The Global Convergence of the Nonlinear Power Method for Mixed-Subordinate Matrix Norms

Antoine Gautier¹ · Matthias Hein² · Francesco Tudisco³

Received: 19 March 2020 / Revised: 10 November 2020 / Accepted: 11 May 2021 / Published online: 5 June 2021
© The Author(s) 2021

Abstract

We analyze the global convergence of the power iterates for the computation of a general mixed-subordinate matrix norm. We prove a new global convergence theorem for a class of entrywise nonnegative matrices that generalizes and improves a well-known results for mixed-subordinate ℓ^p matrix norms. In particular, exploiting the Birkoff–Hopf contraction ratio of nonnegative matrices, we obtain novel and explicit global convergence guarantees for a range of matrix norms whose computation has been recently proven to be NP-hard in the general case, including the case of mixed-subordinate norms induced by the vector norms made by the sum of different ℓ^p -norms of subsets of entries.

Keywords Matrix norm · Power method · Perron–Frobenius theorem · Nonlinear eigenvalues · Duality mapping

Mathematics Subject Classification 65F35 · 15B48 · 47H09 · 47H10

The work of A.G. and M.H. has been founded by the ERC Grant ‘NOLEPRO’ Number 307793. The work of F.T. was funded by the European Union’s Horizon 2020 research and innovation programme under the Marie Skłodowska–Curie individual fellowship “MAGNET” Grant agreement No. 744014.

✉ Francesco Tudisco
francesco.tudisco@gssi.it

Antoine Gautier
antoine.gautier@quantpi.com

Matthias Hein
matthias.hein@uni-tuebingen.de

¹ QuantPi, Saarbrücken, Germany

² University of Tuebingen, Tübingen, Germany

³ Gran Sasso Science Institute, L’Aquila, Italy

1 Introduction

Let A be an $m \times n$ matrix and consider the matrix norm

$$\|A\|_{\beta \rightarrow \alpha} = \max_{x \neq 0} \frac{\|Ax\|_{\alpha}}{\|x\|_{\beta}},$$

where $\|\cdot\|_{\alpha}$ and $\|\cdot\|_{\beta}$ are vector norms.

Computing $\|A\|_{\beta \rightarrow \alpha}$ is a classical problem in computational mathematics, as norms of this kind arise naturally in many situations, such as approximation theory, estimation of matrix condition numbers and approximation of relative residuals [26]. However, attention around the problem of computing $\|A\|_{\beta \rightarrow \alpha}$ has been growing in recent years. In fact, for example, matrix norms of this type can be used in combinatorial optimization and sparse data recovery, to approximate generalized Grothendieck and restricted isometry constants [1,6,16,30], in scientific computing, to estimate the largest entries of large matrices [27], in data mining and learning theory, to minimize empirical risks or obtain robust nonnegative graph embeddings [9,41], or in quantum information theory and the study of Khot's unique game conjecture where the computational complexity of evaluating $\|A\|_{\beta \rightarrow \alpha}$ plays an important role [2]. Moreover, it was observed by Lim [33] that the notion of tensor norm and tensor spectrum relates to $\|A\|_{\beta \rightarrow \alpha}$ in a very natural way and thus relevant advances on the problem of computing $\|A\|_{\beta \rightarrow \alpha}$ when A is entrywise nonnegative and $\|\cdot\|_{\alpha}$, $\|\cdot\|_{\beta}$ are ℓ^p norms have been recently obtained as a consequence of a number of new nonlinear Perron–Frobenius-type theorems for higher-order maps [15,19,21,22].

Closed form solutions and efficient algorithms are known for some special ℓ^p norms, as for instance the case where $\|\cdot\|_{\alpha} = \|\cdot\|_{\beta}$ and they coincide with either the ℓ^1 , the ℓ^2 , or the ℓ^{∞} norm, or the case where $p \leq 1 \leq q$ and $\|\cdot\|_{\alpha}$ and $\|\cdot\|_{\beta}$ are ℓ^p and ℓ^q (semi) norms, respectively (c.f. [10,32,36]). However, the computation of $\|A\|_{\beta \rightarrow \alpha}$ is generally NP-hard [23,38].

The best known method for the computation of $\|A\|_{\beta \rightarrow \alpha}$ is the (nonlinear) power method, essentially introduced by Boyd [4] and then further analyzed and extended for instance in [3,15,25,39]. When the considered vector norms are ℓ^p norms, the power method can count on a very fundamental global convergence result which ensures convergence to the matrix norm $\|A\|_{\beta \rightarrow \alpha}$ for a class of entry-wise nonnegative matrices A and for a range of ℓ^p norms. We discuss in detail the method and its convergence in Sect. 2.

The convergence of the method is a consequence of an elegant fixed point argument that involves a nonlinear operator S_A and its Lipschitz contraction constant. However, the convergence analysis of this method has two main uncovered points: on the one hand, all the work done so far addresses only the case of ℓ^p norms whereas almost nothing is known about the global convergence behavior of the power iterates for more general norms. On the other hand, even for the case of ℓ^p norms, known upper-bounds on the contraction constant of S_A are not sharp, especially for positive matrices. In this work we provide novel results that address and improve both these directions.

Consider for example the case where $\|\cdot\|_{\alpha}$ is defined as

$$\|x\|_{\alpha} = \|(x_1, \dots, x_k)\|_{p_1} + \|(x_{k+1}, \dots, x_n)\|_{p_2} \quad (1)$$

where k is a positive integer not larger than the dimension of x and $\|\cdot\|_{p_i}$ are ℓ^p norms. Of course one can extend this idea by looking at any family of subsets of entries of x and any set of ℓ^p norms, in order to generate arbitrarily new norms. Norms of this form are natural modifications of ℓ^p norms and are used for instance to define the generalized Grothendieck constants as in [30] or in graph matching problems to build continuous relaxation of the set

of matrix permutations [11,34]. However, even for this case, extending the result of Boyd is not straightforward.

In this work we consider general pairs of monotonic and differentiable vector norms and provide a thorough convergence analysis of the power method for the computation of the corresponding induced matrix norm $\|A\|_{\beta \rightarrow \alpha}$. Our result is based on a novel nonlinear Perron–Frobenius theorem for this kind of norms and ensures global convergence of the power method provided that the Birkhoff contraction ratio of the power iterator is smaller than one.

When applied to the case $\|A\|_{q \rightarrow p}$ of ℓ^p norms, our result does not only imply the current convergence result, but actually significantly improves the range of values of p and q for which global convergence can be ensured. This is particularly interesting from a complexity viewpoint. In fact, for example, although the computation of $\|A\|_{q \rightarrow p}$ is well known to be NP-hard for $p > q$, we show that for a non-trivial class of nonnegative matrices the power method converges to $\|A\|_{q \rightarrow p}$ in polynomial time even for p sensibly larger than q . To our knowledge this is the first global optimality result for this problem that does not require the condition $p \leq q$.

In the general case $\|A\|_{\beta \rightarrow \alpha}$, a main computational drawback of the power method is related with the computation of the dual norm $\|\cdot\|_{\beta^*}$. In fact, if $\|\cdot\|_{\beta}$ is not an ℓ^p norm, the corresponding dual norm may be challenging to compute [14]. In practice, evaluating $\|\cdot\|_{\alpha^*}$ from $\|\cdot\|_{\alpha}$ can be done via convex optimization and Corollary 7 of [14] proves that $\|\cdot\|_{\alpha^*}$ can be evaluated in polynomial time (resp. is NP-hard) if and only if $\|\cdot\|_{\alpha}$ can be evaluated in polynomial time (resp. is NP-hard). There are norms for which an explicit expression in terms of arithmetic operations for $\|\cdot\|_{\alpha}$ is given by construction (resp. modelisation), but such an expression is not available for the dual $\|\cdot\|_{\alpha^*}$. As we discuss in Sect. 5.1, examples of this type include for instance $\|x\|_{\alpha} = (\|x\|_p^2 + \|x\|_q^2)^{1/2}$. A further main result of this work addresses this issue for the particular case of norms of the type (1). For this family of norms we provide an explicit convergence bound and an explicit formula for the power iterator for the computation of the corresponding matrix norm $\|A\|_{\beta \rightarrow \alpha}$. To illustrate possible applications of the result, we list in Corollaries 3–8 relatively sophisticated and non-standard matrix norms together with an explicit condition for their computability.

We organize the discussion as follows: In Sect. 2 we review the nonlinear power method and its main convergence properties. In Sect. 3 we review relevant preliminary cone-theoretic results and notation. Then, in Sect. 4, we propose a novel and detailed global convergence analysis of the method based on a Perron–Frobenius type result for the map $x \mapsto \|Ax\|_{\alpha}/\|x\|_{\beta}$, in the case of entry-wise nonnegative matrices and monotonic norms $\|\cdot\|_{\alpha}$, $\|\cdot\|_{\beta}$. We derive new conditions for the global convergence to $\|A\|_{\beta \rightarrow \alpha}$ that, in particular, help shedding new light on the NP-hardness of the problem, and we propose a new explicit bound on the linear convergence rate of the power iterates. In Sect. 5 we focus on the particular case of norms of the same form as (1). We show how to practically implement the power method for this type of norms, we prove a specific convergence criterion that gives a-priori global convergence guarantees and we discuss the complexity of the method. Finally, in Sect. 6 we illustrate the behaviour of the nonlinear power method on some example matrix norm.

2 Boyd's Nonlinear Power Method

Let $\|\cdot\|_p$, $\|\cdot\|_q$ be the usual ℓ^p and ℓ^q vector norms and consider the induced matrix norm $\|A\|_{q \rightarrow p} = \max_{x \neq 0} \|Ax\|_p / \|x\|_q$. A well known explicit formula holds for the ℓ^1 and ℓ^∞ matrix norms $\|A\|_{1 \rightarrow 1}$, $\|A\|_{\infty \rightarrow \infty}$. However, while the mixed norm $\|A\|_{1 \rightarrow \infty}$ equals $\max_{ij} |a_{ij}|$, the computation of $\|A\|_{\infty \rightarrow 1}$ is NP-hard [36]. More generally, when p is any rational number $p \neq 1, 2$, computing the norm $\|A\|_{p \rightarrow p}$ is NP-hard for a general matrix A [23], and the same holds for any norm $\|A\|_{q \rightarrow p}$, for $1 \leq p < q \leq \infty$ [38]. The best known technique to compute $\|A\|_{q \rightarrow p}$ is a form of nonlinear power method that we review in what follows.

Consider the nonnegative function $f_A(x) = \|Ax\|_p / \|x\|_q$. The norm $\|A\|_{q \rightarrow p}$ is the global maximum of f_A by analyzing the optimality conditions of f_A , for differentiable ℓ^p -norms $\|\cdot\|_p$ and $\|\cdot\|_q$, we note that

$$\nabla f_A(x) = 0 \iff A^T J_p(Ax) = f_A(x) J_q(x),$$

where, for $1 < p < \infty$, we denote by $J_p(x)$ the gradient of the norm $\nabla \|x\|_p = J_p(x) = \|x\|_p^{1-p} \Phi_p(x)$, with $\Phi_p(x)$ entrywise defined as $\Phi_p(x)_i = |x_i|^{p-2} x_i$. Let p^* be the dual exponent such that $1/p + 1/p^* = 1$. As $J_{p^*}(J_p(x)) = x/\|x\|_p$ for all $x \neq 0$ and $J_p(\lambda x) = J_p(x)$ for any coefficient $\lambda > 0$, we have that $\nabla f_A(x) = 0$ if and only if $J_{q^*}(A^T J_p(Ax)) = x/\|x\|_q$. Thus, x with $\|x\|_q = 1$ is a critical point of $f_A(x)$ if and only if it is a fixed point of the map $J_{q^*}(A^T J_p(Ax))$. The associated fixed point iteration

$$x_0 = x_0 / \|x_0\|_q, \quad x_{k+1} = J_{q^*}(A^T J_p(Ax_k)) \quad \text{for } k = 0, 1, 2, 3, \dots \quad (2)$$

defines what we call (nonlinear) power method for $\|A\|_{q \rightarrow p}$.

Although, in practice, the method applied to $\|A\|_{p \rightarrow p}$ for $p = 1, \infty$ often seems to converge to the global maximum (see e.g. [24]), no guarantees exist for the general case. For differentiable ℓ^p norms and nonnegative matrices, instead, conditions can be established in order to guarantee that the power iterates always converge to a global maximizer of f_A . The idea is that when the power method is started in the positive orthant then, provided A has an appropriate non-zero pattern, each iterate of the method will stay in this orthant until convergence. Then, a nonlinear Perron–Frobenius type result is proved to guarantee that there exists only one critical point of f_A in this region and this point is a global maximizer of f_A . While this idea was already known by Perron himself in the Euclidean ℓ^2 case, to our knowledge, the first version of this result for norms different than the Euclidean norm, has been proved by Boyd [4]. However, Boyd did not prove the uniqueness of positive critical points but only that they are global maximizer of f_A under the assumption that $A^T A$ is irreducible and $1 < p \leq q < \infty$. This work is then revisited by Bhaskara and Vijayaraghavan [3] who proved uniqueness for positive matrices A and $1 < p \leq q < \infty$. Independently Friedland, Gaubert and Han proved in [15] similar results for $1 < p \leq 2 \leq q < \infty$ and any nonnegative A such that the matrix $\begin{bmatrix} 0 & A \\ A^T & 0 \end{bmatrix}$ is irreducible. Their result was then extended to $1 < p \leq q < \infty$ in [18] under the assumption that $A^T A$ is irreducible. Finally, all these results have been improved in [22], leading to the following

Theorem 1 (Theorems 3.2 and 3.3, [22]) *Let $A \in \mathbb{R}^{m \times n}$ be a matrix with nonnegative entries and suppose that $A^T A$ has at least one positive entry per row. If $1 < p \leq q < \infty$, then, every positive critical point of f_A is a global maximizer. Moreover, if either $p < q$ or $A^T A$ is irreducible, then f_A has a unique positive critical point x^+ and the power sequence (2) converges to x^+ for every positive starting point.*

In this work we consider the case of a matrix norm defined in terms of arbitrary vector norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ and we prove Theorem 4 below, which is a new version of Theorem 1, holding for general vector norms, provided that suitable and mild differentiability and monotonicity conditions are satisfied. We stress that Theorems 1 and 4 are not corollaries of each other in the sense that there are cases where exactly one, both or none apply. However, when both apply, then Theorem 4 is more informative. We discuss in detail these discrepancies in Sect. 4.1 and give there examples to illustrate them. In particular, a noticeable difference is that, for positive matrices A , the newly proposed Theorem 4 ensures uniqueness and maximality for choices of $1 < p, q < \infty$ that include the range $p > q$. This is, to our knowledge, the first global optimality result for this problem that includes such range of values.

The key of our approach is the use of cone geometry techniques and the Birkhoff–Hopf theorem, which we recall below.

3 Cone–theoretic Background

We start by recalling concepts from conic geometry. Let $\mathbb{R}_+^n$ be the nonnegative orthant in $\mathbb{R}^n$, that is $x \in \mathbb{R}_+^n$ if $x_i \geq 0$ for every $i = 1, \dots, n$. The cone $\mathbb{R}_+^n$ induces a partial ordering on $\mathbb{R}^n$ as follows: for every $x, y \in \mathbb{R}^n$ we write $x \leq y$ if $y - x \in \mathbb{R}_+^n$, i.e. $x_i \leq y_i$ for every i . Furthermore, $x, y \in \mathbb{R}_+^n$ are comparable, and we write $x \sim y$, if there exist $c, C > 0$ such that $cy \leq x \leq Cy$. Clearly, $\sim$ is an equivalence relation and the equivalence classes in $\mathbb{R}_+^n$ are called the parts of $\mathbb{R}_+^n$. For example, if $n = 2$ and $x = (1, 0)$, then the equivalence class of x in $\mathbb{R}_+^2$ is given by $\{(y_1, 0) : y_1 > 0\}$.

For simplicity, from now on we will say that a vector is nonnegative (resp. positive) if its entries are nonnegative (resp. positive). The same nomenclature will be used for matrices.

We recall that a norm $\|\cdot\|$ on $\mathbb{R}^n$ is monotonic if for every $x, y \in \mathbb{R}^n$ such that $|x| \leq |y|$, where the absolute value is taken componentwise, it holds $\|x\| \leq \|y\|$ and it is strongly monotonic if for every $x, y \in \mathbb{R}^n$ with $|x| \neq |y|$ and $|x| \leq |y|$ it holds $\|x\| < \|y\|$.

One of the key tools for our main result is the Hilbert’s projective metric $d_H : \mathbb{R}_+^n \times \mathbb{R}_+^n \rightarrow [0, \infty]$, defined as follows:

$$d_H(x, y) = \begin{cases} \ln(M(x/y)M(y/x)) & \text{if } x \sim y, \\ 0 & \text{if } x = y = 0, \\ \infty, & \text{otherwise} \end{cases}$$

where $M(x/y) = \inf\{C > 0 : x \leq Cy\}$. We collect in the following lemma some useful properties of d_H . Most of these results are known and can be found in [31]. Moreover, similarly to what is observed in Theorem 3 of [20], we prove a direct relation between the infinity norm and the Hilbert metric, which is useful for deriving explicitly computable convergence rates for the power method.

Lemma 1 *For every $x, y \in \mathbb{R}_+^n$, it holds $d_H(x, y) = 0$ if and only if $x = \lambda y$ for some $\lambda > 0$ and $d_H(cx, \tilde{c}y) = d_H(x, y)$ for every $c, \tilde{c} > 0$. Moreover, let $\|\cdot\|$ be a monotonic norm on $\mathbb{R}^n$, P a part of $\mathbb{R}_+^n$ and define $\mathbb{M} = P \cap \{x \in \mathbb{R}_+^n : \|x\| = 1\}$. Then, $(\mathbb{M}, d_H)$ is a complete metric space and*

$$\|x - y\|_\infty \leq r d_H(x, y) \quad \forall x, y \in \mathbb{M}, \quad (3)$$

where $r = \inf\{t > 0 : x_i \leq t \forall x \in \mathbb{M}, i = 1, \dots, n\}$.

Proof Proposition 2.1.1 in [31] implies that $d_H(x, y) = 0$ if and only if $x = \lambda y$ and that $(\mathbb{M}, d_H)$ is a metric space. The property $d_H(cx, \tilde{c}y) = d_H(x, y)$ for every $c, \tilde{c} > 0$ follows directly from the definition of d_H . The completeness of $(\mathbb{M}, d_H)$ is a consequence of Proposition 2.5.4 in [31]. We prove (3). If $P = \{0\}$, the result is trivial so we assume $P \neq \{0\}$ and let $i_1, \dots, i_m$ be such that for any $z \in \mathbb{R}_+^n$, $z \in P$ if and only if $z_{i_1}, \dots, z_{i_m} > 0$. Let $x, y \in \mathbb{M}$, then $x \leq M(x/y)y$ and, by monotonicity of $\|\cdot\|$, it follows $1 = \|x\| \leq M(x/y)\|y\| = M(x/y)$. Similarly $M(y/x) \geq 1$, so that $M(x/y)M(y/x) \geq \max\{M(x/y), M(y/x)\}$. It follows that

$$d_H(x, y) \geq \ln(\max\{M(x/y), M(y/x)\}) = \|\bar{x} - \bar{y}\|_\infty,$$

where $\bar{x} = (\ln(x_{i_1}), \dots, \ln(x_{i_m}))$ and $\bar{y} = (\ln(y_{i_1}), \dots, \ln(y_{i_m}))$. By definition of $r > 0$, we have $\ln(x_{i_j}), \ln(y_{i_j}) \in (-\infty, \ln(r)]$ for every $j = 1, \dots, m$. Furthermore, by the mean value theorem, we have

$$|e^s - e^t| \leq |s - t| \max_{\xi \in (-\infty, \ln(r)]} e^\xi = r|s - t| \quad \forall s, t \in (-\infty, \ln(r)].$$

Finally, with $\tilde{x} = (x_{i_1}, \dots, x_{i_m})$ and $\tilde{y} = (y_{i_1}, \dots, y_{i_m})$, we obtain

$$d_H(x, y) \geq \|\bar{x} - \bar{y}\|_\infty \geq r^{-1} \|\tilde{x} - \tilde{y}\|_\infty = r^{-1} \|x - y\|_\infty$$

which concludes the proof. $\square$

Observe that if r is defined as in Lemma 1 and $\|\cdot\|$ is strongly monotonic, then

$$r \leq \tilde{r} = \max_{i=1, \dots, n} \frac{1}{\|e_i\|}. \quad (4)$$

Indeed, if $y \in \mathbb{M}$ is such that there exists $j \in \{1, \dots, n\}$ with $y_j > \tilde{r}$, then $1 = \|y\| > \|\tilde{r}e_j\| = \tilde{r}\|e_j\|$, which is not possible.

The proof of our main theorem is based on the Banach contraction principle. Thus, for a map $F: \mathbb{R}_+^n \rightarrow \mathbb{R}_+^n$ we consider the Birkhoff contraction ratio $\kappa_H(F) \in [0, \infty]$ of F , defined as the smallest Lipschitz constant of F with respect to d_H :

$$\kappa_H(F) = \inf \{C > 0 : d_H(F(x), F(y)) \leq Cd_H(x, y), \forall x, y \in \mathbb{R}_+^n \text{ s.t. } x \sim y\}.$$

Clearly, if there exist $x, y \in \mathbb{R}_+^n$ such that $x \sim y$ and $F(x) \not\sim F(y)$, then $\kappa_H(F) = \infty$. However, such a situation never happens when F is a linear map in which case $\kappa_H(F) \leq 1$ always holds. Indeed, if $A \in \mathbb{R}^{m \times n}$ is a nonnegative matrix, $x, y \in \mathbb{R}_+^n$ and $x \sim y$, then $x \leq M(x/y)y$ implies $Ax \leq M(x/y)Ay$. Similarly, we have $Ay \leq M(y/x)Ax$ and thus $Ax \sim Ay$. These inequalities also imply that $\kappa_H(A) \leq 1$. This upper bound is not tight in many cases. However, thanks to the Birkhoff–Hopf theorem, a better estimate of $\kappa_H(A)$ can be obtained by computing the projective diameter $\Delta(A) \in [0, \infty]$ of A , defined as

$$\Delta(A) = \sup \{d_H(Ax, Ay) : x, y \in \mathbb{R}_+^n \text{ with } x \sim y\}. \quad (5)$$

This is formalized in the following theorem whose proof can be found in Theorems 3.5 and 3.6 of [12].

Theorem 2 (Birkhoff–Hopf) *Let $A \in \mathbb{R}^{m \times n}$ be a matrix with nonnegative entries, then*

$$\kappa_H(A) = \tanh(\Delta(A)/4),$$

where $\tanh(t) = (e^{2t} - 1)/(e^{2t} + 1)$ and with the convention $\tanh(\infty) = 1$.

The above theorem is particularly useful when combined with the following Theorem 6.2 in [12] and Theorem 3.12 in [37]:

Theorem 3 *Let $A \in \mathbb{R}^{m \times n}$ be a matrix with nonnegative entries and $e_1, \dots, e_n$ the canonical basis of $\mathbb{R}^n$. If there exists $\mathcal{I} \subset \{1, \dots, n\}$ such that $Ae_i \sim Ae_j$ for all $i, j \in \mathcal{I}$ and $Ae_i = 0$ for all $i \notin \mathcal{I}$, then*

$$\Delta(A) = \max_{i,j \in \mathcal{I}} d_H(Ae_i, Ae_j) < \infty.$$

In particular, if all the entries of A are positive, then $\Delta(A) = \ln \left(\max_{i,j,k,l} \frac{a_{ki} a_{lj}}{a_{kj} a_{li}} \right)$ and $\Delta(A) = \Delta(A^T)$. Moreover, if A has at least one positive entry per row and per column but A is not positive, then $\Delta(A) = \infty$.

Unfortunately, such simple formulas for the Birkhoff contraction ratio are, to our knowledge, not known for general nonlinear mappings. We refer however to Corollary 2.1 in [35] and Corollary 3.9 in [17] for general characterizations of this ratio.

4 Nonlinear Perron–Frobenius Theorem for $\|A\|_{\beta \rightarrow \alpha}$

Given $A \in \mathbb{R}^{m \times n}$, consider the matrix norm $\|A\|_{\beta \rightarrow \alpha} = \max_{x \neq 0} \|Ax\|_{\alpha} / \|x\|_{\beta}$, where $\|\cdot\|_{\alpha}$ and $\|\cdot\|_{\beta}$ are arbitrary vector norms on $\mathbb{C}^m$ and $\mathbb{C}^n$, respectively. Then, as for the case of ℓ^p norms, consider the function

$$f_A(x) = \frac{\|Ax\|_{\alpha}}{\|x\|_{\beta}}. \quad (6)$$

For an arbitrary possibly non-differentiable vector norm $\|\cdot\|$ it holds ([13, e.g.]

$$\partial \|x\| = \{y : \langle y, x \rangle = \|x\|, \|y\|_* = 1\}, \quad (7)$$

where ∂ denotes the subdifferential and $\|\cdot\|_*$ is the dual norm of $\|\cdot\|$, defined as $\|y\|_* = \max_{x \neq 0} \langle x, y \rangle / \|x\|$. Again, for notational convenience, given the vector norm $\|x\|_{\alpha}$, we introduce the set-valued operator J_{α} such that

$$J_{\alpha}(x) = \partial \|x\|_{\alpha}, \quad \forall x \neq 0 \quad \text{and} \quad J_{\alpha}(0) = 0.$$

The definition of dual norm implies the generalized Hölder inequality $\langle x, y \rangle \leq \|x\| \|y\|_*$. Thus, for a vector x and a norm $\|\cdot\|_{\alpha}$, the set of vectors $J_{\alpha}(x)$ coincides with the set of vectors in the unit sphere of the dual norm of $\|\cdot\|_{\alpha}$, for which equality holds in the Hölder inequality. In fact, the subdifferential of a norm J_{α} is strictly related with the duality mapping $\mathcal{J}_{\alpha}$ induced by that norm. Precisely, by Asplund's theorem (see e.g. [7]), we have that

$$\mathcal{J}_{\alpha}(x) = \frac{1}{2} \partial \|x\|_{\alpha}^2 = \|x\|_{\alpha} J_{\alpha}(x). \quad (8)$$

It is well known that the subgradient of a convex function f is single valued if and only if f is Fréchet differentiable. Therefore J_{α} is single valued if and only if $\|\cdot\|_{\alpha}$ is a Fréchet differentiable norm. The assumption that the duality maps involved are single valued will be crucial for our main result. For this reason, throughout we make the following assumptions on the norms we are considering:

Assumption 1 The norms $\|\cdot\|_{\alpha}$ and $\|\cdot\|_{\beta}$ we consider are such that

1. $\|\cdot\|_\alpha$ is Fréchet differentiable.
2. The dual norm $\|\cdot\|_{\beta^*}$ is Fréchet differentiable.
3. Both $\|\cdot\|_\alpha$ and $\|\cdot\|_{\beta^*}$ are strongly monotonic.

Remark 1 Recall that every monotonic norm $\|\cdot\|$ is also absolute (see e.g. [28, Thm. 1]), that is $\| |x| \| = \|x\|$ for every x , where $|x|$ denotes the entrywise absolute value. This implies, in particular, that a monotonic norm is Fréchet differentiable at every $x \in \mathbb{R}^n \setminus \{0\}$ if and only if it is Fréchet differentiable at every $x \in \mathbb{R}_+^n \setminus \{0\}$.

Points (1) and (1) of Assumption 1 ensure that the following nonlinear mapping

$$S_A(x) = J_{\beta^*}(A^T J_\alpha(Ax)) \quad (9)$$

is single valued. Point (1) ensures that for nonnegative matrices the maximum of f_A is attained on a nonnegative vector and that if $A^T A$ is irreducible, then this maximizer has positive entries. Overall, they allow us to prove the following fundamental preliminary Lemmas 2–6.

First, we discuss the critical points of f_A . If $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ satisfy Assumption 1, then f_A may not be differentiable. Indeed, the differentiability of $\|\cdot\|_{\beta^*}$ does not imply that of $\|\cdot\|_\beta$ (see for instance [7, Chapter II]). Hence, in the following, we use Clarke's generalized gradient [8] to discuss the critical points of f_A . In particular, let us recall that, by [8, Prop. 2.2.7], the generalized gradient of a convex function coincides with its subgradient. Moreover, it can be verified that f_A is locally Lipschitz near every $x \in \mathbb{R}^n \setminus \{0\}$ so that its generalized gradient $\partial f_A(x) \subset \mathbb{R}^n$ is well defined and x is a critical point of f_A if $0 \in \partial f_A(x)$. Moreover, if f_A attains a local minimum or maximum at $x \neq 0$, then $0 \in \partial f_A(x)$ by [8, Prop. 2.3.2].

Lemma 2 Let $\|\cdot\|_\alpha, \|\cdot\|_\beta$ satisfy Assumption 1 and let $x \in \mathbb{R}_+^n$ with $\|x\|_\beta = 1$ and $f_A(x) \neq 0$. If x is a critical point of f_A , then it is a fixed point of S_A . Conversely, if x is a fixed point of S_A and $\|\cdot\|_\beta$ is differentiable, then x is a critical point of f_A .

Proof First, assume that $0 \in \partial f_A(x)$. As $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are Lipschitz functions and $\|x\|_\beta = 1$, Proposition 2.3.14 of [8] implies that

$$\partial f_A(x) \subset A^T J_\alpha(Ax) - f_A(x) J_\beta(x). \quad (10)$$

J_α is single valued since $\|\cdot\|_\alpha$ is differentiable. Hence, $0 \in \partial f_A(x)$ implies that $f_A(x)^{-1} A^T J_\alpha(Ax) \in J_\beta(x)$. Now, as $\|\cdot\|_{\beta^*}$ is differentiable, we have that, for the duality mapping $\mathcal{J}_\beta$, it holds $y \in \mathcal{J}_\beta(x)$ if and only if $x = \mathcal{J}_{\beta^*}(y)$ (c.f. [7, Prop. 4.7]). It follows, with (8), that $\lambda J_{\beta^*}(A^T J_\alpha(Ax)) = x$ with $\lambda > 0$. Finally, as $\|J_{\beta^*}(A^T J_\alpha(Ax))\|_\beta = 1 = \|x\|_\beta$, we have $\lambda = 1$ which implies that $S_A(x) = x$.

Now, suppose that x is a fixed point of S_A . Then, we have $J_\beta(S_A(x)) = J_\beta(x)$. Again, by [7, Prop. 4.7] and (8), we deduce the existence of $\lambda > 0$ such that $\lambda A^T J_\alpha(Ax) \in J_\beta(x)$. The definition of J_β implies that $\langle x, \lambda A^T J_\alpha(Ax) \rangle = \|x\|_\beta = 1$ and thus $\lambda^{-1} = \langle Ax, A J_\alpha(Ax) \rangle = f_A(x)$. It follows that $0 \in A^T J_\alpha(Ax) - f_A(x) J_\beta(x)$. If $\|\cdot\|_\beta$ is differentiable, then f_A is differentiable at x and the sets in (10) are equal (and singletons). It follows that $0 \in \partial f_A(x)$. $\square$

Lemma 3 Let $A \in \mathbb{R}^{m \times n}$ be a matrix with nonnegative entries and P , a part of $\mathbb{R}_+^n$ such that $A^T Ax \in P$ for every $x \in P$. Furthermore, let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ satisfy Assumption 1. If $\kappa_H(S_A) \leq \tau < 1$, then S_A has a unique fixed point z in P and for every positive integer k and every $x \in P$, it holds

$$\|S_A^k(x) - z\|_\infty \leq \tau^k (r/(1-\tau)) d_H(x, S_A(x)),$$

where $r = \inf\{t > 0 : x_i \leq t \ \forall i = 1, \dots, n, \ x \in P, \|x\|_\beta = 1\}$.

Proof By assumption $\mathcal{S}_A$ is a strict contraction on the metric space $(\mathbb{M}, d_H)$ where $\mathbb{M} = P \cap \{x \in \mathbb{R}_+^n : \|x\|_\beta = 1\}$. As $(\mathbb{M}, d_H)$ is complete by Lemma 1, it follows from the Banach fixed point theorem (see for instance Theorem 3.1 in [29]) that $\mathcal{S}_A$ has a unique fixed point z in $\mathbb{M}$ and for every $y \in \mathbb{M}$ it holds

$$d_H(\mathcal{S}_A^k(y), z) \leq \frac{\tau^k}{1 - \tau} d_H(y, \mathcal{S}_A(y)).$$

As $\mathcal{S}_A(\lambda y) = \mathcal{S}_A(y)$ and $d_H(\lambda y, \mathcal{S}_A(y)) = d_H(y, \mathcal{S}_A(y))$ for every $\lambda > 0$, the convergence rate is a direct consequence of the above inequality and Lemma 1. $\square$

We remark that this result does not guarantee that the unique fixed point z of $\mathcal{S}_A$ in P is a global maximizer of f_A and in fact this is not always true. Indeed, if A is a 2×2 diagonal matrix which is not a multiple of the identity and $\|\cdot\|_\alpha = \|\cdot\|_2$, $\|\cdot\|_\beta = \|\cdot\|_3$, then $\kappa_H(\mathcal{S}_A) \leq 1/2$ and $\mathcal{S}_A$ leaves all the parts of $\mathbb{R}_+^2$ invariant but some of them do not contain a global maximizer of f_A . Moreover, as $\mathbb{R}_+^n$ has 2^n parts, testing each part of the cone is computationally too expensive for large n . Therefore, in the remaining part of the section, we derive conditions in order to ensure that the power iterates converge to a global maximizer of f_A .

Lemma 4 Let $A \in \mathbb{R}^{m \times n}$ be a matrix with nonnegative entries and let $\|\cdot\|_\alpha, \|\cdot\|_\beta$ satisfy Assumption 1. Then it holds $f_A(x) \leq f_A(|x|)$ for any $x \in \mathbb{C}^n \setminus \{0\}$ and the maximum of f_A is attained in $\mathbb{R}_+^n$.

Proof Let $x \neq 0$, since A has nonnegative entries, it holds $|Ax| \leq A|x|$. Thus, as monotonic norms are also absolute, we have

$$f_A(x) = \frac{\|Ax\|_\alpha}{\|x\|_\beta} = \frac{\| |Ax| \|_\alpha}{\| |x| \|_\beta} \leq \frac{\|A|x|\|_\alpha}{\| |x| \|_\beta} = f_A(|x|).$$

Now, if y is a global maximizer of f_A , then $f_A(y) \leq f_A(|y|) \leq f_A(y)$ which concludes the proof. $\square$

In the forthcoming Lemma 6, we use the strong monotonicity required in Point (1) of Assumption 1 to prove that if $A^T A$ is irreducible, then the nonnegative maximizer of Lemma 4 has positive entries. To this end, however, we need one additional preliminary result that characterizes strongly monotonic norms in terms of the zero pattern of J and which we prove in the following:

Lemma 5 Let $\|\cdot\|_\gamma$ be a differentiable monotonic norm on $\mathbb{R}^n$, then $\|\cdot\|_\gamma$ is strongly monotonic if and only if $x \sim J_\gamma(x)$ for every $x \in \mathbb{R}_+^n$.

Proof Suppose that $\|\cdot\|_\gamma$ is strongly monotonic. Let $x \in \mathbb{R}_+^n$. If $x = 0$, $J_\gamma(0) = 0$ by construction. Suppose that $x \neq 0$. We use the strong monotonicity to prove the existence of $c > 0$ such that $c x \leq J_\gamma(x)$. Let i be such that $x_i > 0$ and define $f(t) = \|x + (t - x_i)e_i\|_\gamma$ for all $t > 0$. Then, f is differentiable and $f'(t) = J_\gamma(x + (t - x_i)e_i)_i$ for all $t > 0$. Furthermore, f is strictly increasing on $(0, \infty)$ since $\|\cdot\|$ is strongly monotonic. It follows that $J_\gamma(x)_i = f'(x_i) > 0$. As this is true for all i such that $x_i > 0$, we conclude that there exists $c > 0$ such that $c x \leq J_\gamma(x)$. The existence of $C > 0$ such that $J_\gamma(x) \leq C x$ follows from Proposition 5.2 of [7, Chapter 1]. Hence, we have $J_\gamma(x) \sim x$.

For the reverse implication, suppose that $J_\gamma(x) \sim x$ for all $x \in \mathbb{R}_+^n$. Let $x, y \in \mathbb{R}_+^n$ be such that $x \leq y$ and $x \neq y$. If $x = 0$, then $\|x\|_\gamma = 0 < \|y\|_\gamma$. Suppose that $x \neq 0$. As $x \leq y$

and $x \neq 0$, there exists i and $t_0 > 0$ such that $x + te_i \leq y$ for all $t \in (0, t_0)$. For $t \in (0, t_0)$, we have

$$\begin{aligned} \|y\|_\gamma &\geq \|x + \tfrac{1}{2}(t_0 + t)e_i\|_\gamma \\ &\geq \|x + \tfrac{t_0}{2}e_i\|_\gamma + \langle J_\gamma(x + \tfrac{t_0}{2}e_i), \tfrac{t}{2}e_i \rangle \geq \|x\|_\gamma + \tfrac{t}{2}J_\gamma(x + \tfrac{t_0}{2}e_i)_i, \end{aligned}$$

where the second inequality follows from the convexity of $\|\cdot\|_\gamma$. By assumption, we have $J_\gamma(x + \frac{t_0}{2}e_i) \sim x + \frac{t_0}{2}e_i$ and thus $J_\gamma(x + \frac{t_0}{2}e_i)_i > 0$. It follows that $\|y\|_\gamma > \|x\|_\gamma$, i.e. $\|\cdot\|_\gamma$ is strongly monotonic. $\square$

Lemma 6 *Let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ satisfy Assumption 1. Let A be a matrix with nonnegative entries and suppose that $A^T A$ is irreducible. Then, $S_A(x)$ is positive for every positive x and every nonnegative critical point of f_A is positive.*

Proof Lemma 5 implies that $S_A(x) \sim A^T A x$. It follows that S_A maps positive vectors to positive vectors since the irreducibility of $A^T A$ implies that $A^T A$ is positive for all positive x . Finally, note that $A^T A$ is symmetric positive semi-definite and therefore all its eigenvalues are nonnegative. It follows that $A^T A$ is primitive (see e.g. Theorem 1 in [40]). By the same theorem, there exists a positive integer k such that $(A^T A)^k$ is a matrix with positive entries. Since $S_A^k(x) \sim (A^T A)^k x$ for every $x \in \mathbb{R}_+^n \setminus \{0\}$, we deduce that $S_A^k(x)$ is strictly positive for every nonzero, nonnegative x . Finally, suppose that $y \in \mathbb{R}_+^n$ is a critical point of f_A , then y is a fixed point of S_A by Lemma 2 and thus $y = S_A^k(y)$ is strictly positive. $\square$

We are now ready to state our main theorem of this section. This theorem provides conditions on A , $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ that ensure the existence of a unique positive maximizer x^+ such that $\|Ax^+\|_\beta/\|x^+\|_\alpha = \|A\|_{\beta \rightarrow \alpha}$ and that govern the convergence of the power sequence

$$x_0 = x_0/\|x_0\|_\beta, \quad x_{k+1} = J_{\beta^*}(A^T J_\alpha(Ax_k)) \quad \text{for } k = 0, 1, 2, 3, \dots \quad (11)$$

to such x^+ . As announced, this result is essentially a fixed point theorem for S_A and thus the Birkhoff contraction ratio $\kappa_H(S_A)$ and any τ that well-approximate $\kappa_H(S_A)$ from above play a central role.

Theorem 4 *Let $A \in \mathbb{R}^{m \times n}$ be a matrix with nonnegative entries and suppose that $A^T A$ is irreducible. Let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ satisfy Assumption 1.*

If $\kappa_H(S_A) \leq \tau < 1$, then:

1. *f_A has a unique critical point x^+ in $\mathbb{R}_+^n$. Moreover, $f_A(x^+) = \|A\|_{\beta \rightarrow \alpha}$ and x^+ is positive.*
2. *If x_0 is positive and $x_{k+1} = S_A(x_k)$ is the power sequence, then*

$$\|x_k - x^+\|_\infty \leq \tau^k C \quad \text{with } C = \max_{i=1, \dots, n} \frac{d_H(x_0, x_1)}{(1 - \tau)\|e_i\|_\beta}$$

where $e_1, \dots, e_n$ is the canonical basis of $\mathbb{R}^n$. Furthermore, it holds

$$(1 - \tau^k \tilde{C})\|A\|_{\beta \rightarrow \alpha} \leq \|Ax_k\|_\alpha \leq \|A\|_{\beta \rightarrow \alpha}$$

with $\tilde{C} = C \max_{x \neq 0} \frac{\|x\|_\alpha}{\|x\|_\infty}$. In particular, $x_k \rightarrow x^+$ as $k \rightarrow \infty$.

Proof Lemma 4 implies that f_A has a maximizer $x^+ \in \mathbb{R}_+^n$. Lemma 6 implies that x^+ is positive and that the interior of $\mathbb{R}_+^n$ is left invariant by S_A . Hence, all statements except the bounds on $\|Ax_k\|_\alpha$ follow by a direct application of Lemma 3 and Eq. (4). We conclude with

a proof of the estimates for $\|Ax_k\|_\alpha$. Clearly, $\|Ax_k\|_\alpha \leq \|A\|_{\beta \rightarrow \alpha}$ always hold. For the lower bound, let $\gamma = \max_{x \neq 0} \frac{\|x\|_\beta}{\|x\|_\infty}$. The estimate on $\|x_k - x^+\|_\infty$ implies that

$$\begin{aligned}\|A\|_{\beta \rightarrow \alpha} - \|Ax_k\|_\alpha &= \|Ax^+\|_\alpha - \|Ax_k\|_\alpha \leq \|A(x^+ - x_k)\|_\alpha \\ &\leq \|A\|_{\beta \rightarrow \alpha} \|x^+ - x_k\|_\beta \leq \gamma \|A\|_{\beta \rightarrow \alpha} \|x^+ - x_k\|_\infty \leq \tau^k C \gamma \|A\|_{\beta \rightarrow \alpha}\end{aligned}$$

which concludes the proof. $\square$

Note that the condition that requires $A^T A$ to be irreducible is in general weaker than requiring the initial matrix A to be irreducible itself, as $A^T A$ may be irreducible even if A is reducible. This is also observed in the numerical examples in Sect. 6.

Theorem 4 holds for any upperbound τ of $\kappa_H(S_A)$ and a somewhat natural choice for such a τ is the following

$$\tau(S_A) = \kappa_H(A^T) \kappa_H(J_{\beta^*}) \kappa_H(A) \kappa_H(J_\alpha). \quad (12)$$

This coefficient is particularly useful in practice as, thanks to the Birkhoff–Hopf theorem, in many circumstances one can provide explicit bounds for $\tau(S_A)$. Although in principle $\tau(S_A)$ can be larger than $\kappa_H(S_A)$, in the forthcoming Sect. 4.2 we show that there are cases where the equality $\tau(S_A) = \kappa_H(S_A)$ holds. Moreover, we discuss the sharpness of the condition $\kappa_H(S_A) < 1$ required by our main result. In the following Sect. 4.1, instead, we discuss the particular case where $\|\cdot\|_\alpha, \|\cdot\|_\beta$ are ℓ^p norms and we give examples showing how Theorem 4 improves the existing theory for this problem.

4.1 Examples and Comparison with Previous Work

When $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ are ℓ^p norms, Theorem 4 implies the following:

Corollary 1 *Let $A \in \mathbb{R}^{m \times n}$ be a matrix with nonnegative entries and suppose that $A^T A$ is irreducible. Let $1 < p, q < \infty$ and consider*

$$\|A\|_{q \rightarrow p} = \max_{x \neq 0} \frac{\|Ax\|_p}{\|x\|_q}, \quad \text{and} \quad \tau = \kappa_H(A) \kappa_H(A^T) \frac{p-1}{q-1}.$$

If $\tau < 1$, then $\|A\|_{q \rightarrow p}$ can be approximated to an arbitrary precision with the fixed point iteration (2).

In the case of ℓ^p norms, both Theorem 1 and Corollary 1 apply. In order to compare them let us compute the Birkhoff contraction ratio for some simple but explanatory cases. Let $\varepsilon \geq 0$ and $A \in \mathbb{R}^{3 \times 2}$, $B \in \mathbb{R}^{2 \times 2}$, $C \in \mathbb{R}^{3 \times 3}$ be defined as

$$A = \begin{bmatrix} 1 & 2 \\ 3 & 4 \\ 0 & 0 \end{bmatrix}, \quad B = \begin{bmatrix} \varepsilon & 1 \\ 1 & \varepsilon \end{bmatrix}, \quad C = \begin{bmatrix} 0 & 1 & 1 \\ 2 & 2 & 2 \\ 3 & 3 & 0 \end{bmatrix}.$$

Due to Theorem 3, it is easy to see that

$$\begin{aligned}\kappa_H(A) &= \tanh(3/8) \leq 9/25, \\ \kappa_H(A^T) &= \tanh(1/16) \leq 1/16, \\ \kappa_H(B) &= \kappa_H(B^T) = (1 - \varepsilon)^2 / (1 + \varepsilon)^2 \\ \kappa_H(C) &= \kappa_H(C^T) = 1.\end{aligned}$$

Fig. 1 Each line shows the distribution of $\kappa_H(A)$ over 1000 random matrices $A \in \mathbb{R}_+^{10 \times 10}$ with entries between k and 10. Different curves correspond to different values of $k \in \{1, 2, \dots, 5\}$

Note that $A^T A$ and $C^T C$ are positive matrices and $B^T B$ is positive if and only if $\varepsilon > 0$. If $\varepsilon = 0$, then $B^T B$ is the identity matrix. We first discuss the implications of Theorem 1 for the computation of $\|X\|_{q \rightarrow p}$ where $X \in \{A, B, C\}$.

If $p \leq q$ and $\varepsilon > 0$, then Theorem 1 implies that f_X has a unique positive maximizer x^+ , which is global, and the power sequence (11) will converge to x^+ . However, if $\varepsilon = 0$ then Theorem 1 ensures that every positive critical point of f_B is a global maximizer but uniqueness and convergence are only guaranteed under the assumption $p < q$. Now, we look at the implications of Theorem 4. By noting that $\kappa_H(J_p) = p - 1$ and $\kappa_H(J_{q^*}) = 1/(q - 1)$, we have

$$\tau(S_A) \leq \frac{9}{400} \frac{p-1}{q-1}, \quad \tau(S_B) = \left(\frac{1-\varepsilon}{1+\varepsilon} \right)^2 \frac{p-1}{q-1}, \quad \tau(S_C) = \frac{p-1}{q-1}.$$

Hence, for instance, uniqueness and global maximality of a positive maximizer of f_A is guaranteed by Theorem 4 under the assumption $9(p-1) < 400(q-1)$ which includes the known global convergence range of values $p < q$, but is of course a much weaker assumption.

Now, note that for $\varepsilon \geq 1$ we have $\tau(S_B) < 1$ if and only if $(\varepsilon-1)^2(p-1) < (\varepsilon+1)^2(q-1)$. This assumption is less restrictive than $p \leq q$ for every $\varepsilon \geq 1$ as $p \leq q$ correspond to the asymptotic case $\varepsilon \rightarrow \infty$. If $\varepsilon = 1$, Theorem 4 applies for every $1 < p, q < \infty$. The analysis for $0 < \varepsilon < 1$ is similar. However, we note that if $\varepsilon = 0$, then Theorem 4 does not provide any information about f_B for the case $p = q$ in contrast with Theorem 1. When $\varepsilon = 0$ and $p < q$, both theorems imply the same result. Finally, note that $\tau(S_C) < 1$ if and only if $p < q$ and so Theorem 1 is more useful as it also covers the case $p = q$.

More in general, when the considered matrix A has finite projective diameter $\Delta(A)$, then Theorem 2 implies that $\kappa_H(A) < 1$ and thus Theorem 4 ensures that for any $p > 1$, the matrix norm $\|A\|_{q \rightarrow p}$ can be approximated in polynomial time to an arbitrary precision for any choice of $q > \kappa_H(A)^2(p-1) + 1$, without the requirement $q > p$.

Figure 1 shows that the value of $\kappa_H(A)$ for matrices with positive entries is often substantially smaller than one, enhancing the relevance of Theorem 4.

4.2 On the Sharpness of the New Convergence Condition

As we observed earlier, the key property behind the global convergence of the power iterates relies on the fact that, when $\kappa_H(S_A) < 1$, the mapping S_A has a unique positive fixed point x^+ . Due to Lemma 2, this is equivalent to observing that, in this case, x^+ is the unique

positive critical point of f_A , up to scalar multiples. In what follows we show that this is not anymore the case if $\kappa_H(S_A) > 1$. In particular, we limit our attention to the case of ℓ^p norms and we exhibit a one-parameter family of 2×2 positive and symmetric matrices A_ε for which a unique positive critical point of f_{A_ε} exists if and only if $\kappa_H(S_{A_\varepsilon}) \leq 1$. Moreover, we show that for such a family of matrices it holds $\tau(S_A) = \kappa_H(S_A)$ where $\tau(S_A)$ is the estimate of $\kappa_H(S_A)$ discussed in Eq. (12). As f_A is scale invariant, here and in the rest of this section, uniqueness of the critical point is meant up to scalar multiples.

For $\varepsilon > 0$ and $p, q \in (1, \infty)$, let $A_\varepsilon \in \mathbb{R}^{2 \times 2}$ and $f_{A_\varepsilon} : \mathbb{R}^2 \rightarrow \mathbb{R}_+$ be defined as

$$A_\varepsilon = \begin{bmatrix} \varepsilon & 1 \\ 1 & \varepsilon \end{bmatrix} \quad \text{and} \quad f_{A_\varepsilon}(x) = \frac{\|A_\varepsilon x\|_p}{\|x\|_q}. \quad (13)$$

The main result of this section is the following theorem, whose proof is postponed to the end of the section

Theorem 5 *It holds $\kappa_H(S_{A_\varepsilon}) = \tau(S_{A_\varepsilon})$. Furthermore, f_{A_ε} has a unique critical point in $\mathbb{R}_+^2$ if and only if $\tau(S_{A_\varepsilon}) \leq 1$.*

This result shows that, unlike the previous Theorem 1, Theorem 4 is tight in the sense that when $\kappa_H(S_A) > 1$ there might be multiple distinct fixed points of S_A in $\mathbb{R}_+^2$, and thus convergence of the power sequence to a prescribed fixed point cannot be ensured globally without restrictions on the starting point $x_0 \in \mathbb{R}_+^2$.

We subdivide the proof of Theorem 5 above into a number of preliminary results. Before proceeding, we recall that for $p \in (1, \infty)$, $\Phi_p : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is entrywise defined as $\Phi_p(x)_i = |x_i|^{p-2}x_i$ for all i . We compute $\tau(S_{A_\varepsilon})$ and $\kappa_H(S_A)$.

Lemma 7 *For every $\varepsilon > 0$, we have $\kappa_H(S_{A_\varepsilon}) = \tau(S_{A_\varepsilon}) = \left(\frac{1-\varepsilon}{1+\varepsilon}\right)^{\frac{p-1}{q-1}}$.*

Proof As $\kappa_H(A) = \left|\frac{\varepsilon-1}{1+\varepsilon}\right|$ by Theorem 3, we have $\tau(S_{A_\varepsilon}) = \left(\frac{1-\varepsilon}{1+\varepsilon}\right)^{\frac{p-1}{q-1}}$. Now, we show that $\kappa_H(S_{A_\varepsilon}) = \tau(S_{A_\varepsilon})$. Clearly, $\kappa_H(S_{A_\varepsilon}) \leq \tau(S_{A_\varepsilon})$, for the reverse inequality consider $x = (1, 1)^T$ and $y(t) = (1, t)^T$. Furthermore, define $h : (1, \infty) \rightarrow \mathbb{R}$ as

$$h(t) = \frac{d_H(S_{A_\varepsilon}(x), S_{A_\varepsilon}(y(t)))}{d_H(x, y(t))}.$$

Then, we have $h(t) \leq \kappa_H(S_{A_\varepsilon})$ for every $t > 0$. To conclude the proof, we show that $\lim_{t \rightarrow 1^+} h(t) = \tau(S_{A_\varepsilon})$. A direct computation shows that $d_H(x, y(t)) = \ln(t)$ and $A_\varepsilon \Phi_p(A_\varepsilon x) = (1 + \varepsilon)^p(1, 1)^T$. Recalling that $S_{A_\varepsilon}(z) = \Phi_{q^*}(A_\varepsilon \Phi_p(A_\varepsilon z))$, we have

$$d_H(S_{A_\varepsilon}(x), S_{A_\varepsilon}(y(t))) = (q^* - 1)d_H(A_\varepsilon \Phi_p(A_\varepsilon x), A_\varepsilon \Phi_p(A_\varepsilon y(t))).$$

So if we let $f_1, f_2 : (1, \infty) \rightarrow \mathbb{R}$ be such that $A_\varepsilon \Phi_p(A_\varepsilon y(t)) = (f_1(t), f_2(t))^T$ for all $t > 1$, we get

$$\exp\left((q-1)d_H(S_{A_\varepsilon}(x), S_{A_\varepsilon}(y(t)))\right) = \max\left\{\frac{f_1(t)}{f_2(t)}, \frac{f_2(t)}{f_1(t)}\right\}.$$

With

$$g(t) = \frac{f_1(t)}{f_2(t)} = \frac{\varepsilon(t + \varepsilon)^{p-1} + (t\varepsilon + 1)^{p-1}}{(t + \varepsilon)^{p-1} + \varepsilon(t\varepsilon + 1)^{p-1}},$$

the above computations, imply

$$(q-1) \lim_{t \rightarrow 1^+} h(t) = \lim_{t \rightarrow 1^+} \frac{\max\{\ln(g(t)), -\ln(g(t))\}}{\ln(t)} = \left| \lim_{t \rightarrow 1^+} \frac{\ln(g(t))}{\ln(t)} \right|,$$

where the last equality follows by continuity. As $\ln(1) = \ln(g(1)) = 0$, L'Hopital's rule implies that

$$\lim_{t \rightarrow 1^+} \frac{\ln(g(t))}{\ln(t)} = \lim_{t \rightarrow 1^+} \frac{t g'(t)}{g(t)} = \lim_{t \rightarrow 1^+} - \frac{(p-1)t(\varepsilon^2-1)^2(t+\varepsilon)^p(t\varepsilon+1)^p}{\zeta_1(t)\zeta_2(t)}$$

where

$$\zeta_1(t) = (t\varepsilon^2(t+\varepsilon)^p + t(t\varepsilon+1)^p + \varepsilon((t+\varepsilon)^p + (t\varepsilon+1)^p))$$

and

$$\zeta_2(t) = (\varepsilon^2(t\varepsilon+1)^p + (t+\varepsilon)^p + t\varepsilon((t+\varepsilon)^p + (t\varepsilon+1)^p)).$$

As $\zeta_1(1)\zeta_2(1) = (1+\varepsilon)^{2p}(1+\varepsilon)^4$, after rearrangement, we finally obtain

$$\lim_{t \rightarrow 1^+} h(t) = \left| \frac{(p-1)(\varepsilon^2-1)^2(1+\varepsilon)^{2p}}{(q-1)\zeta_1(1)\zeta_2(1)} \right| = \tau(S_{A_\varepsilon}),$$

which implies $\tau(S_{A_\varepsilon}) \leq \kappa_H(S_{A_\varepsilon})$ and thus concludes the proof. $\square$

Now, we prove that the nonnegative critical points of f_{A_ε} are positive and we then characterize them in terms of a real parameter t . As critical points are defined up to multiples, we restrict our attention to the line $\{x \in \mathbb{R}^2 : x_1 + x_2 = 1\}$.

Lemma 8 *Let $x \in \mathbb{R}_+^2$ with $x_1 + x_2 = 1$. Then x is a critical point of f_{A_ε} if and only if there exists $t \in (0, 1)$ such that $x = (t, 1-t)^T$ and $\psi(t) = \psi(1-t)$ where $\psi: [0, 1] \rightarrow \mathbb{R}_+$ is defined as*

$$\psi(t) = t^{q-1}[(t\varepsilon+1-t)^{p-1} + \varepsilon(\varepsilon+t-t\varepsilon)^{p-1}]. \quad (14)$$

Proof As we already observed, f_{A_ε} attains a global maximum in $\mathbb{R}_+^2$. Furthermore, the critical points of f_{A_ε} satisfy

$$A_\varepsilon \Phi_p(A_\varepsilon x) = \lambda \Phi_q(x) \quad x \in \mathbb{R}^2 \setminus \{0\}. \quad (15)$$

As A_ε is positive, (15) implies that every nonnegative critical point of f_{A_ε} is positive. It follows that, for positive vectors x , (15) is equivalent to

$$\begin{cases} (A_\varepsilon \Phi_p(A_\varepsilon x))_1 x_2^{q-1} = (A_\varepsilon \Phi_p(A_\varepsilon x))_2 x_1^{q-1} \\ \lambda = (A_\varepsilon \Phi_p(A_\varepsilon x))_1 / x_1^{q-1} \end{cases} \quad (16)$$

Thus, $x_1 + x_2 = 1$ and $x_1, x_2 > 0$ imply the existence of $t \in (0, 1)$ such that $x_1 = t$ and $x_2 = 1-t$. Substituting $x = (t, 1-t)^T$ in (16) we finally obtain the claimed result. $\square$

A direct consequence of Lemma 8 is that $(1, 1)^T/2$ is a critical point of f_{A_ε} . Moreover, by symmetry, we see that $(t, 1-t)^T$ is a critical point of f_{A_ε} if and only if $(1-t, t)^T$ is also a critical point. This observation implies the following

Lemma 9 *If $\tau(S_{A_\varepsilon}) > 1$, then f_{A_ε} has at least three distinct positive critical points.*

Proof Note that if $\tau(S_{A_\varepsilon}) > 1$, then $\left(\frac{1+\varepsilon}{1-\varepsilon}\right)^2 < \frac{p-1}{q-1}$. Let $h: [0, 1] \rightarrow \mathbb{R}$ be defined as $h(t) = \psi(1-t) - \psi(t)$, where ψ is defined as in (14). The critical points of f_{A_ε} correspond to zeros of h in $(0, 1/2]$. Indeed, by Lemma 8, we know that these points are in bijection with the zeros of h on $(0, 1)$ and $h(t) = -h(1-t)$ for every $t \in (0, 1)$. We have already observed that $h(t_0) = 0$ with $t_0 = 1/2$. We now show that there exists $t_1 \in (0, t_0)$ such that $h(t_1) = 0$. The existence of such t_1 implies that $(t_1, 1-t_1)^T, (1-t_1, t_1)^T, (t_0, t_0)^T$ are three distinct positive critical points of f_{A_ε} , since $h(1-t_1) = h(t_1) = 0$. To construct t_1 , we first prove that our assumption $\tau(S_{A_\varepsilon}) > 1$ is equivalent to the condition $h'(t_0) > 0$. We have

$$\begin{aligned} \psi'(t) &= (q-1)t^{q-2}[(t\varepsilon+1-t)^{p-1} + \varepsilon(\varepsilon+t-t\varepsilon)^{p-1}] \\ &\quad + (p-1)t^{q-1}(1-\varepsilon)[\varepsilon(\varepsilon+t-t\varepsilon)^{p-2} - (t\varepsilon+1-t)^{p-2}]. \end{aligned}$$

With $(\varepsilon+t_0-t_0\varepsilon) = (t_0\varepsilon+1-t_0) = (\varepsilon+1)/2$ we get

$$\begin{aligned} \psi'(t_0) &= (q-1)2^{2-q}(1+\varepsilon)\left(\frac{\varepsilon+1}{2}\right)^{p-1} + (p-1)2^{1-q}(1-\varepsilon)(\varepsilon-1)\left(\frac{\varepsilon+1}{2}\right)^{p-2} \\ &= 2^{3-q-p}(1+\varepsilon)^{p-2}\left[(q-1)(1+\varepsilon)^2 - (p-1)(1-\varepsilon)^2\right]. \end{aligned}$$

As $h'(t_0) = -\psi'(t_0) - \psi'(1-t_0) = -2\psi'(t_0)$, we have $h'(t_0) > 0$ if and only if $(q-1)(1+\varepsilon)^2 < (p-1)(1-\varepsilon)^2$ i.e. $h'(t_0) > 0$ if and only if $\tau(S_{A_\varepsilon}) > 1$.

Now, as $h'(t_0) > 0$, there exists a neighborhood U of t_0 such that h is strictly increasing on U . Since $h(t_0) = 0$, this implies that there exists $s \in (0, t_0) \cap U$ such that $h(s) < 0$. As $\lim_{t \rightarrow 0} h(t) = \varepsilon^{p-1} + \varepsilon > 0$, the intermediate value theorem implies the existence of $t_1 \in (0, s)$ such that $h(t_1) = 0$. As observed above, this concludes the proof. $\square$

Finally, we address the case $\tau(S_{A_\varepsilon}) = 1$.

Lemma 10 *If $\tau(S_{A_\varepsilon}) = 1$, then f_{A_ε} has a unique nonnegative critical point.*

Proof Let $F: \mathbb{R}_+^2 \rightarrow \mathbb{R}_+^2$ be defined as $F(x) = \Phi_{q^*}(A_\varepsilon \Phi_p(A_\varepsilon x))$, where $q^* = q/(q-1)$ denotes the Hölder conjugate of q . Then, for $\mathbf{1} = (1, 1)^T$ and $u = \mathbf{1}/2$, we have $F(u) = \lambda u$ for some $\lambda > 0$. Hence, u is a fixed point of S_{A_ε} and, $\|\cdot\|_q$ is differentiable, by Lemma 2, it follows that u is a critical point of f_{A_ε} . Moreover, it is a fixed point of $G: D_+ \rightarrow D_+$ defined by $G(x) = \langle F(x), \mathbf{1} \rangle^{-1} F(x)$, where $D_+ = \{(t, 1-t) : t \in [0, 1]\}$. Note that the fixed points of G coincide, up to scaling, with those of S_{A_ε} . To conclude, we prove that u is the unique fixed point of G .

As $\tau(S_{A_\varepsilon}) = 1$, we have $d_H(G(x), G(y)) = d_H(F(x), F(y)) \leq d_H(x, y)$ and so G is non-expansive with respect to d_H . Now, Theorem 6.4.1 in [31] implies that u is the unique fixed point of G , if

$$z - G'(u)z \neq 0 \quad \forall z \in \mathbb{R}^2 \setminus \{0\} \quad \text{with} \quad z_1 + z_2 = 0.$$

where $G'(u)$ denotes the Jacobian matrix of G evaluated at u . Moreover, as $F(u) = \lambda u$, Lemma 6.4.2 in [31] implies that $F'(u)u = \lambda u$ and

$$G'(u)z = \frac{1}{\lambda}(F'(u)z - \langle F'(u)z, \mathbf{1} \rangle u).$$

Suppose by contradiction that there exists a $z \in \mathbb{R}^2 \setminus \{0\}$ with $z_1 + z_2 = 0$, such that $z - G'(u)z = 0$. A direct computation shows that $\langle z, F'(u)^T u \rangle = 0$. Then,

$$\begin{aligned} 0 &= z - G'(u)z = z - \frac{1}{\lambda}F'(u)z + \frac{1}{\lambda}\langle F'(u)z, \mathbf{1} \rangle u \\ &= z - \frac{1}{\lambda}F'(u)z + \frac{2}{\lambda}\langle z, F'(u)^T u \rangle u = z - \frac{1}{\lambda}F'(u)z. \end{aligned}$$

It follows that $F'(u)z = \lambda z$ and, as $F'(u)$ is entry-wise positive, the classical Perron–Frobenius theorem implies that $z = \pm u$. However, $u_1 + u_2 > 0$ which contradicts the assumption $z_1 + z_2 = 0$. So $0 \neq z - G'(u)z$ for every $z \neq 0$ such that $z_1 + z_2 = 0$. Hence, u is the unique fixed point of G , which concludes the proof. $\square$

Combining the last two lemmas allows us to conclude:

Proof of Theorem 5 Due to Lemmas 9 and 10 we only need to address the case $\tau(S_{A_\varepsilon}) < 1$. However this is a direct consequence of Lemma 3. In fact, as A_ε is entry-wise positive, the nonnegative fixed points of S_{A_ε} are positive and, if $\tau(S_{A_\varepsilon}) < 1$, then S_{A_ε} is a strict contraction with respect to d_H and so it has a unique fixed point which also is the unique positive maximizer of f_{A_ε} on $\mathbb{R}_+^2$. $\square$

5 Matrix Norms Induced by Sum of Weighted ℓ^p Norms

The Birkhoff contraction ratios $\kappa_H(J_\alpha)$ and $\kappa_H(J_{\alpha^*})$ are easy to compute when $\|\cdot\|_\alpha$ is a weighted ℓ^p norm. More precisely, we have the following

Proposition 1 *Let $\|x\|_\alpha = \|Dx\|_p$ for some $p \in (1, \infty)$ and some diagonal matrix D with positive diagonal entries, then $\|x\|_{\alpha^*} = \|D^{-1}x\|_{p^*}$ where $p^* = p/(p-1)$. Furthermore, it holds $\kappa_H(J_\alpha) = \kappa_H(J_{\alpha^*})^{-1} = p-1$.*

Proof The equality $\|x\|_{\alpha^*} = \|D^{-1}x\|_{p^*}$ follows from Theorem 6 below. To conclude, note that $J_\alpha(x) = \|Dx\|_p^{1-p} D^p \Phi_p(x)$ and therefore $\kappa_H(J_\alpha) = \kappa_H(\Phi_p) = p-1$. The same argument shows that $\kappa_H(J_{\alpha^*}) = \kappa_H(\Phi_{p^*}) = p^*-1 = (p-1)^{-1}$. $\square$

While the above Proposition 1 makes the computation of the Birkhoff constant of weighted ℓ^p -norms particularly easy, computing $\kappa_H(J_\alpha)$ or $\kappa_H(J_{\alpha^*})$ for a general strongly monotonic norm $\|\cdot\|_\alpha$ can be a difficult task. There are norms for which an explicit expression in terms of arithmetic operations for $\|\cdot\|_\alpha$ is given by construction (resp. modelisation), but such an expression is not available for the dual $\|\cdot\|_{\alpha^*}$. Examples include $\|x\|_\alpha = (\|x\|_p^3 + \|x\|_q^3)^{1/3}$ as shown by Theorem 6 below. On the other hand, as discussed in the introduction, monotonic norms different than the standard ℓ^p norms arise quite naturally in several applications.

Motivated by the above observations, we devote the rest of the section to the study of a particular class of monotonic norms of the form $\|x\|_\alpha = \|(\|x\|_{\alpha_1}, \dots, \|x\|_{\alpha_d})\|_\gamma$ where all the norms are monotonic and where we also allow $\|x\|_{\alpha_i}$ to measure only a subset of the coordinates of x .

5.1 Composition of Monotonic Norms and its Dual

Let d be a positive integer. We consider norms of the following form

$$\|x\|_\alpha = \|(\|P_1 x\|_{\alpha_1}, \dots, \|P_d x\|_{\alpha_d})\|_\gamma \quad (17)$$

where $\|\cdot\|_\gamma$ is a monotonic norm on $\mathbb{R}^d$, $\|\cdot\|_{\alpha_i}$ is a norm on $\mathbb{R}^{n_i}$ and $P_i \in \mathbb{R}^{n_i \times n}$ is a “weight matrix” for all $i = 1, \dots, d$. For $\|\cdot\|_\alpha$ to be a norm, we assume that $M = [P_1^T, \dots, P_d^T]^T \in \mathbb{R}^{(n_1 + \dots + n_d) \times n}$ has rank n . Note that the monotonicity of $\|\cdot\|_\gamma$ implies that $\|\cdot\|_\alpha$ satisfies the triangle inequality.

Let us first discuss particular cases of (17). First, note that for two norms $\|\cdot\|_{\alpha_1}, \|\cdot\|_{\alpha_2}$ on $\mathbb{R}^n$, the norm

$$\|x\|_{\alpha_+} = (\|x\|_{\alpha_1}^p + \|x\|_{\alpha_2}^p)^{1/p}$$

can be obtained from (17) with $d = 2$, $\|\cdot\|_{\gamma} = \|\cdot\|_p$, and $P_1 = P_2 = I$, with $I \in \mathbb{R}^{n \times n}$ being the identity matrix. It is also possible to model norms acting on different coordinates of the vectors. For example, if $(x, y) \in \mathbb{R}^{2n}$, then

$$\|(x, y)\|_{\alpha_{\times}} = (\|x\|_{\alpha_1}^p + \|y\|_{\alpha_2}^p)^{1/p}$$

can be obtained from (17) with $d = 2$, $\|\cdot\|_{\gamma} = \|\cdot\|_p$, $P_1 = \text{diag}(1, \dots, 1, 0, \dots, 0) \in \mathbb{R}^{2n \times 2n}$ and $P_2 = \text{diag}(0, \dots, 0, 1, \dots, 1) \in \mathbb{R}^{2n \times 2n}$. The dual of $\|\cdot\|_{\alpha_{\times}}$ is discussed in Lemma 11 below and has a particularly elegant description. More complicated weight matrices P_i can also be used. For example if $\tilde{n}$ is an integer not smaller than n and $P \in \mathbb{R}^{\tilde{n} \times n}$ has rank n , then the norm

$$\|x\|_{\alpha_P} = \|Px\|_p$$

can be obtained with $d = 1$, $\|\cdot\|_{\gamma} = \|\cdot\|_p$, $\|\cdot\|_{\alpha_1} = \|\cdot\|_p$ and $P_1 = P$. Note that if $\tilde{n} = n$, then P is square and invertible and this property can be used to simplify the evaluation of the dual norm of $\|\cdot\|_{\alpha_P}$. Consequences of such additional structure are discussed in Corollary 2.

In the next Theorem 6 we provide a characterization of the dual norm of $\|\cdot\|_{\alpha}$ in its general form as defined in (17). We first need the following lemma that addresses the particular case where $P_1, \dots, P_d$ are projections.

Lemma 11 *Let $n_1, \dots, n_d$ be positive integers and for $i = 1, \dots, d$ let $\|\cdot\|_i$ be a norm on $\mathbb{R}^{n_i}$. Furthermore, let $\|\cdot\|_{\gamma}$ be a monotonic norm on $\mathbb{R}^d$. Let $V = \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_d}$ and for all $(u_1, \dots, u_d) \in V$ define*

$$\|(u_1, \dots, u_d)\|_V = \|(\|u_1\|_{\alpha_1}, \dots, \|u_d\|_{\alpha_d})\|_{\gamma}.$$

Then $\|\cdot\|_V$ is a norm on V and the induced dual norm $\|\cdot\|_{V^}$ satisfies*

$$\|(u_1, \dots, u_d)\|_{V^*} = \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*} \quad \forall (u_1, \dots, u_d) \in V.$$

Proof The fact that $\|\cdot\|_V$ is a norm follows from a direct verification. Let $(u_1, \dots, u_d) \in V$. Then, for every $(y_1, \dots, y_d) \in V$, we have

$$\begin{aligned} \langle (u_1, \dots, u_d), (y_1, \dots, y_d) \rangle &= \sum_{i=1}^d \langle u_i, y_i \rangle \leq \sum_{i=1}^d \|u_i\|_{\alpha_i^*} \|y_i\|_{\alpha_i} \\ &\leq \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*} \|(\|y_1\|_{\alpha_1}, \dots, \|y_d\|_{\alpha_d})\|_{\gamma}, \end{aligned}$$

which shows that

$$\|(u_1, \dots, u_d)\|_{V^*} \leq \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*}. \quad (18)$$

For the reverse inequality, let $v = (\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})$. As $\|\cdot\|_{\gamma}$ is monotonic, by Proposition 5.2 in [7, Chapter 1], there exists $w \in \mathbb{R}_+^d$ such that $\|w\|_{\gamma} \leq 1$ and $\langle v, w \rangle = \|v\|_{\gamma^*}$. Let us denote by $w_1, \dots, w_d \in \mathbb{R}_+$ and $v_1, \dots, v_d \in \mathbb{R}$ respectively the components of w and v in the canonical basis of $\mathbb{R}^d$. Now, let $\bar{y}_1 \in \mathbb{R}^{n_1}, \dots, \bar{y}_d \in \mathbb{R}^{n_d}$ be such that $\|\bar{y}_i\|_{\alpha_i} \leq 1$ and $\langle \bar{y}_i, u_i \rangle = \|u_i\|_{\alpha_i^*}$ for all $i = 1, \dots, d$. Then, as $\|\cdot\|_{\gamma}$ is monotonic with respect to $\mathbb{R}_+^d$ and $\|\bar{y}_i\|_{\alpha_i} \leq 1$ for all i , we have

$$\|(w_1 \bar{y}_1, \dots, w_d \bar{y}_d)\|_{\gamma} = \|(w_1 \|\bar{y}_1\|_{\alpha_1}, \dots, w_d \|\bar{y}_d\|_{\alpha_d})\|_{\gamma} \leq \|w\|_{\gamma} \leq 1.$$

Note that

$$\begin{aligned}\langle (u_1, \dots, u_d), (w_1 \bar{y}_1, \dots, w_d \bar{y}_d) \rangle &= \sum_{i=1}^d w_i \langle u_i, \bar{y}_i \rangle = \sum_{i=1}^d w_i \|u_i\|_{\alpha_i^*} = \langle v, w \rangle \\ &= \|v\|_{\gamma^*} = \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*}.\end{aligned}$$

It follows that $\|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*} \leq \|(u_1, \dots, u_d)\|_{V^*}$, which, together with (18), concludes the proof. $\square$

Theorem 6 Let d be a positive integer. For $i = 1, \dots, d$, let $P_i \in \mathbb{R}^{n_i \times n}$ and let $\|\cdot\|_{\alpha_i}$ be a norm on $\mathbb{R}^{n_i}$. Suppose that $M = [P_1^T, \dots, P_d^T]^T \in \mathbb{R}^{(n_1+\dots+n_d) \times n}$ has rank n . Furthermore, let $\|\cdot\|_{\gamma}$ be a monotonic norm on $\mathbb{R}^d$. For every $x \in \mathbb{R}^n$, define

$$\|x\|_{\alpha} = \|(\|P_1 x\|_{\alpha_1}, \dots, \|P_d x\|_{\alpha_d})\|_{\gamma}.$$

Then, $\|\cdot\|_{\alpha}$ is a norm on $\mathbb{R}^n$ and the induced dual norm is given by

$$\|x\|_{\alpha^*} = \inf_{\substack{u_1 \in \mathbb{R}^{n_1}, \dots, u_d \in \mathbb{R}^{n_d} \\ P_1^T u_1 + \dots + P_d^T u_d = x}} \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*},$$

where $\|\cdot\|_{\alpha_i^*}$ is the dual norm induced by $\|\cdot\|_{\alpha_i}$ and $\|\cdot\|_{\gamma^*}$ is the dual norm induced by $\|\cdot\|_{\gamma}$.

Proof Let $u_1 \in \mathbb{R}^{n_1}, \dots, u_d \in \mathbb{R}^{n_d}$ be such that $P_1^T u_1 + \dots + P_d^T u_d = x$. Such vectors always exists as M has full rank. Then, for every $y \in \mathbb{R}^n$, it holds

$$\begin{aligned}\langle x, y \rangle &= \sum_{i=1}^d \langle P_i^T u_i, y \rangle = \sum_{i=1}^d \langle u_i, P_i y \rangle \\ &\leq \sum_{i=1}^d \|u_i\|_{\alpha_i^*} \|P_i y\|_{\alpha_i} \leq \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*} \|y\|_{\alpha}.\end{aligned}$$

It follows that

$$\|x\|_{\alpha^*} \leq \inf_{\substack{u_1 \in \mathbb{R}^{n_1}, \dots, u_d \in \mathbb{R}^{n_d} \\ P_1^T u_1 + \dots + P_d^T u_d = x}} \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*}.$$

Now, we prove the reverse inequality. To this end, consider the vector space $V = \mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_d}$ endowed with the norm $\|\cdot\|_V$ defined as

$$\|(u_1, \dots, u_d)\|_V = \|(\|u_1\|_{\alpha_1}, \dots, \|u_d\|_{\alpha_d})\|_{\gamma} \quad \forall (u_1, \dots, u_d) \in V.$$

As V is a finite product of finite dimensional vector spaces, we can identify V^* with V and by Lemma 11, we know that the dual norm of $\|\cdot\|_{V^*}$ induced by $\|\cdot\|_V$ satisfies

$$\|(u_1, \dots, u_d)\|_{V^*} = \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*} \quad \forall (u_1, \dots, u_d) \in V.$$

Consider now the vector subspace $W = \{(P_1 y, \dots, P_d y) \mid y \in \mathbb{R}^n\} \subset V$. Note that, we can identify W with the image of M , i.e. $W = \{M y \mid y \in \mathbb{R}^n\}$. Let $M^\dagger \in \mathbb{R}^{n \times (n_1+\dots+n_d)}$ be the Moore–Penrose inverse of M . Then, as M is full rank, we have $M^\dagger M y = y$ for all $y \in \mathbb{R}^n$. Let $\phi: W \rightarrow \mathbb{R}$ be defined as

$$\phi(u_1, \dots, u_d) = \langle M^\dagger (u_1, \dots, u_d), x \rangle \quad \forall (u_1, \dots, u_d) \in W.$$

For every $(u_1, \dots, u_d) \in W$, there exists $y \in \mathbb{R}^n$ such that $(u_1, \dots, u_d) = My$, i.e. $u_i = P_i y$ for all $i = 1, \dots, d$, and thus

$$\begin{aligned} |\phi(u_1, \dots, u_d)| &= |\phi(My)| = |\langle M^\dagger My, x \rangle| = |\langle y, x \rangle| \\ &\leq \|y\|_\alpha \|x\|_{\alpha^*} = \|(\|P_1 y\|_{\alpha_1}, \dots, \|P_d y\|_{\alpha_d})\|_\gamma \|x\|_{\alpha^*} = \|(u_1, \dots, u_d)\|_V \|x\|_{\alpha^*}. \end{aligned}$$

By the Hahn–Banach theorem (see e.g. Corollary 1.2 of [5]), there exists $(u'_1, \dots, u'_d) \in V$ such that

$$\phi(u_1, \dots, u_d) = \sum_{i=1}^d \langle u'_i, u_i \rangle \quad \forall (u_1, \dots, u_d) \in W, \quad (19)$$

and

$$\|(\|u'_1\|_{\alpha_1^*}, \dots, \|u'_d\|_{\alpha_d^*})\|_{\gamma^*} = \|(u'_1, \dots, u'_d)\|_{V^*} \leq \|x\|_{\alpha^*}.$$

Next, let $y \in \mathbb{R}^n$, then $My = (P_1 y, \dots, P_d y) \in W$ and with (19), we have

$$\langle y, x \rangle = \langle M^\dagger My, x \rangle = \sum_{i=1}^d \langle u'_i, P_i y \rangle = \sum_{i=1}^d \langle P_i^T u'_i, y \rangle.$$

As the above is true for all $y \in \mathbb{R}^n$, it follows that $P_1^T u'_1 + \dots + P_d^T u'_d = x$. Hence, we have

$$\inf_{\substack{u_1 \in \mathbb{R}^{n_1}, \dots, u_d \in \mathbb{R}^{n_d} \\ P_1^T u_1 + \dots + P_d^T u_d = x}} \|(\|u_1\|_{\alpha_1^*}, \dots, \|u_d\|_{\alpha_d^*})\|_{\gamma^*} \leq \|(\|u'_1\|_{\alpha_1^*}, \dots, \|u'_d\|_{\alpha_d^*})\|_{\gamma^*} \leq \|x\|_{\alpha^*},$$

which concludes the proof of the formula for $\|\cdot\|_{\alpha^*}$. $\square$

As a consequence of the above Theorem 6, we have that the dual of the norms $\|\cdot\|_{\alpha_+}$, $\|\cdot\|_{\alpha_\times}$, $\|\cdot\|_{\alpha_p}$ considered at the beginning of this section are respectively given by

$$\begin{aligned} \|x\|_{\alpha_+^*} &= \inf_{\substack{u_1 + u_2 = x \\ u_1, u_2 \in \mathbb{R}}} (\|u_1\|_{\alpha_1^*}^{p^*} + \|u_2\|_{\alpha_2^*}^{p^*})^{1/p^*}, \\ \|(x, y)\|_{\alpha_\times^*} &= (\|x\|_{\alpha_1^*}^{p^*} + \|y\|_{\alpha_2^*}^{p^*})^{1/p^*}, \quad \|x\|_{\alpha_p^*} = \inf_{u \in \mathbb{R}^n: P^T u = x} \|u\|_{p^*}, \end{aligned}$$

with $p^* = p/(p-1)$. Note that the $\|\cdot\|_{\alpha_\times^*}$ does not involve an infimum. The infimum can also be removed in $\|x\|_{\alpha_p^*}$, if P is square and invertible and in that case it holds $\|x\|_{\alpha_p^*} = \|P^{-T}x\|_{p^*}$.

We discuss more general examples in the next result.

Corollary 2 *Under the same assumptions as Theorem 6, we have:*

1. If $P_1, \dots, P_d$ are all square invertible matrices and

$$\|x\|_{\alpha^*} = \min_{\substack{x = u_1 + \dots + u_d \\ u_1, \dots, u_d \in \mathbb{R}^n}} \|(\|(P_1^T)^{-1}u_1\|_{\alpha_1^*}, \dots, \|(P_d^T)^{-1}u_d\|_{\alpha_d^*})\|_{\gamma^*}$$

2. If every $x \in \mathbb{R}^n$ can be uniquely written as $x = x_{P_1} + \dots + x_{P_d}$ with $x_{P_i} \in \text{Im}(P_i^T)$ for all $i = 1, \dots, d$ (i.e. $\mathbb{R}^n$ is the direct sum of the range of $P_1, \dots, P_d$), then

$$\|x\|_{\alpha^*} = \left\| \left(\inf_{\substack{u_1 \in \mathbb{R}^{n_1} \\ P_1^T u_1 = x_{P_1}}} \|u_1\|_{\alpha_1^*}, \dots, \inf_{\substack{u_d \in \mathbb{R}^{n_d} \\ P_d^T u_d = x_{P_d}}} \|u_d\|_{\alpha_d^*} \right) \right\|_{\gamma^*}.$$

If, additionally, $n_i = \dim(\text{Im}(P_i^T))$ for all $i = 1, \dots, d$, then

$$\|x\|_{\alpha^*} = \|(\|(P_1^T)^\dagger x\|_{\alpha_1^*}, \dots, \|(P_d^T)^\dagger x\|_{\alpha_d^*})\|_{\gamma^*},$$

where $(P_i^T)^\dagger$ is the Moore–Penrose inverse of P_i^T .

5.2 The Power Method for Compositions of ℓ^p -Norms

We discuss here consequences of Theorems 4 and 6 when applied to a special family of norms defined in terms of subsets of entries of the initial vector, i.e. the case where P_i is a nonnegative diagonal matrix.

For some nonnegative weight vector $\omega \in \mathbb{R}^m$ and coefficient $p \in (1, \infty)$, let $\|\cdot\|_{\omega,p}$ be the ω -weighted ℓ^p -(semi)norm on $\mathbb{R}^m$, defined as

$$\|x\|_{\omega,p} = \|\text{diag}(\omega)^{1/p} x\|_p = \left(\sum_{j=1}^m \omega_j |x_j|^p \right)^{1/p}. \quad (20)$$

To express the dual of $\|x\|_{\omega,p}$ and their compositions, let

$$p^* = \frac{p}{p-1} \quad \text{and} \quad \omega_i^* = \begin{cases} \omega_j^{1-p^*} & \text{if } \omega_j > 0, \\ 0 & \text{if } \omega_j = 0, \end{cases} \quad \forall j = 1, \dots, m. \quad (21)$$

If ω is positive, then $\|x\|_{\omega,p}$ is a norm and it holds $(\|x\|_{\omega,p})_* = \|x\|_{\omega^*,p^*}$ by Proposition 1.

Let $\omega_1, \dots, \omega_d \in \mathbb{R}^m$ be nonzero vectors of nonnegative weights such that $\omega_1 + \dots + \omega_d$ is a positive vector. Further let $s \in [1, \infty)$, $p_1, \dots, p_d \in (1, \infty)$ and define

$$\|x\|_{\alpha} = \left(\sum_{k=1}^d \|x\|_{\omega_k, p_k}^s \right)^{1/s}, \quad (22)$$

The fact that $\omega_1 + \dots + \omega_d$ is positive ensures that $\|\cdot\|_{\alpha}$ is a norm. Note that $\|\cdot\|_{\alpha}$ is strongly monotonic.

The differentiability of $\|\cdot\|_{\alpha}$ is discussed in the following lemma.

Lemma 12 *Let $\|\cdot\|_{\alpha}$ be as in (22), then $\|\cdot\|_{\alpha}$ is differentiable if either $s > 1$ or $s = 1$ and ω_i has at least two positive entries for every $i = 1, \dots, d$.*

Proof As $p_k > 1$, $\|\cdot\|_{\omega_k, p_k}$ is differentiable if ω_k has at least two positive entries. If it has only one positive entry then $\|\cdot\|_{\omega_k, p_k}$ is just a weighted absolute value. Hence, if $s > 1$, then the differentiability of $\|\cdot\|_{\alpha}$ follows from that of the ℓ^s -norm. While if $s = 1$ and ω_i has at least two positive entries for every $i = 1, \dots, d$, then $\|\cdot\|_{\alpha}$ is just a sum of differentiable norms. $\square$

If $\|\cdot\|_{\alpha}$ is differentiable, we have

$$J_{\alpha}(x) = \|x\|_{\alpha}^{1-s} \sum_{k=1}^d \|x\|_{\omega_k, p_k}^{s-p_k} \text{diag}(\omega_k) \Phi_{p_k}(x) \quad (23)$$

and the following lemma provides an upper bound for $\kappa_H(J_{\alpha})$.

Lemma 13 Let $\|\cdot\|_\alpha$ be as in (22). If $\|\cdot\|_\alpha$ is differentiable then

$$\kappa_H(J_\alpha) \leq (s-1) + \sum_{k=1}^d \max\{0, p_k - s\}.$$

Proof Let $\delta = \sum_{k=1}^d \max\{0, p_k - s\}$. We have $J_\alpha(x) = \|x\|_\alpha^{1-s}(F(x) + G(x))$ where for all $x \in \mathbb{R}_+^m \setminus \{0\}$ we let $F(x) = \sum_{p_k \leq s} \|x\|_{\omega_k, p_k}^{s-p_k} \text{diag}(\omega_k) \Phi_{p_k}(x)$ and $G(x) = \sum_{p_k > s} \|x\|_{\omega_k, p_k}^{s-p_k} \text{diag}(\omega_k) \Phi_{p_k}(x)$. Note that if $p_k > s$ for all k then $F(x) = 0$, whereas $G(x) = 0$ when $p_k \leq s$ for all k . Moreover, note that F is order-preserving and homogeneous of degree $s-1$. Now let us set $\tau(x) = 1$ if $p_j \leq s$ for all j and $\tau(x) = \prod_{p_j > s} \|x\|_{\omega_j, p_j}^{p_j-s}$ otherwise. Then τ is order-preserving and homogeneous of degree δ and $x \mapsto \tau(x)F(x)$ is order-preserving and homogeneous of degree $\delta + (s-1)$. Finally, note that

$$x \mapsto \tau(x)G(x) = \sum_{p_k > s} \prod_{\substack{p_j > s \\ j \neq k}} \|x\|_{\omega_j, p_j}^{p_j-s} \text{diag}(\omega_k) \Phi_{p_k}(x)$$

is order-preserving as well and homogeneous of degree $\delta + (s-1)$. This implies that $\delta + (s-1)$ is a Lipschitz constant of $H(x) = \tau(x)(F(x) + G(x))$ with respect to the Hilbert metric μ . Hence, for any $x, y \in \mathbb{R}_+^m \setminus \{0\}$ with $x \sim y$, we finally obtain

$$\mu(J_\alpha(x), J_\alpha(y)) = \mu(H(x), H(y)) \leq (\delta + s - 1)\mu(x, y),$$

which concludes the proof. $\square$

If $s > 1$, by Theorem 6, we have

$$\begin{aligned} \|x\|_{\alpha^*} &= \min_{\substack{u_1, \dots, u_d \in \mathbb{R}^m \\ \text{diag}(\omega_1)u_1 + \dots + \text{diag}(\omega_d)u_d = x}} \left(\sum_{k=1}^d \|u_k\|_{p_k^*}^{s^*} \right)^{1/s^*} \\ &= \min_{\substack{u_1 + \dots + u_d = x \\ u_1, \dots, u_d \in \mathbb{R}^m}} \left(\sum_{k=1}^d \|u_k\|_{\omega_k^*, p_k^*}^{s^*} \right)^{1/s^*}. \end{aligned} \quad (24)$$

It is not difficult to realize that the case $s = 1$ has a similar form, where the sum is replaced by a maximum. We henceforth omit that case, for the sake of brevity.

Now, consider a norm $\|\cdot\|_\beta$ defined as the dual norm of a norm of the type (22)

$$\|x\|_\beta = \min_{\substack{u_1 + \dots + u_h = x \\ u_1, \dots, u_h \in \mathbb{R}^n}} \left(\sum_{k=1}^h \|u_k\|_{\varpi_k, q_k}^t \right)^{1/t} \quad (25)$$

where h is some positive integer, ϖ_i are nonnegative weight vectors whose sum $\varpi_1 + \dots + \varpi_h$ is positive and $q_1, \dots, q_h, t \in (1, \infty)$. As $\min_x f(x) = (\max_x f(x)^{-1})^{-1}$ for continuous positive f , we deduce that for this choice of norm we have

$$\|A\|_{\beta \rightarrow \alpha} = \max_{x \neq 0} \frac{\|Ax\|_\alpha}{\|x\|_\beta} = \max_{\substack{u_1 + \dots + u_h \neq 0 \\ u_1 \in \mathbb{R}^n, \dots, u_h \in \mathbb{R}^n}} \frac{\left(\sum_{k=1}^d \sum_{j=1}^h \|Au_j\|_{\omega_k, p_k}^s \right)^{1/s}}{\left(\sum_{k=1}^h \|u_k\|_{\varpi_k, q_k}^t \right)^{1/t}}$$

Algorithm 1: Computation of $J_\alpha(x)$ for $\|\cdot\|_\alpha$ defined in (22).

Input: Vector $x \in \mathbb{R}^n$; nonnegative weight vectors $\omega_1, \dots, \omega_d \in \mathbb{R}^n$;
inner powers $p_1, \dots, p_d \in (1, \infty)$; outer power $s \in (1, \infty)$

Output: $J_\alpha(x)$ and $\|x\|_\alpha$.

```

1  $y \leftarrow 0$  # Zero vector in  $\mathbb{R}^n$ 
2  $v \leftarrow 0$  # Zero scalar in  $\mathbb{R}$ 
3 #  $u \circ w$  denotes the vector  $(u_1 w_1, \dots, u_n w_n)^T$ 
4 #  $u^p$  denotes the vector  $(u_1^p, \dots, u_n^p)^T$ 
5 for  $\ell = 1, \dots, d$  do
6    $\gamma \leftarrow \|\omega_\ell^{1/p_\ell} \circ x\|_p$ 
7    $y \leftarrow y + (\omega_\ell \circ |x|^{p_\ell-1} \circ \text{sign}(x)) \gamma^{s-p_\ell}$ 
8    $v \leftarrow v + \gamma^s$ 
9 end
10  $\|x\|_\alpha \leftarrow v^{1/s}$ 
11  $J_\alpha(x) \leftarrow \|x\|_\alpha^{1-s} y$ 

```

for any matrix $A \in \mathbb{R}^{m \times n}$.

We emphasize that, while the norm $\|\cdot\|_\beta$ is defined implicitly in the general case, when the weight vectors ϖ_i have disjoint support Corollary 2 yields the following explicit formula

$$\|x\|_\beta = \left(\sum_{k=1}^h \|x\|_{\varpi_k, q_k}^t \right)^{1/t}$$

which also simplifies the definition of $\|A\|_{\beta \rightarrow \alpha}$.

The advantage of choosing $\|\cdot\|_\beta$ as in (25) relies on the fact that both $\|x\|_{\beta^*}$ and J_{β^*} admit an explicit expression analogous to (22) and (23), precisely

$$\|x\|_{\beta^*} = \left(\sum_{k=1}^h \|x\|_{\varpi_k^*, q_k^*}^{t^*} \right)^{1/t^*} \quad \text{and}$$

$$J_{\beta^*}(x) = \|x\|_{\beta^*}^{1-t^*} \sum_{k=1}^h \|x\|_{\varpi_k^*, q_k^*}^{t^*-q_k^*} \text{diag}(\varpi_k^*) \Phi_{q_k^*}(x),$$

for all choices of the weights ϖ_i such that $\varpi_1 + \dots + \varpi_h > 0$.

Thus, we obtain an explicit formula for the operator

$$S_A(x) = J_{\beta^*}(A^T J_\alpha(Ax))$$

which allows us to easily implement the power method (11) for the matrix norm $\|A\|_{\beta \rightarrow \alpha}$. An efficient implementation of the operator J_α for a norm $\|\cdot\|_\alpha$ of the form (22) is provided by Algorithm 1. If we let $\text{nnz}(X)$ denote the number of nonzero entries in X and we assume arithmetic operations have unit cost, evaluating J_α and J_{β^*} via Algorithm 1 costs $\mathcal{O}\left(\sum_{i=1}^d \text{nnz}(\omega_i)\right)$ and $\mathcal{O}\left(\sum_{i=1}^h \text{nnz}(\varpi_i)\right)$ operations, respectively. So, the total cost of evaluating S_A , i.e. of each iteration of the power method in (11), is $\mathcal{O}(C(S_A))$ where

$$C(S_A) = \sum_{i=1}^d \text{nnz}(\omega_i) + \sum_{i=1}^h \text{nnz}(\varpi_i) + \text{nnz}(A)$$

which boils down to $\mathcal{O}(dn + hn + n^2)$ when all the ω_i , ϖ_i and A are full.

As a consequence, we have

Theorem 7 Let $A \in \mathbb{R}^{m \times n}$ be a nonnegative matrix such that $A^T A$ is irreducible. Let $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ be as in (22) and (25), respectively. Let

$$\tau = \kappa_H(A) \kappa_H(A^T) \left(s - 1 + \sum_{k=1}^d \max\{0, p_k - s\} \right) \left(t - 1 + \sum_{j=1}^h \max\{0, q_j - t\} \right).$$

If $\tau < 1$ and $\|\cdot\|_\alpha, \|\cdot\|_\beta$ are differentiable, then $\|A\|_{\beta^* \rightarrow \alpha}$ can be approximated to ε precision in $\mathcal{O}(C(S_A) \ln(1/\varepsilon))$ arithmetic operations with the power sequence (11).

Proof Besides the complexity bound, the result is a direct consequence of Theorem 4 and the upper bounds for $\kappa_H(J_\alpha)$ and $\kappa_H(J_\beta)$ obtained in Lemma 13. Let us provide estimates for the total number of operations required by the fixed point sequence (11). Let $\tilde{C}$ be as in Theorem 4. We have $\tilde{C}\tau^k < \varepsilon$ if and only if $k > (\ln(\varepsilon) - \ln(\tilde{C}))/\ln(\tau)$. As $(\ln(\varepsilon) - \ln(\tilde{C}))/\ln(\tau) \in \mathcal{O}(-\ln(\varepsilon))$ for $\varepsilon \rightarrow 0$, we deduce that $\|A\|_{q \rightarrow p} - \varepsilon \leq \|Ax_k\|_p$ after $\mathcal{O}(\ln(\varepsilon^{-1}))$ iterations of S_A , leading to a total complexity of $\mathcal{O}(C(S_A) \ln(\varepsilon^{-1}))$. $\square$

We conclude the section by proving a number of corollaries of Theorem 7 that illustrate the richness of the class of problem that can be addressed via that theorem. For simplicity, in the statements we assume that the involved matrices are square and positive. However, more general statements involving irreducible and rectangular matrices can be easily derived by reproducing the proof of the corresponding corollary.

Corollary 3 Let $A \in \mathbb{R}^{n \times n}$ be a positive matrix. Let $\omega, \varpi \in \mathbb{R}^n$ be positive weights and $1 < p, q < \infty$. Let

$$\|A\|_{\beta \rightarrow \alpha} = \max_{x \neq 0} \frac{\|Ax\|_{\omega, p}}{\|x\|_{\varpi, q}} \quad \text{and} \quad \tau = \kappa_H(A)^2 \frac{p-1}{q-1}.$$

If $\tau < 1$, then $\|A\|_{\beta \rightarrow \alpha}$ can be computed to ε precision in $\mathcal{O}(\text{nnz}(A) \ln(1/\varepsilon))$ operations.

Proof As $d = h = 1$, $C(S_A) = \text{nnz}(A)$, $\|y\|_\alpha = \|y\|_{\omega, p}$ and $\|x\|_{\beta^*} = \|x\|_{\varpi^*, q^*}$ in Theorem 7. $\square$

Corollary 4 Let $A, B \in \mathbb{R}^{n \times n}$ be positive matrices. Further, let $1 < p, q, r < \infty$,

$$\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_{\beta \rightarrow \alpha} = \max_{x \neq 0} \frac{2\|Ax\|_p + 3\|Bx\|_q}{\|x\|_r} \quad \text{and} \quad \tau = \kappa_H \left(\begin{bmatrix} A \\ B \end{bmatrix} \right)^2 \frac{p+q-2}{r-1}.$$

If $\tau < 1$, then $\left\| \begin{bmatrix} A \\ B \end{bmatrix} \right\|_{\beta \rightarrow \alpha}$ can be computed to ε precision in $\mathcal{O}(N \ln(1/\varepsilon))$ operations with $N = \text{nnz}(A) + \text{nnz}(B)$.

Proof Let $d = 2, h = 1$, $\omega_i = 2^p, \varpi_i = 3^q$ for $i = 1, \dots, n$, and $\|x\|_{\beta^*} = \|x\|_{r^*}$, $\|(y, z)\|_\alpha = \| \|y\|_{\omega, p}, \|z\|_{\varpi, q} \|_1$ in Theorem 7. Also note that $\mathcal{O}(C(S_A)) = \mathcal{O}(N)$. $\square$

Corollary 5 Let $A \in \mathbb{R}^{n \times n}$ positive, $1 < p < \infty$, $2 \leq q, r < \infty$,

$$\|A\|_{\beta \rightarrow \alpha} = \max_{x+y \neq 0} \frac{\|Ax + Ay\|_p}{\sqrt{\|x\|_q^2 + \|y\|_r^2}} \quad \text{and} \quad \tau = \kappa_H(A)^2(p-1).$$

If $\tau < 1$, then $\|A\|_{\beta \rightarrow \alpha}$ can be computed to ε precision in $\mathcal{O}(\text{nnz}(A) \ln(1/\varepsilon))$ operations.

Proof Let $d = 1, h = 2, \|y\|_\alpha = \|y\|_p, \|x\|_{\beta^*} = \|x\|_{q^*}, \|x\|_{r^*}\|_2$ in Theorem 7. $\square$

Corollary 6 Let $A, B \in \mathbb{R}^{n \times n}$ be positive matrices, $1 < s \leq \theta \leq p, q, r < \infty$,

$$\|[A \ B]\|_{\beta \rightarrow \alpha}^\theta = \max_{(x,y) \neq (0,0)} \frac{\|Ax\|_p^\theta + \|By\|_q^\theta}{\|x\|_r^\theta + \|y\|_s^\theta} \text{ and } \tau = \kappa_H([A \ B])^2 \frac{p+q-\theta-1}{s-1}.$$

If $\tau < 1$, then $\|[A \ B]\|_{\beta \rightarrow \alpha}$ can be computed to ε precision in $\mathcal{O}(N \ln(1/\varepsilon))$ operations with $N = \text{nnz}(A) + \text{nnz}(B)$.

Proof Let $d = 2, h = 2, \|(y, z)\|_\alpha = \|y\|_p, \|z\|_q\|_\theta$ and $\|x\|_{\beta^*} = \|x\|_{r^*}, \|x\|_{s^*}\|_{\theta^*}$ in Theorem 7. $\square$

Corollary 7 Let $A, B \in \mathbb{R}^{n \times n}$ be positive matrices and $1 < p, q, r < \infty$, let

$$\phi = \max_{(x,y,u+v) \neq (0,0,0)} \min \left\{ \frac{\|A(x+y) + B(u+v)\|_p}{\|(x, u)\|_q}, \frac{\|A(x+y) + B(u+v)\|_p}{\|(y, v)\|_r} \right\}.$$

If $\tau = \frac{p-1}{q-1} + \frac{p-1}{r-1} < 1$, then ϕ can be computed to ε precision in $\mathcal{O}(N \ln(1/\varepsilon))$ operations with $N = \text{nnz}(A) + \text{nnz}(B)$.

Proof Let $M = \begin{bmatrix} A & A & 0 \\ 0 & B & B \end{bmatrix}$, $d = 2, h = 2, \|y\|_\alpha = \|y\|_p$ and $\|(x, y, z)\|_{\beta^*} = \| \|(x, z)\|_{r^*}, \|(y, z)\|_{s^*}\|_1$ in Theorem 7. Note that $\kappa_H(M) = 1$ by Theorem 3. $\square$

Corollary 8 Let $A, B \in \mathbb{R}^{n \times n}$ be positive matrices and $1 < p, q, r < \infty$. Let $\sigma_p: \mathbb{R}^n \rightarrow \mathbb{R}_+^n$ be defined as $\sigma_p(x) = (|x_1|^p, \dots, |x_n|^p)^T$ and let

$$\|B\|_{\beta \rightarrow \alpha}^p = \max_{\|x\|_r=1} \|A\sigma_p(Bx)\|_q \quad \tau = \frac{pq-1}{r-1} \kappa(B)^2.$$

If $\tau < 1$ then $\|B\|_{\beta \rightarrow \alpha}$ can be computed to ε precision in $\mathcal{O}(N \ln(1/\varepsilon))$ operations with $N = \text{nnz}(A) + \text{nnz}(B)$.

Proof As $x \mapsto A\sigma_p(Bx)$ is positively homogeneous of degree p , we have

$$\max_{\|x\|_r=1} \|A\sigma_p(Bx)\|_q = \max_{x \neq 0} \frac{\|A\sigma_p(Bx)\|_q}{\|x\|_r^p} = \left(\max_{x \neq 0} \frac{\|A\sigma_p(Bx)\|_q^{1/p}}{\|x\|_r} \right)^p.$$

Let $\|\cdot\|_{\beta^*} = \|\cdot\|_{r^*}$ and $\|x\|_\alpha = \|A\sigma_p(x)\|_q^{1/p}$. Then, $\|Bx\|_\alpha = \|A\sigma_p(Bx)\|_q^{1/p}$ and with $\omega_i = (A_{i,1}, \dots, A_{i,n})$, it holds $\|x\|_\alpha = \|(\|x\|_{\omega_1,p}, \dots, \|x\|_{\omega_n,p})\|_{pq}$ for every x . The proof is now a direct consequence of Theorem 3 with $s = n$ and $t = 1$. $\square$

6 Numerical Experiments

In this section we illustrate the numerical behaviour of the power sequence (11) on some example matrices and some choices of the norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$. In particular, we consider both a dense matrix and sparse matrix example.

Fig. 2 Comparison between the true error $\|x_k - x^+\|_\infty$ and the upper bound $(\tau_{p,q,\varepsilon})^k C$ in (26) against number of iterations for the power method (11) applied to $\|A_\varepsilon\|_{q \rightarrow p}$ with $\varepsilon = 1/3$, $q = 2$ and for five values of p ranging within the interval $[1/15, 1/\kappa_H(A_\varepsilon)] \cdot (q - 1) + 1$, i.e. chosen so that $\tau_{p,q,\varepsilon} = \kappa_H(S_{A_\varepsilon}) < 1$

6.1 Dense Matrix Example: Tightness of the Convergence Bound

We verify the convergence bound of Theorem 4 for the family of matrix norms analyzed in Sect. 4.2, that is we verify the bound of Theorem 4 on the convergence of the power sequence (11) for the computation of $\|A_\varepsilon\|_{q \rightarrow p}$ for various $1 < p, q < \infty$ and A_ε defined as in (13).

By Lemma 7, we have $\tau_{p,q,\varepsilon} = \kappa_H(S_{A_\varepsilon}) = \kappa_H(A_\varepsilon)^2 \frac{p-1}{q-1}$. Moreover, with $x^+ = 2^{-1/q}(1, 1)^T$ it holds $S_{A_\varepsilon}(x^+) = x^+$, hence the power sequence converges to x^+ when $\tau_{p,q,\varepsilon} < 1$. By Theorem 4, if $\tau_{p,q,\varepsilon} < 1$, then

$$\|x_k - x^+\|_\infty \leq (\tau_{p,q,\varepsilon})^k C \quad \text{with} \quad C = \frac{d_H(x_0, x_1)}{1 - \tau_{p,q,\varepsilon}}. \quad (26)$$

We use $\delta = 10^{-12}$ and $x_0 = (\delta, 1 - \delta)^T$ in our experiments. The choice of x_0 is motivated by the fact that it is far from the limit point x^+ in the Hilbert metric, so to model a worst-case-scenario setting. In Fig. 2, we plot the true error $\|x_k - x^+\|_\infty$ against the number of iterations k and compare it with the upper bound $(\tau_{p,q,\varepsilon})^k C$, for the choice $\varepsilon = 1/3$, $q = 2$ and five increasing values of p chosen so that $\tau_{p,q,\varepsilon} = \kappa_H(S_{A_\varepsilon}) < 1$. We can observe that the method converges linearly as expected and that the upper bound well captures the decay slope. Moreover, even though larger values of p yield larger values of the contraction constant $\tau_{p,q,\varepsilon}$, the upper bound still behaves as the true error when p grows up to a multiplicative constant.

6.2 Sparse Matrix Example

For this experiment, we consider two families of matrices with growing size which are not irreducible but satisfy the requirement of Theorem 4, that is $A^T A$ is irreducible. More precisely, let $A_1 \in \mathbb{R}^{3 \times 3}$ and $B_1 \in \mathbb{R}^{4 \times 2}$ be given by

$$A_1 = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 0 \\ 1 & 1 & 1 \end{bmatrix} \quad \text{and} \quad B_1 = \begin{bmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \\ 1 & 1 \end{bmatrix}.$$

Fig. 3 Number of iterations k until guaranteed convergence such that $\|x_k - x^+\|_\infty < 10^{-10}$ for the computation of $\|A_s\|_{q \rightarrow p}$ and $\|B_s\|_{q \rightarrow p}$ for $s = 1, \dots, 10$ and two different choices (p, q) such that $\tau(S_{A_s}) = \tau(S_{B_s}) = 3/4$. Note that $A_{10} \in \mathbb{R}^{3^{10} \times 3^{10}}$ with $3^{10} \approx 6 \cdot 10^6$ and $B_{10} \in \mathbb{R}^{4^{10} \times 2^{10}}$ with $4^{10} \approx 10^6$

Clearly, neither A_1 nor B_1 is irreducible, however $A_1^T A_1$ and $B_1^T B_1$ have strictly positive entries and are therefore irreducible. Then, for $s \geq 2$, we consider the matrices $A_s \in \mathbb{R}^{3^s \times 3^s}$, $B_s \in \mathbb{R}^{4^s \times 2^s}$ obtained by taking s times the Kronecker product of A_1 , resp. B_1 , with itself, i.e. $A_s = A_1 \otimes A_{s-1}$ and $B_s = B_1 \otimes B_{s-1}$. Note that for all $s \geq 1$, A_s has at least one full row of only zero entries and therefore cannot be irreducible. On the other hand, it holds $A_s^T A_s = (A_1^T A_1) \otimes \dots \otimes (A_1^T A_1)$ and thus $A_s^T A_s$ has positive entries since it is the Kronecker product of s positive matrices. The same observation holds for the sequence B_s . Furthermore, Theorem 3 implies that $\kappa_H(A_s) = \kappa_H(B_s) = 1$ for all $s \geq 1$. Hence, we have $\tau(S_{A_s}) = \tau(S_{B_s}) = \frac{p-1}{q-1}$ for all s .

In our experiments we analyze the number of iterations until convergence of the power sequences associated to the computation of $\|A_s\|_{q \rightarrow p}$ and $\|B_s\|_{q \rightarrow p}$, where p, q are fixed so that $\tau = \tau(S_{A_s}) = \tau(S_{B_s}) = 3/4$. For each fixed p, q and s , we try 5000 different starting points drawn uniformly from $(0, 1)^n$ with $n = 3^s$ or $n = 2^s$. The boxplots in Fig. 3 show the number of iterations required until the stopping criterion

$$\tau^k \frac{d_H(x_0, x_1)}{(1 - \tau)} < \delta \quad (27)$$

is met, for both A_s (the two panels in the top row) and B_s (the two panels in the row at the bottom), and for $\delta = 10^{-10}$. Note that, due to Theorem 4, if (27) holds for k then we are guaranteed to be δ -close to the true solution x^+ , i.e. the computed approximation x_k is such that $\|x_k - x^+\|_\infty < \delta$. Moreover, since $\|x\|_p \leq n^{1/p} \|x\|_\infty$ for all $x \neq 0$, we have

$$(1 - \delta n^{1/p}) \|M\|_{q \rightarrow p} \leq \|M x_k\|_p \leq \|M\|_{q \rightarrow p}$$

for both $M = A_s$ and $M = B_s$. While Fig. 3 shows the steps required to guarantee approximation to the true solution, we emphasize that in practice the required number of steps to reach floating point precision on two consecutive iterates is typically much smaller.

7 Conclusions

On top of being a classical problem in numerical analysis, computing the norm of a matrix $\|A\|_{\beta \rightarrow \alpha}$ is a problem that appears in many recent applications in data mining and optimiza-

tion. However, except for a few choices of $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$, computing such a matrix norm to an arbitrary precision is generally unfeasible for large matrices as this is known to be an NP-hard problem. The situation is different when the matrix has nonnegative entries, in which case $\|A\|_{q \rightarrow p}$ is known to be computable for ℓ^p norms such that $q \leq p$. In this paper we have both (a) refined this result, by showing that the condition $p < q$ is not necessarily required and (b) extended this result to much more general vector norms $\|\cdot\|_\alpha$ and $\|\cdot\|_\beta$ than ℓ^p norms. In particular, we have shown how to compute matrix norms induced by monotonic norms of the form $\|x\|_\alpha = \|(\|x\|_{\alpha_1}, \dots, \|x\|_{\alpha_d})\|_\gamma$, where we also allow $\|x\|_{\alpha_i}$ to measure only a subset of the coordinates of x . Using these kinds of norms we can globally solve in polynomial time quite sophisticated nonconvex optimization problems, as we discuss in the examples corollaries at the end of Sect. 5.

Funding Open access funding provided by Gran Sasso Science Institute - GSSI within the CRUI-CARE Agreement.

Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons licence, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons licence, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons licence and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this licence, visit <http://creativecommons.org/licenses/by/4.0/>.

References

1. Allen-Zhu, Z., Gelashvili, R., Razenshteyn, I.: Restricted isometry property for general p -norms. *IEEE Trans. Inf. Theory* **62**, 5839–5854 (2016)
2. Barak, B., Brandao, F.G., Harrow, A.W., Kelner, J., Steurer, D., Zhou, Y.: Hypercontractivity, sum-of-squares proofs, and their applications. In: *Proceedings of the Forty-fourth Annual ACM Symposium on Theory of Computing, STOC'12*, pp. 307–326. ACM (2012)
3. Bhaskara, A., Vijayaraghavan, A.: Approximating matrix p -norms. In: *Twenty-Second Annual ACM-SIAM Symposium on Discrete Algorithms*, pp. 497–511. SIAM (2011)
4. Boyd, D.W.: The power method for ℓ^p norms. *Linear Algebra Appl.* **9**, 95–101 (1974)
5. Brezis, H.: *Functional Analysis, Sobolev Spaces and Partial Differential Equations*. Springer, New York (2010)
6. Candes, E.J.: The restricted isometry property and its implications for compressed sensing. *C.R. Math.* **346**, 589–592 (2008)
7. Cioranescu, I.: *Geometry of Banach Spaces, Duality Mappings and Nonlinear Problems*, vol. 62. Springer, New York (2012)
8. Clarke, F.H.: *Optimization and Nonsmooth Analysis*, classics, I SIAM (1990)
9. Combettes, P.L., Salzo, S., Villa, S.: Regularized learning schemes in feature Banach spaces. *Anal. Appl.* **16**(01), 1–54 (2018)
10. Drakakis, K., Pearlmutter, B.A.: On the calculation of the $\ell^2 \rightarrow \ell^1$ induced matrix norm. *Int. J. Algebra* **3**, 231–240 (2009)
11. Duchenne, O., Bach, F., Kweon, I.S., Ponce, J.: A tensor-based algorithm for high-order graph matching. *IEEE PAMI* **33**(12), 2383–2395 (2011)
12. Eveson, S.P., Nussbaum, R.D.: An elementary proof of the Birkhoff–Hopf theorem. *Math. Proc. Camb. Philos. Soc.* **117**, 31–54 (1995)
13. Fletcher, R.: *Practical Methods of Optimization*. Wiley, New York (2013)
14. Friedland, S., Lim, L.H.: The computational complexity of duality. *SIAM J. Optim.* **26**, 2378–2393 (2016)
15. Friedland, S., Gaubert, S., Han, L.: Perron–Frobenius theorem for nonnegative multilinear forms and extensions. *Linear Algebra Appl.* **438**, 738–749 (2013)
16. Friedland, S., Lim, L.H., Zhang, J.: Grothendieck constant is norm of Strassen matrix multiplication tensor. *Numer. Math.* **143**(4), 905–922 (2019)

17. Gaubert, S., Zheng, Q.: Dobrushin ergodicity coefficient for Markov operators on cones, and beyond (2013). <https://hal.inria.fr/hal-00935272>. [arXiv:1302.5226](https://arxiv.org/abs/1302.5226)
18. Gautier, A., Hein, M.: Tensor norm and maximal singular vectors of nonnegative tensors—a Perron–Frobenius theorem, a Collatz–Wielandt characterization and a generalized power method. *Linear Algebra Appl.* **505**, 313–343 (2016)
19. Gautier, A., Tudisco, F.: The contractivity of cone-preserving multilinear mappings. *Nonlinearity* **32**, 4713 (2019)
20. Gautier, A., Nguyen, Q.N., Hein, M.: Globally optimal training of generalized polynomial neural networks with nonlinear spectral methods. In: *Advances in Neural Information Processing Systems (NIPS)* (2016)
21. Gautier, A., Tudisco, F., Hein, M.: The Perron–Frobenius theorem for multihomogeneous mappings. *SIAM J. Matrix Anal. Appl.* **40**(3), 1179–1205 (2019)
22. Gautier, A., Tudisco, F., Hein, M.: A unifying Perron–Frobenius theorem for nonnegative tensors via multihomogeneous maps. *SIAM J. Matrix Anal. Appl.* **40**(3), 1206–1231 (2019)
23. Hendrickx, J.M., Olshevsky, A.: Matrix p -norms are NP-hard to approximate if $p \neq 1, 2, \infty$. *SIAM J. Matrix Anal.* **31**, 2802–2812 (2010)
24. Higham, N.J.: Experience with a matrix norm estimator. *SIAM J. Sci. Stat. Comput.* **11**, 804–809 (1990)
25. Higham, N.J.: Estimating the matrix p -norm. *Numer. Math.* **62**, 539–555 (1992)
26. Higham, N.J.: *Accuracy and Stability of Numerical Algorithms*. SIAM (2002)
27. Higham, N.J., Relton, S.D.: Estimating the largest elements of a matrix. *SIAM J. Sci. Comput.* **38**, C584–C601 (2016)
28. Johnson, C.R., Nysten, P.: Monotonicity properties of norms. *Linear Algebra Appl.* **148**, 43–58 (1991)
29. Khamsi, M.A., Kirk, W.A.: *An Introduction to Metric Spaces and Fixed Point Theory*. Wiley-Interscience, New York (2001)
30. Khot, S., Naor, A.: Grothendieck-type inequalities in combinatorial optimization. *Commun. Pure Appl. Math.* **65**, 992–1035 (2012)
31. Lemmens, B., Nussbaum, R.D.: *Nonlinear Perron–Frobenius Theory*, vol. 189. Cambridge University Press, Cambridge (2012)
32. Lewis, A.D.: A top nine list: Most popular induced matrix norms (2010)
33. Lim, L.: Singular values and eigenvalues of tensors: a variational approach. In: *1st IEEE International Workshop on Computational Advances in Multi-Sensor Adaptive Processing CAMSAP'05*, pp. 129–132 (2005). <https://doi.org/10.1109/CAMAP.2005.1574201>
34. Nguyen, Q., Tudisco, F., Gautier, A., Hein, M.: An efficient multilinear optimization framework for hypergraph matching. *IEEE Trans. Pattern Anal. Mach. Intell.* **39**(6), 1054–1075 (2017)
35. Nussbaum, R.D.: Finsler structures for the part metric and Hilbert’s projective metric and applications to ordinary differential equations. *Differ. Integral Equ.* **7**, 1649–1707 (1994)
36. Rohn, J.: Computing the norm $\|A\|_{\infty,1}$ is NP-hard. *Linear Multilinear Algebra* **47**, 195–204 (2000)
37. Seneta, E.: *Inhomogeneous Products of Non-negative Matrices*, pp. 80–111. Springer, New York (1981)
38. Steinberg, D.: *Computation of matrix norms with applications to robust optimization* (2005)
39. Tao, P.D.: Convergence of a subgradient method for computing the bound norm of matrices (in French). *Linear Algebra Appl.* **62**, 163–182 (1984)
40. Tudisco, F., Cardinali, V., Di Fiore, C.: On complex power nonnegative matrices. *Linear Algebra Appl.* **471**, 449–468 (2015)
41. Zhang, H., Zha, Z.J., Yan, S., Wang, M., Chua, T.S.: Robust non-negative graph embedding: Towards noisy data, unreliable graphs, and noisy labels. In: *2012 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 2464–2471 (2012)