A projected subgradient method for the computation of adapted metrics for dynamical systems

Article  (Accepted Version)

Louzeiro, Mauricio, Kawan, Christoph, Hafstein, Sigurdur, Giesl, Peter and Yuan, Jinyun (2022) A projected subgradient method for the computation of adapted metrics for dynamical systems. SIAM Journal on Applied Dynamical Systems, 21 (4). pp. 2297-2696. ISSN 1536-0040

This version is available from Sussex Research Online: http://sro.sussex.ac.uk/id/eprint/107368/

This document is made available in accordance with publisher policies and may differ from the published version or from the version of record. If you wish to cite this item you are advised to consult the publisher’s version. Please see the URL above for details on accessing the published version.

Copyright and reuse:
Sussex Research Online is a digital repository of the research output of the University.

Copyright and all moral rights to the version of the paper presented here belong to the individual author(s) and/or other copyright owners. To the extent reasonable and practicable, the material made available in SRO has been checked for eligibility before being made available.

Copies of full text items generally can be reproduced, displayed or performed and given to third parties in any format or medium for personal research or study, educational, or not-for-profit purposes without prior permission or charge, provided that the authors, title and full bibliographic details are credited, a hyperlink and/or URL is given for the original metadata page and the content is not changed in any way.

http://sro.sussex.ac.uk
A PROJECTED SUBGRADIENT METHOD FOR THE
COMPUTATION OF ADAPTED METRICS FOR DYNAMICAL
SYSTEMS∗

MAURICIO LOUZEIRO†, CHRISTOPH KAWAN‡, SIGURDUR HAFSTEIN§,
PETER GIESL¶, AND JINYUN YUAN∥

Abstract. In this paper, we extend a recently established subgradient method for the computa-
 tion of Riemannian metrics that optimizes certain singular value functions associated with dynamical
 systems. This extension is threefold. First, we introduce a projected subgradient method which re-
sults in Riemannian metrics whose parameters are confined to a compact convex set and we can thus
 prove that a minimizer exists; second, we allow inexact subgradients and study the effect of the errors
 on the computed metrics; and third, we analyze the subgradient algorithm for three different choices
 of step sizes: constant, exogenous and Polyak. The new methods are illustrated by application to
dimension and entropy estimation of the Hénon map.

Key words. Nonlinear systems, singular value optimization, optimization on manifolds, sub-
 gradient algorithm, adapted metrics, dimension estimates, restoration entropy

AMS subject classifications. 93B07, 93B53, 93B70

1. Introduction. Some of the central problems in the theory of smooth dy-
namical systems are related to the optimization of certain singular value functions
of the derivative cocycle associated with the system. These problems include the
computation of Lyapunov exponents, the computation of contraction metrics for ex-
ponentially stable equilibria and the upper estimation of the dimension of attractors.
A more recent objective, motivated by problems in information-based control, is the
computation of the minimal data rate at which sensory data has to be transmitted
to an observer at a remote location such that it can produce reliable state estimates.
Given a discrete-time system induced by a smooth map \( \phi : \mathbb{R}^n \to \mathbb{R}^n \), in most of these
problems one is interested in the minimization of a function of the type

\[
\Sigma_k : P \mapsto \max_{x \in K} \sum_{i=1}^{k} \log \alpha^P_i(x),
\]

where \( K \) is a compact forward-invariant set of \( \phi \) and \( \alpha^\phi_1(x) \geq \alpha^\phi_2(x) \geq \ldots \geq \alpha^\phi_n(x) \)
denote the singular values of the derivative \( D\phi(x) \), computed with respect to a Rie-
mannian metric \( P(\cdot) \) on \( K \) (see Subsection 2.2 for the technical definition). For
instance, in the case of contraction analysis for equilibria, one is interested in finding
a metric \( P(\cdot) \) such that \( \Sigma_1(P) < 1 \), and in the case of the remote estimation problem,
the smallest data rate is given by the minimum of \( P \mapsto \max\{\Sigma_k(P) \cdot 0 \leq k \leq n\} \).

∗Submitted to the editors DATE.
Funding: This work was funded by the German Research Foundation (DFG) through the grant
ZA 873/4-1 and by through the grant NSFC 12171087.
†Dongguan University of Technology, School of Computer Science and Technology, Dongguan,
Guangdong, China (mauriciosilvalouzeiro@gmail.com)
‡Institute of Informatics, LMU Munich, Oettingenstraße 67, 80538 München, Germany
(christoph.kawan@lmu.de).
§The Science Institute, University of Iceland, Dunhagi 5, 107 Reykjavik, Iceland (shafstein@hi.is).
¶Department of Mathematics, University of Sussex, Falmer, BN1 9QH, United Kingdom
(P.A.Giesl@sussex.ac.uk).
∥Dongguan University of Technology, School of Computer Science and Technology, Dongguan,
Guangdong, China (yuanjy@gmail.com).

This manuscript is for review purposes only.
In [20], we established a convexity property of the functions $\Sigma_k$, which allowed us
to solve a restricted version of the minimization problem via the subgradient algorithm
on manifolds with lower bounded sectional curvature [14, 37, 12]. Namely, we had
34 to restrict the class of Riemannian metrics to metrics of the form $P(x) = e^{r(x)} p$
with a polynomial $r(x)$ and a positive definite matrix $p$. The problem can then be
37 formulated as a geodesically convex optimization problem on a manifold of the form
$\mathbb{R}^N \times \mathcal{S}_n^+$, where the Euclidean component $\mathbb{R}^N$ serves as the parameter space for the
polynomials and $\mathcal{S}_n^+$ is the Riemannian manifold of positive definite matrices. We
demonstrated the efficiency of the algorithm by applying it to the computation of the
minimal data rate in the remote state estimation problem (described by the notion of
restoration entropy) for three nonlinear systems - the Hénon system, a bouncing ball
system and the Lorenz system. In all three examples, we observed a fast convergence
and, where the exact value of the entropy was known, the same value was obtained
by our algorithm with a high accuracy.

Three drawbacks of the algorithm developed in [20] are the following: first, the
computation of a subgradient, which has to be performed in each step of the algorithm,
50 involves a maximization problem that can, in general, only be solved by brute force.
53 This slows the algorithm down and additionally introduces an error in the computed
subgradient that is hard to quantify. Second, we do not know general conditions under
56 which the optimization problem admits a solution. This implies that the algorithm
does not necessarily converge. In particular, if no minimizer exists, the computed
59 metric could become asymptotically singular. Finally, using the exogenous step size
rule, we cannot provide an explicit estimate on the rate of convergence even if a
62 minimizer exists. In this paper, we attempt to overcome all these drawbacks to a
65 certain extent.

A version of the classical (Euclidean) subgradient algorithm is the projected sub-
68 gradient algorithm in which the search for a minimizer is only carried out within a
compact convex subset of the given domain. The compactness guarantees that for this
61 restricted problem a minimizer exists. Moreover, any upper bound on the diameter of
67 the compact set together with an estimate on the maximal norm of the subgradient
allows us to provide explicit estimates on the number of steps necessary to achieve a
69 given accuracy. In this paper, we analyze an inexact projected subgradient algorithm
66 on Hadamard manifolds and provide convergence results for three different choices of
60 step sizes: exogenous, constant and Polyak. In particular, we prove that errors in the
63 computation of the subgradient do not accumulate as the number of steps grows.

We then study order intervals on $\mathcal{S}_n^+$ as candidates for compact convex subsets,
which can be used in the projected subgradient algorithm for singular value optimization.
We further analyze the error in the computation of the subgradient for singular
76 value optimization, which turns out to depend sensitively on a spectral gap in the
singular value spectrum. Finally, we compute a Lipschitz constant for the functions
$\Sigma_k$, which only depends on the dimension $n$. This analysis allows us to apply the
subgradient algorithm with different types of step sizes to the problems of dimension
66 and data rate estimation for the Hénon system.

The advantages of the projected subgradient algorithm thus are the following:
78 the existence of a minimizer is guaranteed and we can prove the convergence; we can
ensure that the computed Riemannian metrics do not have values that are too close
70 to the boundary of $\mathcal{S}_n^+$, which is important in practical applications; we can provide
75 explicit estimates on the rate of convergence and we can use Polyak step sizes, which
delivered the best upper bound on the restoration entropy for our example system,
80 see Section 5.2.
Let us give an overview of the paper: In Section 2 we introduce notations as well as the subgradient algorithm. In Section 3 we introduce the projected subgradient method on Hadamard manifolds with inexact subgradients and perform a convergence analysis with three choices of step sizes. Section 4 applies the projected inexact subgradient method to the singular value optimization problems and presents further results, in particular an error analysis. In Section 5 the methods are applied to derive upper bounds on the dimension and entropy of the Hénon system, before conclusions in Section 5. In the appendix, we determine the tightest lower bound for the sectional curvature of $S^+_n$ equipped with the trace metric.

2. Preliminaries.

2.1. General notation and definitions. We write $\mathbb{N} = \{1, 2, 3, \ldots\}$ for the natural numbers and $\mathbb{N}_0 := \mathbb{N} \cup \{0\}$. By $\mathbb{R}$ and $\mathbb{R}_+$, we denote the reals and non-negative reals, respectively. By $\ln$ (log), we denote the natural logarithm (logarithm to the base 2).

Matrices. We write $\mathbb{R}^{n \times n}$ for the vector space of all $n \times n$ real matrices. The notations $A^\top$, $\det(A)$ and $\tr(A)$ stand for the transpose, the determinant and the trace of $A \in \mathbb{R}^{n \times n}$, respectively. For the singular values of $A$, i.e. the eigenvalues of $\sqrt{A^\top A}$, we write $\alpha_1(A) \geq \ldots \geq \alpha_n(A)$. By $I_{n \times n}$ and $0_{n \times n}$, we denote the identity and the zero matrix in $\mathbb{R}^{n \times n}$, respectively. If the dimension is clear from the context, we also omit the subscript. The general linear group is denoted by $\GL(n, \mathbb{R}) = \{A \in \mathbb{R}^{n \times n} : A^{-1} \text{ exists}\}$. The notation $\text{diag}(\lambda_1, \ldots, \lambda_n)$ is used for the diagonal matrix with entries $\lambda_1, \ldots, \lambda_n$. We write $\langle \cdot, \cdot \rangle_F$ for the standard inner product on $\mathbb{R}^{n \times n}$ which is given by $\langle X, Y \rangle_F = \tr(XY^\top)$ and induces the Frobenius norm $\|X\|_F = \sqrt{\tr(XX^\top)}$.

Manifolds. Next, we present some notations, definitions and basic properties of Riemannian manifolds used throughout the paper; for more details, we refer the reader to [11, 31, 24, 2].

If $\mathcal{M}$ is a smooth manifold, we write $T_p\mathcal{M}$ for the tangent space at $p \in \mathcal{M}$. The derivative of a smooth mapping $f : \mathcal{M} \to \mathcal{N}$ between manifolds at $p \in \mathcal{M}$ is denoted by $Df(p) : T_p\mathcal{M} \to T_{f(p)}\mathcal{N}$. If $K \subset \mathcal{M}$ is any subset, we write $C^k(K, \mathcal{N})$ for the set of $k$ times continuously differentiable maps from $K$ into $\mathcal{N}$, where $k = 0$ means continuous. We write $\text{int}(A)$ and $\text{cl}(A)$ for the interior and closure of a subset $A \subset \mathcal{M}$, respectively. If $\mathcal{M}$ is equipped with a Riemannian metric, we write $\langle \cdot, \cdot \rangle_p$ and $\| \cdot \|_p$ for the associated inner product and norm on $T_p\mathcal{M}$, respectively. If the point $p$ is clear from the context, we also omit the subscript $p$ in the inner product and the norm.

By $\exp_p$, we denote the Riemannian exponential map at $p \in \mathcal{M}$. The diameter of a subset $\mathcal{X} \subset \mathcal{M}$ is given by $\text{diam}(\mathcal{X}) = \sup_{x, y \in \mathcal{X}} d(x, y)$, where $d(\cdot, \cdot)$ is the Riemannian distance function on $\mathcal{M}$. The sectional curvature of the two-dimensional subspace of $T_p\mathcal{M}$, spanned by two linearly independent vectors $x, y \in T_p\mathcal{M}$, is defined by

$$
K(x, y) := \frac{\langle R(x, y)y, x \rangle}{\|x\|^2\|y\|^2 - \langle x, y \rangle^2},
$$

where $R$ is the curvature tensor of $\mathcal{M}$, evaluated at the point $p$.

A Riemannian manifold $\mathcal{M}$ is called a Hadamard manifold if it is simply connected, complete and its sectional curvature is non-positive everywhere. In this case, for each $p \in \mathcal{M}$, the exponential map $\exp_p : T_p\mathcal{M} \to \mathcal{M}$ is a diffeomorphism, which implies that the geodesics connecting any two distinct points exist and are unique. We write $\log_p : \mathcal{M} \to T_p\mathcal{M}$ for the inverse of $\exp_p$.

The geodesic $\gamma : [0, 1] \to \mathcal{M}$ connecting $p$ and $q$ is denoted by $\gamma_{p,q}$. A nonempty subset $\mathcal{X} \subset \mathcal{M}$ is called (geodesically) convex if the image of $\gamma_{p,q}$ lies completely in $\mathcal{X}$.
\[ \mathcal{D} \] for any two points \( p, q \in \mathcal{X} \). If \( \mathcal{D} \) is also closed, then the mapping \( \mathcal{P}_\mathcal{D}(p) := \{ q \in \mathcal{X} : d(p, q) = \inf_{q' \in \mathcal{X}} d(p, q') \} \) is single-valued and non-expansive, i.e.

\[
(2.2) \quad d(\mathcal{P}_\mathcal{D}(p), \mathcal{P}_\mathcal{D}(q)) \leq d(p, q) \quad \text{for all } p, q \in \mathcal{M}.
\]

In this case, \( \mathcal{P}_\mathcal{D}(p) \) is called the projection of \( p \) onto \( \mathcal{D} \).

If \( \mathcal{M} \) is a Hadamard manifold, a function \( f: \mathcal{M} \to \mathbb{R} := \mathbb{R} \cup \{ \pm \infty \} \) is called

(i) proper if \( \text{dom } f := \{ p \in \mathcal{M} : f(p) < \infty \} \neq \emptyset \) and \( f(p) > -\infty \) for all \( p \in \mathcal{M} \).

(ii) convex if, for all \( p, q \in \mathcal{M} \), the composition \( f \circ \gamma_{p,q} : [0,1] \subset \mathbb{R} \to \mathbb{R} \) is a convex function on \([0,1]\) in the classical sense.

(iii) lower semicontinuous at \( p \in \text{dom } f \) if for each sequence \( \{ p_k \} \) converging to \( p \), we have \( \lim \inf_{k \to \infty} f(p_k) \geq f(p) \).

The subdifferential of a proper convex function \( f: \mathcal{M} \to \mathbb{R} \) at \( p \in \text{dom } f \) is defined by

\[
(2.3) \quad \partial f(p) := \{ s \in T_p \mathcal{M} : f(q) \geq f(p) + \langle s, \log_p q \rangle, \forall q \in \text{dom } f \}.
\]

Each element of \( \partial f(p) \) is called a subgradient of \( f \) at \( p \). We remark that \( \partial f(p) \) is nonempty for all \( p \in \text{int}(\text{dom } f) \); see [36, Prop. 2.5]

Positive definite matrices. Next, we recall some facts about the space of positive definite matrices. For a comprehensive treatment, we refer to [5, Ch. 6].

We define \( S_n := \{ A \in \mathbb{R}^{n \times n} : A = A^\top \} \) and write \( A \leq B \) \( (A < B) \) if \( A, B \in S_n \) and \( B - A \) is positive semidefinite (positive definite). The eigenvalues of any \( A \in S_n \) are denoted by \( \lambda_1(A) \geq \ldots \geq \lambda_n(A) \). The space of all positive definite matrices is denoted by \( S_n^+ := \{ A \in S_n : A > 0 \} \). As an open subset of the Euclidean space \( S_n \), it is a smooth manifold. The tangent space of \( S_n^+ \) at any point \( p \) can be canonically identified with \( S_n \). We equip \( S_n^+ \) with the so-called trace metric

\[
(\langle v, w \rangle)_p := \text{tr}(p^{-1}v p^{-1}w) \quad \text{for all } p \in S_n^+, \quad v, w \in T_p S_n^+ = S_n.
\]

Some facts about the Riemannian manifold \( S_n^+ \) (see [5, Ch. 6] and [23, Ch. 12], for instance), used in this paper, are the following:

- \( S_n^+ \) is a Hadamard manifold as well as a symmetric space.
- The group \( \text{GL}(n, \mathbb{R}) \) acts transitively on \( S_n^+ \) by isometries. This action is given by \((g, p) \mapsto g p g^{-1}\).
- The inversion \( g \mapsto g^{-1} \) is an isometry of \( S_n^+ \).
- The unique geodesic joining two points \( p, q \in S_n^+ \) is given by

\[
\gamma_{p,q}(\theta) = p \#_{\theta} q := p^{\frac{1}{2}} (p^{-\frac{1}{2}} q p^{-\frac{1}{2}})^{\theta} p^{\frac{1}{2}}, \quad \theta \in [0, 1].
\]

We also write \( p \#_{\frac{1}{2}} q \) for the midpoint of the geodesic.

- The unique geodesic \( \gamma \) with \( \gamma(0) = p \) and \( \gamma(1) = q \) is given by

\[
\gamma(\theta) = p^{\frac{\theta}{2}} \exp(\theta p^{-\frac{1}{2}} v p^{-\frac{1}{2}}) p^{\frac{1}{2}}, \quad \theta \in \mathbb{R}.
\]

- The distance function on \( S_n^+ \) can be written as

\[
(2.4) \quad d(p, q) = \left( \sum_{i=1}^{n} \ln^2 \lambda_i(p^{-1} q) \right)^{\frac{1}{2}}.
\]

Observe that although \( p^{-1} q \) is not necessarily symmetric, it has real eigenvalues, since it is similar to the symmetric matrix \( p^{-1/2} q p^{-1/2} \).
2.2. The subgradient method for singular value optimization. For any 
\( g \in \text{GL}(n, \mathbb{R}) \), we write
\[
\tilde{\sigma}(g) := (\log_2 \alpha_1(g), \ldots, \log_2 \alpha_n(g)),
\]
which defines a mapping from \( \text{GL}(n, \mathbb{R}) \) into the cone
\[
\mathfrak{a}^+ := \{ \xi \in \mathbb{R}^n : \xi_1 \geq \xi_2 \geq \ldots \geq \xi_n \}.
\]
On \( \mathfrak{a}^+ \), we define the partial order
\[
\xi \preceq \eta :\Leftrightarrow \left\{ \begin{array}{ll}
\xi_1 + \cdots + \xi_k \leq \eta_1 + \cdots + \eta_k & \text{for } k = 1, \ldots, n - 1, \\
\xi_1 + \cdots + \xi_k = \eta_1 + \cdots + \eta_k & \text{for } k = n.
\end{array} \right.
\]
Now, consider a discrete-time dynamical system
\[ x_{t+1} = \phi(x_t) \]
with a \( C^1 \)-map \( \phi : \mathbb{R}^n \to \mathbb{R}^n \). We assume that \( K \subset \mathbb{R}^n \) is a compact forward-invariant set, i.e. \( \phi(K) \subseteq K \). Moreover, we assume that \( K = \text{cl}(\text{int}K) \) and
\[ A(x) := D\phi(x) \in \text{GL}(n, \mathbb{R}) \quad \text{for all } x \in K. \]
A Riemannian metric on \( K \) can be regarded as a continuous map \( P : K \to S^+_n \). We define the singular values \( \alpha_1(x) \geq \ldots \geq \alpha_n(x) \) of \( A(x) \) with respect to the metric \( P \) as the eigenvalues of \( [B(x)B(x)^\top]^{\frac{1}{2}} \), where
\[ B(x) := P(\phi(x))^\frac{1}{2}A(x)P(x)^{-\frac{1}{2}}. \]
That is, \( \alpha_i^P(x) \) are the ordinary singular values of \( B(x) \), or the singular values of \( A(x) \) regarded as a linear operator between the inner product spaces \( (\mathbb{R}^n, \langle P(x) \cdot, \cdot \rangle) \) and \( (\mathbb{R}^n, \langle P(\phi(x)) \cdot, \cdot \rangle) \), respectively, see [21, Lem. 5].

In [20, Lem. 3.2], we have proved that for any two metrics \( P \) and \( Q \), the relation
\[
\tilde{\sigma}([P(\phi(x))^\#_\theta Q(\phi(x))]^{\frac{1}{2}}A(x)[P(x)^\#_\theta Q(x)]^{-\frac{1}{2}}) \\
\leq (1 - \theta)\tilde{\sigma}(P(\phi(x))^\frac{1}{2}A(x)P(x)^{-\frac{1}{2}}) + \theta\tilde{\sigma}(Q(\phi(x))^\frac{1}{2}A(x)Q(x)^{-\frac{1}{2}})
\]
holds for all \( \theta \in [0, 1] \) and \( x \in K \). This can be regarded as a form of (geodesic) convexity. For each \( k \in \{1, \ldots, n\} \) and \( x \in K \), we introduce the function
\[ \Sigma_{k,x} : C^0(K, S^+_n) \to \mathbb{R}, \quad \Sigma_{k,x}(P) := \sum_{i=1}^k \log_2 \alpha_i^P(x). \]
Then, (2.5) implies that for all \( P, Q \in C^0(K, S^+_n) \) and \( \theta \in [0, 1] \) the inequality
\[ \Sigma_{k,x}(P^\#_\theta Q) \leq (1 - \theta)\Sigma_{k,x}(P) + \theta\Sigma_{k,x}(Q) \]
holds, where \( P^\#_\theta Q \) denotes the Riemannian metric
\[ (P^\#_\theta Q)(x) := P(x)^\#_\theta Q(x) \quad \text{for all } x \in K. \]
The same inequality then also holds for the functions

\[ \Sigma_k : C^0(K, S^+_{n}) \to \mathbb{R}, \quad \Sigma_k(P) := \max_{x \in K} \Sigma_{k,x}(P). \]

We are interested in minimizing one of the functions \( \Sigma_k \) or their maximum \( \max_{0 \leq k \leq n} \Sigma_k \), where \( \Sigma_0(P) \equiv 0 \). To attack this problem via convex optimization, we have to restrict the domain to a finite-dimensional geodesically convex space. This can be achieved by considering only metrics of the form \( P(x) = e^{r(x)P} \) with a polynomial function \( r(x) \) satisfying \( \deg r(x) \leq d \) for a given \( d \in \mathbb{N} \) and \( p \in S^+_{n} \). The space of all such metrics will be denoted by \( C_d(K) \) and can be identified with \( \mathbb{R}^N \times S^+_{n} \), where \( N = \binom{d+n}{n} \) is the number of coefficients in the polynomial. We note that \( \mathbb{R}^N \times S^+_{n} \) is a Hadamard manifold whose sectional curvature is bounded below by \( -\frac{1}{2} \), see Proposition A.2. This allows us to use different variants of the subgradient algorithm on manifolds for convex optimization problems on \( \mathbb{R}^N \times S^+_{n} \).

We consider two problems in this paper: dimension and entropy estimation. In the first case, it is well-known (see, e.g., [6, Thm. 9.1.1]) that the Lyapunov, and thus the Hausdorff dimension of \( K \) can be bounded by \( \dim_L(K) \leq k + s \) for an integer \( k \in \{0, 1, \ldots, n - 1\} \) and \( s \in [0, 1] \) provided that \( K \) is backward-invariant and for some Riemannian metric \( P \) the following inequality holds:

\[ \alpha^P_1(x) \alpha^P_2(x) \ldots \alpha^P_k(x) \alpha^P_{k+1}(x)^s < 1 \quad \text{for all } x \in K. \]

Note that it is sufficient to verify this condition for all \( x \in \tilde{K} \) in a larger set \( \tilde{K} \supseteq K \), that is compact and forward invariant and thus contains an invariant set \( K = \omega(\tilde{K}) \).

To verify such a condition with our methods, we introduce the functions

\[ \Sigma_{k+s,x}(P) := \sum_{i=1}^{k} \log_2 \alpha^P_i(x) + s \log_2 \alpha^P_{k+1}(x) \]

for any \( k, s \) as above and \( x \in K \). We can write \( \Sigma_{k+s,x}(P) = s \Sigma_{k+1,x}(P) + (1 - s) \Sigma_{k,x}(P) \), implying that \( \Sigma_{k+s,x} \) satisfies the same convexity properties as described above. Hence, given any guess on \( k \) and \( s \), we can use the subgradient algorithm to minimize \( P \mapsto \max_{x \in K} \Sigma_{k+s,x}(P) \) over \( C_d(K) \) for some \( d \in \mathbb{N} \) and check if the minimum is less than 0. If this is the case, we can do the same for smaller choices of \( k \) and \( s \) in order to optimize the dimension estimate. If the minimum happens to be larger than 1, we have to increase \( d \) or \( s \).

The second problem is the upper estimation of restoration entropy, which measures the smallest data rate above which the system is regularly or finely observable on \( K \) over a digital channel operating at this rate, see [29] for precise definitions. By [21], the restoration entropy on \( K \) satisfies

\[ h_{\text{res}}(\phi, K) = \inf_{P \in C_d(K, S^+_{n})} \max_{0 \leq k \leq n} \Sigma_k(P). \]

Hence, we can again restrict the class of Riemannian metrics to \( C_d(K) \) for some \( d \) and minimize the convex function \( \max \{ \Sigma_k(P) : k = 0, 1, \ldots, n \} \) over \( C_d(K) \).

We fix a parametrization of the space of polynomials in \( n \) variables and degree \( \leq d \) over \( \mathbb{R}^N \) with \( N = \binom{d+n}{n} \), such that \( r_a(x) \) denotes the polynomial with coefficient vector \( a \in \mathbb{R}^N \). We then introduce the functions

\[ J_{k+s,x}(a, p) := \Sigma_{k+s,x}(e^{r_a(\cdot)p}), \quad J_{k+s,x} : \mathbb{R}^N \times S^+_{n} \to \mathbb{R}. \]
for $k \in \{0, 1, \ldots, n-1\}$ and $s \in [0, 1)$. We further introduce

$$J_{k+s}(a, p) := \max_{x \in K} J_{k+s,x}(a, p), \quad J_{k+s} : \mathbb{R}^N \times S_n^+ \to \mathbb{R}.$$  

Via the following proposition, we can reduce the computation of subgradients for our objective functions to the computation of subgradients of the functions $J_{k,x}$ with integer $k$. Its proof follows directly from the definition of the subdifferential and the definitions of the functions involved.

**Proposition 2.1.** The following statements hold:

(i) Assume that the maximum of $J_{k+s,x}(a, p)$ over $x \in K$ is attained at $x_*$. Then any subgradient $v$ of $J_{k+s,x_*}$ at $(a, p)$ is also a subgradient of $J_{k+s}$ at $(a, p)$.

(ii) If $v, w \in \mathbb{R}^N \times S_n$ are subgradients of $J_{k,x}$ and $J_{k+1,x}$ at $(a, p)$, respectively, then $(1 - s)v + sw$ is a subgradient of $J_{k+s,x}$ at $(a, p)$.

(iii) If the maximum of $J_{k,x}(a, p)$, $k \in \{0, 1, \ldots, n\}$, is attained at $k_*$, then any subgradient $v$ of $J_{k,x}$ at $(a, p)$ is a subgradient of $(a, p) \mapsto \max\{J_{k,x}(a, p) : 0 \leq k \leq n\}$ at $(a, p)$.

**Proof.** (i) If $v$ is a subgradient of $J_{k+s,x_*}$ at $(a, p)$ then

$$J_{k+s}(b, q) = \max_{x \in K} J_{k+s,x}(b, q),$$

$$\geq J_{k+s,x_*}(b, q),$$

$$\geq J_{k+s,x_*}(a, p) + \langle v, \log_{(a,p)}(b, q) \rangle,$$

$$= J_{k+s}(a, p) + \langle v, \log_{(a,p)}(b, q) \rangle, \quad \forall (b, q) \in \mathbb{R}^N \times S_n^+,$$

which implies that $v$ is a subgradient of $J_{k+s}$ at $(a, p)$.

(ii) If $v, w \in \mathbb{R}^N \times S_n$ are subgradients of $J_{k,x}$ and $J_{k+1,x}$ at $(a, p)$, respectively, then

$$J_{k,x}(b, q) \geq J_{k,x}(a, p) + \langle v, \log_{(a,p)}(b, q) \rangle, \quad \forall (b, q) \in \mathbb{R}^N \times S_n^+,$$

and

$$J_{k+1,x}(b, q) \geq J_{k+1,x}(a, p) + \langle w, \log_{(a,p)}(b, q) \rangle, \quad \forall (b, q) \in \mathbb{R}^N \times S_n^+.$$

Multiplying (2.6) by $1 - s$ and (2.7) by $s$, and adding these new inequalities, we get

$$((1 - s)J_{k,x} + sJ_{k+1,x})(b, q) \geq ((1 - s)J_{k,x} + sJ_{k+1,x})(a, p)$$

$$+ ((1 - s)v + sw, \log_{(a,p)}(b, q)), \quad \forall (b, q) \in \mathbb{R}^N \times S_n^+,$$

Hence, the conclusion of the proof follows from the definition of subgradient and the fact that $J_{k+s,x} = (1 - s)J_{k,x} + sJ_{k+1,x}$.

(iii) If $v$ is a subgradient of $J_{k,x}$ at $(a, p)$ then

$$\max\{J_{k,x}(b, q) : 0 \leq k \leq n\} \geq J_{k,x}(b, q),$$

$$\geq J_{k,x}(a, p) + \langle v, \log_{(a,p)}(b, q) \rangle,$$

$$= \max\{J_{k,x}(a, p) : 0 \leq k \leq n\} + \langle v, \log_{(a,p)}(b, q) \rangle,$$

for all $(b, q) \in \mathbb{R}^N \times S_n^+$. This completes the proof. 

This manuscript is for review purposes only.
A straightforward computation shows that
\[
J_{k+s,x}(a,p) = \frac{k+s}{2\ln 2} (r_a(\phi(x)) - r_a(x)) + \sum_{k+s,x}(p).
\]

It thus suffices to compute subgradients of \( a \mapsto J^1_{k,x}(a) \) and \( p \mapsto J^2_{k,x}(p) \) separately.

Since \( J^1_{k+s,x} \) is a linear function, the only subgradient is its constant gradient which is a polynomial in the components of \( \phi(x) \) and \( x \). The computation of a subgradient for \( J^2_{k,x} \) is explained in [20]. We briefly summarize the procedure how to compute it. Here, we use the notation \( \zeta_x(p) := p^\frac{1}{2} A(x)p^{-\frac{1}{2}} \):

1. Compute an orthonormal basis \( \{e_i\} \) of the tangent space \( (S_n, \langle \cdot, \cdot \rangle_p) \).
2. Determine the solutions \( X_i \in S_n \) of the Lyapunov equations
   \[
p^\frac{1}{2} X_i + X_i p^{-\frac{1}{2}} = e_i.
   \]
3. Compute the matrices
   \[
   Z_i := D\zeta_x(p)e_i = X_i A(x)p^{-\frac{1}{2}} - p^\frac{1}{2} A(x)p^{-\frac{1}{2}} X_i p^{-\frac{1}{2}}.
   \]
4. Compute a singular value decomposition
   \[
   \zeta_x(p) = U^\top \text{diag}(\alpha_1(\zeta_x(p)), \ldots, \alpha_n(\zeta_x(p))) V.
   \]
5. Compute
   \[
   S_x := \frac{1}{\ln 2} U^\top \text{diag}(\alpha_1(\zeta_x(p))^{-1}, \ldots, \alpha_k(\zeta_x(p))^{-1}, 0, \ldots, 0) V.
   \]
6. Compute a subgradient \( v \) of \( J^2_{k,x} \) by \( v = \sum_i \text{tr}[S_x^T Z_i] e_i \).

To be precise, we can only guarantee that the above procedure works for \( k \) for which there is a spectral gap, i.e. \( \alpha_k(\zeta_x(p)) > \alpha_{k+1}(\zeta_x(p)) \).

3. Projected subgradient method on Hadamard manifolds. In this section, we introduce the projected subgradient method on Hadamard manifolds with inexact subgradients and analyze its convergence properties for different choices of step sizes. Later, we apply the method to the problem of singular value optimization for dynamical systems, where the Hadamard manifold is of the form \( M = R^N \times S_n^+ \) for some \( n, N \in N \) and the objective function is one of the functions \( J_{k+s} \) or

\[
\max\{J_k : 0 \leq k \leq n\},
\]
as introduced in the preceding section.

Let \( f : M \to R \) be a proper convex function. We consider the subgradient method to solve the optimization problem defined by

\[
\min\{f(p) : p \in X\},
\]

where \( M \) is a Hadamard manifold with lower bounded sectional curvature and \( X \subseteq M \) is a compact convex set. For background material about optimization on manifolds, we refer the reader to, e.g., [1, 8, 32, 2, 13, 35, 10, 27, 33, 16].

Throughout this section, we denote by \( D \) an upper bound for the diameter of \( X \), i.e., \( \text{diam}(X) \leq D \). The analysis presented here is based on [15, 14].
3.1. Preliminary results. In this section, all functions \( f : \mathcal{M} \to \mathbb{R} \) are assumed to be convex and lower semicontinuous on \( \mathcal{M} \).

The following result is proved in [36, Prop. 2.5].

**Proposition 3.1.** Let \( \{p_k\} \subset \mathcal{M} \) be a bounded sequence. If \( \{s_k\} \) is a sequence satisfying \( s_k \in \partial f(p_k) \) for each \( k \in \mathbb{N} \), then \( \{s_k\} \) is bounded as well.

Throughout this section, we let \( \kappa \) be a lower bound on the sectional curvature of \( \mathcal{M} \) and define \( \tilde{\kappa} := \sqrt{|\kappa|} \).

The following lemma plays an important role in the next subsections. Its proof will be omitted here, but it can be obtained, with some minor technical adjustments, by using Toponogov’s theorem [31, p. 161, Thm. 4.2] and following the ideas of [36, Lem. 3.2].

**Lemma 3.2.** Let \( \mathcal{M} \) be a Hadamard manifold with sectional curvature \( \geq \kappa \). Let \( p, q \in \mathcal{M}, p \neq q, v \in T_p \mathcal{M}, v \neq 0 \), and \( \gamma : [0, \infty) \to \mathcal{M} \) be defined by \( \gamma(t) := \exp_p(-tv/\|v\|) \). Then, for any \( t \in [0, \infty) \), it holds that

\[
d^2(\gamma(t), q) \leq d^2(p, q) + \frac{\sinh (t\tilde{\kappa})}{\kappa} \left( \frac{\tilde{\kappa}d(p, q)}{\tanh (t\tilde{\kappa}d(p, q))} t + 2 \langle v, \log_p q \rangle \right).
\]

3.2. Inexact projected subgradient method. We denote by \( f^* := \inf_{p \in \mathcal{X}} f(p) \) the optimal value of problem (3.1) and by \( \Omega^* := \{ p \in \mathcal{X} : f(p) = f^* \} \) its solution set. Since \( f \) is lower semicontinuous and \( \mathcal{X} \) is compact, it is not difficult to see that \( \Omega^* \neq \emptyset \).

3.2.1. Algorithm. We propose to solve problem (3.1) with the *inexact Riemannian subgradient algorithm*, stated below.

**Algorithm 3.1**

**Step 0.** Let \( p_0 \in \mathcal{X} \subset \text{int}(\text{dom} f) \). Set \( k = 0 \).

**Step 1.** Choose a step size \( t_k > 0 \) and compute

\[
p_{k+1} = \mathcal{P}_\mathcal{X}(\tilde{p}_{k+1}), \quad \tilde{p}_{k+1} := \exp_{p_k} \left( -t_k \frac{v_k}{\|v_k\|} \right),
\]

where \( v_k := s_k + u_k \) (\( v_k \neq 0 \)) is an approximation for some \( s_k \in \partial f(p_k) \).

**Step 2.** Set \( k \leftarrow k + 1 \) and proceed to Step 1.

**Remark 3.3.** Algorithm 3.1 can be presented with a retraction \( R \) ([1, Definition 4.1.1]) instead of \( \exp \). This alternative is interesting because it would no longer be necessary to know the geodesics of \( \mathcal{M} \) to apply the algorithm, which can be quite useful in cases where the calculation of geodesics is difficult. In this paper, we chose to approach only the case where \( R \) is equal to \( \exp \) because the geodesics of the Riemannian manifold \( S^*_n \) are already well known in the literature. Also, our convergence analysis may not be easily adapted to Algorithm 3.1 with \( R \) arbitrary, since a version of Lemma 3.2 with \( R \) instead of \( \exp \) would be required. An imminent challenge is to overcome the fact that the curve \( R_p(-tv/\|v\|) : [0, \infty) \to \mathcal{M}, p \in \mathcal{M}, v \in T_p \mathcal{M}, \) (\( R_p \) denotes the restriction of \( R \) to \( T_p \mathcal{M} \)) is not necessarily a geodesic, and therefore dealing with the inability to use important geometric tools such as Toponogov’s The-
it follows that

\[
\text{Lemma 3.4. In Algorithm 3.1, let } \sigma > 0 \text{ and } \zeta := D \sinh (\hat{\kappa} \sigma) / (\kappa \tanh (\hat{\kappa} D)). \text{ If }
\]

\[
\sup_{k} t_{k} \leq \sigma, \text{ then }
\]

\[
d^{2}(p_{k+1}, q) \leq d^{2}(p_{k}, q) + \zeta t_{k}^{2} + 2 \frac{\sinh (\hat{\kappa} t_{k})}{\hat{\kappa}} \cdot \frac{f(q) - f(p_{k}) + \|u_{k}\| D}{\|u_{k} + s_{k}\|}
\]

for any \( q \in \mathcal{X} \) and \( k \in \mathbb{N}_{0} \).

Proof. Take \( q \in \mathcal{X} \) and \( k \in \mathbb{N}_{0} \). Applying Lemma 3.2 with \( p = p_{k}, v = s_{k} + u_{k} \) and \( t = t_{k} \), we obtain

\[
d^{2}(\hat{p}_{k+1}, q) \leq d^{2}(p_{k}, q) + \frac{\sinh (\hat{\kappa} t_{k})}{\hat{\kappa}} \left( \frac{\hat{\kappa} d(p_{k}, q)}{\tanh (\hat{\kappa} d(p_{k}, q))} t_{k} + 2 \left( s_{k}, \log_{p_{k}} q \right) + \left( u_{k}, \log_{p_{k}} q \right) \right)
\]

where \( \hat{p}_{k+1} \) is defined in (3.2). Using (2.2), (2.3) and the Cauchy-Schwarz inequality, it follows that

\[
d^{2}(p_{k+1}, q) \leq d^{2}(p_{k}, q) + \frac{\sinh (\hat{\kappa} t_{k})}{\hat{\kappa}} \cdot \frac{\hat{\kappa} d(p_{k}, q)}{\tanh (\hat{\kappa} d(p_{k}, q))} t_{k}^{2} + 2 \frac{\sinh (\hat{\kappa} t_{k})}{\hat{\kappa}} \cdot \frac{f(q) - f(p_{k}) + \|u_{k}\| d(p_{k}, q)}{\|s_{k} + u_{k}\|}.
\]

The maps \((0, \infty) \ni t \mapsto \sinh(t)/t\) and \((0, \infty) \ni t \mapsto t/\tanh(t)\) are positive and increasing. Thus, the last inequality above implies

\[
d^{2}(p_{k+1}, q) \leq d^{2}(p_{k}, q) + \frac{\sinh (\hat{\kappa} \sigma)}{\hat{\kappa} \sigma} \cdot \frac{\hat{\kappa} D}{\tanh (\hat{\kappa} D)} t_{k}^{2} + 2 \frac{\sinh (\hat{\kappa} t_{k})}{\hat{\kappa}} \cdot \frac{f(q) - f(p_{k}) + \|u_{k}\| D}{\|s_{k} + u_{k}\|}.
\]

It is not difficult to see that this inequality concurs with (3.3). Therefore, the proof is complete.

3.3. Convergence analysis.

3.3.1. Exogenous step size. In this subsection, we assume that the sequence \( \{p_{k}\} \) is generated by Algorithm 3.1 with exogenous step size, i.e., \( t_{k} \) satisfies

\[
t_{k} > 0, \quad \sum_{k=0}^{\infty} t_{k} = \infty, \quad \sum_{k=0}^{\infty} t_{k}^{2} < \infty.
\]

Definition 3.5. For any \( \delta > 0 \), a point \( q \in \mathcal{X} \) is said to be a \( \delta \)-solution of (3.1) if \( f(q) \leq f^{*} + \delta \). The set of all \( \delta \)-solutions of (3.1) is denoted by \( \Omega^{*}_{\delta} \).

Theorem 3.6. Take \( \epsilon > 0 \). Suppose there is \( \bar{k} \in \mathbb{N} \) such that \( \|u_{k}\| \leq \epsilon \) for all \( k \geq \bar{k} \). Then

\[
\liminf_{k \to \infty} f(p_{k}) \leq f^{*} + \epsilon D.
\]

In particular, \( \{p_{k}\} \) has a cluster point that belongs to \( \Omega^{*}_{\epsilon D} \).
Proof. Assume towards a contradiction that \( \liminf_k f(p_k) > f^* + \epsilon D \). Then, there is \( \beta > 1 \) such that \( \liminf_k f(p_k) > f^* + \epsilon D \beta \). Hence, using the definition of \( f^* \) and \( \liminf \), we know that there are \( q' \in \mathcal{X} \) and \( k' \geq k \) with

\[
f(q') < -\epsilon D + [\epsilon D (1 - \beta) + \liminf_{k \geq k'} f(p_k)] \leq -\epsilon D + \inf_{k \geq k'} f(p_k).
\]

For better readability of the rest of the proof, we define

\[
a := -f(q') - \epsilon D + \inf_{k \geq k'} f(p_k) > 0.
\]

Since \( \{p_k\} \) is bounded, by Proposition 3.1 the sequence \( \{s_k\} \) is also bounded. Consider a constant \( \iota > 0 \) such that \( \|s_k\| \leq \iota \) for all \( k \geq k' \), we have

\[
\|s_k + u_k\| \leq \|s_k\| + \|u_k\| \leq \iota + \epsilon, \quad \forall k \geq k'.
\]

On the other hand, Lemma 3.4 ensures that

\[
d^2(p_{k+1}, q') \leq d^2(p_k, q') + \zeta t^2_k + 2t_k \frac{\sinh(\hat{\kappa} t_k)}{\hat{\kappa} t_k} \cdot f(q') - f(p_k) + \epsilon D \|u_k + s_k\|, \quad k \in \mathbb{N}_0.
\]

Using (3.6), (3.7) and the fact that the function \( (0, \infty) \ni t \mapsto \sinh(t)/t \) is bounded below by 1, the last inequality implies

\[
d^2(p_{k+1}, q') \leq d^2(p_k, q') + \zeta t^2_k - t_k \frac{2a}{\iota + \epsilon}, \quad k \geq k'.
\]

Consider \( \ell \in \mathbb{N} \). Thus, from the last inequality, summing over \( k \), we obtain

\[
\frac{2a}{\iota + \epsilon} \sum_{k=k'}^{k'+\ell} t_k \leq d^2(p_{k'}, q') - d^2(p_{k'+\ell+1}, q') + \zeta \sum_{k=k'}^{k'+\ell} t^2_k \leq d^2(p_{k'}, q') + \zeta \sum_{k=k'}^{k'+\ell} t^2_k.
\]

Since the last inequality holds for all \( \ell \in \mathbb{N} \), by using the inequality in (3.4), we arrive at a contradiction. Therefore, (3.5) holds.

Note that inequality (3.5) implies that \( \{f(p_k)\} \) has a monotonically decreasing subsequence \( \{f(p_{k_j})\} \) such that \( \lim_{j \to \infty} f(p_{k_j}) \leq f^* + \epsilon D \). Since \( \{p_{k_j}\} \) is bounded, there are \( p_\ast \in \mathcal{X} \) and a subsequence \( \{p_{k_j}\} \subseteq \{p_{k_j}\} \) such that \( \lim_{j \to \infty} p_{k_j} = p_\ast \), which by the lower semicontinuity of \( f \) implies \( f(p_\ast) \leq \liminf_{j \to \infty} f(p_{k_j}) \leq f^* + \epsilon D \), and then \( p_\ast \in \Omega^*_{\epsilon D} \).

The next result allows us to obtain a number of iterations \( N \) that guarantees the existence of a point \( p_k \) sufficiently close to an \( \epsilon D \)-solution of (3.1).

**Theorem 3.7.** Take \( \epsilon > 0 \), \( \iota > 0 \), \( D > 0 \) and \( N \in \mathbb{N} \). Assume that \( \text{diam} \mathcal{X} < D \). If \( \|u_k\| \leq \epsilon \) and \( \|s_k\| \leq \iota \) for all \( k \in \{0, 1, \ldots, N\} \), then

\[
\min \{f(p_k) - (f^* + \epsilon D) : k = 0, 1, \ldots, N\} \leq (\epsilon + \iota) \frac{D^2 + \zeta \sum_{k=0}^N t^2_k}{2 \sum_{k=0}^N t_k}.
\]

Proof. If \( p_k \in \Omega^*_{\epsilon D} \) for some \( k \in \{0, 1, \ldots, N\} \), the definition of \( \Omega^*_{\epsilon D} \) implies that \( f(p_k) - (f^* + \epsilon D) \leq 0 \), and (3.8) is trivially satisfied. Therefore, we only need to consider the case that

\[
p_k \notin \Omega^*_{\epsilon D}, \quad \forall k \in \{0, 1, \ldots, N\}.
\]
Pick $q \in \Omega^*$. Using that $\|u_k\| \leq \epsilon$ for all $k \in \{0, 1, \ldots, N\}$ and applying Lemma 3.4, we obtain
\[
d^2(p_{k+1}, q) \leq d^2(p_k, q) + \zeta^2_k + 2t_k \frac{\sinh(\kappa t_k)}{\kappa t_k} \cdot (f^* + \epsilon D) - f(p_k)
\]
for $k = 0, 1, \ldots, N$. By (3.9), $(f^* + \epsilon D) - f(p_k) < 0$ holds for any $k \in \{0, 1, \ldots, N\}$. Hence, using that the function $(0, \infty) \ni t \mapsto \sinh(t)/t$ is bounded below by 1, we obtain
\[
2t_k \frac{f(p_k) - (f^* + \epsilon D)}{\|u_k + s_k\|} \leq d^2(p_k, q) - d^2(p_{k+1}, q) + \zeta^2_k, \quad k = 0, 1, \ldots, N.
\]
Summing over $k = 0, 1, \ldots, N$, we obtain
\[
2 \sum_{k=0}^{N} t_k \frac{f(p_k) - (f^* + \epsilon D)}{\|u_k + s_k\|} \leq d^2(p_0, q) - d^2(p_{N+1}, q) + \zeta \sum_{k=0}^{N} t_k^2.
\]
Since $\|u_k + s_k\| \leq \|u_k\| + \|s_k\| < \epsilon + \epsilon$ holds for any $k \in \{0, 1, \ldots, N\}$, the last inequality implies
\[
\frac{2}{\epsilon + \epsilon} \min \{ f(p_k) - (f^* + \epsilon D) : k = 0, 1, \ldots, N \} \sum_{k=0}^{N} t_k \leq D^2 + \zeta \sum_{k=0}^{N} t_k^2,
\]
which is equivalent to the desired inequality.

**3.3.2. Constant step size.** Here, the sequence $\{p_k\}$ is generated by Algorithm 3.1 with constant step size, i.e., $t_k$ satisfies
\[
(3.10) \quad \bar{t} = \text{positive constant}, \quad t_k = \bar{t}, \quad \forall k.
\]
Since Lemma 3.4 holds for $t_k$ given in (3.10), the proof of the next result is very similar to the proof of Theorem 3.7, so it will be omitted. In [39, Thm. 9], the authors propose a version of this theorem using the exact subgradient.

**Theorem 3.8.** Take $\epsilon > 0$, $\iota > 0$ and $N \in \mathbb{N}$. If $\|u_k\| \leq \epsilon$ and $\|s_k\| \leq \iota$ for all $k \in \{0, 1, \ldots, N - 1\}$, then
\[
(3.11) \quad \min \{ f(p_k) - (f^* + \epsilon D) : k = 0, 1, \ldots, N - 1 \} \leq (\epsilon + \iota) \frac{D^2 + \zeta N^2}{2N \bar{t}}.
\]
In particular, if $\bar{t} = D/(\sqrt{\zeta N})$, we have
\[
(3.12) \quad \min \{ f(p_k) - (f^* + \epsilon D) : k = 0, 1, \ldots, N - 1 \} \leq (\epsilon + \iota) \frac{D\sqrt{\zeta}}{\sqrt{N}}.
\]

**3.3.3. Polyak step size.** The so-called Polyak step size was introduced in [30] and has been used in the Riemannian context in [3, 4, 36, 14], for instance.

Take $\epsilon > 0$. Define the inexact version of the Polyak step size by
\[
(3.13) \quad t_k = \alpha \frac{f(p_k) - f^* - \epsilon D}{\|v_k\|}, \quad 0 < \alpha < 2 \frac{\tanh(\hat{\kappa} D)}{\hat{\kappa} D}.
\]
THEOREM 3.9. Take \( \epsilon > 0 \) and \( N \in \mathbb{N} \). If \( \|u_k\| \leq \epsilon, \|s_k\| \leq \epsilon \) and \( t_k \geq 0 \) for all \( k \in \{0, 1, \ldots, N - 1\} \), then
\[
\min \{ f(p_k) - (f^* + \epsilon D): k = 0, 1, \ldots, N - 1 \} \leq \frac{1}{\sqrt{\Gamma N}}.
\]
where \( \Gamma \) is the constant defined by
\[
\Gamma = \frac{1}{(\epsilon + \epsilon)^2 D^2} \left( 2\alpha - \frac{\hat{k} D}{\tanh(\hat{k} D)} \alpha^2 \right).
\]

Proof. Pick \( q \in \Omega^* \) and \( k \in \{0, 1, \ldots, N - 1\} \). Applying Lemma 3.2 with \( p = p_k \), \( v = v_k := s_k + u_k \) and \( t = t_k \), we obtain
\[
d^2(p_{k+1}, q) \leq d^2(p_k, q)
\]
\[
+ \frac{\sinh(\hat{k} t_k)}{\hat{k} t_k} \left( \frac{\hat{k} d(p_k, q)}{\tanh(\hat{k} d(p_k, q))} t_k^2 + 2t_k \frac{(s_k, \log p_k q) + (u_k, \log p_k q)}{\|v_k\|} \right),
\]
where \( p_{k+1} \) is defined in (3.2). Hence, using (2.2), (2.3) and the Cauchy–Schwarz inequality, it follows that
\[
d^2(p_{k+1}, q) \leq d^2(p_k, q)
\]
\[
+ \frac{\sinh(\hat{k} t_k)}{\hat{k} t_k} \left( \frac{\hat{k} d(p_k, q)}{\tanh(\hat{k} d(p_k, q))} t_k^2 - 2t_k \frac{f(p_k) - f^* - \|u_k\| d(p_k, q)}{\|v_k\|} \right).
\]
The function \((0, \infty) \ni t \mapsto t/\tanh(t)\) is positive and increasing, and \( d(p_k, q) \leq D \) holds. Thus, the last inequality above implies
\[
d^2(p_{k+1}, q) \leq d^2(p_k, q) + \frac{\sinh(\hat{k} t_k)}{\hat{k} t_k} \left( \frac{\hat{k} D}{\tanh(\hat{k} D)} t_k^2 - 2t_k \frac{f(p_k) - f^* - \epsilon D}{\|v_k\|} \right).
\]
Since \( \sinh(\hat{k} t_k)/(\hat{k} t_k) \geq 1 \) and \( \|v_k\| \leq \epsilon + \epsilon \), after rearranging the terms of the last inequality and using (3.13) and (3.15), we obtain
\[
\Gamma D^2 \frac{(f(p_k) - f^* - \epsilon D)^2}{\|v_k\|^2} \leq \Gamma D^2 (\epsilon + \epsilon)^2 \left( \frac{f(p_k) - f^* - \epsilon D}{\|v_k\|} \right)^2
\]
\[
\leq d^2(p_k, q) - d^2(p_{k+1}, q).
\]
Summing over \( k = 0, 1, \ldots, N - 1 \), we arrive at
\[
\Gamma D^2 \sum_{k=0}^{N-1} (f(p_k) - f^* - \epsilon D)^2 \leq d^2(p_0, q) - d^2(p_N, q) \leq D^2.
\]

Therefore,
\[
\min \{(f(p_k) - f^* - \epsilon D)^2: k = 0, 1, \ldots, N - 1\} \leq \frac{1}{\Gamma N}.
\]
which is equivalent to the desired inequality.

4. Some auxiliary results. In this section, we provide some auxiliary results needed for the application of the projected inexact subgradient method to the singular value optimization problems described in Subsection 2.2.
4.1. Projections to order intervals. For any $p_1, p_2 \in S_n^+$ with $p_1 \leq p_2$, we define the order interval

$$[p_1, p_2] := \{ p \in S_n^+ : p_1 \leq p \leq p_2 \}.$$  

**Lemma 4.1.** Every order interval $[p_1, p_2]$ is a compact and convex subset of $S_n^+$.

**Proof.** If $p, q \in [p_1, p_2]$, then by [26, Prop. 4.1] we have $p_1 = p_1 \# p_1 \leq p \# q \leq p_2 \# p_2 = p_2$, which shows that $[p_1, p_2]$ is convex. It is easy to see that $[p_1, p_2]$ is closed. To prove compactness, it hence suffices to show boundedness (since $S_n^+$ is a complete Riemannian manifold). To this end, take an arbitrary $p \in [p_1, p_2]$. By [6, Cor. I.2.1.1], the relation $p_1 \leq p \leq p_2$ implies $\lambda_i(p) \in [\lambda_i(p_1), \lambda_i(p_2)]$ for $i = 1, \ldots, n$.

It follows that

$$d(p, I)^2 = \sum_{i=1}^{n} \ln^2 \lambda_i(p) = \sum_{\lambda_i(p) < 1} \ln^2 \lambda_i(p) + \sum_{\lambda_i(p) > 1} \ln^2 \lambda_i(p)$$

$$\leq \sum_{\lambda_i(p) < 1} \ln^2 \lambda_i(p_1) + \sum_{\lambda_i(p) > 1} \ln^2 \lambda_i(p_2) \leq d(p_1, I)^2 + d(p_2, I)^2. \quad \square$$

Here, we use that $\ln^2$ is decreasing on $(0, 1]$ and increasing on $[1, \infty)$.

From now on, we only consider order intervals of the form $[\alpha I, \beta I]$ for $0 < \alpha < \beta$.

**Lemma 4.2.** The projection onto the set $[\alpha I, \beta I]$, denoted by $P_{\alpha, \beta}$, is given by

$$P_{\alpha, \beta}(p) = U_p^\top \text{diag}(\zeta_1(p), \ldots, \zeta_n(p)) U_p,$$

where $U_p$ is an orthogonal matrix chosen such that $U_p p U_p^\top$ is diagonal and

$$\zeta_i(p) := \begin{cases} \alpha & \text{if } \lambda_i(p) < \alpha, \\ \lambda_i(p) & \text{if } \lambda_i(p) \in [\alpha, \beta], \\ \beta & \text{if } \lambda_i(p) > \beta. \end{cases}$$

**Proof.** Let $p \in S_n^+$ be arbitrary. To compute $P_{\alpha, \beta}(p) \in [\alpha I, \beta I]$, we have to solve the minimization problem

$$\min_{\alpha I \leq q \leq \beta I} d(p, q)^2.$$  

This problem can equivalently be written as

$$\min_{\lambda_i(q) \in [\alpha, \beta]} \sum_{i=1}^{n} \ln^2 \lambda_i(p^{-1} q).$$

We write $p = U^\top \text{diag}(\gamma_1, \ldots, \gamma_n) U$ for some orthogonal $U$ with $\gamma_1 \geq \cdots \geq \gamma_n$. Now, we will prove that the matrix $q^*$, defined as follows, is the unique minimizer:

$$q^* := U^\top \text{diag}(\zeta_1, \ldots, \zeta_n) U, \quad \zeta_i := \begin{cases} \alpha & \text{if } \gamma_i < \alpha, \\ \gamma_i & \text{if } \gamma_i \in [\alpha, \beta], \\ \beta & \text{if } \gamma_i > \beta. \end{cases}$$

By construction, $q^*$ is an element of $[\alpha I, \beta I]$. Its distance to $p$ satisfies

$$d(p, q^*)^2 = \sum_{i=1}^{n} \ln^2 \lambda_i(p^{-1} q^*) = \sum_{\gamma_i < \alpha} \ln^2 \frac{\alpha}{\gamma_i} + \sum_{\gamma_i > \beta} \ln^2 \frac{\beta}{\gamma_i}.$$
To show that $q^*$ is the (unique) minimizer, it suffices to prove that $d(p, q) \geq d(p, q^*)$ for every $q \in [\alpha I, \beta I]$. To this end, pick $q \in [\alpha I, \beta I]$ arbitrary and define $\tilde{q}_1 := q - \alpha I \geq 0$, $\tilde{q}_2 := \beta I - q \geq 0$. Then

$$d(p, q)^2 = \sum_{i=1}^n \ln^2 \lambda_i(p^{-1}q)$$

$$\geq \sum_{\lambda_i(p^{-1}) > \alpha^{-1}} \ln^2 \lambda_i(p^{-1}(\tilde{q}_1 + \alpha I)) + \sum_{\lambda_i(p^{-1}) < \beta^{-1}} \ln^2 \lambda_i(p^{-1}(\beta I - \tilde{q}_2))$$

$$= \sum_{\lambda_i(p^{-1}) > \alpha^{-1}} \ln^2 \lambda_i(\alpha p^{-1} + p^{-\frac{1}{2}}\tilde{q}_1 p^{-\frac{1}{2}})$$

$$+ \sum_{\lambda_i(p^{-1}) < \beta^{-1}} \ln^2 \lambda_i(\beta p^{-1} - p^{-\frac{1}{2}}\tilde{q}_2 p^{-\frac{1}{2}})$$

$$\geq \sum_{\lambda_i(p^{-1}) > \alpha^{-1}} \ln^2 \{\alpha \lambda_i(p^{-1})\} + \sum_{\lambda_i(p^{-1}) < \beta^{-1}} \ln^2 \{\beta \lambda_i(p^{-1})\}$$

$$= \sum_{\lambda_{n-i}(p) < \alpha} \frac{\alpha}{\lambda_{n-i}(p)} + \sum_{\lambda_{n-i}(p) > \beta} \frac{\beta}{\lambda_{n-i}(p)} = d(p, q^*)^2.$$

In the last inequality, we used that $p^{-\frac{1}{2}}\tilde{q}_1 p^{-\frac{1}{2}} \geq 0$, $p^{-\frac{1}{2}}\tilde{q}_2 p^{-\frac{1}{2}} \geq 0$ in combination with [6, Cor. 1.2.1.1] and the monotonicity properties of $\ln^2$.

**Lemma 4.3.** The diameter of the set $[\alpha I, \beta I]$ is given by

$$\text{diam } [\alpha I, \beta I] = \sqrt{n} \ln \frac{\beta}{\alpha}.$$

**Proof.** Let $p, q \in [\alpha I, \beta I]$ be chosen arbitrarily. Then

$$d(p, q)^2 = \sum_{i=1}^n \ln^2 \lambda_i(p^{-1}q)$$

$$= \sum_{\lambda_i(p^{-1}) < 1} \ln^2 \lambda_i(p^{-1}q) + \sum_{\lambda_i(p^{-1}) > 1} \ln^2 \lambda_i(p^{-1}q)$$

$$\leq \sum_{\lambda_i(p^{-1}) < 1} \ln^2 \lambda_1(p^{-1}q) + \sum_{\lambda_i(p^{-1}) > 1} \ln^2 \lambda_1(p^{-1}q)$$

$$= \sum_{\lambda_i(p^{-1}) < 1} \ln^2 \lambda_1(q^{-1}p) + \sum_{\lambda_i(p^{-1}) > 1} \ln^2 \lambda_1(p^{-1}q).$$

Using that $\lambda_1(AB) \leq \lambda_1(A) \lambda_1(B)$ for positive definite $A, B$ ([38, Lem. 4.2]), we get

$$d(p, q)^2 \leq \sum_{\lambda_i(p^{-1}) < 1} \ln^2 \{\lambda_1(q^{-1})\lambda_1(p)\} + \sum_{\lambda_i(p^{-1}) > 1} \ln^2 \{\lambda_1(p^{-1})\lambda_1(q)\}$$

$$\leq \sum_{\lambda_i(p^{-1}) < 1} \ln^2 \frac{\beta}{\alpha} + \sum_{\lambda_i(p^{-1}) > 1} \ln^2 \frac{\beta}{\alpha} \leq n \ln^2 \frac{\beta}{\alpha}. \quad \square$$

Since $d(\alpha I, \beta I) = \sqrt{n} \ln \frac{\beta}{\alpha}$, the proof is complete.
4.2. Lipschitz constant. In this subsection, we derive a Lipschitz constant of the functions $J_{k+s,x}^2$.

Lemma 4.4. For each $x \in K$, $k \in \{1, \ldots, n-1\}$ and $s \in [0,1)$, the function $J_{k+s,x}^2 : S_n^+ \to \mathbb{R}$, introduced in (2.8), is globally Lipschitz continuous with Lipschitz constant $L = \sqrt{n}/\ln(2)$.

Proof. Since $J_{k+s,x}^2 = (1-s)J_{k,x}^2 + sJ_{k+1,x}^2$, it suffices to prove the statement for $s = 0$. Using the identities

$\lambda_i(p^{-\frac{1}{2}}q) = \lambda_i((p^{-\frac{1}{2}}q^\frac{1}{2})(p^{-\frac{1}{2}}q^\frac{1}{2})^\top) = \alpha_i(p^{-\frac{1}{2}}q^\frac{1}{2})^2$

and formula (2.4), we can rewrite $d(p,q)$ as

$d(p,q) = 2\left(\sum_{i=1}^n \|\alpha_i(p^{-\frac{1}{2}}q^\frac{1}{2})\|^2\right)^{\frac{1}{2}}$.

Let $\omega_0(g) := 1$ and $\omega_k(g) := \alpha_1(g) \cdots \alpha_k(g)$ for any $g \in \text{GL}(n, \mathbb{R})$ and $k = 1, \ldots, n$.

Then

$J_{k,x}^2(p) - J_{k,x}^2(q) = \sum_{i=1}^k \log_2 \alpha_i(p^{\frac{1}{2}}A(x)p^{-\frac{1}{2}}) - \sum_{i=1}^k \log_2 \alpha_i(q^{\frac{1}{2}}A(x)q^{-\frac{1}{2}})$

$= \log_2 \omega_k(p^{\frac{1}{2}}A(x)p^{-\frac{1}{2}}) - \log_2 \omega_k(q^{\frac{1}{2}}A(x)q^{-\frac{1}{2}})$

$= \log_2 \omega_k(p^{\frac{1}{2}}q^{-\frac{1}{2}}A(x)q^{\frac{1}{2}}) - \log_2 \omega_k(q^{\frac{1}{2}}A(x)q^{-\frac{1}{2}})$.

Using Horn’s inequality $\omega_k(AB) \leq \omega_k(A)\omega_k(B)$ (see [6, Prop. I.2.3.1]), we thus obtain

$J_{k,x}^2(p) - J_{k,x}^2(q) \leq \frac{1}{\ln(2)} \ln[\omega_k(p^{\frac{1}{2}}q^{-\frac{1}{2}})\omega_k(q^{\frac{1}{2}}p^{-\frac{1}{2}})]$.

Writing $\|\cdot\|_1$ and $\|\cdot\|_2$ for the 1-norm and 2-norm in $\mathbb{R}^n$, respectively, we can further estimate

$\ln \omega_k(p^{\frac{1}{2}}q^{-\frac{1}{2}}) = \sum_{i=1}^k \ln \alpha_i(p^{\frac{1}{2}}q^{-\frac{1}{2}}) \leq \sum_{i=1}^k |\ln \alpha_i(p^{\frac{1}{2}}q^{-\frac{1}{2}})| \leq \ln(2)\|\tilde{\alpha}(p^{\frac{1}{2}}q^{-\frac{1}{2}})\|_1$

$\leq \sqrt{n} \ln(2)\|\tilde{\alpha}(p^{\frac{1}{2}}q^{-\frac{1}{2}})\|_2 = \sqrt{n} \ln(2)\frac{1}{2}d(p^{-1}, q^{-1}) = \frac{\sqrt{n}}{2}d(p,q)$.

The analogous estimates hold for $\omega_k(q^{\frac{1}{2}}p^{-\frac{1}{2}})$, and hence

$|J_{k,x}^2(p) - J_{k,x}^2(q)| \leq \frac{\sqrt{n}}{\ln(2)}d(p,q)$.

The proof is complete.

4.3. Error analysis. Assume that $s_2(x)$ is the subgradient of $J_{k,x}^2$ at $p \in S_n^+$ for some $k \in \{1, \ldots, n-1\}$ and $x \in K$ as computed via the procedure explained at the end of Subsection 2.2. Our aim is to understand how far $s_2(x)$ can be from a real subgradient in the case that $x$ is not a maximizer of $x \mapsto J_{k,x}^2(p)$. Recall that $\zeta_x(p) = p^{\frac{1}{2}}A(x)p^{-\frac{1}{2}}$.

The following theorem helps to compute the error in case that $x$ is close to a maximizer.
Theorem 4.5. Let \( x, y \in K \) and assume that

\[
\delta := \alpha_k(\zeta_y^p) - \alpha_k+1(\zeta_x(p)) > 0.
\]

Then

\[
\|s_2(x) - s_2(y)\|_p \leq \frac{\sqrt{2n(n+1)}}{\ln 2} \frac{1}{\delta} \|p^\frac{1}{2}\|_F \|p^{-\frac{1}{2}}\|_F \|A(x) - A(y)\|_F.
\]

Proof. By linearity of the trace function and orthonormality of \( \{e_i\} \), we have

\[
\|s_2(x) - s_2(y)\|_p^2 = \left\| \sum_i \text{tr}[S_x^+ \zeta_x(p)e_i - S_y^+ \zeta_y(p)e_i]e_i \right\|_p^2.
\]

Using again the linearity of the trace function and the fact that \( \text{tr}[AB] = \text{tr}[BA] \) for any \( A, B \), we find that

\[
\text{tr}[S_x^+ \zeta_x(p)e_i] = \text{tr}[S_x^+(X_iA(x)p^{-\frac{1}{2}} - p^{\frac{1}{2}}A(x)p^{-\frac{1}{2}}X_i)p^{-\frac{1}{2}}]
\]

\[
= \text{tr}[X_iA(x)p^{-\frac{1}{2}}S_x^+ - X_iS_x^+p^{\frac{1}{2}}A(x)p^{-\frac{1}{2}}]
\]

\[
= \text{tr}[X_iS_x^+(\zeta_x(p)S_x^+ - \zeta_x^p(p))].
\]

Now write \( \alpha_i := \alpha_i(\zeta_x(p)) \) and observe that

\[
\zeta_x^p(p)S_x^+ = \frac{1}{\ln 2} U^\top \text{diag}(\alpha_1, \ldots, \alpha_n)V V^\top \text{diag}(\alpha_1^{-1}, \ldots, \alpha_k^{-1}, 0, \ldots, 0)U
\]

\[
= \frac{1}{\ln 2} U^\top (I_{k \times k} \oplus 0_{(n-k) \times (n-k)})U,
\]

and analogously

\[
S_x^+ \zeta_x(p) = \frac{1}{\ln 2} V^\top (I_{k \times k} \oplus 0_{(n-k) \times (n-k)})V.
\]

We observe that \( U^\top (I_{k \times k} \oplus 0_{(n-k) \times (n-k)})U \) is the orthogonal projection onto the \( k \)-dimensional subspace of \( \mathbb{R}^n \) spanned by the first \( k \) left singular vectors of \( \zeta_x(p) \), that we denote by \( u_1(x) := U_x^e_1, \ldots, u_k(x) := U_x^e_k \). Introducing the notation

\( \tilde{U}_x := [u_1(x) \cdots u_k(x)] \in \mathbb{R}^{n \times k} \), we further observe that

\[
U^\top (I_{k \times k} \oplus 0_{(n-k) \times (n-k)})U = \tilde{U}_x \tilde{U}_x^\top.
\]

Altogether, we have obtained the identity

\[
\|s_2(x) - s_2(y)\|_p^2 = \frac{1}{(\ln 2)^2} \sum_i \text{tr}^2[X_i\gamma^{\frac{1}{2}}(\tilde{U}_x \tilde{U}_x^\top - \tilde{V}_x \tilde{V}_x^\top - \tilde{V}_y \tilde{V}_y^\top)].
\]

To estimate the trace, we use the Cauchy–Schwarz inequality in \( \mathbb{R}^{n \times n} \) which yields

\[
\|s_2(x) - s_2(y)\|_p^2 \leq \frac{1}{(\ln 2)^2} \sum_i \|X_i\gamma^{\frac{1}{2}}\|_F^2 \|\tilde{U}_x \tilde{U}_x^\top - \tilde{V}_x \tilde{V}_x^\top - \tilde{V}_y \tilde{V}_y^\top\|_F^2.
\]
To estimate the term $\|X_ip^{-\frac{1}{2}}\|_F$, we use that

$$1 = \|e_i\|^2 = \text{tr}[p^{-1}e_ie_i] = \text{tr}[p^{-1}(p^{\frac{1}{2}}X_ip^{\frac{1}{2}})p^{-1}(p^{\frac{1}{2}}X_ip^{\frac{1}{2}})]$$

$$= \text{tr}[(p^{\frac{1}{2}}X_ip^{\frac{1}{2}})(p^{-1}X_ip^{-1}(p^{\frac{1}{2}}X_ip^{\frac{1}{2}}))]$$

$$= \text{tr}[p^{-\frac{1}{2}}X_ip^{\frac{1}{2}}X_i] + \text{tr}[p^{-\frac{1}{2}}X_ip^{-1}X_ip^{\frac{1}{2}}] + \text{tr}[p^{-1}X_i^2]$$

$$+ \text{tr}[p^{-1}X_ip^{\frac{1}{2}}X_ip^{\frac{1}{2}}]$$

$$= 2\text{tr}[p^{-\frac{1}{2}}X_ip^{\frac{1}{2}}X_i] + 2\text{tr}[X_ip^{-1}X_i]$$

$$= 2\|X_i\|^2 + 2\text{tr}[(X_ip^{-\frac{1}{2}})^\top(X_ip^{-\frac{1}{2}})]$$

which implies

$$\|X_ip^{-\frac{1}{2}}\|^2_F = \text{tr}[(X_ip^{-\frac{1}{2}})^\top(X_ip^{-\frac{1}{2}})] = \frac{1}{2} - \|X_i\|^2_{p^{\frac{1}{2}}} \leq \frac{1}{2}.$$  

Hence, we obtain

$$\|s_2(x) - s_2(y)\|_p \leq \frac{n(n+1)}{4(\ln 2)^2} \|\bar{U}_x \bar{V}_x^\top - \bar{U}_y \bar{V}_y^\top + \bar{V}_x \bar{V}_y^\top\|^2_F$$

$$\leq \frac{n(n+1)}{4(\ln 2)^2} \|\bar{U}_x \bar{V}_x^\top - \bar{U}_y \bar{V}_y^\top\|^2_F + \|\bar{V}_x \bar{V}_y^\top - \bar{V}_y \bar{V}_y^\top\|^2_F.$$  

Next, we use that $\|\bar{U}_x \bar{V}_x^\top - \bar{U}_y \bar{V}_y^\top\|_F$ is the distance between the two subspaces $L_x^U := \langle u_1(x), \ldots, u_k(x) \rangle$ and $L_y^U := \langle u_1(y), \ldots, u_k(y) \rangle$, regarded as elements of the Grassmannian of $k$-planes in $\mathbb{R}^n$. We further obtain

$$\|s_2(x) - s_2(y)\|_p \leq \frac{\sqrt{n(n+1)}}{2\ln 2}\sqrt{\|\bar{U}_x \bar{V}_x^\top - \bar{U}_y \bar{V}_y^\top\|^2_F + \|\bar{V}_x \bar{V}_y^\top - \bar{V}_y \bar{V}_y^\top\|^2_F}$$

We have

$$\|\bar{U}_x \bar{V}_x^\top - \bar{U}_y \bar{V}_y^\top\|_F = \sqrt{2}\|\Theta(L_x^U, L_y^U)\|_F,$$

where $\Theta$ is the matrix of canonical angles between $L_x$ and $L_y$ (see [34, p. 9]). This leads to

$$\|s_2(x) - s_2(y)\|_p \leq \frac{n(n+1)}{\ln 2}\sqrt{\|\Theta(L_x^U, L_y^U)\|^2_F + \|\Theta(L_y^V, L_y^V)\|^2_F}.$$  

Then Wedin’s theorem (see [9] for a timely reference) states that the inequality $\delta > 0$ with $\delta$ as in (4.1) implies

$$\sqrt{\|\Theta(L_x^U, L_y^U)\|^2_F + \|\Theta(L_y^V, L_y^V)\|^2_F} \leq \frac{1}{\delta}\sqrt{\|R\|^2_F + \|S\|^2_F},$$

where $R$ and $S$ are matrices built from the singular value decompositions of $\zeta_x(p)$ and $\zeta_y(p)$ that satisfy $\max\{\|R\|_F, \|S\|_F\} \leq \|\zeta_x(p) - \zeta_y(p)\|_F$. Altogether,

$$\|s_2(x) - s_2(y)\|_p \leq \frac{\sqrt{2n(n+1)}}{\ln 2}\frac{1}{\delta}\|\zeta_x(p) - \zeta_y(p)\|_F.$$  

This manuscript is for review purposes only.
Finally, we can estimate
\[
\|\zeta_{x}(p) - \zeta_{y}(p)\|_{F} = \|p^{\frac{1}{2}}(A(x) - A(y))p^{-\frac{1}{2}}\|_{F} \leq \|p^{\frac{1}{2}}\|_{F}\|p^{-\frac{1}{2}}\|_{F}\|A(x) - A(y)\|_{F}.
\]

The proof is complete. 

\[\square\]

\begin{remark}
The assumption \(\alpha_{k}(\zeta_{y}(p)) - \alpha_{k+1}(\zeta_{x}(p)) > 0\) will be satisfied if
\[
\alpha_{k}(\zeta_{x}(p)) - \alpha_{k+1}((\zeta_{x}(p)) > \|\zeta_{x}(p) - \zeta_{y}(p)\|,
\]
because the singular values are 1-Lipschitz with respect to the matrix, implying
\[
\alpha_{k}(\zeta_{y}(p)) - \alpha_{k+1}(\zeta_{x}(p)) \geq -\|\zeta_{x}(p) - \zeta_{y}(p)\| + \alpha_{k}(\zeta_{x}(p)) - \alpha_{k+1}(\zeta_{x}(p)).
\]
\end{remark}

\section{Examples}
In this section we will apply the theory to the Hénon system with standard parameters \(a = 1.4\) and \(b = 0.3\), which is given by
\[
x(t + 1) = 1.4 - x(t)^{2} + 0.3y(t),
g(t + 1) = x(t).
\]

The system has the two equilibria \((x_{+}, x_{+})\) and \((x_{-}, x_{-})\), where \(x_{\pm} = \frac{1}{2}(b - 1 \pm \sqrt{(b - 1)^{2} + 4a})\). The quadrilateral \(K\) with the following corners is a trapping region \[\cite{18}:\]

\[
A = (-1.862, 1.96), \quad B = (1.848, 0.6267)
\]

\[
C = (1.743, -0.6533), \quad D = (-1.484, -2.3333)
\]

In particular, \(K\) is a compact forward-invariant set.

We adapted the software from \[\cite{19}\] for our computations; see also \[\cite{20}\]. The maximization of \(x \mapsto J_{k+s,x}(a, p)\) is performed on a 1000 \(\times\) 1000 grid and is then refined on a grid of the same size around the maximizer found in the first step. In all computations reported we used a fourth-degree polynomial \(r_{a}\); we also performed computations with sixth- and eight-degree polynomials with very similar results. W.l.o.g. the constant term of the polynomial \(r_{a}\) is always set equal to zero. The exogenous step size used was \(t_{k} = 16/k\) as in \[\cite{20}\] and we always started with the metric \(P = I\), i.e. \(p = I\) and \(r_{a} = 0\). Let us first discuss the estimation of the dimension of the attractor.

\subsection{Dimension of the attractor}
We used exogenous step sizes and considered polynomials \(r_{a}\) of degree 4. With \(k = 1\) we experimented with different \(s \in [0, 1]\), the results are summarized in Table 1. A few comments are in order. For \(s = 0.430\) we did not obtain a negative value for \(\max_{x \in K} \Sigma_{k+s,x}(P)\) in 100,000 iterations. For \(s \geq 0.435\) we obtained a negative value for \(\max_{x \in K} \Sigma_{k+s,x}(P)\) and therefore the upper bound \(\dim_{L} K \leq 1.435\) on the Lyapunov dimension. For larger \(s\) a negative value is obtained in fewer iterations and for \(s \geq 0.450\) no iterations are needed, i.e. the metric \(P = I\) is sufficient to prove the upper bound. The formula for the computed metric that gives the best upper bound \(\dim_{L} K \leq 1.435\) is shown in Table 2.
Upper bounds on the Lyapunov dimension of the Hénon system on $K$

| $\dim L K$ | $s$ | First neg. itr. | $\max_{x \in K} \sum_{k+s,x} (P)$ |
|-----------|-----|----------------|----------------------------------|
| 1.450     | 0.450 | 0              | $-0.02278695484779664$           |
| 1.445     | 0.445 | 641            | $-0.005147904573907709$          |
| 1.440     | 0.440 | 1,061          | $-0.000157423386151721$          |
| 1.435     | 0.435 | 30,633         | $-0.0001650558073724702$         |
| *         | 0.430 | *              | $> 0$                            |

Table 1

Results for the Lyapunov dimension estimate for various $s \in (0, 1)$. In all cases $k = 1$ so $\dim L K \leq 1 + s$. In the third column we write the first iteration where $\max_{x \in K} \sum_{k+s,x} (P)$ is negative and in the fourth column we give the value. For $s = 0.450$ no iterations are needed to obtain a negative value and for $s = 0.430$ a negative value was not obtained in 100,000 iterations.

Metric $P = e^{r_a(x,y)} p$ for the Hénon system on $K$ that gives the lowest upper bound 1.435 on the Lyapunov dimension.

$$p = \begin{pmatrix}
1.826992505777629 & 0.0001682542238040804 \\
0.0001682542238040805 & 0.5473476356577599
\end{pmatrix}$$

Table 2

For this example, the Lyapunov dimension of a bounded invariant set $K^*$ containing the equilibria $(x_-, x_-)$ and $(x_+, x_+)$ is $\dim L K^* = 1.4953$, see [25, Thm. 3] or [22, Sect. 6.5.1]. Note that our set $K$ does contain $(x_+, x_+)$, but not $(x_-, x_-)$. To apply our method to this case, we enlarged the set $K$ by changing the point $D$ to $(-2, -2.333)$. We then obtained the upper bound $\dim L K \leq 1.5 = 1 + s$ with $s = 0.5$; we did not obtain a negative value for $\max_{x \in K} \sum_{k+s,x} (P)$ in 100,000 iterations with $s = 0.49$. Hence, our method found the upper bound 1.5 on the actual value 1.4953.

Other results in the literature include the numerical value of $1.220 \pm 0.019$ for the fractal dimension of the attractor of the Hénon system, see [28]. Note that the Hausdorff dimension is bounded above by the fractal dimension, which in turn is bounded above by the Lyapunov dimension.

Back to the original set $K$ containing only $(x_+, x_+)$, we note that the Lyapunov dimension of the equilibrium $(x_+, x_+)$ is 1.3521, which is thus a lower bound on the
Lyapunov dimension of $K$. An upper bound computed with our method is $1.435$.

### 5.2. Restoration entropy

In the first computation of the restoration entropy we used exogenous step sizes $t_k = 16/k$ and no projection. The results for the estimation of the restoration entropy are shown in Figure 1 on a log-log graph. The initial upper bound with $P = I$ is $1.95114084926661$. The upper bound quickly increases and then settles down and converges. The best upper bound is $1.313884120199142$ obtained in iteration 99,654. The formula for the computed metric that gives the best upper bound is given in Table 3.

![Image](image.png)

**Fig. 1.** Restoration entropy for the Hénon system using exogenous step sizes $t_k = 16/k$. The best upper bound is $1.313884120199142$ obtained in iteration 99,654.

Metric $P = e^{r_a(x,y)p}$ for the Hénon system on $K$ that gives the lowest upper bound $1.313884120199142$ on the restoration entropy (exogeneous step size).

$$P = \begin{pmatrix} 1.717042057691368 & -0.0157998385287614 \\ -0.0157998385287614 & 0.5825423031576957 \end{pmatrix}$$

| term   | coefficient     | $r_a(x,y)$ | term   | coefficient     |
|--------|-----------------|------------|--------|-----------------|
| $x$    | $-0.03728625249859665$ | $xy^2$     | $x^2$  | $0.3852955420533915$     |
| $y$    | $0.7173696947291418$    | $y^3$      | $x^3$  | $-0.1800768847947239$    |
| $x^2$  | $0.3852955420533915$     | $x^4$      | $x^4$  | $0.28824355432863$      |
| $xy$   | $-0.1800768847947239$    | $x^3y$     | $y^2$  | $-0.5726642618756508$    |
| $y^2$  | $-0.5726642618756508$    | $x^2y^2$   | $x^3$  | $-0.1943808570457239$    |
| $x^3$  | $-0.1943808570457239$    | $xy^3$     | $y^4$  | $0.3333574278923304$     |
| $x^2y$ | $-0.3414860502248702$    |            |        |                  |

**Table 3**
Next we used fixed step sizes, \( t_k = 0.01 \) and \( t_k = 0.001 \), respectively. The results are depicted in Figure 2, step size \( t_k = 0.01 \) in blue and step size \( t_k = 0.001 \) in red. The overshot at the beginning is much lower than when using exogenous step sizes and lower for the smaller step size. Further, they do not converge to a value but oscillate in an interval, whose width is determined by the step size. For step size \( t_k = 0.01 \), the best upper bound on the restoration entropy is 1.304961201612717, obtained in iteration 75,754. For step size \( t_k = 0.001 \), the best upper is 1.306393103425693 obtained in iteration 99,279. Note that although the best upper bounds are somewhat better than for the exogenous step sizes, they take longer to settle to reasonable estimates. For example, after the first 1,000 iterations the best upper bound is 1.341017992604001 for the exogenous step sizes but 1.394573310929959 and 1.573822400831881 for the fixed step sizes \( t_k = 0.01 \) and \( t_k = 0.001 \) respectively. In particular, the high overshoot at the beginning when using exogenous step sizes is eliminated in a few hundred iterations. The formulas for the computed metric that give the best upper bounds are presented in Table 4 and 5 for step sizes \( t_k = 0.01 \) and \( t_k = 0.001 \), respectively.

**Fig. 2.** Restoration entropy for the Hénon system using fixed step sizes \( t_k = 0.01 \) (blue) and \( t_k = 0.001 \) (red), respectively. The best upper bound is 1.304961201612717 and 1.306393103425693, respectively. The bounds oscillate in an interval, the size of which depends on the step size.
Metric $P = e^{r_a(x,y)}p$ for the Hénon system on $K$ that gives the lowest upper bound 1.304961201612717 on the restoration entropy (fixed step size $t_k = 0.01$)

$$p = \begin{pmatrix}
1.825723762954203 & -2.675239129430998 \cdot 10^{-6} \\
-2.675239129430998 \cdot 10^{-6} & 0.547727986225381
\end{pmatrix}$$

| term | coefficient | term | coefficient |
|------|-------------|------|-------------|
| $x$  | 0.5573870474721652 | $xy^2$ | 0.0772466464251425 |
| $y$  | 0.7077871346437571  | $y^4$ | 0.06313658395058322 |
| $x^2$ | 0.2275835960549717  | $x^4$ | 0.2599915790953258 |
| $xy$ | -0.149843153473276  | $x^3y$ | 0.06823486825177701 |
| $y^2$ | 0.1021644733650378  | $x^2y^2$ | -0.03085636902009057 |
| $x^3$ | -0.1663567125415807  | $xy^3$ | 0.03118446283824575 |
| $x^2y$ | -0.3651316623899711  | $y^4$ | 0.05509114432195542 |

Table 4

Metric $P = e^{r_a(x,y)}p$ for the Hénon system on $K$ that gives the lowest upper bound 1.306393103425693 on the restoration entropy (fixed step size $t_k = 0.001$)

$$p = \begin{pmatrix}
1.598582176519417 & -0.0340879261290964 \\
-0.03408792612909639 & 0.6262812142912478
\end{pmatrix}$$

| term | coefficient | term | coefficient |
|------|-------------|------|-------------|
| $x$  | 0.4301700400628248 | $xy^2$ | 0.199142773311169 |
| $y$  | 0.5742005669502973  | $y^3$ | 0.09368300218613741 |
| $x^2$ | 0.2138324226849934  | $x^4$ | 0.2663176394736673 |
| $xy$ | -0.2027704610258738  | $x^3y$ | 0.0937467991423779 |
| $y^2$ | 0.0431790264405415  | $x^2y^2$ | 0.007144757712718732 |
| $x^3$ | -0.1495907154564866  | $xy^3$ | 0.05518248936958972 |
| $x^2y$ | -0.2981438582431617  | $y^4$ | 0.0808752884012343 |

Table 5

In Table 6 we list the eigenvalues of $p$ and the norm of the vector $a$ defining the polynomial $r_a$ in the metric that delivers the lowest upper bound on the restoration entropy for the step sizes we used.

From the table we deduce that a projection of $a$ to the ball with radius 1.5 in $\mathbb{R}^{15}$

This manuscript is for review purposes only.
Table 6

| exogenous  | norm $a$ | min eigenvalue $p$ | max eigenvalue $p$ |
|------------|----------|--------------------|--------------------|
| $t_k = 0.01$ | 1.313270369387425 | 0.582322306187126 | 1.717262054661938 |
| $t_k = 0.001$ | 1.069807454767147 | 0.547727986219938 | 1.825723762959803 |
| $t_k = 0.001$ | 0.923314506269372 | 0.625087590080113 | 1.599775800730552 |

The eigenvalues of $p$ and norm of the vector $a$ defining the polynomial $r_a$ for the optimal metrics obtained with the different step sizes.

and of $p$ to the interval $[0.5J, 2I]$ should not be limiting. We repeat the computations above using the projected subgradient algorithm. For the exogenous step sizes the results were slightly different, see Figure 3, and the bound on the restoration entropy was lower, i.e. $1.305419886141552$ obtained in iteration 99,456. The values for the best metric are displayed in Table 7.

For the fixed step sizes they were so similar that we abstain from plotting them. The difference in the bound on the restoration entropy was in the fourth decimal place. For the step size $t_k = 0.001$ we also tried projecting $p$ on $[0.61I, 1.61I]$ and $a$ to the ball with radius 0.95, i.e. very close to the values we had found for the optimal metric. The results were very similar to the results without projection with no notable improvement. Note that the results with fixed step size are what one would expect from the error analysis results for the projected subgradient algorithm with fixed step size, cf. (3.11): the estimates oscillate in an interval and do not converge.

![Fig. 3. Restoration entropy for the Hénon system using exogenous step sizes $t_k = 16/k$, with projection of $a$ to the ball with radius 1.5 in $\mathbb{R}^{15}$ and of $p$ to the interval $[0.5J, 2I]$ (red) and without projection (blue, see also Fig. 1). The best upper bound of the projected version is $1.305419886141552$ obtained in iteration 99,456, while the best result for the one without projection was $1.313884120199142$ obtained in iteration 99,654.](image-url)
Metric $P = e^{r_a(x,y)}p$ for the Hénon system on $K$ that gives the lowest upper bound 1.305419886141552 on the restoration entropy (exogeneous step size with projection to ball of radius 1.5 and $[0.5I, 2I]$).

$$P = \begin{pmatrix} 1.674120023285701 & -0.02266875912007223 \\ -0.02266875912007223 & 0.5976356884178141 \end{pmatrix}$$

| term   | coefficient | term     | coefficient |
|--------|-------------|----------|-------------|
| $x$    | 0.3113339501091084 | $xy^2$   | 0.2845946891080638 |
| $y$    | 0.7114011283722037  | $y^3$    | 0.03537685809298098 |
| $x^2$  | 0.2642178218204402  | $x^4$    | 0.2828190161018292 |
| $xy$   | -0.1555315664585948 | $x^3y$   | 0.06467714057792809 |
| $y^2$  | -0.2072950713728302  | $x^2y^2$ | -0.04965314808703013 |
| $x^3$  | -0.1894657057265681  | $xy^3$   | 0.02054268087684425 |
| $x^2y$ | -0.3660416537330809  | $y^4$    | 0.1771757807926334 |

Table 7

When the projection is to a subset of $S^+_2 \times \mathbb{R}^{15}$, where an optimal metric is not located, the subgradient method converges to a higher value of the bound on the restoration entropy; see Figure 4 where the projection is to a ball with radius 0.5 for $a$ and $[1.5I, 2I]$ for $p$ and the upper bound converges to 1.387955604316810. The metric is given in Table 8.
Metric $P = e^{r_a(x,y)}p$ for the Hénon system on $K$ that gives the lowest upper bound $1.387955604316810$ on the restoration entropy (exogenous step size and projection to a ball of radius 0.5 and $[1.5I, 2I]$; the optimal metric is not in this set).

$$P = \begin{pmatrix}
1.960938619846507 & -0.1342041312602802 \\
-0.1342041312602802 & 1.539064363507252
\end{pmatrix}$$

| term  | coefficient | term  | coefficient |
|-------|-------------|-------|-------------|
| $x$   | 0.1032680736194371 | $xy^2$ | 0.157557752285691 |
| $y$   | 0.18109915872324425 | $y^3$  | 0.191949086085328 |
| $x^2$ | -0.01629057454520433 | $x^4$  | 0.26301214880179 |
| $xy$  | -0.166424112423047 | $x^3y$ | 0.0794455397063254 |
| $y^2$ | 0.0220230639431921 | $x^2y^2$ | 0.0861048023465296 |
| $x^3$ | 0.00208860299479542 | $xy^3$ | 0.1177457076550595 |
| $x^2y$ | -0.1377282628515208 | $y^4$  | 0.02732136529820107 |

Table 8

Finally, we used Polyak step sizes. From our results we deemed it sensible to set $f^* := 1.3$ in (3.13) and use the projection $p$ to the interval $[0.5I, 2I]$ and $\alpha$ to the ball with radius 1.5 in $\mathbb{R}^{15}$. We tried $\alpha$ equal to 0.5, 1, and 1.5 times $\frac{\tanh(\hat{\kappa}D)}{\hat{\kappa}D}$.

The best results were obtained with 1.5: the upper bound $1.300531470950855$ on the restoration entropy in iteration 98,861. The metric is displayed in Table 9. This was the best result we obtained in our computations.

Fig. 5. Restoration entropy for the Hénon system using Polyak step sizes with $\alpha = 1.5\frac{\tanh(\hat{\kappa}D)}{\hat{\kappa}D}$ and using projection of $\alpha$ to the ball with radius 1.5 in $\mathbb{R}^{15}$ and of $p$ to the interval $[0.5I, 2I]$. The best upper bound on the restoration entropy was $1.300531470950855$ obtained in iteration 98,861.
Metric $P = e^{r_a(x,y)}p$ for the Hénon system on $K$ that gives the lowest upper bound $1.300531470950855$ on the restoration entropy (Polyak step size with $f^* = 1.3$, projection to $[0.5I, 2I]$ and $\alpha = 1.5\frac{\tanh(\hat{\kappa}D)}{\hat{\kappa}D}$).

$$p = \begin{pmatrix} 1.738181852184494 & -0.01314292044388416 \\ -0.01314292044388416 & 0.5754131739079722 \end{pmatrix}$$

| term      | coefficient |
|-----------|-------------|
| $x$       | 0.5043455401581707 |
| $y$       | 0.70002957798829 |
| $x^2$     | 0.2393485122804599 |
| $xy$      | -0.1604171490173032 |
| $y^2$     | 0.05239219536145903 |
| $x^3y$    | 0.03270559792757753 |
| $x^2y$    | -0.02920581871325886 |
| $x^3$     | -0.1702108324691254 |
| $x^4$     | 0.06889085991787346 |

Table 9

6. Conclusions and outlook. In this paper we have overcome several shortcomings of the subgradient algorithm with applications to derive upper bounds on the dimension of attractors and the entropy of systems. In particular, we have introduced a projected subgradient method on a compact subset of the domain which guarantees that for this restricted problem a minimizer exists. We demonstrated the applicability of our novel algorithm for upper bounding the Lyapunov dimension and the restoration entropy of the Hénon map.

The dimension of the parameter space could be reduced by considering only positive definite matrices with a fixed value for their determinant. However, this would not significantly improve the numerical performance as the most time-consuming part of the algorithm is the maximization, which drastically suffers from the curse of dimensionality. This maximization, which is now performed by brute force, can potentially be made faster by using a branch-and-bound technique, which is the topic of future work.

Appendix A. Sectional curvature of the space of positive definite matrices. In this section, we compute the tightest lower bound on the sectional curvature of the Riemannian manifold $\mathcal{S}^n_+$. Our analysis is based on the fact that $\mathcal{S}^n_+$ has the structure of a Riemannian symmetric space. We refer to [17, ch. 4] for the theory of symmetric spaces. We use the notation $\text{SO}(n) = \{ A \in \text{GL}(n, \mathbb{R}) : AA^\top = I, \det(A) = 1 \}$ for the special orthogonal group.

By [17, Thm. 3.3], every symmetric space $M$ is a homogeneous space, i.e. there is a Lie group $G$ acting transitively on $M$ such that $M$ is diffeomorphic to the Lie group quotient $G/K$ with $K = \{ g \in G : gp = p \}$ being the isotropy group of a base point $p \in M$ (that can be chosen arbitrarily). In fact, one can take $G$ to be the isometry group of $M$.

To compute the curvature of a homogeneous space, it obviously suffices to compute the curvature at one point. Since the identity matrix $I \in \mathcal{S}^n_+$ is a canonical choice,
we only look at the sectional curvature at \( p = I \).

        We will use the following result, see [17, Prop. 3.4].

        \textbf{Proposition A.1.} Let \( G \) be a connected Lie group, \( \sigma : G \to G \) an involutive automorphism and \( G^\sigma := \{ g \in G : \sigma(g) = g \} \) the associated group of fixed points. Let \( G_0^\sigma \) be the connected component of \( G^\sigma \) which contains the identity and assume that \( G_0^\sigma \) is compact. Then, for any compact subgroup \( K \subset G \) with \( G_0^\sigma \subset K \subset G^\sigma \), the homogeneous space \( G/K \), equipped with a \( G \)-invariant metric, is a symmetric space.

        For a better understanding of the proposition, we recall that a \( G \)-invariant metric on \( G/K \) is a Riemannian metric in which all of the maps \( xK \mapsto gxK, g \in G \), are isometries. To use the proposition for our purposes, let

\[
G := GL^+(n, \mathbb{R}) = \{ g \in GL(n, \mathbb{R}) : \det(g) > 0 \},
\]

which is clearly a connected Lie group. We define

\[
\sigma : G \to G, \quad g \mapsto (g^\top)^{-1}.
\]

This map is an involutive automorphism, since \( \sigma(g_1 g_2) = (g_2^\top g_1^\top)^{-1} = (g_1^\top)^{-1}(g_2^\top)^{-1} = \sigma(g_1)\sigma(g_2) \) and \( \sigma^2(g) = g \). We compute

\[
G^\sigma = \{ g \in G : (g^\top)^{-1} = g \} = \{ g \in G : gg^\top = I \} = SO(n).
\]

Clearly, \( G^\sigma \) is connected, so \( G^\sigma = G_0^\sigma \) and we can only choose \( K = SO(n) \). It follows that \( G/K = GL^+(n, \mathbb{R})/SO(n) \) is a symmetric space when equipped with a \( G \)-invariant metric. We can identify the space \( G/K \) with \( S_n^+ \) via the mapping

\[
\varphi : G/K \to S_n^+, \quad gK \mapsto gg^\top.
\]

Indeed, any matrix of the form \( gg^\top \) with \( g \in G \) is symmetric and positive definite. Conversely, every \( p \in S_n^+ \) can be written as \( p = \varphi(p^\top) \). We have the equivalences

\[
g_1 K = g_2 K \iff g_2^{-1} g_1 \in SO(n)
\]

\[
\iff g_2^{-1} g_1 = (g_2^{-1} g_1)^\top = g_2^\top (g_1^\top)^{-1}
\]

\[
\iff g_1 g_2^\top = g_2 g_1^\top.
\]

We have thus shown that \( \varphi \) is well-defined and bijective. To do the curvature computations in \( G/K \) instead of \( S_n^+ \), the Riemannian metric on \( G/K \) needs to be defined in such a way that \( \varphi \) is an isometry. First, observe that the pullback of the metric that we defined on \( S_n^+ \) to \( G/K \) via \( \varphi \) yields a \( G \)-invariant metric on \( G/K \), since the left translations \( xK \mapsto gxK \) correspond to the maps \( p \mapsto gpg^\top \) on \( S_n^+ \) under the given identification and these maps preserve the trace metric.

The tangent space of \( G/K \) at the base point can be identified with the space of symmetric \( n \times n \) matrices, since \( T_I G = \mathbb{R}^{n \times n} \) and \( T_I K = \{ X \in \mathbb{R}^{n \times n} : X = -X^\top \} \).

To compute the derivative of \( \varphi \) at the base point, consider a smooth curve \( \gamma : \mathbb{R} \to G/K, \gamma(t) = \tilde{\gamma}(t) K \) with \( \tilde{\gamma} : \mathbb{R} \to G \) smooth and \( \tilde{\gamma}(0) = I, v := \tilde{\gamma}'(0) \in S_n \). Then

\[
D \varphi(K)v = \frac{d}{dt} \bigg| _{t=0} \varphi(\gamma(t)) = \frac{d}{dt} \bigg| _{t=0} \tilde{\gamma}(t) \tilde{\gamma}(t)^\top = v + v^\top = 2v.
\]

Since we consider the pullback of the trace metric to \( G/K \) via \( \varphi \), this map needs to be an isometry. That is, the inner product on \( T_K(G/K) = S_n \) has to be defined by

\[
\langle X, Y \rangle = 4 \text{tr}(XY).
\]
According to [17, Thm. 4.2], the Riemannian curvature tensor of $G/K$ at the point $p_0 = K$ can be computed as

$$R(X^*, Y^*)Z^*(p_0) = -[[X, Y], Z]^*(p_0)$$

for all $X, Y, Z \in \mathcal{S}_n$.

Here, $[X, Y] = XY - YX$ is the commutator of matrices and

$$X^*(p) := \left. \frac{d}{dt} \right|_{t=0} \exp(tX) \cdot p,$$

where $\cdot$ denotes the canonical action of $G$ on $G/K$ and $\exp$ is the Lie group exponential map. In our case,

$$X^*(p_0) = \left. \frac{d}{dt} \right|_{t=0} \pi(\exp(tX)),$$

where $\exp(tX) = \sum_{k=0}^{\infty} \frac{1}{k!} t^k X^k$ and $\pi : G \to G/K$, $\pi(g) = gK$, is the canonical projection. It follows that

$$X^*(p_0) = D\pi(I) X.$$

Since $D\pi(I)$ is the identity on $\mathcal{S}_n$, it follows that $X^*(p_0) = X$ and thus

$$R(X^*, Y^*)Z^*(p_0) = -[[X, Y], Z]$$

for all $X, Y, Z \in \mathcal{S}_n$.

Using (A.1), the sectional curvature for a plane spanned by two linearly independent $X, Y \in \mathcal{S}_n$ can then be computed as

$$K(X, Y) = \frac{\langle R(X, Y)Y, X \rangle}{\|X\|^2 \|Y\|^2 - \langle X, Y \rangle^2} = \frac{4\text{tr}(-[[X, Y], Y]X)}{4\text{tr}(X^2) \text{tr}(Y^2) - \text{tr}(XY)^2}$$

$$= \frac{1}{4} \frac{\text{tr}(2Y^2X^2 - 2(XY)^2)}{\text{tr}(X^2) \text{tr}(Y^2) - \text{tr}(XY)^2} = \frac{1}{2} \frac{\text{tr}(X^2Y^2) - \text{tr}((XY)^2)}{\text{tr}(X^2) \text{tr}(Y^2) - \text{tr}(XY)^2}.$$

Here, we used that $\text{tr}(AB) = \text{tr}(BA)$ for arbitrary $A, B$ and the linearity of $\text{tr}$.

Since the sectional curvature of a plane does not depend on the choice of the basis vectors, it suffices to consider orthonormal $X$ and $Y$. Hence, to find the smallest possible curvature, we have so solve the minimization problem

$$\min[\text{tr}((XY)^2) - \text{tr}(X^2Y^2)] \quad \text{s.t. } \text{tr}(X^2) = \text{tr}(Y^2) = 1, \text{ tr}(XY) = 0.$$

Putting $A := XY$, we obtain

$$\text{tr}((XY)^2) - \text{tr}(X^2Y^2) = \text{tr}(A^2) - \text{tr}(XYXY) = \text{tr}(A^2) - \text{tr}(AA^\top)$$

$$= \text{tr}(A(A - A^\top)) = \langle A, A^\top - A \rangle_F.$$

We decompose $A$ into its symmetric and its skew-symmetric part: $A = A_+ + A_-$ with $A_+ = \frac{1}{2}(A + A^\top)$ and $A_- = \frac{1}{2}(A - A^\top)$. Since $\langle A_+, A_- \rangle_F = 0$, we obtain

$$\text{tr}((XY)^2) - \text{tr}(X^2Y^2) = \frac{1}{2} \langle A - A^\top, A^\top - A \rangle_F = -\frac{1}{2} \|A - A^\top\|^2_F.$$
Using \cite[Thm. 2.2]{7}, we can estimate
\begin{equation}
\|A - A^\top\|_F^2 = \|XY - YX\|_F^2 = \|(X, Y)\|_F^2 \leq (\sqrt{2}\|X\|_F\|Y\|_F)^2 = 2.
\end{equation}
Hence, we obtain that $K(X, Y) \geq -\frac{1}{2}$. The results of \cite[Sec. 4]{7} show that this bound is tight. Using \cite[Lem. A.1]{20}, we obtain the following result.

**Proposition A.2.** For each $n \in \mathbb{N}$, the smallest possible lower bound on the sectional curvature of $S^m_n$ is $-\frac{1}{2}$ and the same is true for any space of the form $\mathbb{R}^N \times S^m_n$ with $N \in \mathbb{N}$.

**REFERENCES**

[1] P.-A. Absil, R. Mahony, and R. Sepulchre. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.

[2] M. Bácák. *Convex Analysis and Optimization in Hadamard Spaces*, volume 22 of De Gruyter Series in Nonlinear Analysis and Applications. De Gruyter, 2014.

[3] G. C. Bento, O. P. Ferreira, and J. G. Melo. Iteration-Complexity of Gradient, Subgradient and Proximal Point Methods on Riemannian Manifolds. *J. Optim. Theory Appl.*, 173(2):548–562, 2017.

[4] G. C. Bento and J. G. Melo. Subgradient method for convex feasibility on Riemannian manifolds. *J. Optim. Theory Appl.*, 152(3):773–785, 2012.

[5] R. Bhatia. *Positive definite matrices*. Princeton university press, 2009.

[6] V. A. Boichenko, G. A. Leonov, and V. Reitmann. *Dimension theory for ordinary differential equations*. Teubner Stuttgart, 2005.

[7] A. Böttcher and D. Wenzel. The Frobenius norm and the commutator. *Linear algebra and its applications*, 429(8-9):1864–1885, 2008.

[8] N. Boumal. An introduction to optimization on smooth manifolds. To appear with Cambridge University Press, Apr 2022.

[9] T. T. Cai and A. Zhang. Rate-optimal perturbation bounds for singular subspaces with applications to high-dimensional statistics. *The Annals of Statistics*, 46(1):60–89, 2018.

[10] J. da Cruz Neto, L. De Lima, and P. Oliveira. Geodesic algorithms in Riemannian geometry. *Balkan J. Geom. Appl.*, 3(2):89–100, 1998.

[11] M. P. do Carmo. *Riemannian geometry*. Mathematics: Theory & Applications. Birkhäuser Boston, Inc., Boston, MA, 1992. Translated from the second Portuguese edition by Francis Flaherty.

[12] O. P. Ferreira. Proximal subgradient and a characterization of Lipschitz function on Riemannian manifolds. *J. Math. Anal. Appl.*, 313(2):587–597, 2006.

[13] O. P. Ferreira, M. S. Louzeiro, and L. F. Prudente. First order methods for optimization on riemannian manifolds. In *Handbook of Variational Methods for Nonlinear Geometric Data*, pages 499–525. Springer International Publishing.

[14] O. P. Ferreira, M. S. Louzeiro, and L. F. Prudente. Iteration-complexity of the subgradient method on Riemannian manifolds with lower bounded curvature. *Optimization*, 68(4):713–729, 2019.

[15] O. P. Ferreira and P. R. Oliveira. Subgradient algorithm on Riemannian manifolds. *J. Optim. Theory Appl.*, 97(1):93–104, 1998.

[16] D. Gabay. Minimizing a differentiable function over a differential manifold. *J. Optim. Theory Appl.*, 37(2):177–219, 1982.

[17] S. Helgason. *Differential geometry, Lie groups, and symmetric spaces*. Academic press, 1979.

[18] M. Hénon. A two-dimensional mapping with a strange attractor. In *The theory of chaotic attractors*, pages 94–102. Springer, 1976.

[19] C. Kawan, S. Hafstein, and P. Giesl. Recentrs: Restoration entropy estimation for dynamical systems via riemannian metric optimization. *SoftwareX*, 15:1000743, 2021.

[20] C. Kawan, S. Hafstein, and P. Giesl. A subgradient algorithm for data-rate optimization in the remote state estimation problem. *SIAM J. Appl. Dyn. Syst.*, 20(4):2142–2173, 2021.

[21] C. Kawan, A. S. Matveev, and A. Y. Pogromsky. Remote state estimation problem: Towards the data-rate limit along the avenue of the second lyapunov method. *Automatica*, 125:109467, 2021.

[22] N. Kuznetsov and V. Reitmann. *Attractor dimension estimates for dynamical systems: theory and computation—dedicated to Gennady Leonov*, volume 38 of *Emergence, Complexity and Computation*. Springer, Cham, [2021] ©2021.
This manuscript is for review purposes only.