The Extrinsic Geometry of Dynamical Systems Tracking Nonlinear Matrix Projections

SIAM J. MATRIX ANAL. APPL.
Vol. 40, No. 2, pp. 814–844

THE EXTRINSIC GEOMETRY OF DYNAMICAL SYSTEMS TRACKING NONLINEAR MATRIX PROJECTIONS*

FLORIAN FEPPON† AND PIERRE F. J. LERMUSIAUX†

Abstract. A generalization of the concepts of extrinsic curvature and Weingarten endomorphism is introduced to study a class of nonlinear maps over embedded matrix manifolds. These (nonlinear) *oblique projections* generalize (nonlinear) orthogonal projections, i.e., applications mapping a point to its closest neighbor on a matrix manifold. Examples of such maps include the truncated SVD, the polar decomposition, and functions mapping symmetric and nonsymmetric matrices to their linear eigenprojectors. This paper specifically investigates how oblique projections provide their image manifolds with a canonical *extrinsic* differential structure, over which a generalization of the Weingarten identity is available. By diagonalization of the corresponding Weingarten endomorphism, the manifold principal curvatures are explicitly characterized, which then enables us to (i) derive explicit formulas for the differential of oblique projections and (ii) study the global stability of a governing generic ordinary differential equation (ODE) computing their values. This methodology, exploited for the truncated SVD in [Feppon and Lermusiaux, *SIAM J. Matrix Anal. Appl.*, 39 (2018), pp. 510–538], is generalized to non-Euclidean settings and applied to the four other maps mentioned above and their image manifolds: respectively, the Stiefel, the isospectral, and the Grassmann manifolds and the manifold of fixed rank (nonorthogonal) linear projectors. In all cases studied, the oblique projection of a target matrix is surprisingly the unique stable equilibrium point of the above gradient flow. Three numerical applications concerned with ODEs tracking dominant eigenspaces involving possibly multiple eigenvalues finally showcase the results.

Key words. Weingarten map, principal curvatures, polar decomposition, dynamic dominant eigenspaces, Isospectral, normal bundle, Grassmann, and bi-Grassmann manifolds

AMS subject classifications. 65C20, 53B21, 65F30, 15A23, 53A07

DOI. 10.1137/18M1192780

1. Introduction. Continuous time matrix algorithms have been receiving a growing interest in a wide range of applications including data assimilation [43], data processing [66], machine learning [26], and matrix completion [65]. In many applications, a time-dependent matrix $\mathfrak{R}(t)$ is given, for example, in the form of the solution of an ODE, and one is interested in continuous algorithms tracking the value of an algebraic operation $\Pi_{\mathscr{M}}(\mathfrak{R}(t))$: in other words one wants to compute efficiently the update $\Pi_{\mathscr{M}}(\mathfrak{R}(t+\Delta t))$ at a later time $t+\Delta t$ from the knowledge of $\Pi_{\mathscr{M}}(\mathfrak{R}(t))$.

For such a purpose, a large number of works have focused on deriving dynamical systems that, given an input matrix $\mathfrak{R}$, compute an algebraic operation $\Pi_{\mathscr{M}}(\mathfrak{R})$, such as eigenvalues, singular values, or polar decomposition [16, 9, 14, 58, 19, 31]. Typical examples of maps $\Pi_{\mathscr{M}}$ specifically considered in this paper include the following:

1. The truncated singular value decomposition (SVD) mapping an l-by-m matrix $\mathfrak{R}\in\mathcal{M}_{l,m}$ to its best rank r approximation. Denoting $\sigma_1(\mathfrak{R})\geq\sigma_2(\mathfrak{R})\geq\cdots\geq\sigma_{\mathrm{rank}(\mathfrak{R})}>0$ the singular values of $\mathfrak{R}$, and (u_i), (v_i) the corresponding

*Received by the editors June 7, 2018; accepted for publication (in revised form) by P.-A. Absil April 10, 2019; published electronically June 27, 2019.
http://www.siam.org/journals/simax/40-2/M119278.html
Funding: The work of the authors was supported by the Office of Naval Research grants N00014-14-1-0476 (Science of Autonomy–LEARNS) and N00014-14-1-0725 (Bays-DA) to the Massachusetts Institute of Technology.
†MSEAS, Massachusetts Institute of Technology, Cambridge, MA, 02139 (feppon@mit.edu, pierrel@mit.edu).

orthonormal basis of right and left singular vectors, $\Pi_{\mathscr{M}}$ is the map

$$\mathfrak{R} = \sum_{i=1}^{\mathrm{rank}(\mathfrak{R})} \sigma_i(\mathfrak{R}) u_i v_i^T \longmapsto \Pi_{\mathscr{M}}(\mathfrak{R}) = \sum_{i=1}^{r} \sigma_i(\mathfrak{R}) u_i v_i^T. \tag{1}$$

2. Given $p \leq n$, the application mapping a full rank n-by-p matrix $\mathfrak{R} \in \mathcal{M}_{n,p}$ to its polar part, i.e., the unique matrix $P \in \mathcal{M}_{n,p}$ such that $P^T P = I$ and $\mathfrak{R} = PS$ with $S \in \mathrm{Sym}_p$, a symmetric positive definite p-by-p matrix:

$$\mathfrak{R} = \sum_{i=1}^{p} \sigma_i(\mathfrak{R}) u_i v_i^T \longmapsto \Pi_{\mathscr{M}}(\mathfrak{R}) = \sum_{i=1}^{p} u_i v_i^T. \tag{2}$$

3. The application replacing all eigenvalues $\lambda_1(\mathbf{S}) \geq \lambda_2(\mathbf{S}) \geq \cdots \geq \lambda_n(\mathbf{S})$ of a n-by-n symmetric matrix $\mathbf{S} \in \mathrm{Sym}_n$ with a prescribed sequence $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_n$. Denoting (u_i) an orthonormal basis of eigenvectors of $\mathbf{S}$,

$$\mathbf{S} = \sum_{i=1}^{n} \lambda_i(\mathbf{S}) u_i u_i^T \longmapsto \Pi_{\mathscr{M}}(\mathbf{S}) = \sum_{i=1}^{n} \lambda_i u_i u_i^T. \tag{3}$$

4. The application mapping a real n-by-n matrix $\mathfrak{R} \in \mathcal{M}_{n,n}$ to the linear orthogonal projector $UU^T \in \mathcal{M}_{n,n}$ on the p dimensional, dominant invariant subspace of $\mathfrak{R}$ (invariant in the sense $\mathrm{Span}(\mathfrak{R}U) = \mathrm{Span}(U)$). Denote $\lambda_i(\mathfrak{R})$ the eigenvalues of $\mathfrak{R}$ ordered according to their real parts, $\Re(\lambda_1(\mathfrak{R})) \geq \Re(\lambda_2(\mathfrak{R})) \geq \cdots \geq \Re(\lambda_n(\mathfrak{R}))$, and (u_i) and $(\overline{v}_i^T)$ the bases of corresponding right and left eigenvectors. Then, the p dimensional dominant invariant subspace is $\mathrm{Span}(u_i)_{1\leq i\leq p}$, i.e., the space spanned by the p eigenvectors of maximal real parts. $\Pi_{\mathscr{M}}$ is the map

$$\mathfrak{R} = \sum_{i=1}^{n} \lambda_i(\mathfrak{R}) u_i \overline{v}_i^T \longmapsto \Pi_{\mathscr{M}}(\mathfrak{R}) = UU^T \tag{4}$$

for any $U \in \mathcal{M}_{n,p}$ satisfying $\mathrm{Span}(U) = \mathrm{Span}(u_i)_{1\leq i\leq p}$ and $U^T U = I$.

5. The application mapping a real n-by-n matrix $\mathfrak{R}$ to the linear projector whose image is the p dimensional dominant invariant subspace $\mathrm{Span}(u_i)_{1\leq i\leq p}$ and whose kernel is the complement invariant subspace $(\mathrm{Span}(v_i)_{1\leq i\leq p})^{\perp} = \mathrm{Span}(u_{p+j})_{1\leq j\leq n-p}$:

$$\mathfrak{R} = \sum_{i=1}^{n} \lambda_i(\mathfrak{R}) u_i \overline{v}_i^T \longmapsto \Pi_{\mathscr{M}}(\mathfrak{R}) = \sum_{i=1}^{p} u_i \overline{v}_i^T = UV^T \tag{5}$$

for matrices $U, V \in \mathcal{M}_{n,p}$ such that $V^T U = I$ $\mathrm{Span}(U) = \mathrm{Span}(u_i)_{1\leq i\leq p}$, and $\mathrm{Span}(V)^{\perp} = \mathrm{Span}(u_{p+j})_{1\leq j\leq n-p}$.

Dynamical systems tracking the truncated SVD (1) have been used for efficient ensemble forecasting and data assimilation [42, 7] and for realistic large-scale stochastic field variability analysis [44]. Closed form differential systems were initially proposed for such purposes in [40, 60] and further investigated in [22, 23] for dynamic model order reduction of high dimensional matrix ODEs. Tracking the polar decomposition (2) (with $p = n$) has been the interest of works in continuum mechanics [11, 59, 29]. The map (3) has been initially investigated by Brockett [10] and used in adaptive signal filtering [42]. Recently, a dynamical system that computes the map (4) has been found in the fluid mechanics community by Babaee and Sapsis [6]. As for (5), differential systems computing this map in the particular case of $p = 1$ have been proposed by [28] for efficient evaluations of ϵ-pseudospectra of real matrices.

A main contribution of this paper is to develop a unified view for analyzing the above maps $\Pi_{\mathscr{M}}$ and deriving dynamical systems for computing or tracking their values. As we shall detail, each of the above maps $\Pi_{\mathscr{M}} : E \to \mathscr{M}$ can be geometrically interpreted as a (nonlinear) projection from an ambient Euclidean space of matrices E onto a matrix submanifold $\mathscr{M} \subset E$: for instance, $E = \mathcal{M}_{l,m}$ and $\mathscr{M} = \{R \in \mathcal{M}_{l,m} \,|\, \operatorname{rank}(R) = r\}$ is the fixed rank manifold for the map (1). The maps (1)–(3) turn to be *orthogonal projections*, in that $\Pi_{\mathscr{M}}(\mathfrak{R})$ minimizes some Euclidean distance $\|\cdot\|$ from $\mathfrak{R} \in E$ to $\mathscr{M}$:

$$\|\mathfrak{R} - \Pi_{\mathscr{M}}(\mathfrak{R})\| = \min_{R \in \mathscr{M}} \|\mathfrak{R} - R\|. \tag{6}$$

The maps (4) and (5) do not satisfy such a property but still share common mathematical structures; in this paper, we more generally refer to them as (nonlinear) *oblique* projections.

We are concerned with two kinds of ODEs. For a smooth trajectory $\mathfrak{R}(t)$, it is first natural to look for the dynamics satisfied by $\Pi_{\mathscr{M}}(\mathfrak{R}(t))$ itself:

$$\begin{cases} \dot{R} = \dfrac{\mathrm{d}}{\mathrm{d}t}\Pi_{\mathscr{M}}(\mathfrak{R}(t)), \\ R(0) = \Pi_{\mathscr{M}}(\mathfrak{R}(0)). \end{cases} \tag{7}$$

The explicit computation of the right-hand side $\mathrm{d}\Pi_{\mathscr{M}}(\mathfrak{R}(t))/\mathrm{d}t$ requires the differential of $\Pi_{\mathscr{M}}$. In the literature, its expression is most often sought from algebraic manipulations, using, e.g., derivatives of eigenvectors that unavoidably require simplicity assumptions for the eigenvalues [18, 14, 15]. As will be detailed further on, it is, however, possible to show that (1)–(5) are differentiable on the domains where they are nonambiguously defined, including cases with multiple eigenvalues: for example, (1) is differentiable as soon as $\sigma_r(\mathfrak{R}) > \sigma_{r+1}(\mathfrak{R})$, even if $\sigma_i(\mathfrak{R}) = \sigma_j(\mathfrak{R})$ for some $i < j \leq r$ [22]. Differentiating eigenvectors is expected to be an even more difficult strategy for (4), since it includes implicit reorthonormalization of the basis $(u_i)_{1 \leq i \leq p}$.

Second, if a fixed input matrix $\mathfrak{R}$ is given, we shall see that $\Pi_{\mathscr{M}}(\mathfrak{R})$ can be obtained as an asymptotically stable equilibrium of the following dynamical system:

$$\dot{R} = \Pi_{\mathcal{T}(R)}(\mathfrak{R} - R), \tag{8}$$

where $\Pi_{\mathcal{T}(R)} : E \to \mathcal{T}(R)$ is a relevant linear projection operator onto the tangent space $\mathcal{T}(R)$ at $R \in \mathscr{M}$. If $\Pi_{\mathscr{M}}$ is an orthogonal projection, then (8) coincides with a gradient flow solving the minimization problem (6), and hence $R(t)$ converges asymptotically to $\Pi_{\mathscr{M}}(\mathfrak{R})$ for sufficiently close initializations $R(0)$. In the general, oblique case, (8) is not a gradient flow but we shall show that $\Pi_{\mathscr{M}}(\mathfrak{R})$ still remains a stable equilibrium point. A question of practical interest regarding the robustness of (8) lies in determining whether $\Pi_{\mathscr{M}}(\mathfrak{R})$ is globally stable.

In this work, we highlight that both the explicit derivation of (7) and the stability analysis of (8) can be obtained from the spectral decomposition of a single linear operator $L_R(N)$ called the Weingarten map (respectively, in Propositions 5 and 6 below). In the Euclidean case, $L_R(N)$ is a standard object of differential geometry whose eigenvalues $\kappa_i(N)$ characterize the extrinsic curvatures of the embedded manifold $\mathscr{M}$ [39, 62]. The relevance of the Weingarten maps $L_R(N)$ for matrix manifolds in relation to the minimum distance problem (6) has been initially observed by Hendriks and Landsman [32, 33] and later by Absil, Mahony, and Trumpf [3] for computing Riemannian Hessians. In the less standard, non-Euclidean case, it turns out that an oblique projection $\Pi_{\mathscr{M}}$ provides *intrinsically* a differential structure on its image manifold $\mathscr{M}$ and a generalization of $L_R(N)$ sharing analogous properties.

In fact, this paper is an extension of our recent work concerned with the truncated SVD (1): in [22], the eigendecomposition of $L_R(N)$ is computed explicitly for the fixed rank manifold, which yields an explicit expression for (7) and a global stability result for (8). Here, we further investigate explicit spectral decompositions of the Weingarten map for relevant matrix manifolds related to the maps (2)–(5), and we use these to obtain (i) the Fréchet derivatives of the corresponding matrix decompositions as well as (ii) the stability analysis of the dynamical system (8) for computing them. We shall highlight in particular how this unified view sheds new light on some previous convergence results [9, 31, 6, 28] or previous formulas available for the differential of matrix decompositions [11, 18], which become elementary consequences of the curvature analysis of their related manifolds.

The paper is organized as follows. Definitions and properties of abstract oblique projections are stated in section 2 and the differences with the more standard orthogonal case are highlighted. We introduce our generalization of the Weingarten map to non-Euclidean ambient spaces before making explicit the link between its spectral decompositions and the ODEs (7) and (8).

The subsequent three sections then examine the value of the Weingarten maps $L_R(N)$ and of the curvatures $\kappa_i(N)$ more specifically for the each of the four maps above. Sections 3 and 4 are, respectively, concerned with (2) and (3), which are orthogonal projections onto their manifold $\mathscr{M}$. These are, respectively, the Stiefel manifold (the set of n-by-p orthogonal matrices) and the isospectral manifold (the set of symmetric matrices with a prescribed spectrum). The application of Proposition 5 below then allows obtaining explicit expressions for the Fréchet derivatives of the polar decomposition (2) and of the map (3) or equivalently of the projectors over the invariant spaces spanned by a selected number of eigenvectors. The gradient flow (8) for computing these maps is then made explicit, and global convergence is obtained for almost every initial data (located in the right-connected component for (2) with $n = p$). We relate our analysis of the isospectral manifold to the popular Brockett flow introduced in the seminal paper [9] and to some works of Chu and Driessel [13] and Absil and Malick [4].

The non-Euclidean framework is then applied in section 5 in order to study the maps (4) and (5). The image manifold $\mathscr{M}$ of (4) is the set of orthogonal linear rank p linear projectors, which is again the Grassmann manifold, but embedded in $\mathcal{M}_{n,n}$ instead of Sym_n. For (5), $\mathscr{M}$ is the set of rank p linear projectors (not necessarily orthogonal), referred to as "bi-Grassmann" manifold in this paper, since it can also be interpreted as the set of all possible *pairs* of two supplementary p dimensional subspaces. Generalized Weingarten maps and their spectral decompositions are obtained explicitly, yielding fully explicit formulas for their differential. The flow (8) is then derived and again found to admit only $\Pi_{\mathscr{M}}(\mathfrak{R})$ as a locally stable equilibrium point.

Finally, three numerical applications are investigated in section 6. First, we show how gradient flows on the isospectral manifold can be used for tracking symmetric eigenspaces involving possibly eigenvalue crossings. Then we examine a reduced method which generalizes the dynamical low rank or DO method of [40, 22] on the isospectral manifold. This method allows approximating the dynamic of eigenspaces of clustered eigenvalues of a symmetric matrix $\mathbf{S}(t)$. Finally, we discuss the correspondence with iterative algorithms and the loss of global convergence issues for the non-Euclidean setting of the map (5).

Notation used in this paper is summarized in Appendix A. It is important, though, to state here those used for differentials. As in [22], the differential of a smooth function f at the point R belonging to some manifold $\mathscr{M}$ embedded in a space E (this includes $\mathscr{M} = E$) in the direction $X \in T(R)$, is denoted $\mathrm{D}_X f(R)$:

$$\mathrm{D}_X f(R) = \left.\frac{\mathrm{d}}{\mathrm{d}t} f(R(t))\right|_{t=0} = \lim_{\Delta t \to 0} \frac{f(R(t+\Delta t)) - f(R(t))}{\Delta t},$$

where $R(t)$ is a curve of $\mathscr{M}$ such that $R(0) = R$ and $\dot{R}(0) = X$. The differential of a linear projection operator $R \mapsto \Pi_{T(R)}$ at $R \in \mathscr{M}$, in the direction $X \in T(R)$ and applied to $Y \in E$, is denoted $\mathrm{D}\Pi_{T(R)}(X) \cdot Y$:

$$\mathrm{D}\Pi_{T(R)}(X) \cdot Y = \left[\left.\frac{\mathrm{d}}{\mathrm{d}t} \Pi_{\mathcal{T}(R(t))}\right|_{t=0}\right](Y) = \left[\lim_{\Delta t \to 0} \frac{\Pi_{\mathcal{T}(R(t+\Delta t))} - \Pi_{\mathcal{T}(R(t))}}{\Delta t}\right](Y).$$

2. Oblique projections. This section develops oblique projections and their main properties in relation with the differential geometry of their image manifold. A smooth manifold $\mathscr{M} \subset E$ embedded in a finite dimensional vector space E is given, where E is not necessarily assumed to be Euclidean (i.e., equipped with a scalar product).

Definition 1. *An application $\Pi_{\mathscr{M}} : \mathcal{V} \to \mathscr{M}$ defined on an open neighborhood $\mathcal{V}$ such that $\mathcal{M} \subset \mathcal{V} \subset E$ is said to be an* oblique *projection onto $\mathscr{M}$ if at each point $R \in \mathscr{M}$ is attached a (normal) vector space $\mathcal{N}(R) \subset E$ such that the portion of affine subspace $\mathcal{V} \cap (R + \mathcal{N}(R))$ is invariant by $\Pi_{\mathscr{M}}$:*

$$\forall \mathfrak{R} \in \mathcal{V} \text{ such that } \mathfrak{R} = R + N \text{ with } N \in \mathcal{N}(R), \Pi_{\mathscr{M}}(\mathfrak{R}) = R.$$

The concepts of oblique projection, normal space $\mathcal{N}(R)$, and neighborhood $\mathcal{V}$, where $\Pi_{\mathscr{M}}$ is defined, are illustrated on Figure 1. Geometrically, $\Pi_{\mathscr{M}}$ maps all points of the portion of affine subspace $R + \mathcal{N}(R)$ sufficiently close to R onto R. Formally, the bundle of normal spaces $\mathcal{N}(R)$ can be understood as a set of straight "hairs" on the manifolds, and $\Pi_{\mathscr{M}}$ maps a point $\mathfrak{R}$ of the hair to its root R on the manifold. When two (affine) normal spaces intersect (i.e., on the *skeleton* in the Euclidean case; see [17, 22]), there is an ambiguity in the definition of $\Pi_{\mathscr{M}}(\mathfrak{R})$, which explains why the domain where $\Pi_{\mathscr{M}}$ is defined is restricted to a neighborhood $\mathcal{V}$.

In the previous definition, nothing is required regarding the dimension of the normal spaces $\mathcal{N}(R)$. A first elementary but essential remark is that these are necessarily in direct sum with the tangent spaces $\mathcal{T}(R)$. (See [21, Proposition 2.4] for the proof.)

Proposition 2. *If $\Pi_{\mathscr{M}}$ is a differentiable oblique projection, then for any $R \in \mathscr{M}$, the direct sum decomposition $E = \mathcal{T}(R) \oplus \mathcal{N}(R)$ holds, and*

$$\Pi_{\mathcal{T}(R)} : \begin{array}{ccc} \mathfrak{X} & \mapsto & \mathrm{D}_{\mathfrak{X}} \Pi_{\mathscr{M}}(R), \\ E & \to & \mathcal{T}(R) \end{array}$$

is the linear projector whose image is $\mathcal{T}(R)$ and whose kernel is $\mathcal{N}(R)$.

Conversely, it is possible to construct a unique oblique projection $\Pi_{\mathscr{M}}$ associated to any smooth bundle of normal spaces $R \mapsto \mathcal{N}(R)$ satisfying $E = \mathcal{T}(R) \oplus \mathcal{N}(R)$

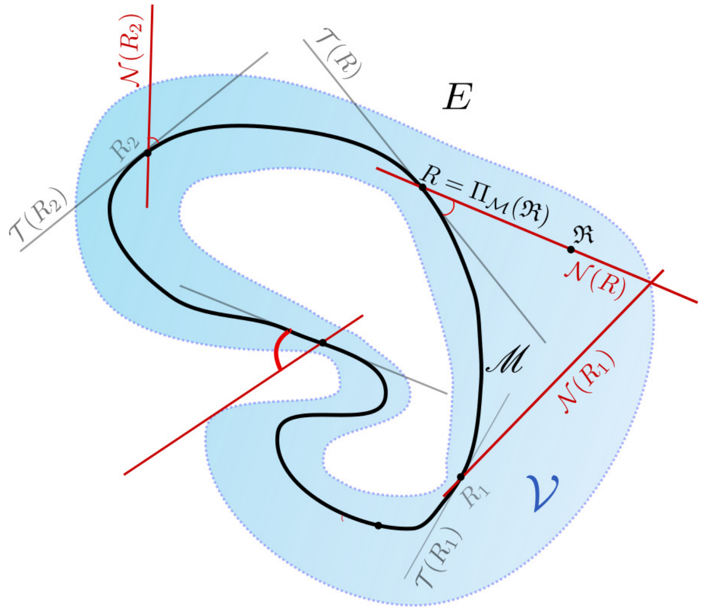


FIG. 1. *An embedded manifold $\mathscr{M} \subset E$ with its tangent spaces $\mathcal{T}(R)$ and a given bundle of normal spaces $\mathcal{N}(R)$. The associated oblique projection $\Pi_{\mathscr{M}}$ defined on an open neighborhood $\mathcal{V}$ maps all points $\mathfrak{R}$ of the portion of the normal affine space (the "hair") $\mathcal{V} \cap (R + \mathcal{N}(R))$ to the point R on the manifold (the "root").*

[21, Proposition 2.7], by using a variant of the tubular neighborhood theorem [8]. The following proposition shows that oblique *linear* projectors $R \mapsto \Pi_{\mathcal{T}(R)}$ onto the decompositions $\mathcal{T}(R) \oplus \mathcal{N}(R)$ define a relevant differential structure on $\mathscr{M}$.

PROPOSITION 3. *Let $\mathscr{M} \subset E$ be an embedded smooth manifold equipped with a differentiable map $R \mapsto \Pi_{\mathcal{T}(R)}$ of linear projectors over the tangent spaces at $\mathscr{M}$. Consider X and Y two differentiable tangent vector fields in a neighborhood of $R \in \mathscr{M}$. Then $\Pi_{\mathcal{T}(R)}$ defines a covariant derivative* [62] *on $\mathscr{M}$ by the formula*

$$\forall X, Y \in \mathcal{T}(R), \quad \nabla_X Y = \Pi_{\mathcal{T}(R)}(\mathrm{D}_X Y). \tag{9}$$

One has the Gauss formula

$$\forall X, Y \in \mathcal{T}(R), \quad \nabla_X Y = \mathrm{D}_X Y + \Gamma(X, Y),$$

where the Christoffel symbol $\Gamma(X, Y)$ depends only on the values of X and Y at R and satisfies

$$\forall X, Y \in \mathcal{T}(R), \quad \Gamma(X, Y) = \Gamma(Y, X) = -\mathrm{D}\Pi_{\mathcal{T}(R)}(X) \cdot Y \in \mathcal{N}(R).$$

For any normal vector $N \in \mathcal{N}(R)$, the Weingarten map *$L_R(N)$ defined by*

$$\forall X \in \mathcal{T}(R), \quad L_R(N)X = \mathrm{D}\Pi_{\mathcal{T}(R)}(X) \cdot N \in \mathcal{T}(R) \tag{10}$$

is a linear application of the tangent space $\mathcal{T}(R)$ into itself.

Proof. These properties, classical for $\Pi_{\mathcal{T}(R)}$ being an orthogonal projection operator [62], are easily obtained for the non-Euclidean case by differentiating $\Pi_{\mathcal{T}(R)}(Y) = Y$ and $\Pi_{\mathcal{T}(R)}(N) = 0$ with respect to X for given tangent and normal vector fields Y and N, and by using the fact that the Lie bracket is a tangent vector. ∎

It is not clear that one can find a Riemannian metric associated with the torsion-free connection ∇ defined from $\Pi_{\mathcal{T}(R)}$ (see [61] about this question), and hence this setting is fundamentally different than the one of fully intrinsic approaches, e.g., [2, 54]. Nevertheless, we find that the duality bracket $\langle \cdot , \cdot \rangle$ over E plays the role of the metric in this embedded setting, as shown in the next proposition. In the following, the dual space of E is denoted E^*, and it is recalled that the adjoint A^* of a linear operator $A : E \to E$ is defined by $\forall v \in E^*, \forall x \in E$, $\langle A^* v, x\rangle = \langle v, Ax\rangle$.

PROPOSITION 4. *For any $R \in \mathscr{M}$, the direct sum $E^* = \mathcal{T}(R)^* \oplus \mathcal{N}(R)^*$ holds where $\mathcal{T}(R)^* = \Pi^*_{\mathcal{T}(R)} E^*$ and $\mathcal{N}(R)^* = (I - \Pi^*_{\mathcal{T}(R)})E^*$. In particular, $\Pi^*_{\mathcal{T}(R)}$ is the linear projector whose image is $\mathcal{T}(R)^*$ and kernel is $\mathcal{N}(R)^*$. The map of projections $R \mapsto \Pi^*_{\mathcal{T}(R)}$ induces a connection over the dual bundle $R \mapsto \mathcal{T}(R)^*$ by the formula*

$$\forall V \in \mathcal{T}(R)^* \text{ and } \forall X \in \mathcal{T}(R), \quad \nabla_X V = \Pi^*_{\mathcal{T}(R)}(\mathrm{D}_X V) \ . \tag{11}$$

The connection ∇ defined by (9) *and* (11) *is compatible with the duality bracket:*

$$\forall X, Y \in \mathcal{T}(R) \text{ and } \forall V \in \mathcal{T}(R)^*, \quad \mathrm{D}_X\langle V, Y\rangle = \langle \nabla_X V, Y\rangle + \langle V, \nabla_X Y\rangle.$$

One has the Gauss formula

$$\forall V \in \mathcal{T}(R)^* \text{ and } \forall X \in \mathcal{T}(R), \quad \nabla_X V = \mathrm{D}_X V + \Gamma(X, V) \ ,$$

where the Christoffel symbol depends only on the value of the tangent vector and dual fields X and V at R and satisfies

$$\forall X \in \mathcal{T}(R) \text{ and } \forall V \in \mathcal{T}(R)^*, \quad \Gamma(X, V) = -\mathrm{D}\Pi^*_{\mathcal{T}(R)}(X) \cdot V \in \mathcal{N}(R)^*.$$

For any normal dual vector $N \in \mathcal{N}(R)^*$, the dual *Weingarten map $L^*_R(N)$ defined by*

$$\forall X \in \mathcal{T}(R), \quad L^*_R(N)X = \mathrm{D}\Pi^*_{\mathcal{T}(R)}(X) \cdot N \tag{12}$$

defines a linear application of the tangent space $\mathcal{T}(R)$ into its dual $\mathcal{T}(R)^*$ and the following Weingarten identities *hold:*

$$\forall X, Y \in \mathcal{T}(R), N \in \mathcal{N}(R)^*, \quad \langle N, \mathrm{D}\Pi_{\mathcal{T}(R)}(X) \cdot Y\rangle = \langle \mathrm{D}\Pi^*_{\mathcal{T}(R)}(X) \cdot N, Y\rangle, \tag{13}$$

$$\forall V \in \mathcal{T}(R)^*, X \in \mathcal{T}(R), N \in \mathcal{N}(R), \ \langle \mathrm{D}\Pi^*_{\mathcal{T}(R)}(X) \cdot V, N\rangle = \langle V, \mathrm{D}\Pi_{\mathcal{T}(R)}(X) \cdot N\rangle. \tag{14}$$

Proof. The proof is a straightforward adaptation of the one of [62, Theorem 8]. For example, the Weingarten identities (13) and (14) are obtained by differentiating the relations $\langle N, X\rangle = 0$ for $N \in \mathcal{N}(R)^*$, $X \in \mathcal{T}(R)$, and $\langle V, N\rangle = 0$ for $V \in \mathcal{T}(R)^*$ and $N \in \mathcal{N}(R)$. ∎

If E is Euclidean, then the dual space E^* can be identified to E by replacing the duality bracket with the scalar product over E. Then Propositions 3 and 4 are redundant and express the classical Euclidean setting [62, 39, 32], where identity (13) states that the Weingarten map $L_R(N) = L^*_R(N)$ (equations (10) and (12)) is symmetric. In that case, it admits an orthonormal basis of real eigenvectors (Φ_i) and eigenvalues $\kappa_i(N)$ called principal directions and principal curvatures in the normal direction $N \in \mathcal{N}(R)$.

The following two propositions state how the Weingarten map $L_R(N)$ (equation (10)) is related to (i) the differential of the oblique projection $\Pi_{\mathscr{M}}$ and (ii) the stability analysis of a dynamical system for which $\Pi_{\mathscr{M}}(\mathfrak{R})$ is a stable equilibrium point.

PROPOSITION 5. *If $\mathscr{M}$ is compact, $\Pi_{\mathscr{M}}$ is continuous, and $R \mapsto \Pi_{\mathcal{T}(R)}$ is differentiable, then there exists an open neighborhood $\mathcal{V} \subset E$ of $\mathscr{M}$ over which $\Pi_{\mathscr{M}}$ is differentiable with $1 \notin sp(L_R(N))$ for any $R \in \mathscr{M}, N \in \mathcal{N}(R)$ such that $R + N \in \mathcal{V}$. The differential of $\Pi_{\mathscr{M}}$ at $\mathfrak{R} = R + N \in \mathcal{V}$ is given by*

$$\mathfrak{X} \mapsto \mathrm{D}_{\mathfrak{X}}\Pi_{\mathscr{M}}(\mathfrak{R}) = (I - L_R(N))^{-1}\Pi_{\mathcal{T}(R)}(\mathfrak{X}). \tag{15}$$

In particular, if $L_R(N)$ is diagonalizable in $\mathbb{C}$, and if one denotes by $\kappa_i(N)$, (Φ_i), and (Φ_i^*), respectively, the eigenvalues, a basis of respective eigenvectors, and its dual basis, then

$$\forall \mathfrak{X} \in E, \quad \mathrm{D}_{\mathfrak{X}}\Pi_{\mathscr{M}}(\mathfrak{R}) = \sum_i \frac{1}{1 - \kappa_i(N)} \langle \Phi_i^*, \Pi_{\mathcal{T}(R)}\mathfrak{X}\rangle \Phi_i. \tag{16}$$

Proof. The proof is done in three steps; see [21, Proposition 2.7 and Theorem 2.2] for details. First, one obtains from the constant rank theorem that $\mathcal{Q} = \{(R, N) \in \mathscr{M} \times E | N \in \mathcal{N}(R)\}$ is a manifold. Then one checks that the differential of the map

$$\begin{array}{cccc} \Phi & : & \mathcal{Q} & \to & E \\ & & (R, N) & \mapsto & R + N \end{array}$$

is invertible provided $1 \notin \mathrm{sp}(I - L_R(N))$. The local inversion theorem and an adaptation of the proof of the tubular neighborhood theorem (see, e.g., [48, Lemma 4]) allows one to obtain that Φ is a diffeomorphism from a neighborhood of $\mathcal{Q}$ onto a neighborhood $\mathcal{V}$ of $\mathscr{M}$ in E. The continuity of $\Pi_{\mathscr{M}}$ implies that $\mathfrak{R} \mapsto (\Pi_{\mathscr{M}}(\mathfrak{R}), \mathfrak{R} - \Pi_{\mathscr{M}}(\mathfrak{R}))$ is the inverse map of this diffeomorphism and hence is differentiable. Finally, (16) follows by differentiating $\Pi_{\mathcal{T}(\Pi_{\mathscr{M}}(\mathfrak{R}))}(\mathfrak{R} - \Pi_{\mathscr{M}}(\mathfrak{R})) = \Pi_{\mathscr{M}}(\mathfrak{R})$ with respect to $\mathfrak{X}$, from which (15) is obtained exactly as in [32, 22]. ∎

Remark 1. Stronger results hold in the Euclidean case, for which $\mathcal{N}(R)$ and $\mathcal{T}(R)$ are mutually orthogonal. In that case, $\Pi_{\mathscr{M}}$ is equivalently defined by the minimization principle (6) at all points $\mathfrak{R} \in E$ yielding a unique minimizer $R = \Pi_{\mathscr{M}}(\mathfrak{R})$ and is automatically continuous on its domain (see [32, 5, 22]).

PROPOSITION 6. *Consider a given $\mathfrak{R} \in E$. The dynamical system*

$$\dot{R} = \Pi_{\mathcal{T}(R)}(\mathfrak{R} - R) \tag{17}$$

satisfies the following properties:

1. *Trajectories of* (17) *lie on the manifold $\mathscr{M}$: $R(0) \in \mathscr{M} \Rightarrow \forall t \geq 0, R(t) \in \mathscr{M}$.*
2. *Equilibrium points of* (17) *are all $R \in \mathscr{M}$ such that $N = \mathfrak{R} - R \in \mathcal{N}(R)$. The linearized dynamics around such equilibria reads*

   $$\dot{X} = (L_R(N) - I)X.$$

   Hence R is stable if $\Re(\kappa_i(N)) < 1$ holds for all eigenvalues $\kappa_i(N)$ of $L_R(N)$.
3. *There exists an open neighborhood $\mathcal{V}$ of $\mathscr{M}$ such that for any $\mathfrak{R} \in \mathcal{V}$, $\Pi_{\mathscr{M}}(\mathfrak{R})$ is a* stable *equilibrium point of* (17).

Proof. Item 1 is a consequence of $\Pi_{\mathcal{T}(R)}(\mathfrak{R} - R) \in \mathcal{T}(R)$. Item 2 is a restatement of $\Pi_{\mathcal{T}(R)}(\mathfrak{R} - R) = 0 \Leftrightarrow \mathfrak{R} - R \in \mathcal{N}(R)$. The linearized dynamics is obtained by differentiating (17) with respect to R along a tangent vector $X \in \mathcal{T}(R)$. Item 3 is a mere consequence of the continuity of the eigenvalues of $L_R(N)$ with respect to N, noticing that for $N = 0$, the linearized dynamics is $\dot{X} = -X$ and hence is stable. ∎

The local stability of the dynamical system (17) is a powerful result, as it yields systematically a continuous time algorithm to find the value of $\Pi_{\mathscr{M}}(\mathfrak{R}+\delta\mathfrak{R})$ given the knowledge of $\Pi_{\mathscr{M}}(\mathfrak{R})$ and a small perturbation $\delta\mathfrak{R}$ (see subsection 6.1 for numerical applications). It may become global if $R = \Pi_{\mathscr{M}}(\mathfrak{R})$ is the only point satisfying

$$N = \mathfrak{R} - R \in \mathcal{N}(R) \text{ and } \Re(\kappa_i(N)) < 1 \text{ for all eigenvalues } \kappa_i(N) \text{ of } L_R(N). \tag{18}$$

Indeed, in that case $\Pi_{\mathscr{M}}(\mathfrak{R})$ is the only asymptotically stable equilibrium among all stationary points of (17). In the Euclidean case, (17) is a gradient descent and Morse theory [36] ensures then the global convergence of the trajectories to $\Pi_{\mathscr{M}}(\mathfrak{R})$ for almost every initialization $R(0)$. In the non-Euclidean case, such a result is not an automatic consequence of (18) and an additional boundedness assumption on $\mathscr{M}$ must hold, as we shall illustrate in section 5.

In the next sections, the above results are utilized to study the maps (2)–(5) and the matrix manifolds they are related to. The Euclidean case is used to study the maps (2) and (3): the relevant manifolds are introduced *first*, and it is shown that (2) and (3) are indeed the applications defined by the minimization principle (6). For the maps (4) and (5), there is no ambient scalar product making the bundle of normal spaces $\mathcal{N}(R)$ orthogonal to the tangent spaces $\mathcal{T}(R)$, and hence the generalization of oblique projections is used: the manifold $\mathscr{M}$ and the decomposition $\mathcal{T}(R)\oplus\mathcal{N}(R)$ are first identified, which allows obtaining the corresponding linear projection $\Pi_{\mathcal{T}(R)}$ and the Weingarten map $L_R(N)$.

3. Stiefel manifold and differentiability of the polar decomposition. In the following, n and p are two given integers satisfying $n \geq 2$ and $p \leq n$. The Stiefel manifold is the set $\mathscr{M}$ of orthonormal n-by-p matrices embedded in $E = \mathcal{M}_{n,p}$:

$$\mathscr{M} = \{U \in \mathcal{M}_{n,p} | U^T U = I\}.$$

The extrinsic geometry of this manifold has been previously studied by a variety of authors [20, 3, 32]. $\mathscr{M}$ is a smooth manifold of dimension $np - p^2 + p(p-1)/2$. Its tangent spaces $\mathcal{T}(U)$ at $U \in \mathscr{M}$ are the sets

$$\begin{aligned}\mathcal{T}(U) &= \{X \in \mathcal{M}_{n,p} | X^T U + X U^T = 0\} \\ &= \{\Delta + U\Omega | \Delta \in \mathcal{M}_{n,p},\, \Omega \in \mathcal{M}_{p,p} \text{ and } \Delta^T U = 0,\, \Omega^T = -\Omega\}.\end{aligned} \tag{19}$$

The orthogonal projection $\Pi_{\mathcal{T}(U)}$ on $\mathcal{T}(U)$ is the map

$$\begin{array}{rccl}\Pi_{\mathcal{T}(U)} & : & \mathcal{M}_{n,p} & \longrightarrow & \mathcal{T}(U), \\ & & \mathfrak{X} & \longmapsto & (I - UU^T)\mathfrak{X} + U\text{skew}(U^T\mathfrak{X}),\end{array}$$

where $\text{skew}(\mathfrak{X}) = (\mathfrak{X} - \mathfrak{X}^T)/2$. The normal space at $U \in \mathscr{M}$ is

$$\mathcal{N}(U) = \{UT | T \in \text{Sym}_p\}.$$

It is well known since Grioli [27, 64, 32, 56] that the map $\Pi_{\mathscr{M}}$ defined by (2) is the nonlinear orthogonal projection operator on the Stiefel manifold $\mathscr{M}$. In other words, if $\mathfrak{R} \in \mathcal{M}_{n,p}$ is a full-rank n-by-p matrix, the matrix $P \in \mathscr{M}$ in the polar decomposition $\mathfrak{R} = PS$ with $S \in \mathcal{M}_{p,p}$ symmetric positive definite minimizes the distance $U \mapsto \|\mathfrak{R} - U\|$ for $U \in \mathscr{M}$ (see also [21, Proposition 2.22] for a geometric proof). The Weingarten map $L_R(N)$ has been computed in [3] and even diagonalized in the case $p = n$ in [32, 33]. The following proposition provides its expression and its spectral decomposition for the general case $p \leq n$. In the following, it is assumed in the notation $X = \Delta + U\Omega \in \mathcal{T}(U)$ (equation (19)) that $\Delta^T U = 0$ and $\Omega^T = -\Omega$.

PROPOSITION 7. *Let $N = UT \in \mathcal{N}(U)$ be a normal vector where the eigenvalue decomposition of the matrix $T \in \mathrm{Sym}_p$ is given by $T = \sum_{i=1}^{p} \lambda_i(T) v_i v_i^T$. The Weingarten map of $\mathscr{M}$ with respect to the normal direction N is the application*

$$
\begin{array}{llll} L_U(N) & : & \mathcal{T}(U) & \longrightarrow \mathcal{T}(U), \\ & & \Delta + U\Omega & \longmapsto -\Delta T - U(\Omega T + T\Omega)/2. \end{array} \tag{20}
$$

The principal curvatures in the direction N are the $p(p-1)/2$ real numbers

$$\forall \{i,j\} \subset \{1,\dots,p\},\ \kappa_{ij}(T) = -\frac{\lambda_i(T) + \lambda_j(T)}{2},$$

associated with the normalized eigenvectors

$$\Phi_{ij} = \frac{U}{\sqrt{2}}(v_i v_j^T - v_j v_i^T),$$

and, if $p < n$, the p real numbers

$$\forall 1 \le i \le p,\ \kappa_i(T) = -\lambda_i(T),$$

associated with the $n-p$ dimensional eigenspaces

$$\{v v_i^T | v \in \mathrm{Span}(U)^\perp\}.$$

Proof. One differentiates $\Pi_{\mathcal{T}(U)}\mathfrak{X}$ with respect to U in the direction $X = \Delta + U\Omega$ before setting $\mathfrak{X} = N$ to obtain

$$
\begin{aligned}
&\mathrm{D}\Pi_{\mathcal{T}(U)}(X)\cdot N \\
&\quad = -2\mathrm{sym}((\Delta + U\Omega)U^T)N + (\Delta + U\Omega)\mathrm{skew}(U^T N) + U\mathrm{skew}((\Delta + U\Omega)^T N) \\
&\quad = -(\Delta U^T + U\Delta^T)N - U\mathrm{skew}(\Omega T),
\end{aligned}
$$

which yields (20) by setting $N = UT$. (This expression also coincides with the one found in [3].) Therefore an eigenvector $X = \Delta + U\Omega \in \mathcal{T}(U)$ of $L_U(N)$ with eigenvalue λ satisfies $-\Delta T = \lambda\Delta$ and $-\frac{1}{2}(\Omega T + T\Omega) = \lambda\Omega$. One then checks that the solutions (Δ, Ω) are $(v v_i^T, 0)$ with v a vector in $\mathrm{Span}(U)^\perp$ and $(0, (v_i v_j^T - v_j v_i^T)/\sqrt{2})$ with the eigenvalues claimed. Because the total dimension formed by these eigenspaces coincides with the dimension of the tangent space, there exist no other eigenvalues. ∎

As a direct application, a fully explicit expression for the differential of the polar decomposition is obtained. The proposition below provides a generalization of the formula initially obtained by Chen and Wheeler [11] and later by Hendriks and Landsman [32] for the particular case of the orthogonal group ($p = n$).

PROPOSITION 8. *Let $\mathfrak{R} = PS$ denote the polar decomposition of a full rank matrix $\mathfrak{R} \in \mathcal{M}_{n,p}$ with $S \in \mathrm{Sym}_p$ positive definite and $P \in \mathscr{M}$, and $S = \sum_{i=1}^{p} \sigma_i(\mathfrak{R}) v_i v_i^T$ the eigendecomposition of S. The orthogonal projection $\Pi_{\mathscr{M}}$, namely, the application $\mathfrak{R} \mapsto P$, is differentiable at $\mathfrak{R}$ and the derivative in the direction $\mathfrak{X}$ is given by the formula*

$$
\begin{aligned}
\mathrm{D}_{\mathfrak{X}}\Pi_{\mathscr{M}}(\mathfrak{R}) = &\sum_{i<j} \frac{2}{\sigma_i(\mathfrak{R}) + \sigma_j(\mathfrak{R})} \left(v_i^T \mathrm{skew}(P^T \mathfrak{X}) v_j\right) P(v_i v_j^T - v_j v_i^T) \\
&+ \sum_{i=1}^{p} \frac{1}{\sigma_i(\mathfrak{R})}(I - PP^T)\mathfrak{X} v_i v_i^T.
\end{aligned} \tag{21}
$$

Proof. The result is immediately obtained by applying (16) with the normal vector $N = \mathfrak{R} - P = P(S - I)$, for which one finds

$$\begin{aligned} 1 - \kappa_i(N) &= 1 - (1 - \sigma_i(\mathfrak{R})) = \sigma_i(\mathfrak{R}), \\ 1 - \kappa_{ij}(N) &= 1 - \big(1 - (\sigma_i(\mathfrak{R}) + \sigma_j(\mathfrak{R}))/2\big) = (\sigma_i(\mathfrak{R}) + \sigma_j(\mathfrak{R}))/2, \\ \langle \Phi_{ij}, \mathfrak{X}\rangle \Phi_{ij} &= \langle P\mathrm{skew}(v_i v_j^T), \mathfrak{X}\rangle P(v_i v_j^T - v_j v_i^T) \\ &= \langle v_i v_j^T, \mathrm{skew}(P^T\mathfrak{X})\rangle P(v_i v_j^T - v_j v_i^T), \end{aligned}$$

and denoting (e_j) an orthonormal basis of $\mathrm{Span}(P)^\perp$,

$$\sum_j \langle \mathfrak{X}, e_j v_i^T\rangle e_j v_i^T = \sum_j \big((e_j^T \mathfrak{X} v_i) e_j\big) v_i^T = (I - PP^T)\mathfrak{X} v_i v_i^T.$$

∎

Remark 2. The derivative (21) has already been obtained in some previous works (e.g., [18, equation (2.19)] or [35, equation (10.2.7)]), albeit in a less explicit form featuring the solution Ω of the Sylvester equation $(S \oplus S)\Omega = S\Omega + \Omega S = P^T\mathfrak{X} - \mathfrak{X}^T P$. Let us observe that the operator $S \oplus S$ is enclosed in the Weingarten map (20), and its explicit inverse is found from the spectral decomposition of Proposition 7.

Applying Proposition 6, we obtain a dynamical system that achieves the polar decomposition and that satisfies global convergence for the gradient descent. Other dynamical systems satisfying related properties can also be found in [24] (without global convergence) and [31] (for $p = n$ with a gradient flow on the larger manifold $\mathscr{M} \times S_n$).

PROPOSITION 9. *Consider a full rank matrix $\mathfrak{R} \in \mathcal{M}_{n,p}$ whose singular value decomposition is written $\mathfrak{R} = \sum_{i=1}^p \sigma_i(\mathfrak{R}) u_i v_i^T$.*

- *If $p < n$, then $\Pi_{\mathscr{M}}(\mathfrak{R})$ is the unique local minimum of the distance function $J : U \mapsto \frac{1}{2}\|\mathfrak{R} - U\|^2$, and therefore, for almost any initial data $U(0) \in \mathscr{M}$, the solution $U(t)$ of the gradient flow*

$$\dot{U} = \mathfrak{R} - \frac{1}{2}(UU^T\mathfrak{R} + U\mathfrak{R}^T U) \tag{22}$$

  converges to the polar part $\Pi_{\mathscr{M}}(\mathfrak{R}) = \sum_{i=1}^p u_i v_i^T$ of $\mathfrak{R}$.
- *If $n = p$, then J admits other local minima that are the matrices*

$$U = \sum_{i=1}^{n-1} u_i v_i^T - u_n v_n^T \in \mathscr{M}, \tag{23}$$

  where u_n is an arbitrary singular vector corresponding to the smallest singular value $\sigma_n(\mathfrak{R})$. Therefore any solution $U(t)$ of the gradient flow (22) converges almost surely to the polar part $\Pi_{\mathscr{M}}(\mathfrak{R})$ provided the initial data $U(0)$ lies in the same connected component of $\mathcal{O}_n$. Otherwise, $U(t)$ converges almost surely toward an element $U \in \mathcal{O}_n$ of the form (23).

Proof. A necessary condition for $U \in \mathscr{M}$ to be a minimizer is that $N = \mathfrak{R} - U \in \mathcal{N}(U)$, i.e., $N = UT$ with $T \in \mathrm{Sym}_p$. Then $\mathfrak{R} = (I + U)T$ and the eigenvalues of T satisfy $\lambda_i(T) = \sigma_i(\mathfrak{R}) - 1$ or $\lambda_i(T) = -(\sigma_i(\mathfrak{R}) + 1)$. If $p < n$, then the condition, $\forall 1 \le i \le p$, $\kappa_i(N) = -\lambda_i(T) \le 1$, required for U to be a local minimum cannot be satisfied if there exists i such that $\lambda_i(T) = -(\sigma_i(\mathfrak{R}) + 1)$. This proves that the only local minimum is achieved by $\Pi_{\mathscr{M}}(\mathfrak{R})$.

If $p = n$, then this condition reads $\kappa_{ij}(N) = -\frac{1}{2}(\lambda_i(T) + \lambda_j(T)) \leq 1$ for all pairs $\{i, j\}$, which cannot be satisfied if there exists at least two indices i and j such that $\lambda_i(T) = -(\sigma_i(\mathfrak{R}) + 1)$ and $\lambda_j(T) = -(\sigma_j(\mathfrak{R}) + 1)$. If i is an index such that $\lambda_i(T) = -(\sigma_i(\mathfrak{R}) + 1)$, then $\kappa_{ij}(N) \leq 1$ implies $\forall j \neq i$, $\sigma_i(\mathfrak{R}) \leq \sigma_j(\mathfrak{R})$ and therefore $i = n$ and U is of the form (23). Finally, the gradient flow is obtained by making (8) explicit with $\Pi_{\mathcal{T}(U)}(\mathfrak{R} - U) = (I - UU^T)(U - \mathfrak{R}) + U\text{skew}(U^T(U - \mathfrak{R}))$. ∎

4. The isospectral manifold, the Grassmannian, and the geometry of mutually orthogonal subspaces.

This section now considers the map (3) and its image manifold which is the set of symmetric matrices $S \in \text{Sym}_n$ having m prescribed eigenvalues $\lambda_1 > \cdots > \lambda_m$ with multiplicities $n_1, \ldots, n_m$. The set of such symmetric matrices has been called an isospectral or "spectral" manifold by [9, 14, 13, 4]. Denoting Λ a reference matrix with such a spectrum, the isospectral manifold $\mathscr{M}$ admits the following parameterizations:

$$\begin{aligned} \mathscr{M} &= \{P\Lambda P^T | P \in \mathcal{O}_n\} \\ &= \left\{ \sum_{i=1}^{m} \lambda_i U_i U_i^T | U_i \in \mathcal{M}_{n,n_i},\, U_i^T U_j = \delta_{ij} \right\}. \end{aligned} \tag{24}$$

There is a motivation for examining $\mathscr{M}$ in its own right: identifying the linear eigenprojector $U_i U_i^T$ with the eigenspace $V_i = \text{Span}(U_i)$, the isospectral manifold models the set of all collections $V_1, V_2, \ldots, V_m \subset V$ of m subspaces of a n dimensional Euclidean space V with prescribed dimensions $n_1, \ldots, n_m$, and orthogonal to each other ($n = n_1 + \cdots + n_m$ and $V_1 \oplus \cdots \oplus V_m = V$). This allows one to include in this analysis an embedded definition of the Grassmann manifold (the set of all p dimensional subspace embedded in a n dimensional space), as the set $\{UU^T \in \text{Sym}_n | U^T U = I \text{ and } U \in \mathcal{M}_{n,p}\}$ of all rank p orthogonal linear projectors. This approach, which has also been favored by some other authors [53, 30], stands in contrast with the more usual intrinsic definitions of the Grassmannian via quotient manifolds [20, 1, 2].

In the following, the set of m matrices $U_i \in \mathcal{M}_{n,n_i}$ is used to describe points on the manifold $\mathscr{M}$, where each U_i represents the eigenspace $\text{Span}(U_i)$. The time derivative of a trajectory $U_i(t)$ can be decomposed along the basis given by the union of the U_k as $\dot{U}_i = \sum_{j=1}^m U_j \Delta_i^j$ with $\Delta_i^j \in \mathcal{M}_{n_j,n_i}$. The matrix Δ_i^j can be interpreted as the magnitude of the rotation of the subspace $\text{Span}(U_i)$ around the axis given by the subspace $\text{Span}(U_j)$. To remain orthogonal to one another, the antisymmetry condition $\Delta_i^j = -(\Delta_i^j)^T$ must be satisfied by the Δ_i^j.

Proposition 10. *The tangent space $\mathcal{T}(S)$ at $S \in \mathscr{M}$ is the set*

$$\begin{aligned} \mathcal{T}(S) &= \{[\Omega, S] = \Omega S - S\Omega | \Omega \in \mathcal{M}_{n,n},\, \Omega^T = -\Omega\} \\ &= \left\{ \sum_{i \neq j} (\lambda_i - \lambda_j) U_j \Delta_i^j U_i^T \middle| \Delta_i^j \in \mathcal{M}_{n_j,n_i},\, \Delta_j^i = -(\Delta_i^j)^T \right\}. \end{aligned} \tag{25}$$

The Δ_i^j defined in the above expressions for each pair $\{i, j\} \subset \{1, \ldots, m\}$ parameterize uniquely the tangent space $\mathcal{T}(S)$. Therefore $\mathscr{M}$ is a smooth manifold of dimension $(n^2 - \sum_{i=1}^m n_i^2)/2$.

Proof. The first equality and the dimension of $\mathscr{M}$ can be found in [14]. Consider $S = \sum_{i=1}^m \lambda_i U_i U_i^T$ with $U_i \in \mathcal{M}_{n,n_i}$ and $U_i^T U_j = \delta_{ij}$. Differentiating the constraint

$U_i^T U_j = 0$ yields $(\Delta_i^j)^T = -\Delta_j^i$, which implies that tangent vectors are of the form $X = \sum_{i,j=1}^m \lambda_i (U_j \Delta_i^j U_i^T - U_i \Delta_j^i U_j^T)$, giving the other equality. Finally, if $X \in \mathcal{T}(S)$, the formula $\Delta_i^j = U_j^T X U_i/(\lambda_i - \lambda_j)$ determines uniquely Δ_i^j. ◻

By analogy to the vocabulary of quotient manifolds [2], we call *horizontal space* the set $\mathcal{H} = \{(\Delta_i^j)_{i,j} | \Delta_i^j \in \mathcal{M}_{n_j,n_i}, \Delta_i^j = -(\Delta_i^j)^T\}$ as it parameterizes uniquely $\mathcal{T}(S)$. In the following, one denotes $(\Delta_X)_i^j \in \mathcal{H}$ the coordinates of a tangent vector $X \in \mathcal{T}(S)$.

PROPOSITION 11. *The projection $\Pi_{\mathcal{T}(S)}$ on the tangent space $\mathcal{T}(S)$ is the map*

$$\begin{array}{rcll} \Pi_{\mathcal{T}(S)} & : & \mathrm{Sym}_n & \longrightarrow & \mathcal{T}(S), \\ & & \mathfrak{X} & \longmapsto & \sum_{\{i,j\}\subset\{1,\dots,m\}} (U_j U_j^T \mathfrak{X} U_i U_i^T + U_i U_i^T \mathfrak{X} U_j U_j^T), \end{array}$$

that is, with the coordinates of the horizontal space, $\Delta_i^j = U_j^T \mathfrak{X} U_i/(\lambda_i - \lambda_j)$. Therefore the normal space $\mathcal{N}(S)$ at S is the set of all symmetric matrices N that let invariant each eigenspace $\mathrm{Span}(U_i)$ *of S:*

$$\mathcal{N}(S) = \left\{ \sum_{i=1}^m U_i U_i^T \mathfrak{X} U_i U_i^T | \mathfrak{X} \in \mathrm{Sym}_n \right\}.$$

In other words, it is the set of all matrices $N \in \mathrm{Sym}_n$ of the form

$$N = \sum_{i=1}^m \sum_{a=1}^{n_i} \lambda_{i,a}(N) u_{i,a} u_{i,a}^T, \tag{26}$$

where for each $1 \le i \le m$, $\lambda_{i,a}(N)_{1\le a \le n_i}$ is a set of n_i real eigenvalues associated with n_i eigenvectors $(u_{i,a})_{1 \le a \le n_i}$ forming a basis of the eigenspace $\mathrm{Span}(U_i)$.

Proof. This is obtained by differentiating $\|\mathfrak{X} - X\|^2$ with respect to Δ_i^j for a tangent vector $X \in \mathcal{H}$ written with the coordinates Δ_i^j of the horizontal space. The normal space is obtained from the equality $\mathcal{N}(S) = \{(I - \Pi_{\mathcal{T}(S)})\mathfrak{X} | \mathfrak{X} \in \mathrm{Sym}_n\}$. ◻

Absil and Malick proved that $\Pi_{\mathscr{M}}(\mathbf{S})$ as defined by (3) is the orthogonal projection operator on $\mathscr{M}$ for matrices $\mathbf{S}$ in a small neighborhood around $\mathscr{M}$ [4, Theorem 3.9]. We propose below a short proof showing this result holds in fact for almost any $\mathbf{S} \in \mathrm{Sym}_n$ using the above geometric analysis of normal and tangent spaces.

PROPOSITION 12. *Let $\mathbf{S} \in \mathrm{Sym}_n$ and denote $\mathbf{S} = \sum_{i=1}^m \sum_{a=1}^{n_i} \lambda_{i,a}(\mathbf{S}) u_{i,a} u_{i,a}^T$ its eigenvalue decomposition, where the eigenvalues have been ordered decreasingly, i.e.,*

$$\forall 1 \le a_i \le n_i, \lambda_{1,a_1} \ge \lambda_{2,a_2} \ge \cdots \ge \lambda_{m,a_m},$$
$$\forall 1 \le i \le m, \lambda_{i,1} \ge \lambda_{i,2} \ge \cdots \ge \lambda_{i,n_i}.$$

If for any $1 \le i \le m-1$, $\lambda_{i+1,1}(\mathbf{S}) > \lambda_{i,n_i}(\mathbf{S})$, i.e., if the eigenspaces of $\mathbf{S}$ are well separated relative to the ordering given by Λ, then the matrix $\Pi_{\mathscr{M}}(\mathbf{S})$ obtained by replacing the eigenvalues of $\mathbf{S}$ by those of Λ in the same order,

$$\Pi_{\mathscr{M}}(\mathbf{S}) = \sum_{i=1}^m \sum_{a=1}^{n_i} \lambda_i u_{i,a} u_{i,a}^T,$$

minimizes the distance $S \mapsto \|\mathbf{S} - S\|$ from $\mathbf{S}$ to $\mathscr{M}$. Furthermore, the minimum distance is given by $\|\mathbf{S} - \Pi_{\mathscr{M}}(\mathbf{S})\|^2 = \sum_{i=1}^m \sum_{a=1}^{n_i} (\lambda_{i,a} - \lambda_i)^2$.

Proof. For a given $S \in \mathscr{M}$, $N = \mathbf{S} - S$ is a normal vector at S if $\mathbf{S} = S + N$, where S and N can be diagonalized by a similar orthonormal basis. Denoting by $\mathbf{S} = \sum_{l=1}^{n} \lambda_l(\mathbf{S}) u_l u_l^T$ the eigendecomposition of $\mathbf{S}$ (no ordering assumed), one has $N = \sum_{l=1}^{n} (\lambda_l(\mathbf{S}) - \Lambda_{\sigma(l)}) u_l u_l^T$, where the $\Lambda_1 \geq \Lambda_2 \geq \cdots \geq \Lambda_n$ are the eigenvalues λ_i of Λ (including multiplicities), and σ a permutation. Since for any given numbers satisfying $a < b$ and $c < d$, $(a-c)^2 + (b-d)^2 < (a-d)^2 + (b-c)^2$ holds, the norm of N is minimized by selecting the permutation σ to be the identity. ∎

The Weingarten map and its spectral decomposition are now explicitly derived. This allows one to obtain an explicit formula for the differential of the map (4) in Proposition 14 and Corollary 15, and the global convergence property of gradient flows associated with the minimization principle (6) in Proposition 16.

PROPOSITION 13. *Let $N = \sum_{i=1}^{m} \sum_{a=1}^{n_i} \lambda_{i,a}(N) u_{i,a} u_{i,a}^T \in \mathcal{N}(S)$ be a normal vector decomposed as in* (26)*. The Weingarten map at $S \in \mathscr{M}$ in the direction N is*

$$\begin{array}{rcccc} L_S(N) & : & \mathcal{H} & \longrightarrow & \mathcal{H}, \\ & & (\Delta_i^j) & \longmapsto & \left(\frac{1}{\lambda_i - \lambda_j} (U_j^T N U_j \Delta_i^j - \Delta_i^j U_i^T N U_i) \right)_i^j . \end{array} \tag{27}$$

The principal curvatures are the real numbers

$$\kappa_{i,a}^{j,b} = \frac{\lambda_{j,b}(N) - \lambda_{i,a}(N)}{\lambda_i - \lambda_j}$$

for all pairs $\{i,j\} \subset \{1, \ldots, m\}$ and couples (a,b) with $1 \leq a \leq n_i$ and $1 \leq b \leq n_j$. Corresponding normalized eigendirections are the tangent vectors

$$\Phi_{i,a}^{j,b} = \frac{1}{\sqrt{2}} (u_{i,a} u_{j,b}^T + u_{j,b} u_{i,a}^T).$$

Proof. Differentiating $\Pi_{\mathcal{T}(S)} N$ with respect to $\Delta_i^j \in \mathcal{H}$ yields

$$\mathrm{D}\Pi_{\mathcal{T}(S)}(X) \cdot N = \sum_{i \neq j} \left[(U_k \Delta_j^k U_j^T - U_j \Delta_k^j U_k^T) N U_i U_i^T + U_j U_j^T N (U_k \Delta_i^k U_i^T - U_i \Delta_k^i U_k^T) \right]$$

with summation over repeated indices k. The fact that N is a normal vector implies

$$\mathrm{D}\Pi_{\mathcal{T}(S)}(X) \cdot N = \sum_{i \neq j} \left[- U_j \Delta_i^j U_i^T N U_i U_i^T + U_j U_j^T N U_j \Delta_i^j U_i^T \right].$$

Expression (27) follows from $(\Delta_{\mathrm{D}\Pi_{\mathcal{T}(S)}(X) \cdot N})_i^j = U_j^T (\mathrm{D}\Pi_{\mathcal{T}(S)}(X) \cdot N) U_i / (\lambda_i - \lambda_j)$. One checks that $\Delta_{i,a}^{j,b} = U_j^T u_{j,b} u_{i,a}^T U_i$ is a basis of eigenvectors with eigenvalues $\kappa_{i,a}^{j,b}$. ∎

PROPOSITION 14. *Let $\mathbf{S} \in \mathrm{Sym}_n$ be a symmetric matrix satisfying the conditions of Proposition* 12*. The projection onto $\mathscr{M}$ is differentiable at $\mathbf{S}$ and the derivative in a direction $\mathfrak{X} \in \mathrm{Sym}_n$ is given by*

$$\mathrm{D}_{\mathfrak{X}} \Pi_{\mathscr{M}}(\mathbf{S}) = \sum_{\substack{\{i,j\} \subset \{1, \ldots, m\} \\ 1 \leq a \leq n_i \\ 1 \leq b \leq n_j}} \frac{\lambda_i - \lambda_j}{\lambda_{i,a}(\mathbf{S}) - \lambda_{j,b}(\mathbf{S})} (u_{i,a}^T \mathfrak{X} u_{j,b})(u_{i,a} u_{j,b}^T + u_{j,b} u_{i,a}^T). \tag{28}$$

Proof. One applies Proposition 5 with $N=\sum_{i=1}^m\sum_{a=1}^{n_i}(\lambda_{i,a}(\mathbf{S})-\lambda_i)u_{i,a}u_{i,a}^T$. The expression claimed is found from the equalities

$$1-\kappa_{i,a}^{j,b}(N)=\frac{\lambda_{j,b}(\mathbf{S})-\lambda_{i,a}(\mathbf{S})}{\lambda_i-\lambda_j},$$
$$\langle\mathfrak{X},\Phi_{i,a}^{j,b}\rangle\Phi_{i,a}^{j,b}=\frac{1}{2}\langle\mathfrak{X},2\,\mathrm{sym}(u_{i,a}u_{j,b}^T)\rangle(u_{i,a}u_{j,b}^T+u_{j,b}u_{i,a}^T).\qquad\square$$

As an application, a formula is found for the derivative of the subspace spanned by the first $p<n$ eigenvectors $\mathrm{Span}(u_i)_{1\le i\le p}$ of a time-dependent matrix $\mathbf{S}(t)\in\mathrm{Sym}_n$. As far as we know, this formula is original in the sense that no simplicity assumption is required for the eigenvalues (only the condition $\lambda_p(t)<\lambda_{p+1}(t)$), although the reader may find related results using resolvents in [37]. It is remarkable that a smooth evolution of $\mathrm{Span}(u_i(t))$ is obtained as long as the eigenvalues of order p and $p+1$ do not cross, although crossing of eigenvalues (and hence discontinuities of the eigenvectors $u_i(t)$ themselves) may occur within the eigenspace.

COROLLARY 15. *Consider* $\mathbf{S}(t)=\sum_{i=1}^n\lambda_i(t)u_i(t)u_i(t)^T$ *the eigendecomposition of a smoothly varying symmetric matrix. Let* $p<n$ *and assume* $\lambda_p(t)<\lambda_{p+1}(t)$. *Then the projector* $\sum_{i=1}^p u_i(t)u_i(t)^T$ *over* $\mathrm{Span}(u_i(t))_{1\le i\le p}$ *is differentiable, and an ODE for the evolution of an orthonormal basis* $U(t)\in\mathcal{M}_{n,p}$ *satisfying* $\mathrm{Span}(U(t))=\mathrm{Span}(u_i(t))_{1\le i\le p}$ *is*

$$\dot U=\sum_{\substack{1\le i\le p\\ p+1\le j\le n}}\frac{1}{\lambda_i(t)-\lambda_j(t)}(u_i^T\dot{\mathbf{S}}u_j)u_ju_i^TU. \tag{29}$$

Proof. This is immediately obtained by applying (28) to the particular case where $\mathscr{M}$ is the Grassmann manifold, i.e., with $m=2$, $\lambda_1=1$, and $\lambda_2=0$. $\square$

The dynamical system (8) that finds the dominant subspaces of a symmetric matrix or equivalently computes the map (3) is now provided.

PROPOSITION 16. *Consider* $\mathbf{S}=\sum_{i=1}^m\sum_{a=1}^{n_i}\lambda_{i,a}(\mathbf{S})u_{i,a}u_{i,a}^T\in\mathrm{Sym}_n$ *satisfying the conditions of Proposition* 18. *The distance functional* $S\mapsto\|\mathbf{S}-S\|^2$ *admits no local minimum on* $\mathscr{M}$ *other than* $\Pi_{\mathscr{M}}(\mathbf{S})$. *Therefore, for almost any initial data* $S(0)\in\mathrm{Sym}_n$, *the solution* $S(t)=\sum_{i=1}^m\lambda_iU_iU_i^T$ *of the gradient flow*

$$\dot U_i=\sum_{j\neq i}\frac{1}{\lambda_i-\lambda_j}U_jU_j^T\mathbf{S}U_i \tag{30}$$

converges to $\Pi_{\mathscr{M}}(\mathbf{S})$, *or in other words, each of the matrices* $U_i(t)$ *converges to a matrix spanning the same subspace as* $\mathrm{Span}(u_{i,a})_{1\le a\le n_i}$.

Proof. Denote $N=\mathbf{S}-S$ the residual normal vector of a critical point S of the distance functional. The condition for S to be a local minimum is that all curvatures in the direction N satisfy $\kappa_{i,a}^{j,b}(N)\le 1$, which is equivalent to the condition $\frac{\lambda_{i,a}-\lambda_{j,b}}{\lambda_i-\lambda_j}\ge 0$. This condition can be satisfied only for $S=\Pi_{\mathscr{M}}(\mathbf{S})$. $\square$

Remark 3. Proposition 16 is a reformulation and an improvement of the convergence result for the Brockett flow $\dot H=[H,[H,\mathbf{S}]]$ as introduced in the seminal paper [9], where the eigenvalues of $\mathbf{S}$ were assumed distinct. The Brockett flow is a gradient descent for the functional $J(P)=\|P\Lambda P^T-\mathbf{S}\|^2$ with respect to $P\in\mathcal{O}_n$ [9, 14]. The corresponding expression in horizontal coordinates is $\dot U_i=\sum_{j\neq i}(\lambda_i-\lambda_j)U_jU_j^T\mathbf{S}U_i$, hence a rescaling of (30) by multiplication of all components of the covariant gradient by the positive numbers $(\lambda_i - \lambda_j)^2$.

Applying this result to the particular case of the Grassmann manifold, one obtains that for almost any initial data $U(0) \in \mathcal{M}_{n,p}$ with $U(0)^T U(0) = I$, the solution $U(t)$ of the gradient flow $\dot{U} = (I - UU^T)\mathbf{S}U$ converges to a matrix U whose columns spans the p dimensional dominant subspace of $\mathbf{S}$. In fact, Babaee and Sapsis have found that this result still holds for the general case of real matrices: the limit U then spans the p dimensional subspace spanned by the eigenvectors associated with the p eigenvalues of maximal real parts [6, Theorem 2.3]. In the next section, it is shown how this result is related to the map (4) and can in fact be obtained and generalized within the framework of oblique projections of section 2.

5. Non-Euclidean Grassmannian, bi-Grassmann manifold, and derivatives of eigenspaces of nonsymmetric matrices. This section focuses on the maps (4) and (5), mapping n-by-n matrices $\mathfrak{R} \in \mathcal{M}_{n,n}$ to, respectively, the orthogonal projector UU^T onto the dominant p dimensional invariant subspace and to the linear projector UV^T onto this subspace whose kernel is the complementary invariant subspace. These two maps are studied, respectively, in subsections 5.1 and 5.2. Following section 2, the normal bundles $\mathcal{N}(R)$ and the respective oblique linear tangent projectors $\Pi_{\mathcal{T}(R)}$ are first identified, which allows one to compute Weingarten maps. In both cases, we are able to provide explicit (complex) spectral decompositions and therefore the differential of these maps and the stability analysis of the dynamical system (17) to compute them. For the map (4), this dynamical system is found to coincide with the one introduced by Babaee and Sapsis [6], and we retrieve a direct proof of their result as a corollary of the computation of (complex) extrinsic curvatures.

Let us note that different dynamical system approaches have also been proposed for solving related nonsymmetric eigenvalue problems [19, 34, 28, 12].

5.1. Oblique projection on the Grassmann manifold. The image manifold of (4) is the Grassmann manifold

$$\mathscr{M} = \{UU^T \in \mathcal{M}_{n,n} | U \in \mathcal{M}_{n,p} \text{ and } U^T U = I\},$$

embedded this time in $\mathcal{M}_{n,n}$ instead of Sym_n. From the previous section (equation (19)), its tangent spaces are given by $\mathcal{T}(UU^T) = \{U\Delta^T + \Delta U^T | \Delta \in \mathcal{M}_{n,p},\, U^T\Delta = 0\}$. Following section 2, the first step to study (4) is to characterize the normal bundle $\mathcal{N}(UU^T)$ of the candidate oblique projection $\Pi_{\mathscr{M}}$.

PROPOSITION 17. *$\Pi_{\mathscr{M}}$ defined as in (4) is an oblique projection on $\mathscr{M}$, and the respective normal space $\mathcal{N}(UU^T)$ at $UU^T \in \mathscr{M}$ is the set of matrices $N \in \mathcal{M}_{n,n}$ under which the subspace* $\mathrm{Span}(U)$ *is invariant:*

$$\begin{aligned}\mathcal{N}(UU^T) &= \{N \in \mathcal{M}_{n,n} | \mathrm{Span}(NU) \subset \mathrm{Span}(U)\} \\ &= \{N \in \mathcal{M}_{n,n} | (I - UU^T)NUU^T = 0\}.\end{aligned}$$

Proof. One checks the conditions of Definition 1. The continuity of the eigenvalues of a matrix implies that $\Pi_{\mathscr{M}}$ is unambiguously defined on an open neighborhood $\mathcal{V} \subset \mathcal{M}_{n,n}$ containing $\mathscr{M}$. It is clear that $\Pi_{\mathscr{M}}(UU^T) = UU^T$ and that if $N \in \mathcal{V}$ satisfies $\Pi_{\mathscr{M}}(UU^T + N) = UU^T$, then the subspace spanned by U must be invariant by $N = (N + UU^T) - UU^T$. Reciprocally consider $N \in \mathcal{N}(UU^T)$ and denote $(\lambda_i(N))_{1 \le i \le n}$ the eigenvalues of N where the first p are associated with the invariant subspace $\mathrm{Span}(U)$. $\mathrm{Span}(U)$ is invariant by $N + UU^T$, with eigenvalues $\lambda_i(N) + 1$ for $1 \le i \le p$, while $\mathrm{Span}(U)^\perp$ is invariant by $N^T + UU^T$, with associated eigenvalues $\lambda_{p+j}(N)$ for $1 \le j-p \le n-p$. Since N^T+UU^T and $N+UU^T$ share the same eigenvalues, it is found by continuity that for $N \in \mathcal{N}(R)$ in a neighborhood of 0, $\Re(1+\lambda_i(N)) > \Re(\lambda_{p+j}(N))$ for $1 \le i \le p$ and $1 \le j \le n-p$. Hence $\Pi_{\mathscr{M}}(UU^T + N) = UU^T$. ∎

Proposition 18. *The linear projector $\Pi_{\mathcal{T}(UU^T)}$ whose image is the tangent space $\mathcal{T}(UU^T)$ and whose kernel is $\mathcal{N}(UU^T)$ is given by*

$$\begin{array}{rcccc} \Pi_{\mathcal{T}(UU^T)} & : & \mathcal{M}_{n,n} & \to & \mathcal{T}(UU^T), \\ & & \mathfrak{X} & \mapsto & (I-UU^T)\mathfrak{X}UU^T + UU^T\mathfrak{X}^T(I-UU^T). \end{array} \tag{31}$$

Proof. The relation $\Pi_{\mathcal{T}(UU^T)} = \Pi_{\mathcal{T}(UU^T)} \circ \Pi_{\mathcal{T}(UU^T)}$ shows that $\Pi_{\mathcal{T}(UU^T)}$ is a linear projector. One then checks that $\mathrm{Ker}(\Pi_{\mathcal{T}(UU^T)}) = \mathcal{N}(UU^T)$ and $\mathrm{Span}(\Pi_{\mathcal{T}(UU^T)}) \subset \mathcal{T}(UU^T)$. The result follows by noticing that $\mathcal{T}(UU^T) \cap \mathcal{N}(UU^T) = \{0\}$, which implies $\mathrm{Span}(\Pi_{\mathcal{T}(UU^T)}) = \mathcal{T}(UU^T)$. ∎

The knowledge of $\Pi_{\mathcal{T}(UU^T)}$ then allows one to compute the Weingarten map and its spectral decomposition.

Proposition 19. *The Weingarten map in a direction $N \in \mathcal{N}(UU^T)$ for the manifold $\mathscr{M}$ equipped with the map of projectors $UU^T \mapsto \Pi_{\mathcal{T}(UU^T)}$ is given by*

$$\begin{array}{rcccc} L_{UU^T}(N) & : & \mathcal{T}(UU^T) & \to & \mathcal{T}(UU^T), \\ & & X & \mapsto & 2 \times \mathrm{sym}((I-UU^T)NXUU^T - XUU^TNUU^T). \end{array}$$

If $N = \sum_{i=1}^n \lambda_i u_i \overline{v}_i^T$ is diagonalizable and $\mathrm{Span}(U) = \mathrm{Span}(u_i)_{1\le i\le p}$, *then $L_{UU^T}(N)$ is also diagonalizable and the $p(n-p)$ eigenvalues are given by*

$$\kappa_{i(p+j)}(N) = \lambda_{p+j}(N) - \lambda_i(N),\ 1 \le i \le p, \text{ and } 1 \le j \le n-p.$$

A corresponding basis of eigenvectors $\Phi_{ij} \in \mathcal{T}(UU^T)$ is given by

$$\Phi_{i(p+j)} = UU^T\overline{v}_i u_{p+j}^T(I-UU^T) + (I-UU^T)u_{p+j}\overline{v}_i^T UU^T,$$

associated with the dual basis of left eigenvectors defined by

$$\forall X \in \mathcal{T}(R), \quad \langle \Phi^*_{i(p+j)}, X\rangle = \overline{v}_{p+j}^T X u_i.$$

Proof. The derivation of the expression of the Weingarten map is analogous to Proposition 13 and is omitted. It is straightforward to check that the Φ_{ij} are indeed eigenvectors of $L_{UU^T}(N)$. The dual basis is found by considering a linear combination $X = \sum_{ij} \alpha_{ij}\Phi_{ij} \in \mathcal{T}(UU^T)$ and by checking that $\alpha_{ij} = \overline{v}_j^T X u_i$ as claimed. ∎

As a corollary, one obtains a dynamical system satisfied by the subspace spanned by a fixed number of dominant eigenvectors of nonsymmetric matrices, which includes instantaneous reorthonormalization of a representing basis.

Corollary 20. *Let $\mathfrak{R}(t) = \sum_{i=1}^n \lambda_i(t) u_i \overline{v}_i^T \in \mathcal{M}_{n,n}$ be the spectral decomposition of a time-dependent diagonalizable matrix with eigenvalues $\lambda_i(t)$ ordered such that $\Re(\lambda_1(t)) \ge \cdots \ge \Re(\lambda_n(t))$. If $\Re(\lambda_p(t)) > \Re(\lambda_{p+1}(t))$, then the p dimensional dominant invariant subspace $\mathcal{U}(t) = \mathrm{Span}(u_i)_{1\le i\le p}$ of $\mathfrak{R}(t)$ is differentiable with respect to t and an ODE for the evolution of a corresponding orthonormal basis of vectors $U(t) \in \mathcal{M}_{n,p}$ satisfying $\mathrm{Span}(U(t)) = \mathrm{Span}(u_i)_{1\le i\le p}$ and $U(t)^TU(t) = I$ is*

$$\dot U = \sum_{\substack{1\le i\le p\\ 1\le j\le n-p}} \frac{1}{\lambda_i - \lambda_{p+j}} \left[\overline{v}_{p+j}^T \dot{\mathfrak{R}} u_i\right] (I-UU^T) f_{p+j}\overline{g}_i^T U, \tag{32}$$

where $(f_i)_{1\le i\le n}$ *and* $(g_i)_{1\le i\le n}$ *are the right and left eigenvectors of* $\Re(t)-U(t)U(t)^T$*, associated with the eigenvalues* $\lambda_i(t)-1$ *for* $1\le i\le p$ *and* $\lambda_{p+j}(t)$ *for* $1\le j\le n-p$*.*

Proof. Proposition 5 (and in particular the implicit function theorem) ensures the existence of a differentiable trajectory $V(t)V^T(t)$ such that $V(t)$ is invariant by $\Re(t)$ and $V(0)=U(0)$. The continuity of eigenvalues implies $V(t)V(t)^T=U(t)U(t)^T=\Pi_{\mathscr{M}}(\Re(t))$. Formula (32) follows identically as in Corollary 15. ∎

Remark 4. Note that in (32), $(u_i)_{1\le i\le p}$ and $(\overline{v}_{p+j}^T)_{1\le j\le n-p}$ are also right and left eigenvectors of $\Re(t)-U(t)U(t)^T$, but not $(u_{p+j})_{1\le j\le n-p}$ and $(\overline{v}_i^T)_{1\le i\le p}$.

Applying now Proposition 6 and examining the eigenvalues of the Weingarten map, we retrieve [6, Theorem 2.3], that is, that $\Pi_{\mathscr{M}}(\Re)$ is the unique stable equilibrium point and the generalization of (30) to nonsymmetric matrices.

COROLLARY 21 (see also [6]). *Let* $\Re=\sum_{i=1}^n\lambda_i(\Re)u_i\overline{v}_i^T\in\mathcal{M}_{n,n}$ *be a diagonalizable matrix satisfying* $\Re(\lambda_p(\Re))>\Re(\lambda_{p+1}(\Re))$*. Then* $\Pi_{\mathscr{M}}(\Re)$ *as defined by* (4) *is the unique stable equilibrium point of the dynamical system* (17)*, which can be written as an ODE for a representing basis* $U(t)$ *as*

$$\dot{U}=(I-UU^T)\Re U. \tag{33}$$

Proof. Equation (33) is obtained by writing (17) with $\Pi_{\mathcal{T}(UU^T)}$ being given by (31). Equilibrium points UU^T are those for which $N=\Re-UU^T\in\mathcal{N}(UU^T)$, i.e., such that $\mathrm{Span}(U)$ is a subspace of $\Re$ spanned by p eigenvectors. Denote $(\lambda_i)_{1\le i\le p}$ the eigenvalues of $\Re$ within $\mathrm{Span}(U)$ and $(\lambda_{p+j})_{1\le j\le n-p}$ the remaining ones. Then the eigenvalues of the Weingarten map are $\kappa_{i(p+j)}(N)=\lambda_{p+j}-(\lambda_i-1)$. The stability condition $\Re(\kappa_{i(p+j)}(N))<1$ is therefore satisfied for all eigenvalues only if $UU^T=\Pi_{\mathscr{M}}(\Re)$. ∎

Note that global convergence, although expected because of the boundedness of $\mathscr{M}$, is not completely clear since (33) is not a gradient flow.

5.2. Oblique projection on the bi-Grassmann manifold. The focus is now on map (5), whose image manifold is the set of linear projectors (not necessary orthogonal) over a p dimensional subspace. Next, we rely on the remark that any rank p projector can be factorized as $R=UV^T$, where $U,V\in\mathcal{M}_{n,p}$ are n-by-p matrices satisfying the orthogonality condition $V^TU=I$. Such matrices U and V can be obtained from any basis of, respectively, right and left eigenvectors of R. UV^T is then the unique projector whose image is $\mathrm{Span}(UV^T)=\mathrm{Span}(U)$ and whose kernel is $\mathrm{Ker}(UV^T)=\mathrm{Span}(V)^\perp$. Since this set identifies a rank p projector to a pair of p-dimensional subspaces of a n dimensional vector space, we refer to it as *bi-Grassmann manifold.*

DEFINITION 22. *The bi-Grassmann manifold is the set* $\mathscr{M}$ *of* rank-p *linear projectors of* $\mathcal{M}_{n,n}$*:*

$$\begin{aligned}\mathscr{M}&=\{R\in\mathcal{M}_{n,n}|R^2=R \textit{ and } \mathrm{rank}(R)=p\}\\&=\{UV^T|U\in\mathcal{M}_{n,p},V\in\mathcal{M}_{n,p},\,V^TU=I\}.\end{aligned} \tag{34}$$

A tangent vector $X\in\mathcal{T}(UV^T)$ can be written as $X=X_UV^T+UX_V^T$, where X_U and X_V can be understood as the time derivatives of the matrices U and V. Similarly to the case of the Grassmann manifold, a gauge condition (analogous to that of the fixed rank manifold; see [22]) is required on the matrices X_U and X_V to uniquely parameterize the tangent spaces of $\mathscr{M}$.

PROPOSITION 23. *The tangent space of $\mathscr{M}$ is*

$$\begin{aligned}\mathcal{T}(UV^T) &= \{X_U V^T + U X_V^T | X_U \in \mathcal{M}_{n,p},\, X_V \in \mathcal{M}_{n,p}\,, V^T X_U + U^T X_V = 0\}\\ &= \{X_U V^T + U X_V^T | X_U \in \mathcal{M}_{n,p},\, X_V \in \mathcal{M}_{n,p}\,, V^T X_U = U^T X_V = 0\}.\end{aligned}$$

The set $\mathcal{H}_{UV^T} = \{(X_U, X_V) \in \mathcal{M}_{n,p} \times \mathcal{M}_{n,p} | X_U^T V = X_V^T U = 0\}$ is referred to as the horizontal space *at $R = UV^T$. The map $(X_U, X_V) \mapsto X_U V^T + U X_V^T$ from $\mathcal{H}_{UV^T}$ to $\mathcal{T}(UV^T)$ is an isomorphism. Hence $\mathscr{M}$ is a smooth manifold of dimension $2p(n-p)$.*

Proof. Only the inclusion $\subset$ is proved, the inclusion $\supset$ being obvious. Consider two matrices $X_U, X_V \in \mathcal{M}_{n,p}$ such that $V^T X_U + U^T X_V = 0$. One can always write

$$\begin{cases} X_U = (I - UV^T) X_U + U\Omega, \\ X_V = (I - VU^T) X_V - V\Omega^T, \end{cases}$$

where $\Omega = V^T X_U = -U^T X_V$. Denote now $X'_U = (I - UV^T) X_U$ and $X'_V = (I - VU^T) X_V$. Since $V^T X'_U = U^T X'_V = 0$ and $X_U V^T + U X_V^T = X'_U V^T + U (X'_V)^T$, one obtains the inclusion $\subset$. Now, if $X = X_U V^T + U X_V^T$ with $U^T X_V = V^T X_U = 0$, one obtains $X_U = XU$ and $X_V = X^T V$ showing the uniqueness of the parameterization by the horizontal space. ∎

The next step is to identify the normal space of the candidate oblique projection (5).

PROPOSITION 24. *Consider $\Pi_{\mathscr{M}}$ as defined by* (5)*. $\Pi_{\mathscr{M}}$ is an oblique projection on $\mathscr{M}$, and the corresponding normal space $\mathcal{N}(UV^T)$ at $R = UV^T \in \mathscr{M}$ is the set of matrices $\mathfrak{R} \in \mathcal{M}_{n,n}$ letting both subspaces* $\mathrm{Span}(U)$ *and* $\mathrm{Span}(V)^\perp$ *be invariant:*

$$\begin{aligned}\mathcal{N}(UV^T) &= \{N \in \mathcal{M}_{n,n} | \mathrm{Span}(NU) \subset \mathrm{Span}(U) \text{ and } N[\mathrm{Span}(V)^\perp] \subset \mathrm{Span}(V)^\perp\}\\ &= \{N \in \mathcal{M}_{n,n} | N = (I - UV^T) N (I - UV^T) + UV^T N UV^T\}.\end{aligned}$$

In the following, results analogous to Proposition 19 and Corollaries 20 and 21 are derived for the bi-Grassmann manifold equipped with the bundle of normal spaces of the map (5). The proofs are almost strictly identical and are omitted.

PROPOSITION 25. *The linear projector $\Pi_{\mathcal{T}(UV^T)}$ whose image is the tangent space $\mathcal{T}(UV^T)$ and whose kernel is $\mathcal{N}(UV^T)$ is given by*

$$\Pi_{\mathcal{T}(UV^T)} \;:\; \begin{array}{ccc} \mathcal{M}_{n,n} & \to & \mathcal{T}(UV^T), \\ \mathfrak{X} & \mapsto & (I - UV^T)\mathfrak{X} UV^T + UV^T \mathfrak{X} (I - UV^T), \end{array} \tag{35}$$

or in the horizontal coordinates, $X_U = (I - UV^T)\mathfrak{X} U$ and $X_V = (I - VU^T)\mathfrak{X}^T V$.

Remark 5. In [28], the word "oblique projection" is used to refer to the linear tangent projection $\Pi_{\mathcal{T}(UV^T)}$ while ours refers to the nonlinear map $\Pi_{\mathscr{M}}$.

PROPOSITION 26. *The Weingarten map $L_{UV^T}(N)$ with respect to a normal vector $N \in \mathcal{N}(UV^T)$ is given by*

$$L_{UV^T}(N) \,:\, X \mapsto NXUV^T + UV^T XN - XUV^T NUV^T - UV^T NUV^T X.$$

If N is diagonalizable over $\mathbb{C}$ and $N = \sum_{i=1}^n \lambda_i(N) u_i \overline{v}_i^T$ denotes its eigendecomposition, written such that $UV^T = \sum_{i=1}^p u_i \overline{v}_i^T$, then $L_{UV^T}(N)$ is diagonalizable and its eigenvalues are the $n(n-p)$ numbers

$$\kappa_{ij} = \lambda_j(N) - \lambda_i(N) \quad \forall 1 \le i \le p,\, p+1 \le j \le n.$$

Each eigenvalue is associated with two independent eigenvectors,

$$\Phi_{ij,U} = u_i \overline{v}_j^T,\ \Phi_{ij,V} = u_j \overline{v}_i^T,$$

with their respective dual eigenvectors

$$\forall X \in \mathcal{T}(UV^T), \quad \langle \Phi^*_{ij,U}, X\rangle = \overline{v}_i^T X u_j,\ \langle \Phi^*_{ij,V}, X\rangle = \overline{v}_j^T X u_i.$$

As a result, another generalization of (29) to nonsymmetric matrices is obtained.

COROLLARY 27. *Let $\mathfrak{R}(t) = \sum_{i=1}^n \lambda_i(t) u_i \overline{v}_i^T \in \mathcal{M}_{n,n}$ be the spectral decomposition of a diagonalizable time-dependent matrix with $\Re(\lambda_1(t)) \geq \cdots \geq \Re(\lambda_n(t))$. If $\Re(\lambda_p(t)) > \Re(\lambda_{p+1}(t))$, then the p dimensional dominant invariant subspace $\mathcal{U}(t)$ of $\mathfrak{R}(t)$ and its invariant complement $\mathcal{V}(t)$ are differentiable with respect to t. An ODE for the evolution of corresponding bases $U(t), V(t) \in \mathcal{M}_{n,p}$ satisfying* $\mathrm{Span}(U(t)) = \mathcal{U}(t)$ *and* $\mathrm{Span}(V(t))^\perp = \mathcal{V}(t)$ *is*

$$\begin{cases} \dot{U} = \displaystyle\sum_{\substack{1\leq i\leq p\\ 1\leq j\leq n-p}} \frac{1}{\lambda_i(t) - \lambda_{p+j}(t)} (\overline{v}_{p+j}^T \dot{\mathfrak{R}} u_i) u_{p+j} \overline{v}_i^T U, \\ \dot{V} = \displaystyle\sum_{\substack{1\leq i\leq p\\ 1\leq j\leq n-p}} \frac{1}{\lambda_i(t) - \lambda_{p+j}(t)} (\overline{v}_i^T \dot{\mathfrak{R}} u_{p+j}) \overline{v}_{p+j} u_i^T V. \end{cases} \tag{36}$$

Finally, the dynamical system (17) that allows one to compute the invariant subspaces U and V is made explicit, before investigating its numerical implementation in section 6.

COROLLARY 28. *If $\mathfrak{R} = \sum_{i=1}^n \lambda_i(\mathfrak{R}) u_i \overline{v}_i^T$ is a real matrix diagonalizable in $\mathbb{C}$ and such that $\Re(\lambda_p(\mathfrak{R})) > \Re(\lambda_{p+1}(\mathfrak{R}))$, then $UV^T = \sum_{i=1}^p u_i \overline{v}_i^T = \Pi_{\mathscr{M}}(\mathfrak{R})$ is the unique asymptotically stable equilibrium point of the dynamical system*

$$\begin{cases} \dot{U} = (I - UV^T)\mathfrak{R}U, \\ \dot{V} = (I - VU^T)\mathfrak{R}^T V. \end{cases} \tag{37}$$

It is important to note that the above corollary does not guarantee, in contrast with the previously derived dynamical systems, global convergence almost everywhere. Indeed, the bi-Grassmann manifold is unbounded (for example, any matrix of the form $R = UU^T + UW^T$ with $U, W \in \mathcal{M}_{n,p}$, $U^T U = I$, and $W^T U = 0$ belongs to $\mathscr{M}$). Nevertheless convergence toward $\Pi_{\mathscr{M}}(\mathfrak{R})$ holds as soon as the initial point is sufficiently close, which may be acceptable in numerical algorithms for smoothly evolving two matrices $U(t)$ and $V(t)$ such that $U(t)V(t)^T = \Pi_{\mathscr{M}}(\mathfrak{R}(t))$.

6. Three numerical applications. We now present numerical examples that illustrate how the extrinsic framework provides tools for numerically tracking the values $\Pi_{\mathscr{M}}(\mathfrak{R}(t))$ of oblique projections (2)–(5) on evolving matrices $\mathfrak{R}(t)$. We showcase generalizable methods on three applications: two on the isospectral manifold concerned, respectively, with the gradient flow (8) and an approximation of the exact dynamic (17), and one on the "bi-Grassmann" manifold concerned with convergence issues and the conversion of the local dynamic (37) to a convergent iterative algorithms.

Note that we do not make any efficiency claim about the above dynamical systems over more classical iterative algorithms [25] which can also make use of good initial guesses. A useful feature of these continuous methods, though, is that they guarantee smooth evolution of their values, for example, when tracking a basis $U(t)$ of eigenvectors representing a subspace as in (30). This property can be very important, e.g., when integrated with high-order time discretizations of dynamical systems which require smooth evolutions of $U(t)$ (see, e.g., [23]).

6.1. Example 1: Revisiting the Brockett flow. We illustrate here the use of the gradient flow (8). Denote $\mathbf{S}(t) = \sum_{i=1}^{n} \lambda_i(t)u_i(t)u_i(t)^T$ the eigendecomposition as in Corollary 15 with $\lambda_1(t) > \cdots > \lambda_n(t)$. The objective is to track an eigenspace decomposition $\mathbb{R}^n = \oplus_i^m V_i(t)$ of m smoothly evolving subspaces $V_i(t)$ spanned by successive eigenvectors $u_i(t)$ and with fixed given dimensions n_i:

$$V_i(t) = \mathrm{Span}(u_j(t))_{j\in I_i} \text{ with } I_i = \{1 + n_1 + \cdots + n_{i-1} \le j \le n_1 + \cdots + n_i\}. \tag{38}$$

Following section 4, we label the eigenspaces $V_i(t)$ with m distinct real numbers $\lambda_1 > \cdots > \lambda_m$, and we consider $\mathscr{M}$ the isospectral manifold of symmetric matrices S with prescribed eigenvalues λ_i of multiplicities n_i (definition (24)). Tracking the subspaces $V_i(t)$ then becomes equivalent to tracking the orthogonal projection $\Pi_{\mathscr{M}}(\mathbf{S}(t)) = \sum_{i=1}^{m} \lambda_i U_i(t)U_i(t)^T$, where $U_i(t)$ are n-by-n_i orthonormal matrices satisfying $\mathrm{Span}(U_i(t)) = V_i(t)$.

Let $n = 10$, $T = 1$, and define, for $t \in [0, T]$,

$$\begin{aligned} d_k(t) &= \begin{cases} \sin\left(\dfrac{3\pi}{4}t + \dfrac{2k\pi}{n}\right) \text{ for } 1 \le k \le 5, \\ 1 - \dfrac{k}{2} - \dfrac{1}{2}\sin(\pi t), \text{ for } 6 \le k \le 10, \end{cases} \\ D(t) &= \mathrm{diag}(d_k(t)), \\ P(t) &= \exp(8\pi t \Omega) \end{aligned} \tag{39}$$

for a given antisymmetric matrix Ω taken as random. Since $P(t)$ is an orthonormal matrix, we obtain a time-dependent symmetric matrix $\mathbf{S}(t)$ whose eigenvalues are the $d_k(t)$ associated with rotating eigenvectors $P(t)$ by setting $\mathbf{S}(t) = P(t)D(t)P(t)^T$. Notably, $\mathbf{S}(t)$ has been especially designed for admitting crossing eigenvalues at various times as visible in Figure 2. We consider the following two settings:

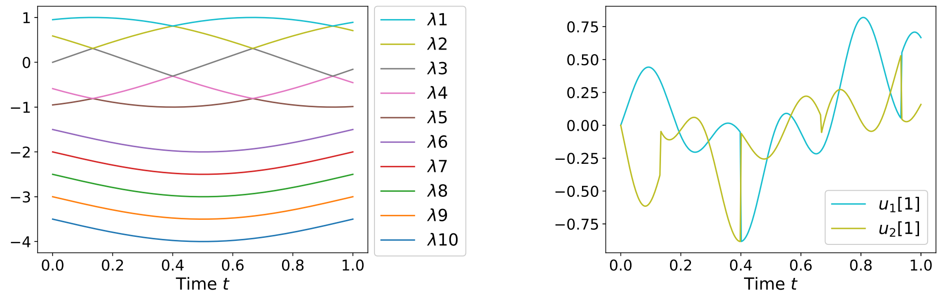
Case 1: $\mathbb{R}^n = V_1(t) \oplus V_2(t) \oplus V_3(t) \oplus V_4(t)$
with $(\lambda_1, \lambda_2, \lambda_3, \lambda_4) = (3, 2, 1, 0)$ and $(n_1, n_2, n_3, n_4) = (1, 1, 3, 5)$:

$$\begin{cases} V_1(t) = \mathrm{Span}(u_1(t)), \\ V_2(t) = \mathrm{Span}(u_2(t)), \\ V_3(t) = \mathrm{Span}(u_3(t), u_4(t), u_5(t)), \\ V_4(t) = \mathrm{Span}(u_6(t), \ldots, u_{,10}(t)). \end{cases} \tag{40}$$

Case 2: $\mathbb{R}^n = V_1(t) \oplus V_2(t) \oplus V_3(t)$
with $(\lambda_1, \lambda_2, \lambda_3) = (2, 1, 0)$ and $(n_1, n_2, n_3) = (2, 3, 5)$:

$$\begin{cases} V_1(t) = \mathrm{Span}(u_1(t), u_2(t)), \\ V_2(t) = \mathrm{Span}(u_3(t), u_4(t), u_5(t)), \\ V_3(t) = \mathrm{Span}(u_6(t), \ldots, u_{10}(t)). \end{cases} \tag{41}$$

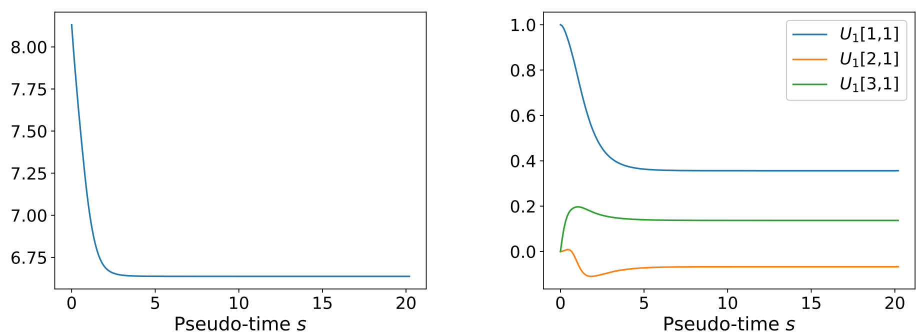
We report in Figure 3 the convergence of a single performance of the gradient flow (30) for fixed $\mathbf{S} = \mathbf{S}(T)$, solved on a pseudotime window $s \in [0, 20]$ with Euler step $\Delta s = 0.1$ and initialized with some orthonormal matrices $U_i(0)$ satisfying $\mathrm{Span}(U_i(0)) = V_i(0)$.



(a) Ordered eigenvalues $\lambda_i(t)$: $\lambda_1(t) \geq \lambda_2(t) \geq \cdots \geq \lambda_{10}(t)$ and $\{\lambda_1(t), \ldots, \lambda_{10}(t)\} = \{d_1(t), \ldots, d_{10}(t)\}$.

(b) Trajectories of the first coordinate of the eigenvectors $u_1(t)$ and $u_2(t)$ (obtained from (39)). These are discontinuous when $\lambda_1(t) = \lambda_2(t)$ or $\lambda_2(t) = \lambda_3(t)$.

FIG. 2. *Variations of the ordered spectrum of* $\mathbf{S}(t) = P(t)D(t)P(t)^T$ *(equation* (39)*). Nonsmooth behaviors occur when eigenvalues become multiple.*



(a) Distance $\|S(s) - \mathbf{S}\|$

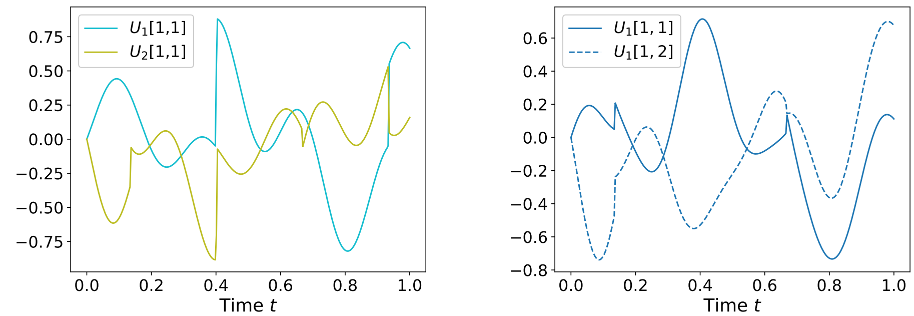
(b) Coefficients $U_1[k,1](s)$ for $k \in \{1,2,3\}$

FIG. 3. *Convergence of the gradient descent* (30) *on the isospectral manifold* $\mathscr{M}$ (24) *for the* fixed *symmetric matrix* $\mathbf{S} = \mathbf{S}(T)$. *The distance function* $S \mapsto \|S - \mathbf{S}\|$ *is minimized on* $\mathscr{M}$, *by evolving smoothly basis matrices* $U_i(s)$ *which align at convergence with eigenspaces of* $\mathbf{S}$.

Figure 3(b) illustrates how the pseudotime solutions $U_i(s)$ smoothly align on the eigendecomposition of $\mathbf{S} = \mathbf{S}(T)$.

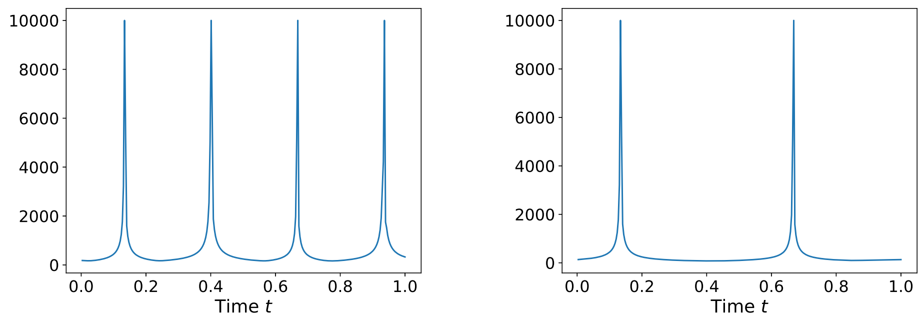
Regarding the numerical discretization of the flow (30), the only difficulty is to maintain the orthogonality $U_i^T(t)U_i(t) = I$ which is not preserved after a standard Euler time step. This issue is addressed by the introduction of suitable *retractions* during the time stepping [2, 4]. For our purpose, maintaining the orthogonality $U_i(t)^T U_i(t) = I$ for $U \in \mathcal{M}_{n,r}$ can be directly addressed by various reorthonormalization procedures preserving the continuous evolution of the matrices $U_i(t)$ [4, 23, 49]. Note that in general, the time step of the discretization of ODEs such as (8) must be selected sufficiently small with respect to the Lipschitz constant of the right-hand side vector field. A simple rule of thumb used, e.g., in [23], is to scale it proportionally to the inverse of the curvature locally experienced on the manifold. More elaborate schemes can also be used to specifically address such issues [52, 38].

We then use this gradient flow to smoothly evolve bases $U_i(t)$ representing the projection $\Pi_{\mathscr{M}}(\mathbf{S}(t))$. The time interval $[0,T]$ is discretized into time steps $t_k = k\Delta T$ for



(a) Case 1: $V_1(t) = \mathrm{Span}(u_1(t))$ and $V_2(t) = \mathrm{Span}(u_2(t))$. Discontinuities occur whenever $\lambda_1(t) = \lambda_2(t)$ or $\lambda_2(t) = \lambda_3(t)$.

(b) Case 2: $V_1(t) = \mathrm{Span}(u_1(t), u_2(t))$. Discontinuities occur only when $\lambda_2(t) = \lambda_3(t)$.



(c) Intermediate gradient iterations: Case 1.

(d) Intermediate gradient iterations: Case 2.

FIG. 4. *Tracking dominant eigenspaces of* $\mathbf{S}(t) = P(t)D(t)P(t)^T$ (39) *by successive converged solutions* $U(t)$ *of the gradient flow* (30)*. The number of iterations required for convergence shown in panels* (c)–(d) *confirms that it is slower at discontinuities of* $V_1(t)/\ V_2(t)$.

a uniform increment $\Delta t = T/300$. From the knowledge at time t^k of orthonormal matrices $U_i(t^k)$ satisfying $\mathrm{Span}(U_i(t^k)) = V_i(t^k)$, we obtain continuous updates $U_i(t^{k+1})$ as converged solutions of inner-iterations of the gradient flow (30) with $\mathbf{S} = \mathbf{S}(t^{k+1})$. Results are reported in Figure 4. As expected, we observe that the matrices $U_i(t)$ (or equivalently, the projection $\Pi_{\mathscr{M}}(\mathbf{S}(t))$) become discontinuous at the exact instants where crossing of specific eigenvalues occurs *in between the eigenspaces* $V_i(t)$, for which the decomposition $\mathbb{R}^n = \oplus_{i=1}^m V_i(t)$ becomes ill-defined. (Eigenvectors $u_i(t)$ corresponding to crossed eigenvalues could be attributed indistinctly to two different $V_i(t)$.) Near these instants, the gradient flow (30) converges more and more slowly. Nevertheless and importantly, crossings of eigenvalues are not an issue when they occur *within the eigenspaces* $V_i(t)$: $V_i(t)$ and its representing basis $U_i(t)$ evolve smoothly, and the convergence of the gradient flow is not altered, as visible in Figure 4(d). For example, the crossing of $\lambda_1(t)$ and $\lambda_2(t)$ near $t = 0.4$ is not felt in Case 2, designed to track the subspace spanned by the first two eigenvectors without tracking specifically the individual trajectories of $u_1(t)$ and $u_2(t)$. Interestingly, global convergence of the gradient flow (30) allows one to recapture correct bases $U_i(t)$ after passing through such discontinuities (which could not have been done using, e.g., the ODE (28)), although the use here of an iterative algorithm reinitializing directly the matrices $U_i(t)$ would certainly prove more efficient computationally.

6.2. Example 2: Dynamic approximation on the isospectral manifold. In this part, we assume the trajectory $\mathbf{S}(t)$ is given by an ODE

$$\dot{\mathbf{S}} = \mathcal{L}(t, \mathbf{S}) \tag{42}$$

for some smooth symmetric vector field $\mathcal{L}$. Suppose it is known a priori that the trajectory $\mathbf{S}(t)$ lies *close* to some isospectral manifold $\mathscr{M}$, in the sense that the eigenvalues $\lambda_i(t)$ of $\mathbf{S}(t)$ are organized in m fixed clusters of size $n_1, \ldots, n_m$ centered around fixed reals $\lambda_1 > \cdots > \lambda_m$. If the objective is to track these clustered eigenspaces $V_i(t)$ but not the precise, potentially discontinuous, trajectories of each eigenvector of $\mathbf{S}(t)$, then a possible solution is to approximate the ODE (29) satisfied by $\Pi_{\mathscr{M}}(\mathbf{S}(t))$ by projected dynamical systems on the isospectral manifold:

$$\text{(a)} \begin{cases} \dot{S} = \Pi_{\mathcal{T}(S(t))}(\mathcal{L}(t, S(t))), \\ S(0) = \Pi_{\mathscr{M}}(\mathbf{S}(0)) \end{cases} \quad \text{or} \quad \text{(b)} \begin{cases} \dot{S} = \Pi_{\mathcal{T}(S(t))}(\dot{\mathbf{S}}), \\ S(0) = \Pi_{\mathscr{M}}(\mathbf{S}(0)). \end{cases} \tag{43}$$

The autonomous dynamical system (43)(a) is known as the *DO approximation* [22] or *dynamical low-rank* [40] method when $\mathscr{M}$ is the fixed rank manifold, while the nonautonomous (43)(b) offers a slightly better approximation if one has direct access to the derivative $\dot{\mathbf{S}}$ of the nonreduced dynamic $\dot{\mathbf{S}} = \mathcal{L}(t, \mathbf{S})$, i.e., (42). Equation (43)(b) is obtained by assuming the normal component can be neglected in (16) ($\|N\| \simeq 0$), and (43)(a) by additionally approximating $\mathcal{L}(t, \mathbf{S})$ by $\mathcal{L}(t, S)$. Each offers a convenient approximation of the true dynamic (7) of $\Pi_{\mathscr{M}}(\mathbf{S}(t))$ which avoids computing and storing the full spectral decomposition $\kappa_i(N), \Phi_i$ of the Weingarten map. For the isospectral manifold, one solves the coupled system of ODEs for the trajectories of representing n-by-n_i matrices $U_i(t)$ such that $S(t) = \sum_{i=1}^m \lambda_i U_i(t) U_i(t)^T$ (Proposition 11), which in the case of (43)(b) writes as

$$\dot{U}_i = \sum_{j \neq i} \frac{1}{\lambda_i - \lambda_j} U_j U_j^T \dot{\mathbf{S}} U_i. \tag{44}$$

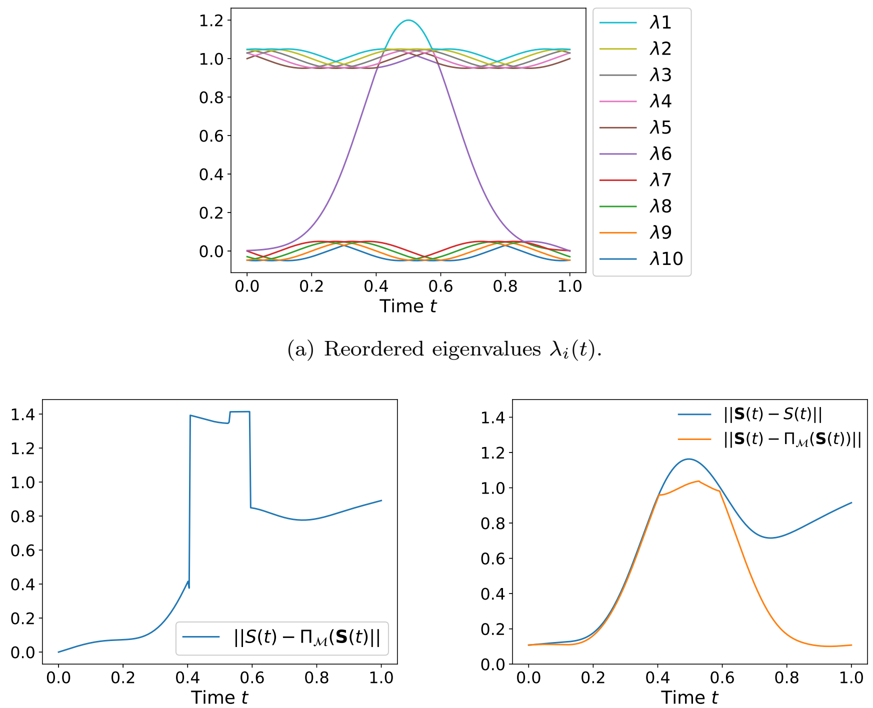
Note that (44) is a generalization of the "OTD" equation proposed by [6]. In this Euclidean setting, it is possible to show that error $\|S(t) - \Pi_{\mathscr{M}}(\mathbf{S}(t))\|$ remains "controlled" as long as the projection $\Pi_{\mathscr{M}}(\mathbf{S}(t))$ stays continuous. (We refer the reader to [22, 21] for precise statements on this topic.) This means that one can expect the approximation $\mathrm{Span}(U_i(t)) \simeq V_i(t)$ to hold if the eigenvalues of $\mathbf{S}(t)$ do not cross inbetween the spaces $V_i(t)$. As a numerical example, we consider the same setting as subsection 6.1, but with

$$d_k(t) = \begin{cases} 1 + 0.05 \sin\left(4\pi t + \dfrac{2k\pi}{n}\right) \text{ for } 1 \leq k \leq 5, \\ \\ 0.05 \sin\left(4\pi t + \dfrac{2k\pi}{n}\right) \text{ for } 6 \leq k \leq 9, \\ \\ 1.2 \exp(-(t - 0.5)^2/0.04) \text{ for } k = 10\ . \end{cases} \tag{45}$$

The corresponding trajectories of the eigenvalues $d_k(t)$, clustered near 0 and 1, are plotted on Figure 5(a). We aim at tracking the decomposition

$$\mathbb{R}^n = V_1(t) \oplus V_2(t) \text{ with } \begin{cases} V_1(t) = \mathrm{Span}(u_1(t), \ldots, u_5(t)), \\ V_2(t) = \mathrm{Span}(u_6(t), \ldots, u_{10}(t)), \end{cases} \tag{46}$$

which corresponds to the setting $m = 2$, $(\lambda_1, \lambda_2) = (1, 0)$, and $(n_1, n_2) = (5, 5)$. In



(a) Reordered eigenvalues $\lambda_i(t)$.

(b) Approximation error with respect to the exact reduction $\Pi_{\mathscr{M}}(\mathbf{S}(t))$: i.e., $\|S(t) - \Pi_{\mathscr{M}}(\mathbf{S}(t))\|$.

(c) Errors with respect to the full $\mathbf{S}(t)$: dynamical error (blue) vs. best achievable error (orange).

FIG. 5. *"Dynamic approximation" method* (43)(b) *on the isospectral manifold* $\mathscr{M}$ *for tracking dynamic eigenspaces associated with clustered eigenvalues around* 0 *and* 1.

our implementation, we use of the approximate dynamics (44) with the analytical value of $\dot{\mathbf{S}}$. We report in Figure 5(b) the time evolution of the approximation error $\|\Pi_{\mathscr{M}}(\mathbf{S}(t)) - S(t)\|$. As observed in [55, 22], for approximate dynamics on the fixed rank manifold, the approximation $S(t) \simeq \Pi_{\mathscr{M}}(\mathbf{S}(t))$ holds till a discontinuity of the projection $\Pi_{\mathscr{M}}(\mathbf{S}(t))$ occurs, which cannot be captured by the continuous evolution of $S(t)$. For this example, this happens at the exact instants t where $\lambda_5(t) = \lambda_6(t)$. Figure 5(c) displays the evolution of the best achievable error $\|\mathbf{S}(t) - \Pi_{\mathscr{M}}(\mathbf{S}(t))\|$ between the nonreduced matrix $\mathbf{S}(t)$ and its projection $\Pi_{\mathscr{M}}(\mathbf{S}(t))$, versus the dynamical error $\|\mathbf{S}(t) - S(t)\|$, which further illustrates the somewhat independent evolution of the approximate solution $S(t)$ after the first discontinuity.

6.3. Application 3: Issues with the non-Euclidean bi-Grassmann manifold. The ODE (37) associated with the map (5) is of interest for tracking the dual subspaces of dominant left and right eigenvectors of a dynamic matrix $\mathfrak{R}(t)$. However, (37) does not behave as smoothly for reasons examined now and related to the nonboundedness of the bi-Grassmann manifold. We will still show how this problem can be fixed, since its discretization allows guessing a convergent iterative algorithm of equivalent complexity. (Note that we will not prove here the "convergence" claim.)

Going back to the setting and the notations of subsection 5.2, a first issue occurring when discretizing the dynamical system (37) for n-by-p matrices $U(t)$ and

$V(t)$ is maintaining the property $V^TU = I$ during the time discretization. For this, we propose a simple retraction for the biorthogonal manifold. (We refer to [4] for a precise definition of retraction).

PROPOSITION 29. *For $U, V \in \mathcal{M}_{n,p}$ with $V^TU = I$, the map*

$$\rho_{UV^T} \;:\; \begin{array}{ccc} \mathcal{H}_{UV^T} & \longrightarrow & \mathscr{M}, \\ (X_U, X_V) & \longmapsto & (U + X_U)[(V + X_V)^T(U + X_U)]^{-1}(V + X_V)^T \end{array} \tag{47}$$

is a retraction on the biorthogonal manifold $\mathscr{M}$.

Proof. First, denoting $U_1 = U + X_U$ and $V_1 = (V + X_V)[(V + X_V)^T(U + X_U)]^{-T}$, one has $\rho_{UV^T}(X_U, X_V) = U_1V_1^T$ with $U_1, V_1 \in \mathcal{M}_{n,p}$ and $V_1^TU_1 = I$, ensuring $\rho_{UV^T}(X_U, X_V) \in \mathscr{M}$. Since $V^TU = I$ and for $(X_U, X_V) \in \mathcal{H}_{UV^T}$, $X_V^TU = X_U^TV = 0$, one can write $\rho_{UV^T}(X_U, X_V) = (U + X_U)(I + X_V^TX_U)^{-1}(V + X_V)^T$. Hence the following asymptotic expansion holds:

$$\begin{aligned} \rho_{UV^T}(tX_U, tX_V) &= (U + tX_U)(I - t^2X_V^TX_U + o(t^2))(V^T + tX_V) \\ &= UV^T + t(X_UV^T + UX_V^T) + o(t^2), \end{aligned} \tag{48}$$

which implies that ρ_{UV^T} is a first-order retraction. ∎

The above retraction ρ_{UV^T} can be used to obtain a discretization of (37) preserving the property $V^TU = I$. At a time step k, the time derivatives $\dot{U}_k$ and $\dot{V}_k$ given by (37) are first computed according to

$$\begin{cases} \dot{U}_k = \mathfrak{R}U_k - U_k(V_k^T\mathfrak{R}U_k), \\ \dot{V}_k = \mathfrak{R}^TV_k - V_k(U_k^T\mathfrak{R}^TV_k)\,, \end{cases} \tag{49}$$

where parentheses highlight products rendering the evaluation efficient (e.g., for n possibly much greater than p; see [44]). A possible discretization of (37) using a first-order Euler scheme and the retraction ρ_{UV^T} (47) could therefore be

$$\begin{cases} U_{k+1} = U_k + \Delta t\dot{U}_k, \\ V_{k+1} = (V_k + \Delta t\dot{V}_k)A_k^{-T}, \\ \quad A_k = (V_k + \Delta t\dot{V}_k)^T(U_k + \Delta t\dot{U}_k) = I + \Delta t^2\dot{V}_k^T\dot{U}_k. \end{cases} \tag{50}$$

We observe that the use of the retraction (47) induces only a second-order correction on the first-order Euler scheme, necessary to ensure consistency and the smooth evolutions of the matrices U and V. The computation of the inverse matrix $A_k^{-T} \in \mathcal{M}_{p,p}$ is not costly for moderate values of p. Nevertheless the implementation of (50) can be numerically unstable for initial values U_0 and V_0 too far from the equilibrium point (confirmed numerically; not shown here). This is no surprise, because in spite of

$\Pi_{\mathscr{M}}(\mathfrak{R})$ being the unique stable equilibrium point of the flow (37), trajectories can possibly escape to infinity. This lack of global convergence emphasizes that the benefit of the approach mainly holds to find small corrections to add to sufficiently good initial guesses U_0, V_0 such that $U_0 V_0^T \simeq \Pi_{\mathscr{M}}(\mathfrak{R})$. For one shot computations or if continuous updates are not required, one may rather rely on direct iterative algorithms such as the one we propose now and detailed in Algorithm 1. The discretization (49) featuring matrix vector products $\mathfrak{R}U_k$ and $\mathfrak{R}^T V_k$ suggests a variant of the power method for computing left and right eigenvectors; our Algorithm 1 implements this idea with a computational complexity analogous to that of (50). Here, numerical stability is obtained by adding the normalization step (51) (results confirmed in examples; not shown here). We note that Guglielmi and Lubich [28] relied on restarting procedures in their implementation of a dynamical system analogous to (37).

Algorithm 1 Computing dominant invariant subspaces of nonsymmetric matrices.

Given a real matrix $\mathfrak{R} \in \mathcal{M}_{n,p}$:

1: Grow the part of the p dominant subspace already present in U_k, V_k:

$$U_{k+1} = \mathfrak{R}U_k, V_{k+1} = \mathfrak{R}^T V_k$$

2: Rotate the columns of V_{k+1} such that $V_{k+1}^T U_{k+1} = I$:

$$V_{k+1} \leftarrow V_{k+1}(U_{k+1}^T V_{k+1})^{-1} \tag{51}$$

3: Normalize U_{k+1} and V_{k+1} such that $U_{k+1}^T U_{k+1} = I$ and $V_{k+1}^T U_{k+1} = I$:

$$\begin{cases} A_{k+1} = U_{k+1}^T U_{k+1}, \\ V_{k+1} \leftarrow V_{k+1} A^{\frac{1}{2}}, \\ U_{k+1} \leftarrow U_{k+1} A^{-\frac{1}{2}}. \end{cases} \tag{52}$$

Hence U_{k+1} and V_{k+1}^T are normalized such that $U_{k+1}V_{k+1}^T$ is a rank-p projector, where $\|U_{k+1}\|^2 = p$ and $\|V_{k+1}\|^2 = \|U_{k+1}V_{k+1}^T\|^2$ ($\|\cdot\|$ is the Frobenius norm).

7. Conclusion. A geometric framework was introduced for studying the differentiability of (nonlinear) *oblique projections* and dynamical systems that allow their efficient tracking on smoothly varying matrices. This was achieved by obtaining explicitly the spectral decomposition of the Weingarten map for various image manifolds equipped with the natural *extrinsic* differential structure associated with the oblique projection. The nonrequirement of the ambient space to be Euclidean allowed studying nonsymmetric matrix maps that are not characterized by a minimization problem. Popular matrix manifolds were studied with respect to a natural embedding yielding new interpretations. Global stability analysis of the dynamical systems that compute oblique projections was performed for all these manifolds. Their relevance and numerical implementations for tracking smooth decompositions was discussed. Possible future applications of the derived dynamic matrix equations abound over a rich spectrum of needs, from dynamic reduced-order modeling [57, 22] and data sciences [41] to adaptive data assimilation [45, 7, 51] and adaptive path planning and sampling [50, 63, 46, 47].

Appendix A. Notation.

E	Finite dimensional ambient space
E^*	Dual space of E
A^*	Dual operator of a linear map A
$\mathrm{sp}(A)$	Set of (complex) eigenvalues of a linear map A
$\mathscr{M}$	Embedded manifold $\mathscr{M} \subset E$
$\mathfrak{R} \in E$	Point of the ambient space
$R \in \mathscr{M}$	Point of the manifold
$\mathcal{T}(R)$	Tangent space at R
$\mathcal{N}(R)$	Normal space at R
$\mathfrak{X} \in E$	Vector of the ambient space E
$X \in \mathcal{T}(R)$	Tangent vector X at R
$\mathrm{Ker}(A)$	Kernel of a linear operator A
$\mathrm{Span}(A)$	Image of a linear operator A
$\Pi_{\mathcal{T}(R)}$	Linear projector with $\mathrm{Span}(\Pi_{\mathcal{T}(R)}) = \mathcal{T}(R)$ and $\mathrm{Ker}(\Pi_{\mathcal{T}(R)}) = \mathcal{N}(R)$
$L_R(N)$	Weingarten map at $R \in \mathscr{M}$ in the direction $N \in \mathcal{N}(R)$
$\langle\, ,\, \rangle$	Duality bracket or scalar product on E
$\\| \\| = \sqrt{\langle\, ,\, \rangle}$	Euclidean norm
$\Pi_{\mathscr{M}}$	Oblique projection onto $\mathscr{M}$
I	Identity mapping
$\mathcal{M}_{n,p}$	Space of n-by-p matrices
Sym_n	Space of n-by-n symmetric matrices
A^T	Transpose of a square matrix A
$\langle A, B\rangle = \mathrm{Tr}(A^T B)$	Frobenius matrix scalar product
$\\|A\\| = \mathrm{Tr}(A^T A)^{1/2}$	Frobenius norm
$\sigma_1(A) \geq \cdots \geq \sigma_{\mathrm{rank}(A)}(A)$	Nonzeros singular values of $A \in \mathcal{M}_{n,p}$
$(\lambda_i(A))_{1\leq i\leq n}$ $\Re(\lambda_1(A)) \geq \Re(\lambda_2(A)) \geq \cdots \geq \Re(\lambda_n(A))$	Eigenvalues of an n-by-n matrix A ordered according to their real parts
$\mathrm{sym}(\mathfrak{X}) = (\mathfrak{X} + \mathfrak{X}^T)/2$	Symmetric part of a square matrix $\mathfrak{X}$
$\mathrm{skew}(\mathfrak{X}) = (\mathfrak{X} - \mathfrak{X}^T)/2$	Skew-symmetric part of a square matrix $\mathfrak{X}$
δ_{ij}	Kronecker symbol; $\delta_{ii} = I$ and $\delta_{ij} = 0$ for $i \neq j$
$\dot{R} = \mathrm{d}R/\mathrm{d}t$	Time derivative of a trajectory $R(t)$
$\mathrm{D}_X f(R)$	Differential of a function f in direction X
$\mathrm{D}\Pi_{T(R)}(X) \cdot Y$	Differential of the projection operator $\Pi_{T(R)}$ applied to Y

Acknowledgment. We thank the members of the MSEAS group at MIT for insightful discussions related to this topic.

REFERENCES

[1] P.-A. Absil, R. Mahony, and R. Sepulchre, *Riemannian geometry of Grassmann manifolds with a view on algorithmic computation*, Acta Appl. Math., 80 (2004), pp. 199–220.

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

[3] P.-A. Absil, R. Mahony, and J. Trumpf, *An extrinsic look at the Riemannian Hessian*, in Geometric Science of Information, Springer, New York, 2013, pp. 361–368.

[4] P.-A. Absil and J. Malick, *Projection-like retractions on matrix manifolds*, SIAM J. Optim., 22 (2012), pp. 135–158.

[5] L. Ambrosio, *Geometric Evolution Problems, Distance Function and Viscosity Solutions*, Springer, New York, 2000.

[6] H. Babaee and T. Sapsis, *A minimization principle for the description of modes associated with finite-time instabilities*, in Proc. A, 472 (2016), 20150779.

[7] K. Bergemann, G. Gottwald, and S. Reich, *Ensemble propagation and continuous matrix factorization algorithms*, Q. J. Roy. Meteorol. Soc., 135 (2009), pp. 1560–1572.

[8] G. E. Bredon, *Topology and Geometry*, Vol. 139, Springer Science, New York, 2013.

[9] R. W. Brockett, *Dynamical systems that sort lists, diagonalize matrices and solve linear programming problems*, in Proceedings of the 27th IEEE Conference on Decision and Control, 1988, IEEE, pp. 799–803.

[10] R. W. Brockett, *Dynamical systems that learn subspaces*, in Mathematical System Theory, Springer, New York, 1991, pp. 579–592.

[11] Y.-C. Chen and L. Wheeler, *Derivatives of the stretch and rotation tensors*, J. Elasticity, 32 (1993), pp. 175–182.

[12] M. T. Chu, *A list of matrix flows with applications*, Fields Inst. Commun., 3 (1994), pp. 87–97.

[13] M. T. Chu and K. R. Driessel, *Constructing symmetric nonnegative matrices with prescribed eigenvalues by differential equations*, SIAM J. Math. Anal., 22 (1991), pp. 1372–1387.

[14] J. Dehaene, *Continuous-Time Matrix Algorithms Systolic Algorithms and Adaptive Neural Networks*, Ph.D. thesis, K. U. Leuven, Belgium, 1995.

[15] J. Dehaene, M. Moonen, and J. Vandewalle, *Analysis of a class of continuous-time algorithms for principal component analysis and subspace tracking*, IEEE Trans. Circuits Syst. I Fundam. Theory Appl., 46 (1999), pp. 364–372.

[16] P. Deift, T. Nanda, and C. Tomei, *Ordinary differential equations and the symmetric eigenvalue problem*, SIAM J. Numer. Anal., 20 (1983), pp. 1–22.

[17] M. C. Delfour and J.-P. Zolésio, *Shapes and Geometries: Metrics, Analysis, Differential Calculus, and Optimization*, Vol. 22, SIAM, Philadelphia, 2011.

[18] L. Dieci and T. Eirola, *On smooth decompositions of matrices*, SIAM J. Matrix Anal. Appl., 20 (1999), pp. 800–819.

[19] C. Ebenbauer, *A dynamical system that computes eigenvalues and diagonalizes matrices with a real spectrum*, in Proceedings of the 46th IEEE Conference on Decision and Control, IEEE, 2007, pp. 1704–1709.

[20] A. Edelman, T. A. Arias, and S. T. Smith, *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Anal. Appl., 20 (1998), pp. 303–353.

[21] F. Feppon, *Riemannian Geometry of Matrix Manifolds for Lagrangian Uncertainty Quantification of Stochastic Fluid Flows*, Master's thesis, MIT, Cambridge, MA, 2017.

[22] F. Feppon and P. F. Lermusiaux, *A geometric approach to dynamical model order reduction*, SIAM J. Matrix Anal. Appl., 39 (2018), pp. 510–538.

[23] F. Feppon and P. F. J. Lermusiaux, *Dynamically Orthogonal Numerical Schemes for Efficient Stochastic Advection and Lagrangian Transport*, SIAM Rev., 60 (2018), pp. 595–625.

[24] N. H. Getz and J. E. Marsden, *Dynamic inversion and polar decomposition of matrices*, in Proceedings of the 34th IEEE Conference on Decision and Control, Vol. 1, IEEE, 1995, pp. 142–147.

[25] G. H. Golub and C. F. Van Loan, *Matrix Computations*, Johns Hopkins University Press, Baltimore, MD, 1996, pp. 374–426.

[26] B. Gong, Y. Shi, F. Sha, and K. Grauman, *Geodesic flow kernel for unsupervised domain adaptation*, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), IEEE, 2012, pp. 2066–2073.

[27] G. Grioli, *Una proprieta di minimo nella cinematica delle deformazioni finite*, Boll. Unione Mat. Ital, 2 (1940).

[28] N. Guglielmi and C. Lubich, *Differential equations for roaming pseudospectra: Paths to extremal points and boundary tracking*, SIAM J. Numer. Anal., 49 (2011), pp. 1194–1209.

[29] G. Haller, *Dynamic rotation and stretch tensors from a dynamic polar decomposition*, J. Mech. Phys. Solids, 86 (2016), pp. 70–93.

[30] U. Helmke, K. Hüper, and J. Trumpf, *Newton's Method on Grassmann Manifolds*, preprint, arXiv:0709.2205, 2007.

[31] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, Springer, New York, 2012.

[32] H. Hendriks and Z. Landsman, *Asymptotic behavior of sample mean location for manifolds*, Statist. Probab. Lett., 26 (1996), pp. 169–178.

[33] H. Hendriks and Z. Landsman, *Mean location and sample mean location on manifolds: Asymptotics, tests, confidence regions*, J. Multivariate Anal., 67 (1998), pp. 227–243.
[34] G. Hori, *Isospectral gradient flows for non-symmetric eigenvalue problem*, Jpn. J. Ind. Appl. Math., 17 (2000), pp. 27–42.
[35] W. Huang, *Optimization Algorithms on Riemannian Manifolds with Applications*, Ph.D. thesis, Florida State University, 2013.
[36] J. Jost, *Riemannian Geometry and Geometric Analysis*, Springer, New York, 2008.
[37] T. Kato, *Perturbation Theory for Linear Operators*, Springer, New York, 2013.
[38] E. Kieri, C. Lubich, and H. Walach, *Discretized dynamical low-rank approximation in the presence of small singular values*, SIAM J. Numer. Anal., 54 (2016), pp. 1020–1038.
[39] S. Kobayashi and K. Nomizu, *Foundations of Differential Geometry*, Vol. 2, Interscience, New York, 1969.
[40] O. Koch and C. Lubich, *Dynamical low-rank approximation*, SIAM J. Matrix Anal. Appl., 29 (2007), pp. 434–454.
[41] J. N. Kutz, *Data-Driven Modeling & Scientific Computation: Methods for Complex Systems & Big Data*, Oxford University Press, Oxford, 2013.
[42] P. F. J. Lermusiaux, *Error Subspace Data Assimilation Methods for Ocean Field Estimation: Theory, Validation and Applications*, Ph.D. thesis, Harvard University, 1997.
[43] P. F. J. Lermusiaux, *Estimation and study of mesoscale variability in the Strait of Sicily*, Dyn. Atmos. Oceans, 29 (1999), pp. 255–303.
[44] P. F. J. Lermusiaux, *Evolving the subspace of the three-dimensional multiscale ocean variability: Massachusetts Bay*, J. Marine Syst., 29 (2001), pp. 385–422.
[45] P. F. J. Lermusiaux, *Adaptive modeling, adaptive data assimilation and adaptive sampling*, Phys. D, 230 (2007), pp. 172–196.
[46] P. F. J. Lermusiaux, T. Lolla, P. J. Haley, Jr., K. Yigit, M. P. Ueckermann, T. Sondergaard, and W. G. Leslie, *Science of Autonomy: Time-Optimal Path Planning and Adaptive Sampling for Swarms of Ocean Vehicles*, in Springer Handbook of Ocean Engineering, T. Curtin, ed., Springer, New York, 2016, pp. 481–498.
[47] P. F. J. Lermusiaux, D. N. Subramani, J. Lin, C. S. Kulkarni, A. Gupta, A. Dutt, T. Lolla, P. J. Haley, Jr., W. H. Ali, C. Mirabito, and S. Jana, *A future for intelligent autonomous ocean observing systems*, J. Marine Res., 75 (2017), pp. 765–813.
[48] A. S. Lewis and J. Malick, *Alternating projections on manifolds*, Math. Oper. Res., 33 (2008), pp. 216–234.
[49] J. Lin and P. F. J. Lermusiaux, *Minimum-Correction Second-Moment Matching: Theory, Algorithms and Applications*, to be submitted.
[50] T. Lolla, P. J. Haley, Jr., and P. F. J. Lermusiaux, *Path planning in multiscale ocean flows: Coordination and dynamic obstacles*, Ocean Model., 94 (2015), pp. 46–66.
[51] T. Lolla and P. F. J. Lermusiaux, *A Gaussian mixture model smoother for continuous nonlinear stochastic dynamical systems: Theory and schemes*, Mon. Wea. Rev., 145 (2017), pp. 2743–2761.
[52] C. Lubich and I. V. Oseledets, *A projector-splitting integrator for dynamical low-rank approximation*, BIT, 54 (2014), pp. 171–188.
[53] A. Machado and I. Salavessa, *Grassmannian manifolds as subsets of Euclidean spaces*, Differential Geom., 131 (1984).
[54] B. Mishra, G. Meyer, S. Bonnabel, and R. Sepulchre, *Fixed-rank matrix factorizations and Riemannian low-rank optimization*, Comput. Statist., 29 (2014), pp. 591–621.
[55] E. Musharbash, F. Nobile, and T. Zhou, *Error analysis of the dynamically orthogonal approximation of time dependent random PDEs*, SIAM J. Sci. Comput., 37 (2015), pp. A776–A810.
[56] P. Neff, J. Lankeit, and A. Madeo, *On Grioli's minimum property and its relation to Cauchy's polar decomposition*, Internat. J. Engrg. Sci., 80 (2014), pp. 209–217.
[57] B. Peherstorfer and K. Willcox, *Dynamic data-driven reduced-order models*, Comput. Methods Appl. Mech. Engrg., 291 (2015), pp. 21–41.
[58] M. Przybylska, *On matrix differential equations and abstract FG algorithm*, Linear Algebra Appl., 346 (2002), pp. 155–175.
[59] J. Salençon, *Mécanique des Milieux Continus: Concepts Généraux*, Vol. 1, Editions Ecole Polytechnique, 2005.
[60] T. P. Sapsis and P. F. J. Lermusiaux, *Dynamically orthogonal field equations for continuous stochastic dynamical systems*, Phys. D, 238 (2009), pp. 2347–2360.
[61] B. Schmidt, *Conditions on a connection to be a metric connection*, Comm. Math. Phys., 29 (1973), pp. 55–59.
[62] M. Spivak, *A Comprehensive Introduction to Differential Geometry*, Publish or Perish, Houston, TX, 1973.

[63] D. N. Subramani and P. F. J. Lermusiaux, *Energy-optimal path planning by stochastic dynamically orthogonal level-set optimization*, Ocean Model., 100 (2016), pp. 57–77.

[64] C. Truesdell and W. Noll, *The non-linear field theories of mechanics*, Encyclopedia Phys., 3 (1965).

[65] B. Vandereycken, *Low-rank matrix completion by Riemannian optimization*, SIAM J. Optim., 23 (2013), pp. 1214–1236.

[66] F. Yger, M. Berar, G. Gasso, and A. Rakotomamonjy, *Adaptive Canonical Correlation Analysis Based on Matrix Manifolds*, preprint, arXiv:1206.6453, 2012.