) - f(\mathbf{x}_i) \ge \frac{c_1(1 - c_2)}{2L} \sum_{i=0}^{k} \cos^2\varphi_i \|\nabla f(\mathbf{x}_i)\|^2.$$

Since $f(\mathbf{x})$ is bounded in (3.28), inequality (4.1) is then deduced.

Next, we show that the angle φ_k is bounded by $\frac{\pi}{2}$. By combining (3.23) and (3.24), we obtain

$$\frac{\nabla f(\mathbf{x}_k)^\top \mathbf{p}_k}{\|\nabla f(\mathbf{x}_k)\|\|\mathbf{p}_k\|} \ge \left(1 - \frac{1}{4\tau}\right) \frac{\|\nabla f(\mathbf{x}_k)\|}{\|\mathbf{p}_k\|} \ge \frac{1}{M_0}\left(1 - \frac{1}{4\tau}\right) \equiv C_0. \tag{4.6}$$

The above inequalities indicate that

$$\cos\varphi_k \ge C_0 > 0.$$

Therefore, from (4.1) we have

$$\lim_{k\to\infty} \|\nabla f(\mathbf{x}_k)\| = 0.$$

□

The conclusions in the rest of this section are given under the condition that the parameter p is a positive integer. Recall that the graph of a function $h(\mathbf{x})$ is defined as

$$\mathrm{Gr}\, h := \{(\mathbf{x}, \lambda) \in \mathbb{R}^n \times \mathbb{R} : f(\mathbf{x}) = \lambda\}.$$

For the function $f(\mathbf{x})$ involved in our problem (3.3), we have

$$\mathrm{Gr}\, f = \left\{ (\mathbf{x}, \lambda) : [(r-1)!\mathcal{A}\mathbf{x}^r]^p = \lambda^p \left(\sum_i |x_i|^p \right)^r \right\}.$$

Since $\mathrm{Gr}\, f$ is a semialgebraic set, $f(\mathbf{x})$ is a semialgebraic function and satisfies the Łojasiewicz inequality [1,4,65]. Thus for a critical point $\mathbf{x}_*$ of $f(\mathbf{x})$, there exist constants $\theta \in [0, 1)$ and $C_1 > 0$, as well as $\mathscr{U}$ being a neighbourhood of $\mathbf{x}_*$ such that for $\mathbf{x} \in \mathscr{U}$

$$|f(\mathbf{x}) - f(\mathbf{x}_*)|^\theta \le C_1 \|\nabla f(\mathbf{x})\|. \tag{4.7}$$

The next theorem shows that if the sequence $\{\mathbf{x}_k\}$ is infinite, it has a unique accumulation point.

Theorem 4.2 *Assume the infinite sequence $\{\mathbf{x}_k\}$ is generated by the CSRH algorithm. Then it converges to a unique point $\mathbf{x}_*$, that is,*

$$\lim_{k\to\infty} \mathbf{x}_k = \mathbf{x}_*,$$

and $\mathbf{x}_*$ is a first-order stationary point.

Proof From (4.5), (3.23) and (3.24) we have

$$\begin{aligned} \alpha_k &\ge \frac{1-c_2}{2L}\left(1 - \frac{1}{4\tau}\right)\frac{\|\nabla f(\mathbf{x}_k)\|^2}{\|\mathbf{p}_k\|^2} \\ &\ge \frac{1-c_2}{2LM_0^2}\left(1 - \frac{1}{4\tau}\right) \\ &\equiv \alpha_{\min} > 0. \end{aligned}$$

Moreover, from (3.29) and (3.23) we obtain

$$\begin{aligned} f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k) &\ge c_1\alpha_k \nabla f(\mathbf{x}_k)^\top \mathbf{p}_k \\ &\ge c_1\alpha_{\min}(1 - \frac{1}{4\tau})\|\nabla f(\mathbf{x}_k)\|^2 \\ &> 0. \end{aligned} \tag{4.8}$$

We exclude the condition $\|\nabla f(\mathbf{x}_k)\| = 0$ under which the algorithm terminates finitely. The above inequality indicates that

$$[f(\mathbf{x}_{k+1}) = f(\mathbf{x}_k)] \Rightarrow [\mathbf{x}_{k+1} = \mathbf{x}_k]. \tag{4.9}$$

Based on (3.24), (3.29) and (4.3), we have

$$\begin{aligned} f(\mathbf{x}_{k+1}) - f(\mathbf{x}_k) &\ge c_1\alpha_k \left(1 - \frac{1}{4\tau}\right)\frac{\|\nabla f(\mathbf{x}_k)\|\|\mathbf{p}_k\|}{M_0} \\ &\ge \left(1 - \frac{1}{4\tau}\right)\frac{c_1}{2M_0}\|\nabla f(\mathbf{x}_k)\|\|\mathbf{x}_{k+1} - \mathbf{x}_k\| \end{aligned} \tag{4.10}$$

From (4.9) and (4.10), as well as the Łojasiewicz inequality (4.7), we conclude the theorem holds based on [1, Theorem 3.2]. □

4.2 Probability of Obtaining the Exact p-Spectral Radius

With the help of the feasibility of Łojasiewicz inequality in (4.7), we obtain the probability of the CSRH method touching the true p-spectral radius.

Proposition 4.1 *Suppose the CSRH algorithm is implemented N times with N uniformly distributed initial points on $\mathbb{S}^{n-1}$. We take the largest result of these trails as the p-spectral radius of the relevant problem. The probability of obtaining the exact p-spectral radius is*

$$1 - (1 - \zeta)^N,$$

where ζ is a constant satisfying $\zeta \in (0, 1]$. If N is large, the probability is close to 1.

Proof This proposition can be proved in the way similar to [8, Theorem 4.9]. We omit the details.

5 Numerical Experiments

In this section, we show the performance of CSRH for both small and large scale hypergraphs. We compare our method with several existing ones for computing eigenvalues of adjacency tensors. Examples of approximating the Lagrangian of a hypergraph are given in Sect.2. All experiments are carried out by using MATLAB version R2015b and Tensor Toolbox version 2.6 [3]. The experiments in Sects 5.1 and 5.2 are terminated when

$$\|\nabla f(\mathbf{x})\| \le 10^{-8} \quad \text{or} \quad \|\lambda^{(p)} - \lambda_*^{(p)}(G)\| \le 10^{-12},$$

where $\lambda^{(p)}$ is our computated p-spectral radius and $\lambda_*^{(p)}(G)$ is the exact result we obtained from theorems or conclusions in existing literature. The maximum number of iterations used in CSRH is set to 1000 for all algorithms except those in Tensor Toolbox. For each experiment in this section, we run CSRH 100 times to obtain 100 estimated values $\lambda_1^{(p)}, \ldots, \lambda_{100}^{(p)}$ and choose the largest one as our computational result of the p-spectral radius related to G. The accuracy of the CSRH algorithm is defined as

$$\text{Accu.} \equiv \left| \left\{ i : \frac{|\lambda_i^{(p)} - \lambda_*^{(p)}(G)|}{|\lambda_*^{(p)}(G)|} \le 10^{-8} \right\} \right| \times 1\%. \tag{5.1}$$

The number of iterations (Iter.) and computational time (Time) we reported in this section are the sum of corresponding quantities for all 100 executions of the experiment. The relative errors (Err.) between the numerical results and the exact solutions are provided.

5.1 Computation of p-Spectral Radii of Hypergraphs

We compare the following three algorithms for computing eigenvalues of adjacency tensors associated with different hypergraphs:

- An adaptive shifted power method [33] SS-HOPM. This method can be invoked by `eig_sshopm` in Tensor Toolbox 2.6 for Z-eigenvalues of symmetric tensors.
- A first-order optimization algorithm CEST [8] which is developed for computing eigenvalues of large scale sparse tensors involving even order hypergraphs.
- CSRH: the method proposed in Sect. 3.

Table 1 Z-Eigenvalues of adjacency tensors of several small hypergraphs

Hypergraph	CSRH				SS-HOPM			
	Iter.	Time (s)	Accu.	Err.	Iter.	Time (s)	Accu.	Err.
G_1	13593	3.35	1.00	5.44×10^{-16}	2668	4.89	1.00	5.44×10^{-16}
G_2	1257	0.78	1.00	3.85×10^{-16}	18610	32.58	0.94	3.85×10^{-16}
G_3	674	0.42	1.00	3.85×10^{-16}	731	1.61	1.00	7.69×10^{-16}
G_4	8901	2.23	0.18	1.48×10^{-16}	2317	4.38	0.22	2.96×10^{-16}

Recall that the weight of any edge in an ordinary hypergraph is regarded as 1 from Definition 2.1.

Example 1 ($\boldsymbol{p = 2}$). First, we compute the largest Z-eigenvalues of adjacency tensors of the following hypergraphs:

$$\begin{cases} G_1 : V = \{1, 2, 3, 4\} \quad \text{and} \quad E = \{123, 234\}; \\ G_2 : V = \{1, 2, 3, 4, 5, 6, 7\} \quad \text{and} \quad E = \{123, 345, 567\}; \\ G_3 : V = \{1, 2, 3, 4, 5\} \quad \text{and} \quad E = \{123, 345\}; \\ G_4 : V = \{1, 2, 3, 4\} \quad \text{and} \quad E = \{123, 124, 134, 234\}. \end{cases}$$

The first hypergraph G_1 is given in [64] as Example 1, while the last three hypergraphs are Examples 4, 7, and 9 in [52]. The hypergraph G_4 is actually a tetrahedron.

In Table 1, we demonstrate results of CSRH and SS-HOPM for computing the largest Z-eigenvalues of adjacency tensors of small hypergraphs. Since the above four hypergraphs are of odd orders, the comparison does not include the CEST method, which is designed for hypergraphs with even orders only. The Err. column shows the relative error between the computational result and the exact largest Z-eigenvalue provided in the corresponding references. It can be seen that CSRH is much more stable and efficient than SS-HOPM.

In the next experiment, we study the probability of CSRH arriving at the true largest Z-eigenvalue of G_4 and show that the probability increases when the number of trials increases. We employ CSRH to compute the largest Z-eigenvalue of the adjacency tensor of G_4 from uniformly distributed and randomly chosen initial points. Once the relative error between the computed largest Z-eigenvalue and its exact value $3/2$ reaches 10^{-8}, the experiment is terminated and we record the number of trails. This experiment is repeated for one thousand times. Let $\sigma(i)$ be the total occurrence of experiments whose trail time is the integer i. The frequency of touching the exact Z-eigenvalue when running i times is

$$v_i = \frac{\sum_{j \le i} \sigma(j)}{1000}. \tag{5.2}$$

In Fig. 2, we display the relation between trail times and success probability. It illustrates that the probability approaches one along with the increase of trail times i, which coincides with the conclusion in Proposition 4.1.

Example 2 ($\boldsymbol{p = r}$). Next, we compare CSRH with CEST for the largest H-eigenvalues problem of adjacency tensors of loose paths. An r-graph with m edges is called a loose path if its vertex set is

$$V = \left\{ i_{(1,1)}, \ldots, i_{(1,r)}, i_{(2,2)}, \ldots, i_{(2,r)}, \ldots, i_{(m,2)}, \ldots, i_{(m,r)} \right\}$$

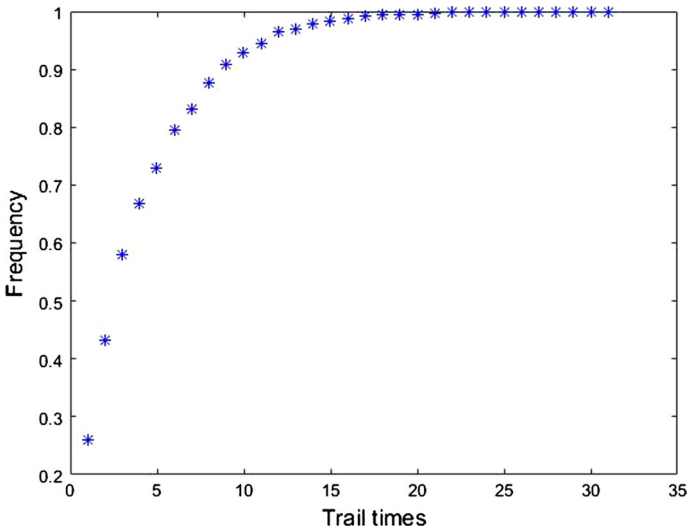

Fig. 2 Probability of touching the exact largest Z-eigenvalue of adjacency of G_4

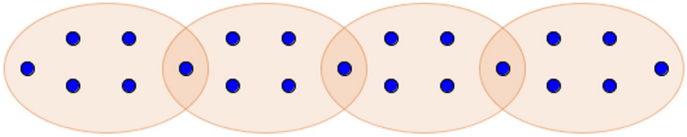

Fig. 3 A 6-uniform loose path with 4 edges

and its edge set is

$$E = \left\{\{i_{(1,1)}, \ldots, i_{(1,r)}\}, \{i_{(1,r)}, i_{(2,2)}, \ldots, i_{(2,r)}\}, \ldots, \{i_{(m-1,r)}, i_{(m,2)}, \ldots, i_{(m,r)}\}\right\}.$$

An r-uniform loose path with m edges has $m(r-1)+1$ vertices. For example, the 6-unform loose path with 4 edges in Fig. 3 has 21 vertices. The following theorem proved in [66] offers a convenient way to compute the largest H-eigenvalues of adjacency tensors of loose paths with $m = 3$ or $m = 4$.

Theorem 5.1 ([66]) *Let G be an r-uniform loose path with m edges and $\lambda_H(G)$ be the largest H-eigenvalue of its adjacency tensor. Then we have*

1. $\lambda_H(G) = \left(\frac{1+\sqrt{5}}{2}\right)^{\frac{2}{r}}$ *for* $m = 3$,
2. $\lambda_H(G) = 3^{\frac{1}{r}}$ *for* $m = 4$.

In Table 2, we compare CSRH and CEST for computing the largest H-eigenvalues of adjacency tensors of different loose paths. The column Err. presents the relative error between our computed result and the exact one given by Theorem 5.1. The comparison between CEST and CSRH verifies that the high efficiency of CSRH does not only rely on the fast computation technique in [8], because CEST uses this technique as well.

Example 3 If all edges of a hypergraph share a same vertex, then it is called a β-star. An r-uniform β-star with m edges have $m(r-1)+1$ vertices. We present a class of 6-uniform β-star in Fig. 4 as an example.

Table 2 H-eigenvalues of adjacency tensors of loose paths

m	r	CSRH				CEST			
		Iter.	Time (s)	Accu.	Err.	Iter.	Time (s)	Accu.	Err.
3	4	38123	9.14	1.00	3.49×10^{-16}	42760	70.28	1.00	3.49×10^{-16}
	6	62780	17.55	0.97	5.67×10^{-16}	65706	105.53	0.99	7.56×10^{-16}
	8	71311	23.38	0.66	3.94×10^{-16}	76778	106.95	0.65	7.88×10^{-16}
4	4	69517	16.92	1.00	5.06×10^{-16}	49331	79.81	1.00	5.06×10^{-16}
	6	86171	24.83	0.96	5.55×10^{-16}	76105	113.11	0.98	5.55×10^{-16}
	8	75907	24.71	0.33	7.74×10^{-16}	91690	106.57	0.42	9.68×10^{-16}

Fig. 4 A 6-uniform β-star

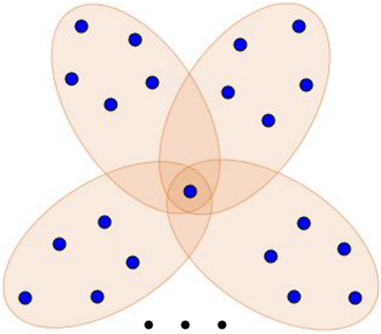

Table 3 The p-spectral radius of r-uniform β-stars

n	$p = 3,\ r = 3\ (p > r - 1)$				n	$p = 4,\ r = 6\ (p < r - 1)$			
	Iter.	Time (s)	Accu.	Err.		Iter.	Time (s)	Accu.	Err.
21	1835	0.34	1.00	5.38×10^{-16}	51	14747	4.79	0.99	1.59×10^{-11}
201	2609	0.60	1.00	3.55×10^{-15}	501	26019	14.52	0.98	9.56×10^{-12}
2,001	3539	1.87	1.00	4.33×10^{-14}	5,001	30108	57.82	0.99	2.01×10^{-11}
20,001	4475	12.93	1.00	6.39×10^{-14}	50,001	32387	426.60	0.95	1.08×10^{-11}
200,001	6038	263.39	0.98	1.93×10^{-11}	500,001	30070	6309.58	0.99	4.49×10^{-11}
2,000,001	20018	15437.99	1.00	1.22×10^{-10}	5,000,001	51609	125869.02	0.97	2.40×10^{-10}
The 3-spectral radius of 3-uniform β-stars ($p > r - 1$)					The 4-spectral radius of 6-uniform β-stars ($p < r - 1$)				

We calculate p-spectral radii of β-stars with various orders and edges and display the results in Table 3. The Err. column presents the relative error between our computational result and the corresponding exact result generated from Theorem 2.1. It can be seen that CSRH outperforms CEST for all tests. In addition, 3-spectral radii and 4-spectral radii of β-stars with millions of vertices can be computed using CSRH with high probability and efficiency.

Table 4 p_ϑ-spectral radius of 3-uniform β-star with 10 edges

p_ϑ	Iter.	Time(s)	Accu.	Err.
$p_\vartheta = \frac{12}{7}$	3037	0.99	1.00	0.00
$p_\vartheta = \frac{14}{9}$	13271	17.88	1.00	3.08×10^{-16}
$p_\vartheta = \frac{10}{7}$	51018	110.53	1.00	1.85×10^{-16}
$p_\vartheta = \frac{4}{3}$	84848	88.85	1.00	3.07×10^{-14}

5.2 Approximation of Lagrangians of Hypergraphs

When $p = 1$, the 1-spectral radius is also known as the Lagrangian of an ordinary hypergraph (2.4). However, $f(\mathbf{x})$ is not smooth at $\mathbf{x}$ that has some zero elements. We use $\lambda^{(p_\vartheta)}(G)$ to approximate $\lambda^{(1)}(G)$, with p_ϑ defined as

$$p_\vartheta = 1 + \frac{1}{2\vartheta + 1}, \text{ for } \vartheta = 1, 2, \ldots. \tag{5.3}$$

Since $\lim_{\vartheta\to\infty} p_\vartheta = 1$, we have $\lim_{\vartheta\to\infty} \lambda^{(p_\vartheta)}(G) = \lambda^{(1)}(G)$ from Proposition 3.1. Therefore, we can use p_ϑ-spectral radius to approximate the Lagrangian of a hypergraph. The function $f_{p_\vartheta}(\mathbf{x})$ is continuous and differentiable and CSRH can be used to compute the p_ϑ-spectral radius of a uniform hypergraph. Let $\mathbf{w}$ be a vector such that its ith element being

$$w_i = x_i^{\frac{1}{2\vartheta+1}}, \quad \text{for } i = 1, \ldots, n.$$

Then function $f_{p_\vartheta}(\mathbf{x}) = f_{p_\vartheta}(\mathbf{w}^{[2\vartheta+1]})$ is also a semialgebraic function and satisfies the Łojasiewicz inequality (4.7). Therefore, the conclusions in Sect. 4 hold for p_ϑ in (5.3).

In this subsection, we show the results of CSRH for approximating Lagrangian of a hypergraph. First we give an example to demonstrate that CSRH is competent to compute the p-spectral radius of a uniform hypergraph when p is the fraction in (5.3). Next, we present the numerical results of approximating the Lagrangians of complete hypergraphs by p_ϑ-spectral radius. The termination criteria of algorithms in the remaining part of this paper is set as $\|\nabla f(\mathbf{x})\| \le 10^{-6}$.

In Table 4, we present the results of the p_ϑ-spectral radius of a 3-uniform β-star with 10 edges, with p_ϑ being fractions. The true p_ϑ-spectral radius can be obtained from Theorem 2.1. All experiments produce the exact p_ϑ-spectral radius with probability 1 and the relative error between our numerical result and the theoretical value obtained from Theorem 2.1 is at most 3.07×10^{-14}.

An r-uniform hypergraph is said to be complete if it contains all possible edges when the number of its vertices is fixed. We use C_n^r to denote a complete r-graph with n vertices. Then the 3-graph C_4^3 is actually the hypergraph G_4 in Example 1. The Lagrangian of a complete uniform hypergraph can be obtained directly from Proposition 2.1.

We compute different p_ϑ-spectral radii of 3 complete hypergraphs C_4^3, C_{10}^3 and C_{20}^3. In Fig. 5, the ordinate reflects the error between the p_ϑ-spectral radius and the true Lagrangian of the corresponding complete hypergraph which is obtained from the Proposition 2.1, while the abscissa means the value of $p_\vartheta - 1$. When p_ϑ approaches to 1, the p_ϑ-spectral radius is close to the exact Lagrangian of the related hypergraph.

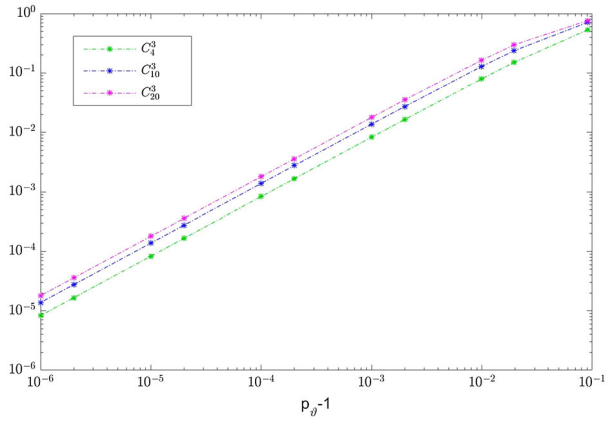

Fig. 5 Approximation of Lagrangian of complete hypergraphs

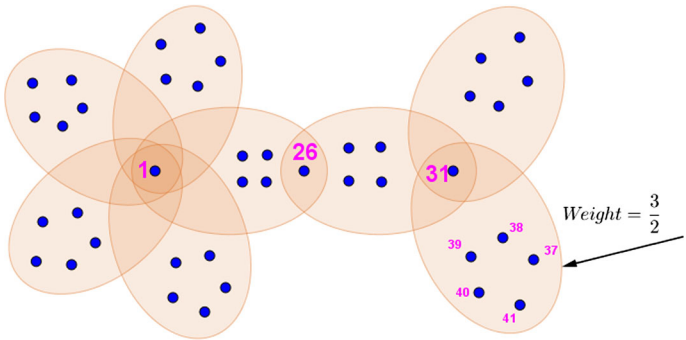

Fig. 6 A 6-uniform hypergraph

6 Network Analysis

In the p-spectral radius model (3.3), not only the p-spectral radius, i.e., the optimal value of $f(\mathbf{x})$ in (3.3), but also the optimal point $\mathbf{x}$ characterize the structure of a hypergraph. Recall (2.3) that an optimal point is called a p-optimal weighting. The elements of the p-optimal weighting reflect the importance of the corresponding vertices in the hypergraph. Therefore, we may call the ith element of the p-optimal weighting the impact factor of the ith vertex. Different selections of parameter p lead to different meanings of the ranking results. When p is relatively large, the p-spectral model tends to evaluate the importance of vertices more individually. When p is relatively small, the ranking result demonstrates the significance of groups of vertices. In this section, we compute each p-spectral radius 10 times and choose the vector corresponding to the largest objective value $f(\mathbf{x})$ as the p-optimal weighting.

6.1 A Toy Problem

We first employ a toy problem to illustrate the impact of p on the ranking results. We construct a 6-uniform weighted hypergraph with 8 edges as shown in Fig. 6. The weights of all edges of this hypergraph are set to 1, except the last one whose weight is $\frac{3}{2}$. The vertices numbered 1, 31, and 26 are distinct from other vertices, and the edge {31, 37, 38, 39, 40, 41} is also distinct from other edges. In Table 5, we show the different ranking of vertices via different p-optimal

Table 5 Top ten vertices in Fig. 6

Ranking	$p = \frac{4}{3}$		$p = 5$		$p = 16$	
	Num.	Val.	Num.	Val.	Num.	Val.
1	39	0.4082483175	41	0.4081204985	1	0.1709715830
2	38	0.4082482858	39	0.4081204985	31	0.1678396311
3	31	0.4082482855	31	0.4081204983	26	0.1618288319
4	41	0.4082482854	38	0.4081204982	39	0.1600192388
5	40	0.4082482849	40	0.4081204973	38	0.1600192387
6	37	0.4082482834	37	0.4081204958	41	0.1600192387
7	24	0.0000000000	28	0.0073198868	40	0.1600192386
8	34	0.0000000000	30	0.0073192175	37	0.1600192385
9	23	0.0000000000	26	0.0073061265	23	0.1550865094
10	3	0.0000000000	29	0.0071906282	22	0.1550865094

weightings. The abbreviation Num. means the number of a vertex and Val. represents the the value of impact factors of the corresponding vertices.

When $p = \frac{4}{3}$, the top 6 vertices are in the edge that has the only largest weight among all edges. The impact factor of the top 6 vertices in the $\frac{4}{3}$-optimal weighting are much greater than others. In fact, the value of all impact factors, except those attached to the top 6 vertices, are less than 5×10^{-10}. Hence, the dominant vertices are the ones from the largest weighted edge and the others can be ignored. That is to say, the ranking in this case offers the most important group of the vertices. When $p = 5$, the vertex numbered 26 appears in the top 10 list and the difference among the top 10 impact factors is not as great as that when $p = \frac{4}{3}$. When $p = 16$, the top 3 vertices are 1, 31, 26, and the impact factors of vertices that have same status in the hypergraph are rather close to each other. Thus, we believe that the ranking results of 16-spectral radius reflects the significance of vertices individually.

6.2 Author Ranking

Ng et al. collected publication information from DBLP[1] and gave different rankings of the authors in accordance with different factors, such as citations of authors, category concepts, collaborations, and papers [46]. In this subsection, we use the same data set and rank the authors based on their collaborations.[2]

In order to store the cooperation information, we construct a weighted 3-uniform multi-hypergraph G_A. The vertex set of G_A is composed of 10305 authors, while the edge set contains 1, 243, 443 edges. Each edge has 3 vertices indicating that these three authors have cooperations under a same topic, and the weight of an edge is decided by the collaboration times among the three authors in this edge. The adjacency tensor of this multi-hypergraph G_A is a sparse tensor with 1.17% nonzero entries.

The example in Sect. 6.1 shows that we can obtain the rank from different viewpoints by computing different p-optimal weighting. Therefore, we compute 2-optimal weighting and 12-optimal weighting of G_A to obtain the author group ranking and the author ranking respectively. In Fig. 7a, the stars stand for the 2-optimal impact factors of vertices of G_A.

[1] http://www.informatik.uni-trier.de/ley/db/.

[2] We would like to thank Dr. Xutao Li for providing the database.

Table 6 Top 10 authors

Ranking	Author name		
	$p = 2$	$p = 12$	MultiRank
1	Zheng Chen	Wei-Ying Ma	C. Lee Giles
2	Wei-Ying Ma	Zheng Chen	Philip S. Yu
3	Qiang Yang	Jiawei Han	Wei-Ying Ma
4	Jun Yan	Philip S. Yu	Zheng Chen
5	Benyu Zhang	C. Lee Giles	Jiawei Han
6	Hua-Jun Zeng	Jian Pei	Christos Faloutsos
7	Weiguo Fan	Christos Faloutsos	Bing Liu
8	Wensi Xi	Yong Yu	Johannes Gehrke
9	Dou Shen	Qiang Yang	Gerhard Weikum
10	Shuicheng Yan	Ravi Kumar	Elke A. Rundensteiner

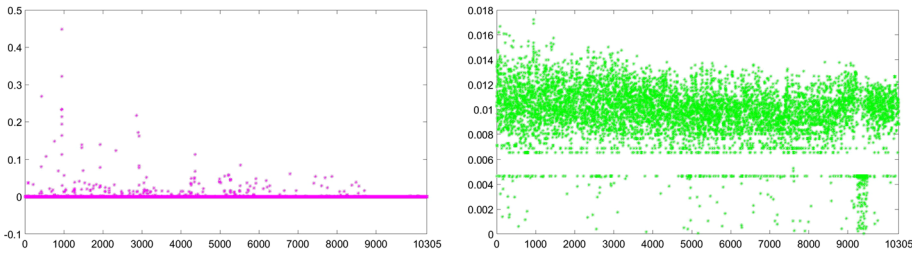

Fig. 7 Optimal points. **a** $p = 2$, **b** $p = 12$

Most elements of 2-optimal weighting are extraordinarily close to zero and only dozens of corresponding stars are above the horizontal line of $y = 0.1$. In fact, 97.2% of the entries in the 2-optimal weighting are less than 10^{-3} and the elements that are greater than 0.1 occupy only 1.8%. On the other hand, the largest impact factor reaches to 0.4481 and the upper stars are considerably larger than others. It means that the 2-optimal weighting is dominated by a small proportion of its components and we regard these leading elements as a group. The top ten authors ranked according to the 2-optimal impact factor are presented in the second column of Table 6. The average collaboration times of each two authors among these top ten authors are 8.533, which is far larger than 9.76×10^{-4}, the average collaboration times of each two authors among the whole 10305 authors. Since these top ten authors have such an intimate cooperation, it is rational to consider them as a group and interpret the ranking in the second column as the most powerful group.

Stars in Fig. 7b are the 12-optimal impact factors of vertices of G_A. The distribution of these stars is totally different from the ones in Fig. 7a. It can be seen in Fig. 7b that the distribution of the 12-optimal impact factors of the 10305 authors are uniform and most of them are concentrated in the internal between 0.006 and 0.014. Because in the original data set, the collaboration times of different authors are mostly one or two and we rank the authors based on their collaborations, the balance and concentration of the impact factors match up with the cooperation information. The top ten authors generated via the 12-optimal impact factors are listed in the third column of Table 6. Ng et al. also ranked the authors in the

light of collaboration times and the influence of category concepts of their publications. We demonstrate the top 10 authors of their experimental result [46] in the MultiRank column in Table 6. Six of the top ten authors in the MultiRank are coincident with results of our 12-optimal rank.

7 Conclusions

We convert the p-norm constraint in the p-spectral radius problem into an orthogonal constraint, and propose a first order iterative algorithm CSRH to compute the radius. It is feasible to obtain a proper step length to satisfy the Wolfe conditions under the curvilinear line search. Convergence analysis shows that the CSRH method is globally convergent. Numerical experiments show that CSRH method is efficient and powerful. In the author ranking application problem, we construct a weighted hypergraph with millions of edges. By computing the p-spectral radius of this hypergraph, the most influential cooperation group and the top ten ranked authors are identified.

References

1. Absil, P.-A., Mahony, R., Andrews, B.: Convergence of the iterates of descent methods for analytic cost functions. SIAM J. Optim. **16**(2), 531–547 (2005)
2. Agarwal, S., Lim, J., Zelnik-Manor, L., Perona, P., Kriegman, D., Belongie, S.: Beyond pairwise clustering. In: 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol 2, pp. 838–845. IEEE (2005)
3. Bader, B.W. , Kolda, T.G. et al.: Matlab tensor toolbox version 2.6. Available online, February (2015) URL http://www.sandia.gov/~tgkolda/TensorToolbox/
4. Bolte, J., Daniilidis, A., Lewis, A.: The Łojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems. SIAM J. Optim. **17**(4), 1205–1223 (2006)
5. Bretto, A., Gillibert, L.: Hypergraph-based image representation. In: International Workshop on Graph-Based Representations in Pattern Recognition, pp. 1–11. Springer (2005)
6. Brown, W., Simonovits, M.: Digraph extremal problems, hypergraph extremal problems, and the densities of graph structures. Discret. Math. **48**(2–3), 147–162 (1984)
7. Caraceni, A.: Lagrangians of hypergraphs, URL http://alessandracaraceni.altervista.org/MyWordpress/wp-content/uploads/2014/05/Hypergraph_Lagrangians.pdf (2011). [Online; Accessed 26 Jan 2017]
8. Chang, J., Chen, Y., Qi, L.: Computing eigenvalues of large scale sparse tensors arising from a hypergraph. SIAM J. Sci. Comput. **38**(6), A3618–A3643 (2016)
9. Chang, K.C., Pearson, K., Zhang, T.: Perron–Frobenius theorem for nonnegative tensors. Commun. Math. Sci. **6**(2), 507–520 (2008)
10. Chen, L., Han, L., Zhou, L.: Computing tensor eigenvalues via homotopy methods. SIAM J. Matrix Anal. Appl. **37**(1), 290–319 (2016a)
11. Chen, Y., Dai, Y.-H., Han, D.: Fiber orientation distribution estimation using a Peaceman–Rachford splitting method. SIAM J. Imaging Sci. **9**(2), 573–604 (2016b)
12. Chen, Y., Qi, L., Wang, Q.: Positive semi-definiteness and sum-of-squares property of fourth order four dimensional Hankel tensors. J. Comput. Appl. Math. **302**, 356–368 (2016c)
13. Cooper, J., Dutle, A.: Spectra of uniform hypergraphs. Linear Algebra Appl. **436**(9), 3268–3292 (2012)
14. Cui, C.-F., Dai, Y.-H., Nie, J.: All real eigenvalues of symmetric tensors. SIAM J. Matrix Anal. Appl. **35**(4), 1582–1601 (2014)
15. Ding, C., He, X., Husbands, P., Zha, H. Simon, H.D.: Pagerank, hits and a unified framework for link analysis. In: Proceedings of the 25th annual international ACM SIGIR conference on Research and development in information retrieval, pp. 353–354. ACM (2002)
16. Ding, W., Qi, L., Wei, Y.: Fast Hankel tensor-vector product and its application to exponential data fitting. Numer. Linear Algebra Appl. **22**(5), 814–832 (2015)

17. Ducournau, A., Rital, S., Bretto, A., Laget, B.: A multilevel spectral hypergraph partitioning approach for color image segmentation. In: 2009 IEEE International Conference on Signal and Image Processing Applications (ICSIPA), pp. 419–424. IEEE (2009)
18. Ducournau, A., Bretto, A., Rital, S., Laget, B.: A reductive approach to hypergraph clustering: an application to image segmentation. Pattern Recognit. **45**(7), 2788–2803 (2012)
19. Erdös, P., Stone, A.H.: On the structure of linear graphs. Bull. Am. Math. Soc. **52**, 1087–1091 (1946)
20. Frankl, P., Füredi, Z.: Extremal problems whose solutions are the blowups of the small Witt-designs. J. Comb. Theory Ser. A **52**(1), 129–147 (1989)
21. Frankl, P., Rödl, V.: Hypergraphs do not jump. Combinatorica **4**(2–3), 149–159 (1984)
22. Frankl, P., Peng, Y., Rödl, V., Talbot, J.: A note on the jumping constant conjecture of Erdős. J. Comb. Theory Ser. B **97**(2), 204–216 (2007)
23. Gunopulos, D., Mannila, H., Khardon, R., Toivonen, H.: Data mining, hypergraph transversals, and machine learning. In: Proceedings of the Sixteenth ACM SIGACT-SIGMOD-SIGART Symposium on Principles of Database Systems, pp. 209–216. ACM (1997)
24. Hager, W.W., Zhang, H.: A new conjugate gradient method with guaranteed descent and an efficient line search. SIAM J. Optim. **16**(1), 170–192 (2005)
25. Hager, W.W., Zhang, H.: A survey of nonlinear conjugate gradient methods. Pac. J. Optim. **2**(1), 35–58 (2006)
26. Hu, S., Qi, L.: The Laplacian of a uniform hypergraph. J. Comb. Optim. **29**(2), 331–366 (2015)
27. Huang, Y., Liu, Q., Zhang, S., Metaxas, D.N.: Image retrieval via probabilistic hypergraph ranking. In: 2010 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), pp. 3376–3383. IEEE (2010)
28. Kang, L., Nikiforov, V., Yuan, X.: The p-spectral radius of k-partite and k-chromatic uniform hypergraphs. Linear Algebra Appl. **478**, 81–107 (2015)
29. Karypis, G., Aggarwal, R., Kumar, V., Shekhar, S.: Multilevel hypergraph partitioning: applications in VLSI domain. IEEE T. VLSI Syst. **7**(1), 69–79 (1999)
30. Keevash, P.: Hypergraph Turán problems. Surv. combinatorics **392**, 83–140 (2011)
31. Keevash, P., Lenz, J., Mubayi, D.: Spectral extremal problems for hypergraphs. SIAM J. Discrete Math. **28**(4), 1838–1854 (2014)
32. Klamt, S., Haus, U.-U., Theis, F.: Hypergraphs and cellular networks. PLoS Comput. Biol. **5**(5), e1000385 (2009)
33. Kolda, T.G., Mayo, J.R.: Shifted power method for computing tensor eigenpairs. SIAM J. Matrix Anal. Appl. **32**(4), 1095–1124 (2011)
34. Kolda, T.G., Mayo, J.R.: An adaptive shifted power method for computing generalized tensor eigenpairs. SIAM J. Matrix Anal. Appl. **35**(4), 1563–1581 (2014)
35. Kolda, T.G., Bader, B.W., Kenny, J.P.: Higher-order web link analysis using multilinear algebra. In: Fifth IEEE International Conference on Data Mining (ICDM'05), p. 8. IEEE (2005)
36. Konstantinova, E.V., Skorobogatov, V.A.: Application of hypergraph theory in chemistry. Discrete Math. **235**(13), 365–383 (2001). Combinatorics (Prague, 1998)
37. Krohn-Grimberghe, A., Drumond, L., Freudenthaler, C., Schmidt-Thieme, L.: Multi-relational matrix factorization using Bayesian personalized ranking for social network data. In: Proceedings of the fifth ACM international conference on Web search and data mining, pp. 173–182. ACM (2012)
38. Li, H., Shao, J.-Y., Qi, L.: The extremal spectral radii of k-uniform supertrees. J. Comb. Optim. **32**(3), 741–764 (2016)
39. Li, X., Hu, W., Shen, C., Dick, A., Zhang, Z.: Context-aware hypergraph construction for robust spectral clustering. IEEE T. Knowl. Data En. **26**(10), 2588–2597 (2014)
40. Liu, Y., Shao, J., Xiao, J., Wu, F., Zhuang, Y.: Hypergraph spectral hashing for image retrieval with heterogeneous social contexts. Neurocomputing **119**, 49–58 (2013)
41. Lu, L., Man, S.: Connected hypergraphs with small spectral radius. Linear Algebra Appl. **509**, 206–227 (2016)
42. Michoel, T., Nachtergaele, B.: Alignment and integration of complex networks by hypergraph-based spectral clustering. Phys. Rev. E **86**(5), 056111 (2012)
43. Motzkin, T.S., Straus, E.G.: Maxima for graphs and a new proof of a theorem of Turán. Can. J. Math. **17**, 533–540 (1965)
44. Mubayi, D.: A hypergraph extension of Turán's theorem. J. Comb. Theory Ser. B **96**(1), 122–134 (2006)
45. Ng, M., Qi, L., Zhou, G.: Finding the largest eigenvalue of a nonnegative tensor. SIAM J. Matrix Anal. Appl. **31**(3), 1090–1099 (2009)
46. Ng, M.K.-P., Li, X., Ye, Y.: Multirank: co-ranking for objects and relations in multi-relational data. In: Proceedings of the 17th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 1217–1225. ACM (2011)

47. Nikiforov, V.: Bounds on graph eigenvalues II. Linear Algebra Appl. **427**(2–3), 183–189 (2007)
48. Nikiforov, V.: An analytic theory of extremal hypergraph problems. arXiv preprint arXiv:1305.1073, (2013)
49. Nikiforov, V.: Analytic methods for uniform hypergraphs. Linear Algebra Appl. **457**, 455–535 (2014)
50. Nocedal, J., Wright, S.J.: Numerical Optimization. Springer Series in Operations Research and Financial Engineering, 2nd edn. Springer, New York (2006)
51. Page, L., Brin, S., Motwani, R., Winograd, T.: The pagerank citation ranking: bringing order to the web. (1999)
52. Pearson, K.J., Zhang, T.: On spectral hypergraph theory of the adjacency tensor. Graphs Comb. **30**(5), 1233–1248 (2014)
53. Peng, Y.: Using Lagrangians of hypergraphs to find non-jumping numbers I. Ann. Comb. **12**(3), 307–324 (2008)
54. Peng, Y., Zhao, C.: Generating non-jumping numbers recursively. Discret. Appl. Math. **156**(10), 1856–1864 (2008)
55. Pliakos, K., Kotropoulos, C.: Weight estimation in hypergraph learning. In: 2015 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1161–1165. IEEE (2015)
56. Qi, L.: Eigenvalues of a real supersymmetric tensor. J. Symb. Comput. **40**(6), 1302–1324 (2005)
57. Qi, L., Luo, Z.: Tensor Analysis: Spectral Theory and Special Tensors. Society for Industrial and Applied Mathematics, Philadephia (2017)
58. Rohe, K., Chatterjee, S., Yu, B.: Spectral clustering and the high-dimensional stochastic blockmodel. Ann. Stat. **39**(4), 1878–1915 (2011)
59. Sidorenko, A.F.: The maximal number of edges in a homogeneous hypergraph containing no prohibited subgraphs. Math. Notes **41**(3), 247–259 (1987)
60. Sun, L., Ji, S. Ye, J.: Hypergraph spectral learning for multi-label classification. In: Proceedings of the 14th ACM SIGKDD international conference on Knowledge discovery and data mining, pp. 668–676. ACM (2008)
61. Talbot, J.: Lagrangians of hypergraphs. Comb. Probab. Comput. **11**(2), 199–216 (2002)
62. Turán, P.: Eine Extremalaufgabe aus der Graphentheorie. Mat. Fiz. Lapok **48**, 436–452 (1941)
63. Turán, P.: Research problems. MTA Mat. Kutató Int. Közl **6**, 417–423 (1961)
64. Xie, J., Chang, A.: On the Z-eigenvalues of the adjacency tensors for uniform hypergraphs. Linear Algebra Appl. **439**(8), 2195–2204 (2013)
65. Xu, Y., Yin, W.: A block coordinate descent method for regularized multiconvex optimization with applications to nonnegative tensor factorization and completion. SIAM J. Imaging Sci. **6**(3), 1758–1789 (2013)
66. Yue, J., Zhang, L., Lu, M.: Largest adjacency, signless Laplacian, and Laplacian H-eigenvalues of loose paths. Front. Math. China **11**(3), 623–645 (2016)
67. Zhou, D., Huang, J., Schölkopf, B.: Learning with hypergraphs: clustering, classification, and embedding. In: Advances in Neural Information Processing Systems, pp. 1601–1608. (2006)
68. Zhuang, Y., Liu, Y., Wu, F., Zhang, Y., Shao, J.: Hypergraph spectral hashing for similarity search of social image. In: Proceedings of the 19th ACM International Conference on Multimedia, pp. 1457–1460. ACM (2011)