AN EFFICIENT RETRACTION MAPPING FOR THE SYMPLECTIC STIEFEL MANIFOLD

HARRY OVIEDO* AND RAFAEL HERRERA†

Abstract. This article introduces a new retraction on the symplectic Stiefel manifold. The operation that requires the highest computational cost to compute the novel retraction is a matrix inversion of size $2p$–by–$2p$, which is much less expensive than those required for the available retractions in the literature. Later, with the new retraction, we design a constraint preserving gradient method to minimize smooth functions defined on the symplectic Stiefel manifold. In order to improve the numerical performance of our approach, we use the non–monotone line–search of Zhang and Hager with an adaptive Barzilai–Borwein type step–size. Our numerical studies show that the proposed procedure is computationally promising and is a very good alternative to solve large–scale optimization problems over the symplectic Stiefel manifold.

Key words. Symplectic Stiefel manifold, Riemannian gradient method, Riemannian optimization, Symplectic matrix.

AMS subject classifications. 65K05, 70G45, 90C48.

1. Introduction.

The aim of this work is to design a fast Riemannian gradient method to find local minimizers of the following optimization problem with skew–symmetric constraints

$$\min_{X\in\mathbb{R}^{2n\times 2p}} \mathcal{F}(X) \quad \text{s.t.} \quad X^\top J_{2n} X = J_{2p}, \tag{1.1}$$

where $p \leq n$, $\mathcal{F} : \mathbb{R}^{2n\times 2p} \to \mathbb{R}$ is a bounded below and continuously differentiable function, and $J_{2m} := [0, I_m; -I_m, 0]$, where I_m denotes the m–by–m identity matrix for any positive integer m. The constraints set $Sp(2n, 2p) = \{X \in \mathbb{R}^{2n\times 2p} : X^\top J_{2n} X = J_{2p}\}$ of problem (1.1) is so–called the *symplectic Stiefel manifold*, [15]. In fact, if we equip $Sp(2n, 2p)$ with a positive–definite inner product $\langle\cdot,\cdot\rangle_X$ on the tangent space $T_X Sp(2n, 2p)$ at each point X, then the pair $(Sp(2n, 2p), \langle\cdot,\cdot\rangle_X)$ becomes a closed embedded Riemannian sub–manifold of $\mathbb{R}^{2n\times 2p}$, whose dimension is equal to $4np - p(2p - 1)$, see Proposition 3.1 in [15]. In the special case $p = n$, the symplectic Stiefel manifold reduces to the *symplectic group* and denoted by $Sp(2n)$, which is a Lie group.

Symplectic matrices arise in many fields, such as optical systems [14], optimal control of quantum symplectic gates [37], optimization problems on the set of symplectic matrices [14, 15, 30, 33], procrustes problem [39], matrix decompositions [6, 32], computation of eigenvalues of (skew–)Hamiltonian matrices [5, 7, 9, 34], trace minimization problems [33] and symplectic principal component analysis [21, 22, 29], which motivates the study of this class of matrices.

Most of the optimization methods on an Euclidean space, perform a line search after computing a descent search direction using a straight line as parameterization [23, 24]. By contrast, in the Riemannian optimization context, the concept of a straight line is substituted with a curve (not necessarily a geodesic) over the determined Riemannian manifold [1, 18]. In particular, the *retractions* are the most pragmatic approach to constructing these curves. In short words, a retraction is a smooth mapping that sends tangent vectors to the manifold, and defines an appropriate curve for searching a next iterate point on the corresponding manifold. The rigorous definition of retraction appears in [1].

The Riemannian gradient method equipped with a non–monotone globalization technique and the Barzilai–Borwein step–size [4], has been a efficient alternative to solve several manifold constrained optimization problems with real applications, for example see [19, 20, 26, 27, 28, 33, 35]. Recently, some constraint preserving gradient methods were proposed based on the geometrical study of the symplectic Stiefel manifold [15, 33]. In particular, in [15] the authors developed two Riemannian gradient procedures to solve the optimization problem (1.1). The first approach uses a global defined quasi–geodesic on the constraint set. Nevertheless, it is necessary to compute two matrix exponential in order to evaluate that quasi–geodesic, which is computationally restrictive. With the goal of designing a less computationally

*Escola de Matemática Aplicada, Fundação Getulio Vargas (FGV/EMAp). Rio de Janeiro, RJ, Brazil. (harry.oviedol@gmail.com).

†Centro de Investigación en Matemáticas, CIMAT A.C. Guanajuato, Gto. Mexico. (rherrera@cimat.mx)

costly iterative scheme, Gao et.al. in [15] build another gradient method based on the Cayley transform by generalizing the works of Wen and Yin in [35]. Specifically, in [15] is introduced the following retraction map

$$R_X(\xi_X) = \left(I_{2n} - \frac{1}{2}A_X J_{2n}\right)^{-1}\left(I_{2n} + \frac{1}{2}A_X J_{2n}\right)X, \tag{1.2}$$

for all $X \in Sp(2n, 2p)$ and $\xi_X = A_X J_{2n} X$ in the tangent space of $Sp(2n, 2p)$ at X. With this retraction, the authors in [15, 33], designed a Riemannian gradient method combined with chosen of Barzilai–Borwein step–size. Additionally, the sophisticated retraction (1.2) is equivalent to

$$R_X(\xi_X) = X + U\left(I_{4p} + \frac{1}{2}V^\top J_{2n}^\top U\right)^{-1} V^\top J_{2n} X, \tag{1.3}$$

where $A_X = UV^\top$, with $U, V \in \mathbb{R}^{2n\times 4p}$ are two matrices associated with $\xi_X = A_X J_{2n} X$, for more details see Proposition 5.4 in [15]. Observe that this last formula is more advantageous than (1.2), when $p < \frac{n}{2}$. However, both formulae (1.2)–(1.3) require inverting a large–size matrix when $p \geq n/2$, and n large.

In this paper, we introduce a more numerically attractive retraction mapping on the symplectic Stiefel manifold, which requires inverting a matrix of smaller size, specifically, it needs to invert a matrix of size $2p \times 2p$. The new retraction is constructed following the descriptions presented in [20]. Thus, the new retraction can be seen as a generalization of the curve developed in [20] for the Stiefel manifold. Using the new retraction, we design a very efficient Riemannian gradient method to address problem (1.1), which uses the Zhang and Hager non–monotone globalization strategy [38] combined with a step–size presented in [40], to speed up the numerical behavior of the proposal. Our preliminary numerical experiments suggest that our proposal is numerically superior to all the existing Riemannian gradient methods in the state–of–the–art.

The following sections are organized as follows: section 2 reviews the geometry of the symplectic Stiefel manifold by summarizing the concepts contained in [15, 33]; section 3 describes the new approach in detail. In section 4 we present some numerical results to illustrate the numerical performance of our proposal. Finally, in section 5 we provide the conclusions of this work.

2. Geometric tools associated with the symplectic Stiefel manifold. This section contains some notations and key concepts for a good understanding of this manuscript. The notions summarized in this section can also be found in [1, 15, 33].

Let m be a positive integer number, the matrix J_{2m}, defined at the beginning of Section 1, has the following properties

$$J_{2m}^\top = -J_{2m}, \quad J_{2m}^\top J_{2m} = I_{2m}, \quad J_{2m}^2 = -I_{2m}, \quad J_{2m}^{-1} = J_{2m}^\top.$$

Now, given a square matrix $A \in \mathbb{R}^{m\times m}$, $\texttt{sym}(A)$ will denotes the symmetric part of A that is, $\texttt{sym}(A) = 0.5(A + A^\top)$. $\mathcal{S}_+(m)$ will denotes the set of m–by–m positive definite matrices with real entries. The trace of A is defined as the sum of the diagonal elements which we denote by $tr(A)$. The standard inner product of two matrices $A, B \in \mathbb{R}^{m\times n}$ is given by $\langle A, B\rangle := \sum_{i,j} a_{ij} b_{ij} = tr(A^\top B)$. The Frobenius norm is defined by $||A||_F = \sqrt{\langle A, A\rangle}$. Let $X \in Sp(2n, 2p)$ an arbitrary matrix, the tangent space of the symplectic Stiefel manifold at X is given by [15]

$$T_X Sp(2n, 2p) = \{Z \in \mathbb{R}^{2n\times 2p} : Z^\top J_{2n} X + X^\top J_{2n} Z = 0\}.$$

A characterization of the tangent space see [15], crucial for this manuscript is

$$T_X Sp(2n, 2p) = \{S J_{2n} X : S^\top = S, \, S \in R^{2n\times 2n}\}.$$

The tangent bundle of $Sp(2n, 2p)$ is defined as the arbitrary union of all the tangent spaces, i.e. $TSp(2n, 2p) = \cup_{X\in Sp(2n,2p)} T_X Sp(2n, 2p)$. Let $\mathcal{F} : \mathbb{R}^{2n\times 2p} \to \mathbb{R}$ be a differentiable function, we denote by $\nabla\mathcal{F}(X) := (\frac{\partial \mathcal{F}(X)}{\partial X_{ij}})$ the matrix of partial derivatives of $\mathcal{F}$ (the Euclidean gradient of $\mathcal{F}$). Let $\Phi : Sp(2n,2p) \to \mathbb{R}$ be a smooth function defined on the symplectic Stiefel manifold, then the Riemannian gradient of Φ at $X \in Sp(2n,2p)$, denote by $\mathrm{grad}\Phi(X)$, is the unique vector in $T_X Sp(2n,2p)$ satisfying

$$\mathcal{D}\Phi(X)[\xi_X] := \lim_{\tau\to 0} \frac{\Phi(\gamma(\tau)) - \Phi(\gamma(0))}{\tau} = \langle \nabla\Phi(X), \xi_X\rangle, \quad \forall \xi_X \in T_X St(n,p),$$

where $\gamma : [0, \tau_{\max}] \to Sp(2n,2p)$ is any curve that verifies $\gamma(0) = X$ and $\dot\gamma(0) = \xi_X$.

In addition, let $X \in Sp(2n,2p)$, the canonical–like metric [33] associated with the symplectic Stiefel manifold is defined as

$$\langle \xi_X, \eta_X\rangle_c := tr\left(\xi_X^\top \left(\frac{1}{\rho} J_{2n}X(J_{2n}X)^\top - (J_{2n}XJ_{2p}X^\top J_{2n}^\top - J_{2n})^2\right)\eta_X\right),$$

where $\xi_X, \eta_X \in T_X Sp(2n,2p)$ and $\rho > 0$. Under the canonical–like metric, the Riemannian gradient of $\mathcal{F}$ has the following closed expression, see Ref. [15],

$$\mathrm{grad}_\rho \mathcal{F}(X) = A_X J_{2n} X, \tag{2.1}$$

where $A_X = 2\mathtt{sym}(H_X \nabla\mathcal{F}(X)(XJ_{2p})^\top)$, $H_X = I_{2p} + \frac{\rho}{2}XX^\top + J_{2n}X(X^\top X)^{-1}X^\top J_{2n}$, and $\rho > 0$. Similar to the unconstrained optimization theory, in the Riemannian setting, the critical points of the Riemannian gradient are candidates to be local minimizers of an optimization problem defined over a corresponding manifold. Therefore, $\mathrm{grad}_\rho\mathcal{F}(X) = 0$ is the first–order necessary optimality condition for the problem of interest (1.1).

3. A computationally efficient retraction on $Sp(2n,2p)$.

In this section, we introduce a new constraint preserving approach in order to tackle the optimization problem (1.1). The new approach can be regarded as a generalization of the feasible curve constructed by Jiang and Dai in [20]. Afterwards, we will introduce a very efficient retraction for the symplectic Stiefel manifold, based on the new feasible curve.

Before introducing the new curve on $Sp(2n,2p)$, let's define by

$$P_X := I_{2n} + XJ_{2p}X^\top J_{2n}, \tag{3.1}$$

for all $X \in Sp(2n,2p)$. Let $X \in Sp(2n,2p)$ and $Z \in \mathbb{R}^{2n\times 2p}$ be an arbitrary matrix, it is easy to prove that $P_X Z \in T_X Sp(2n,2p)$. Hence, the matrix P_X is a projection operator on the tangent space of $Sp(2n,2p)$ at X. In addition, observe that this matrix satisfies that $X^\top J_{2n} P_X = 0$, for all $X \in Sp(2n,2p)$.

Now, given a tangent vector $Z \in T_X Sp(2n,2p)$, we will construct a feasible curve by using the following formulation

$$Y(\tau) = (\,XR(\tau) + \tau P_X Z\,)S(\tau), \tag{3.2}$$

where $S : [0,\tau_{\max}] \to \mathbb{R}^{2p\times 2p}$ is an invertible curve (at least locally), that is, there exists $\tau_{\max} > 0$ such that the matrix $S(\tau)$ is invertible for all $\tau \in [0,\tau_{\max}]$; and $R : [0,\tau_{\max}] \to \mathbb{R}^{2p\times 2p}$.

Since our aim is to build a retraction on $Sp(2n,2p)$, we require that $Y(\tau)$ verify the following properties

(A) The curve must pass through the point X, i.e. $Y(0) = X$.
(B) Local rigidity: $\dot Y(0) = Z$.
(C) Feasibility: $Y(\tau)^\top J_{2n} Y(\tau) = J_{2p}$, for all $\tau \in [0,\tau_{\max}]$.

Note that if we impose the initial conditions $S(0) = I_{2p}$ and $R(0) = I_{2p}$ then the property (A) is guaranteed. By differentiating $Y(\tau)$ and using these initial conditions, we have the following relation

$$\dot Y(0) = X(\,\dot R(0) + \dot S(0)\,) + P_X Z. \tag{3.3}$$

Then, we can assure the second property (B) by imposing that

$$\dot{R}(0) + \dot{S}(0) = J_{2p}^{\top} X^{\top} J_{2n} Z. \tag{3.4}$$

In the rest of this construction, we select the following correlative model between the curves $S(\tau)$ and $R(\tau)$,

$$R(\tau) = 2I_{2p} - S(\tau)^{-1}, \tag{3.5}$$

note that this correspondence is exactly the one used in [20].

Differentiating both sides of (3.5) we obtain $\dot{R}(\tau)S(\tau) + R(\tau)\dot{S}(\tau) = 2\dot{S}(\tau)$, which together with the initial conditions imply that $\dot{R}(0) = \dot{S}(0)$. Combining this last relation with (3.4) we arrive at

$$\dot{S}(0) = \frac{1}{2} J_{2p}^{\top} X^{\top} J_{2n} Z. \tag{3.6}$$

On the other hand, it follows form the third property (C) that

$$S(\tau)^{\top}(\, R(\tau)^{\top} X^{\top} + \tau Z^{\top} P_X^{\top}\,) J_{2n}(\, X R(\tau) + \tau P_X Z\,) S(\tau) = J_{2p}, \tag{3.7}$$

or equivalently

$$R(\tau)^{\top} J_{2p} R(\tau) + \tau^2 Z^{\top} J_{2n} P_X Z = S(\tau)^{-\top} J_{2p} S(\tau)^{-1}. \tag{3.8}$$

In addition, form (3.5) we have

$$R(\tau)^{\top} J_{2p} R(\tau) = 4J_{2p} - 2J_{2p} S(\tau)^{-1} - 2S(\tau)^{-\top} J_{2p} + S(\tau)^{-\top} J_{2p} S(\tau)^{-1}. \tag{3.9}$$

In view of the equations (3.8) and (3.9), we obtain $J_{2p}S(\tau)^{-1} + S(\tau)^{-\top} J_{2p} = 2J_{2p} + \frac{\tau^2}{2} Z^{\top} J_{2n} P_X Z$, which implies that $J_{2p}S(\tau)^{-1} - (J_{2p}S(\tau)^{-1})^{\top} = 2J_{2p} + \frac{\tau^2}{2} Z^{\top} J_{2p} P_X Z$. This last equation suggests that

$$J_{2p} S(\tau)^{-1} = J_{2p} + \frac{\tau^2}{4} Z^{\top} J_{2n} P_X Z + L(\tau), \tag{3.10}$$

where $L : [0, \tau_{\max}] \to \mathbb{R}^{2p\times 2p}$ must satisfy that $L(0) = 0$ and $L(\tau)$ should be a symmetric matrix for all $\tau \in [0, \tau_{\max}]$. Notice that the equation (3.10) is equivalent to

$$S(\tau) = \left(J_{2p} + \frac{\tau^2}{4} Z^{\top} J_{2n} P_X Z + L(\tau) \right)^{-1} J_{2p}. \tag{3.11}$$

Observe that if we find a formula for the curve $L(\tau)$ then the curves $S(\tau)$ and $R(\tau)$ will be completely defined. Therefore, this description reduces to finding a suitable curve $L(\tau)$.

Again, by differentiating (3.11) an evaluating at $\tau = 0$, we get $\dot{S}(0) = -J_{2p}^{\top} \dot{L}(0)$. Substituting this last result in (3.6) we arrive at

$$\dot{L}(0) = -\frac{1}{2} X^{\top} J_{2n} Z. \tag{3.12}$$

Therefore, we simply have to select a curve $L(\tau)$ such that $L(0) = 0$, $\dot{L}(0) = -\frac{1}{2} X^{\top} J_{2n} Z$ and $L(\tau)^{\top} = L(\tau)$, we can achieve these three properties by taking $L(\tau) = -\frac{\tau}{2} X^{\top} J_{2n} Z$.

Clearly with this choice of $L(\tau)$, the initial conditions $L(0) = 0$, $\dot{L}(0) = -\frac{1}{2} X^{\top} J_{2n} Z$ hold. In addition, since $Z \in T_X Sp(2n, 2p)$, then we have $Z^{\top} J_{2n} X = -X^{\top} J_{2n} Z$. This property of Z implies that $L(\tau)^{\top} = L(\tau)$ for all $\tau \in \mathbb{R}$.

In summary, we can write the curve $Y(\tau)$ in a simple equation as follows

$$Y(\tau) = (\,2X + \tau P_X Z\,)\left(I_{2p} + \frac{\tau}{2} J_{2p}^{\top}(J_{2n}X)^{\top} Z + \frac{\tau^2}{4} J_{2p}^{\top} Z^{\top} J_{2n} P_X Z\right)^{-1} - X. \tag{3.13}$$

Finally, this formulation of $Y(\tau)$ suggests the following retraction on $Sp(2n, 2p)$

$$\mathcal{R}_X(\xi_X) = (\,2X + P_X\xi_X\,)\left(I_{2p} + \frac{1}{4}(\xi_X J_{2p})^{\top} J_{2n}(2X + P_X\xi_X)\right)^{-1} - X, \tag{3.14}$$

which is obtained from (3.13) substituting τZ by $\xi_X \in T_X Sp(2n, 2p)$. We confirm right away that (3.14) is indeed a retraction.

LEMMA 3.1. *The map $\mathcal{R} : TSp(2n, 2p) \to Sp(2n, 2p)$ defined in (3.14) is a retraction.*

Proof. In order to prove this lemma, we need to demonstrate that the map (3.14) satisfies the Definition 4.1.1 contained in [1]. Firstly, notice that $\mathcal{R}_X(\cdot)$ is a smooth mapping because it is made up of products and subtractions of differentiable functions. In addition, for the construction of the curve $Y(\tau)$ in (3.13), we have that the mapping $\mathcal{R}(\cdot)$ preserves the manifold structure of $Sp(2n, 2p)$, i.e. $\mathcal{R}_X(\xi_X) \in Sp(2n, 2p)$, for all $X \in Sp(2n, 2p)$ and $\xi_X \in T_X Sp(2n, 2p)$. Secondly, let X an arbitrary matrix in $Sp(2n, 2p)$, then note that

$$\mathcal{R}_X(0_X) = 2X\,(J_{2p})^{-1} J_{2p} - X = X. \tag{3.15}$$

On the other hand, the directional derivative of $\mathcal{R}_X(\cdot)$ at 0_X in direction $Z \in T_X Sp(2n, 2p)$ verifies that

$$\mathcal{D}\mathcal{R}_X(0_X)[Z] = \lim_{\tau\to 0} \frac{\mathcal{R}_X(0_X + \tau Z) - \mathcal{R}_X(0_X)}{\tau} = \lim_{\tau\to 0} \frac{\mathcal{R}_X(\tau Z) - X}{\tau} = \lim_{\tau\to 0} \frac{Y(\tau) - Y(0)}{\tau} = \dot{Y}(0) = Z. \tag{3.16}$$

where $Y(\tau)$ is the curve defined in (3.13), the last equation is obtained by the construction of the curve $Y(\tau)$. Therefore, we conclude that the mapping defined in (3.14) is a retraction. ☐

Similar to the retraction based on the Cayley transform (1.2), the new retraction (3.14) is not globally defined, since the matrix $M(\tau) := I_{2p} + 0.5\tau J_{2p}^{\top}(J_{2n}X)^{\top} Z + 0.25\tau^2 J_{2p}^{\top} Z^{\top} J_{2n} P_X Z$ may be singular for some selections of $\tau \in \mathbb{R}$. However, since J_{2p} is a non–singular matrix, and $M(\tau)$ is a continuous function on τ (it is a polynomial function), then there always exists an neighborhood of $\tau = 0$ such that $M(\tau)$ is non–singular. Therefore, the new retraction (3.14) is well–defined at least locally. In addition, compared to the Cayley based retraction, this new functional provides a more computationally efficient scheme, because the retraction (3.14) requires inverting a small–size matrix $2p \times 2p$, while the Cayley transform needs to invert a matrix of size $2n \times 2n$ (or $4p \times 4p$ in the best case, see equation (1.3)). This interesting property makes retraction (3.14) more attractive for solving large–scale optimization problems over $Sp(2n, 2p)$.

3.1. A Riemannian gradient algorithm.

Focusing on our novel retraction, we propose a non–monotone Riemannian gradient method to address the minimization problem (1.1). Specifically, we propose to construct a feasible sequence $\{X_k\}$ using the following curvilinear search iterative scheme, starting at $X_0 \in Sp(2n, 2p)$

$$X_{k+1} = (\,2X_k - \tau_k P_{X_k} Z_k\,)\left(I_{2p} - \frac{\tau_k}{4}(Z_k J_{2p})^{\top} J_{2n}(2X_k - \tau_k P_{X_k} Z_k)\right)^{-1} - X_k, \tag{3.17}$$

where $Z_k = \mathrm{grad}_\rho \mathcal{F}(X_k)$ and $\tau_k > 0$ is the step–size. To determine the k–th step–size, in this work, we use the adaptive Barzilai–Borwein (ABB), originally introduced in [40], which is given by

$$\tilde{\tau}_k^{ABB} = \begin{cases} \tau_k^{BB2} & \text{if } \tau_k^{BB2} \leq \kappa \tau_k^{BB1}; \\ \tau_k^{BB1} & \text{otherwise,} \end{cases} \tag{3.18}$$

where $\kappa \in (0, 1)$, τ_k^{BB1} and τ_k^{BB2} are the well–known Barzilai–Borwein step–sizes [4],

$$\tau_k^{BB1} = \frac{||S_{k-1}||_F^2}{|\langle S_{k-1}, Y_{k-1}\rangle|} \quad \text{and} \quad \tau_k^{BB1} = \frac{|\langle S_{k-1}, Y_{k-1}\rangle|}{||Y_{k-1}||_F^2},$$

where $S_{k-1} = X_k - X_{k-1}$ and $Y_{k-1} = \text{grad}_\rho\mathcal{F}(X_k) - \text{grad}_\rho\mathcal{F}(X_{k-1})$. As the term $|\langle S_{k-1}, Y_{k-1}\rangle|$ can be equal to zero or very close to zero, in our procedure, we include a safeguard that guarantees that the k–th step–size is neither too small nor too large, in particular we use $\tau_k^{ABB} = \max(\min(\tilde{\tau}^{ABB}, \tau_M), \tau_m)$, where $0 < \tau_m < \tau_M < \infty$. In addition, since the step–size τ_k^{ABB} alone does not guarantee a sufficient decrease in the objective function value at every iteration, it may invalidate the convergence of the proposed method. However, this issue can be solved by incorporating a globalization strategy that regulates the step–size τ_k^{ABB} only when necessary [10, 31]. Particularly in this work, we use the non–monotone line–search globalization technique developed by Zhang and Hager in [38]. Now we are ready to present the proposed iterative algorithm in detail, see Algorithm 3.1.

Algorithm 3.1 Riemannian gradient method.

Require: $X_0 \in Sp(2n, 2p)$, $0 < \tau_m < \tau_M < \infty$, $\eta \in [0, 1)$, $c_1, \rho, \epsilon, \delta, \tau, \kappa \in (0, 1)$, $Q_0 = 1$, $C_0 = \mathcal{F}(X_0)$, $k = 0$.

while $||\text{grad}_\rho\mathcal{F}(X_k)||_F > \epsilon$ **do**
    while $\mathcal{F}(\,\mathcal{R}_{X_k}(-\tau\text{grad}_\rho\mathcal{F}(X_k))\,) > C_k - c_1\tau||\text{grad}_\rho\mathcal{F}(X_k)||_F^2$ **do**
        $\tau = \delta\tau$,
    end while
    $X_{k+1} = \mathcal{R}_{X_k}(-\tau\text{grad}_\rho\mathcal{F}(X_k))$, according to (3.17),
    Compute $\tilde{\tau}^{ABB}$ according to (3.18),
    $\tau = \max(\min(\tilde{\tau}^{ABB}, \tau_M), \tau_m)$,
    $Q_{k+1} = \eta Q_k + 1$ and $C_{k+1} = (\eta Q_k C_k + F(x_{k+1}))/Q_{k+1}$.
    $k \leftarrow k + 1$.
end while

The convergence of the retraction–based Riemannian gradient method combined with the non–monotone Zhang–Hager globalization strategy has been demonstrated in [18, 25]. However, these convergence are valid only for globally defined retractions. Therefore, the theorems contained in [18, 25] do not apply directly for our Algorithm 3.1. Fortunately, the fact that the new retraction is not globally defined does not prevents us from adapting the convergence and complexity results of [1, 18, 25]. In fact, using the same proof as in Gao et al. [15] (see Theorem 5.6 and Corollary 5.7) proposed for the Riemannian gradient method using the retraction based on the Cayley transformation (which is also not globally defined), we have the following theoretical result.

THEOREM 3.2. *Let $\{X_k\}$ be an infinite sequence of matrices generated by Algorithm 3.1. Then any accumulation point X_* of $\{X_k\}$ is a stationary point of $\mathcal{F}$, i.e., $||\text{grad}_\rho\mathcal{F}(X_*)||_F = 0$.*

4. Numerical experiments. In this section, we show the efficiency of Algorithm 3.1 applying to two different groups of experiments, considering the solution of the nearest symplectic matrix problem and also the trace minimization problem over the symplectic Stiefel manifold. We implement all the simulations in Matlab (version 2017b) with double precision on a machine intel(R) CORE(TM) i7–8750H, CPU 2.20 GHz with 1TB HD and 16GB RAM. For comparative purposes, we test our Algorithm 3.1 against the Riemannian gradient method based on the Cayley transformation (Cayley), and with the Riemannian gradient method based on the quasi–geodesic approach (Qgeodesic), these two methods were developed in [15][1]. All the methods use, as search direction, the Riemannian gradient obtained from the canonical metric with $\rho = 1$. In Algorithm 3.1 we use the following defaults values: τ_m = 1e-15, τ_M = 1e+15, $\eta = 0.85$, c_1 = 1e-4, $\delta = 0.2$, τ = 1e-3 and $\kappa = 0.65$. In all the experiments, we adopt the following expressions as the stopping criterion: the iterations stops if the algorithms find a matrix $\hat{X} \in Sp(2n, 2p)$ such that $||\text{grad}_\rho\mathcal{F}(\hat{X})||_F < \epsilon$, or if the corresponding algorithm exceeds N iterations, where the values of N and ϵ will be specified for each experiment in the following subsections. The implementation of our algorithm is found in http://www.optimization-online.org/DB_HTML/2021/07/8478.html, (see the compressed postscript).

[1]The Riemannian gradient methods Cayley and Qgeodesic can be downloaded from https://github.com/opt-gaobin/spopt

TABLE 1
Numerical results related to randomly generated nearest symplectic matrix problems.

p	100	200	300	400	500	600
Cayley						
Iter	33.8	38.6	43.2	50.3	57.1	69
Time	1.73	5.83	14.76	27.78	49.90	77.29
Grad	7.05e-6	5.75e-6	8.35e-6	7.87e-6	5.92e-6	5.77e-6
Feasi	2.79e-13	2.95e-13	4.32e-13	2.70e-12	6.67e-11	7.61e-11
Qgeodesic						
Iter	33.7	38.6	46	51.8	61	70.6
Time	2.34	9.32	28.63	57.61	113.84	203.41
Grad	6.04e-06	6.52e-6	6.87e-6	6.30e-6	6.88e-6	4.79e-6
Feasi	3.99e-12	1.57e-11	2.53e-11	3.32e-11	6.54e-11	8.50e-11
Algorithm 1						
Iter	34.8	37.9	43.1	47.9	56.6	65.2
Time	1.97	5.78	14.17	24.43	43.73	69.87
Grad	6.62e-6	7.51e-6	7.44e-6	7.03e-6	6.18e-6	7.18e-6
Feasi	1.53e-12	5.44e-12	1.59e-11	2.06e-11	3.89e-11	4.49e-11

Throughout this section, we use the following notation: *Time*, *Iter*, *Grad*, *Feasi* denote the average total computing time in seconds, the average number of iterations, the average residual $||\mathrm{grad}_\rho \mathcal{F}(X)||_F$ and the average feasibility error $||\hat{X}^\top J_{2n}\hat{X} - J_{2p}||_F$, respectively. In all the experiment, we solve thirty independent instances for each pair (n,p) and then we report all these mean values. When we do not specify how the initial point $X_0 \in Sp(2n,2p)$ was designed, it will be understood that we randomly generated X_0 following the strategy suggested in the subsection 6.1 in [15].

4.1. The nearest symplectic matrix problem. Given an arbitrary matrix $A \in \mathbb{R}^{2n\times 2p}$, the nearest symplectic matrix problem refers to computing the symplectic matrix $X^* \in Sp(2n,2p)$ closest to A in Frobenius norm. This problem is formulated mathematically as follows

$$\min_{X\in\mathbb{R}^{2n\times 2p}} ||X - A||_F^2 \quad s.t. \qquad X^\top J_{2n}X = J_{2p}. \tag{4.1}$$

As a first experiment, we apply the methods on problem (4.1) considering random data. In particular, given (n,p), the matrix $A \in \mathbb{R}^{2n\times 2p}$ is assembled as $A = \bar{A}/||\bar{A}||_2$, where $\bar{A} \in \mathbb{R}^{2n\times 2p}$ is a matrix whose entries are sampled from the standard Gaussian distribution. Additionally, in this experiments we use $N = 1000$ and $\epsilon = $ 1e-5 for all the algorithms. Table 1 reports the numerical results associated to this first test varying $p \in \{100, 200, 300, 400, 500, 600\}$ and a fixed $n = 1000$. As shown in Tables 1 all the methods obtained estimates of a solution of problem (4.1) with the required accuracy. Furthermore, we notice that as p approaches n our proposal converges more quickly (in terms of computational time) than the other two methods.

Secondly, we test the three methods on the solution of problem (4.1) but now using real data. For this end, we consider 9 large sparse and square matrices taken from the SuiteSparse Matrix Collection [8][2]. Since all these data matrices are square we truncate their columns to obtain the matrices A's of appropriate size. In this experiment, we fix $p = 50$ and the matrix A will be determined as $A = \bar{A}/||\bar{A}||_\infty$, where $\bar{A} = M(:, 1 : 2p)$ using Matlab notation and M is the original matrix taken form the SuiteSparse matrix collection. For this second set of experiments, we use $N = 1000$ and $\epsilon = $ 1e-4 in the stop criteria of the algorithms. In order for this computational comparison to be reproducible, we construct the starting point with the following Matlab commands (omitting the multiplication symbol):

$$\texttt{randn}('\mathrm{seed}', 1);\ W = \texttt{randn}(2p, 2p);\ W = W'W + 0.1\texttt{eye}(2p);\ E = \texttt{expm}([W(p+1 : \mathrm{end}, :); -W(1 : p, :)]);$$

and

$$X_0 = [E(1 : p, :); \texttt{zeros}(n - p, 2p); E(p + 1 : \mathrm{end}, :); \texttt{zeros}(n - p, 2p)].$$

[2]The SuiteSparse Matrix Collection tool–box is available in https://sparse.tamu.edu/

TABLE 2
Solving the nearest symplectic matrix problem for 9 instances in the SuiteSparse matrix collection.

		Cayley				Qgeodesic				Algorithm 1			
Name	$2n$	Iter	Time	Grad	Feasi	Iter	Time	Grad	Feasi	Iter	Time	Grad	Feasi
1138_bus	1138	1000	19.6	4.90e-3	1.06e+0	658	11.5	8.83e-5	1.30e-11	588	7.1	9.80e-5	1.13e-11
bcsstk08	1074	1000	18.1	6.50e-3	7.34e-2	394	6.6	9.77e-5	6.18e-12	415	4.7	8.18e-5	5.28e-11
bcsstk10	1086	1000	17.7	2.80e-3	1.00e-2	334	5.4	9.83e-5	3.31e-12	325	3.8	9.37e-5	5.97e-12
bcsstk16	4884	1000	90.8	1.10e-3	2.13e-1	82	5.6	9.42e-5	7.38e-13	51	3.8	9.64e-5	7.01e-13
bcsstk21	3600	1000	70.4	1.50e-3	2.10e-3	371	20.5	9.82e-5	7.22e-12	389	22.9	8.69e-5	1.35e-11
bcsstk27	1224	1000	20.7	5.20e-3	2.18e-2	307	5.4	8.79e-5	6.45e-12	345	4.4	9.75e-5	8.03e-12
bodyy4	17546	1000	289.9	1.17e-2	3.19e-1	217	45.1	9.74e-5	3.73e-12	75	17.9	7.70e-5	4.39e-12
crystm03	24696	1000	366.8	1.40e-3	5.47e-4	91	25.7	9.03e-5	9.00e-13	59	19.4	2.74e-5	9.51e-13
Trefethen_20000	20000	1000	320.1	5.40e-3	3.00e+0	132	32.0	9.11e-5	5.54e-12	38	11.2	6.15e-5	1.02e-11

The numerical results associated with this second test set are contained in Table 2. On the one hand, we note that the Cayley method does not achieve convergence for any of the instances considered. In particular, we observe that the error in the feasibility of the iterates deteriorates, which possibly occurs due to the numerical instability of the Sherman–Morrison–Woodbury formula used in the matrix inversion of this procedure. In contrast, the geodesic–based gradient method and Algorithm 3.1 achieve the desired precision in the gradient norm for all the instances. Comparing these last two methods, we clearly see that our proposal is more efficient than the Qgeodesic solver, since it obtains local minimum estimates in less computational time than Qgeodesic. In fact, the Qgeodesic method only wins for the instance bcsstk21.

4.2. Symplectic eigenvalues computation via trace minimization. Let $A \in \mathbb{R}^{2n\times 2n}$ be a positive definite matrix. It follows from the Williamson's theorem [36] that there exists a symplectic matrix $V \in Sp(2n)$ such that

$$V^\top AV = \begin{bmatrix} D & 0 \\ 0 & D \end{bmatrix} \tag{4.2}$$

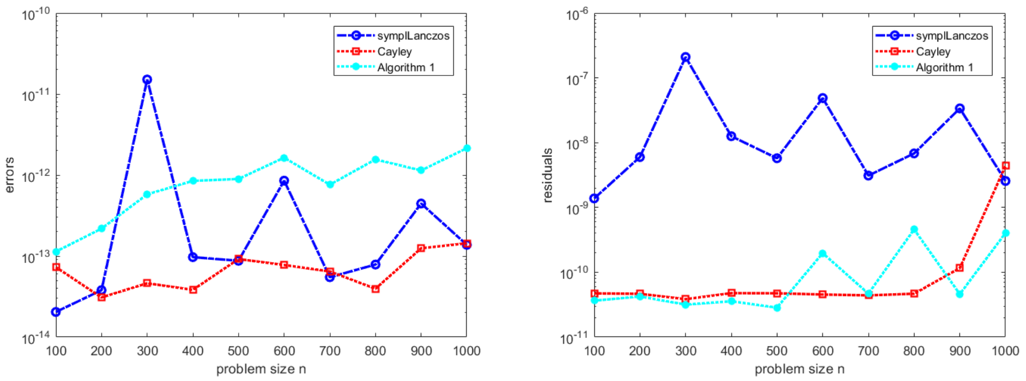
where $D = \mathtt{diag}(d_1, d_2, \ldots, d_n) \in \mathbb{R}^n$ is a diagonal matrix with positive entries. The diagonal entries d_i's are uniquely determined by A and characterise the orbits of $\mathcal{S}_+(2n)$ under the action of the Lie group $Sp(2n)$. These numbers are known as the symplectic eigenvalues of A. This kind of eigenvalues are of great importance in quantum mechanics [11], Hamiltonian dynamics [3, 5, 34], symplectic principal component analysis [21, 29], in symplectic topology [17], and in the more recent subject of quantum information; see e.g., [12, 16]. It is known that the symplectic eigenvalues of A are related to the following constrained optimization problem

$$\min_{X\in\mathbb{R}^{2n\times 2p}} \mathcal{F}(X) := tr(X^\top AX) \quad \text{s.t.} \quad X^\top J_{2n}X = J_{2p}. \tag{4.3}$$

In particular, if X_* is a solution of the trace minimization problem (4.3) then $\mathcal{F}(X_*) = 2\sum_{i=1}^{p} d_i$, this result appears in [33], see Theorem 4.1.

In this subsection, we consider problem (4.3) to evaluate the effectiveness of the following methods: Cayley [15], Algorithm 3.1 and the symplectic Lanczos method developed in [2], denote by *SympLanczos*. In particular, we follow the same design of the experiment reported in [33], Section 6.1. Specifically, the matrix A is given by $A = Q\mathtt{diag}(D, D)Q^\top$, where $D = \mathtt{diag}(d_1, d_2, \ldots, d_n)$ and $Q = KL(n/2, 1.2, -\sqrt{n/5})$, where $L(n/2, 1.2, -\sqrt{n/5}) \in Sp(2n)$ is the symplectic Gauss transformation defined in [13], and $K \in \mathbb{R}^{2n\times 2n}$ is constructed in the same way as in [33]. Observe that by construction of matrix A, its p smallest symplectic eigenvalues are $1, 2, \ldots, p-1, p$. In order to compare the precision of the estimated p smallest symplectic eigenvalues $\tilde{d}_1, \ldots, \tilde{d}_p$ by the algorithms, we compute the absolute error $\sum_{i=1}^{p} = |\tilde{d}_i - i|$ for every method. In addition, we also measure the normalized residual

$$\frac{\left|\left|A\tilde{X}_{1:p} - J_{2n}\tilde{X}_{1:p}\begin{bmatrix} 0 & -\tilde{D}_{1:p} \\ \tilde{D}_{1:p} & 0 \end{bmatrix}\right|\right|_F}{||A\tilde{X}_{1:p}||_F},$$



(a) Absolute error (b) Normalized residual

FIG. 1. *Solving the trace minimization problem* (4.3) *with a matrix A with known symplectic eigenvalues. On the left, the absolute errors of the computed symplectic eigenvalues. On the right, the corresponding normalized residuals.*

TABLE 3
The 5 smallest symplectic eigenvalues of a 2000–by–2000 positive definite matrix A, computed by the three algorithms.

SympLanczos	Cayley	Algorithm 1
1.000000000000028	1.000000000000043	1.000000000000195
1.999999999999998	1.999999999999984	2.000000000000241
3.000000000000045	3.000000000000041	3.000000000000449
3.999999999999971	3.999999999999964	4.000000000000592
5.000000000000033	5.000000000000009	5.000000000000669

where $\tilde{X}_{1:p}$ is the symplectic eigenvector matrix related to the symplectic eigenvalues $\tilde{D}_{1:p} = \texttt{diag}(\tilde{d}_1, \ldots \tilde{d}_p)$ obtained for each method. In this experiments, we set $p = 5$ and solve (4.3) for different values of n in the range between 100 and 1000. For this experiment, we use $N = 5000$ and $\epsilon = $ 1e-9 in the stop criteria of all the algorithms. In Figure 1 we plot the residuals and the errors obtained by the methods for each of the values of n. Additionally, in Table 3, we report the eigenvalues computed by the three methods for $n = 1000$. As shown in Table 3, the three methods obtain estimates of the 5 smallest symplectic eigenvalues very close to the real ones.

5. Final remarks. We have presented a first–order iterative approach for minimizing smooth nonlinear functions defined on the domain of symplectic matrices $Sp(2n, 2p)$. The proposed procedure is a Riemannian gradient method, which uses a new retraction with low computational cost to preserve the feasibility of each point. The designed retraction can be regarded as a generalization of the feasible curve introduced by Jiang and Dai in [20]. The operation that requires the highest computational cost, to evaluate the new retraction, is a matrix inversion of size $2p$–by–$2p$, which is significantly less expensive than the operations required by the existing methods in the literature. In order to improve the numerical performance of the proposal, we consider an adaptive step–size based on the Barzilai–Borwein step–sizes [4]. To guarantee the convergence of the method to critical points of the objective function (in the purely Riemannian sense), we adopt the globalization strategy of Zhang and Hager [38].

The numerical experiments carried out indicate that the new algorithm is suitable for solving large–sale and sparse, as well as small and dense, symplectic Stiefel manifold constrained optimization problems. Moreover, we notice that the our proposal is more efficient than the two Riemannian gradient methods recently developed in [15], solving trace minimization problems and computing the projection of an arbitrary matrix onto the symplectic Stiefel manifold.

Acknowledgements. The first author was financially supported by FGV (Fundação Getulio Vargas) through the excellence post–doctoral fellowship program.

REFERENCES

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization algorithms on matrix manifolds*, Princeton University Press, 2009.

[2] P. Amodio, *A symplectic lanczos-type algorithm to compute the eigenvalues of positive definite hamiltonian matrices*, in International Conference on Computational Science, Springer, 2003, pp. 139–148.

[3] V. I. Arnol, *Mathematical methods of classical mechanics*, Springer-Verlag., 1989.

[4] J. Barzilai and J. M. Borwein, *Two-point step size gradient methods*, IMA journal of numerical analysis, 8 (1988), pp. 141–148.

[5] P. Benner and H. Faβbender, *An implicitly restarted symplectic lanczos method for the hamiltonian eigenvalue problem*, Linear algebra and its applications, 263 (1997), pp. 75–111.

[6] P. Benner, H. Faβbender, and D. S. Watkins, *Sr and sz algorithms for the symplectic (butterfly) eigenproblem*, Linear algebra and its applications, 287 (1999), pp. 41–76.

[7] P. Benner and H. Fassbender, *The symplectic eigenvalue problem, the butterfly form, the sr algorithm, and the lanczos method*, Linear algebra and its applications, 275 (1998), pp. 19–47.

[8] T. A. Davis and Y. Hu, *The university of florida sparse matrix collection*, ACM Transactions on Mathematical Software (TOMS), 38 (2011), pp. 1–25.

[9] N. Del Buono, L. Lopez, and T. Politi, *Computation of functions of hamiltonian and skew-symmetric matrices*, Mathematics and Computers in Simulation, 79 (2008), pp. 1284–1297.

[10] D. Di Serafino, V. Ruggiero, G. Toraldo, and L. Zanni, *On the steplength selection in gradient methods for unconstrained optimization*, Applied Mathematics and Computation, 318 (2018), pp. 176–195.

[11] B. Dutta, N. Mukunda, R. Simon, et al., *The real symplectic groups in quantum mechanics and optics*, Pramana, 45 (1995), pp. 471–497.

[12] J. Eisert, T. Tyc, T. Rudolph, and B. C. Sanders, *Gaussian quantum marginal problem*, Communications in mathematical physics, 280 (2008), pp. 263–280.

[13] H. Fassbender, *The parameterized sr algorithm for symplectic (butterfly) matrices*, Mathematics of computation, 70 (2001), pp. 1515–1541.

[14] S. Fiori, *A riemannian steepest descent approach over the inhomogeneous symplectic group: Application to the averaging of linear optical systems*, Applied Mathematics and Computation, 283 (2016), pp. 251–264.

[15] B. Gao, N. T. Son, P.-A. Absil, and T. Stykel, *Riemannian optimization on the symplectic stiefel manifold*, SIAM Journal on Optimization, 31 (2021), pp. 1546–1575, https://doi.org/https://doi.org/10.1137/20M1348522.

[16] T. Hiroshima, *Additivity and multiplicativity properties of some gaussian channels for gaussian inputs*, Physical Review A, 73 (2006), p. 012330.

[17] H. Hofer and E. Zehnder, *Symplectic invariants and Hamiltonian dynamics*, Birkhäuser, 2012.

[18] J. Hu, X. Liu, Z.-W. Wen, and Y.-X. Yuan, *A brief introduction to manifold optimization*, Journal of the Operations Research Society of China, 8 (2020), pp. 199–248.

[19] B. Iannazzo and M. Porcelli, *The riemannian barzilai–borwein method with nonmonotone line search and the matrix geometric mean computation*, IMA Journal of Numerical Analysis, 38 (2018), pp. 495–517.

[20] B. Jiang and Y.-H. Dai, *A framework of constraint preserving update schemes for optimization on stiefel manifold*, Mathematical Programming, 153 (2015), pp. 535–575.

[21] M. Lei and G. Meng, *Symplectic principal component analysis: A noise reduction method for continuous chaotic systems*, Advances in Noise Analysis, Mitigation and Control, (2016), p. 23.

[22] M. Lei, G. Meng, W. Zhang, J. Wade, and N. Sarkar, *Symplectic entropy as a novel measure for complex systems*, Entropy, 18 (2016), p. 412.

[23] D. G. Luenberger, Y. Ye, et al., *Linear and nonlinear programming*, vol. 2, Springer, 1984.

[24] J. Nocedal and S. Wright, *Numerical optimization*, Springer Science & Business Media, 2006.

[25] H. Oviedo, *Global convergence of riemannian line search methods with a zhang–hager–type condition*, Technical report in http://www.optimization-online.org/DB_HTML/2021/03/8297.html.

[26] H. Oviedo, *Implicit steepest descent algorithm for optimization with orthogonality constraints*, Technical report in http://www.optimization-online.org/DB_HTML/2020/03/7682.html.

[27] H. Oviedo, O. Dalmau, and H. Lara, *Two adaptive scaled gradient projection methods for stiefel manifold constrained optimization*, Numerical Algorithms, (2020), pp. 1–21.

[28] H. Oviedo, H. Lara, and O. Dalmau, *A non-monotone linear search algorithm with mixed direction on stiefel manifold*, Optimization Methods and Software, 34 (2019), pp. 437–457.

[29] L. C. Parra, *Symplectic nonlinear component analysis*, Advances in Neural Information Processing Systems, (1996), pp. 437–443.

[30] L. Peng and K. Mohseni, *Symplectic model reduction of hamiltonian systems*, SIAM Journal on Scientific Computing, 38 (2016), pp. A1–A27.

[31] M. Raydan, *The barzilai and borwein gradient method for the large scale unconstrained minimization problem*, SIAM Journal on Optimization, 7 (1997), pp. 26–33.

[32] A. Salam and E. Al-Aidarous, *Equivalence between modified symplectic gram-schmidt and householder sr algorithms*, BIT Numerical Mathematics, 54 (2014), pp. 283–302.

[33] N. T. Son, P.-A. Absil, B. Gao, and T. Stykel, *Symplectic eigenvalue problem via trace minimization and riemannian optimization*, arXiv preprint arXiv:2101.02618, (2021).

[34] C. Van Loan, *A symplectic method for approximating all the eigenvalues of a hamiltonian matrix*, Linear algebra and its applications, 61 (1984), pp. 233–251.

[35] Z. Wen and W. Yin, *A feasible method for optimization with orthogonality constraints*, Mathematical Programming, 142 (2013), pp. 397–434.

[36] J. Williamson, *On the algebraic problem concerning the normal forms of linear dynamical systems*, American journal of mathematics, 58 (1936), pp. 141–163.
[37] R. Wu, R. Chakrabarti, and H. Rabitz, *Optimal control theory for continuous-variable quantum gates*, Physical Review A, 77 (2008), p. 052303.
[38] H. Zhang and W. W. Hager, *A nonmonotone line search technique and its application to unconstrained optimization*, SIAM journal on Optimization, 14 (2004), pp. 1043–1056.
[39] L. Zhao, *Linear constraint problem of hermitian unitary symplectic matrices*, Linear and Multilinear Algebra, (2020), pp. 1–19.
[40] B. Zhou, L. Gao, and Y.-H. Dai, *Gradient methods with adaptive step-sizes*, Computational Optimization and Applications, 35 (2006), pp. 69–86.