AUTO-REGRESSIVE MOVING-AVERAGE DISCRETE-TIME DYNAMICAL SYSTEMS AND AUTOCORRELATION FUNCTIONS ON REAL-VALUED RIEMANNIAN MATRIX MANIFOLDS

SIMONE FIORI
Dipartimento di Ingegneria dell’Informazione
Università Politecnica delle Marche, Via Brecce Bianche
I-60131, Ancona, Italy

(Communicated by Peter E. Kloeden)

ABSTRACT. The present research paper proposes an extension of the classical scalar Auto-Regressive Moving-Average (ARMA) model to real-valued Riemannian matrix manifolds. The resulting ARMA model on matrix manifolds is expressed as a non-linear discrete-time dynamical system in state-space form whose state evolves on the tangent bundle associated with the underlying manifold. A number of examples are discussed within the present contribution that aim at illustrating the numerical behavior of the proposed ARMA model. In order to measure the degree of temporal dependency between the state-values of the ARMA model, an extension of the classical autocorrelation function for scalar sequences is suggested on the basis of the geometrical features of the underlying real-valued matrix manifold.

1. Introduction. The literature on the subject of random matrix generation is rich in ad-hoc solutions. The paper [3] deals with the generation of random orthogonal matrices according to a Haar measure and discusses methods based on the Gram-Schmidt orthogonalization, the eigenvalue decomposition, the singular value decomposition, the Cayley transform that turns a skew-symmetric matrix into an orthogonal matrix, the Householder transformation and the Givens rotations. Direct methods to construct random orthogonal matrices are explained in the paper [21] with an application to randomizing integration methods over spherically-symmetric integration regions. The problem of generating correlation matrices (a special case of symmetric, positive-semidefinite matrices) is discussed in the paper [11], that makes use of the Bendel-Mickey algorithm to transform a symmetric positive-semidefinite matrix into a correlation matrix. Likewise, the problem of generating random points on the surface of an hypersphere has been long investigated, as illustrated, for example, in the paper [23].

Contributed papers aimed at generalizing the notion of random structured matrix generation. For example, the paper [12] presents a ‘subgroup’ algorithm for generating uniformly-distributed random matrices on compact groups, with application to matrix generation over the orthogonal group, the unitary group, the symplectic
group and the general linear group over a finite field. An account of the problem of generating random compact-group elements was also presented in [30]. It is important to note that not all the manifolds of interest in the literature qualify as algebraic groups. This is the case, for example, of the compact Stiefel manifold of ‘tall-skinny’ orthogonal matrices or the Grassmann manifold. Sampling on shape-spaces is a further example of generation of random elements on a general smooth manifold [28]. For these manifolds, a more general setting than the ones working on algebraic groups is needed.

As to what concerns modeling discrete-time sequences by non-conventional Auto-Regressive Moving-Average (ARMA) discrete-time dynamical systems, models arising from specific applications are known from the scientific literature. The paper [43] suggests the use of an ARMA model on the tangent bundle of ‘shape space’ with application to matching video shape sequences in human movement analysis. The contribution [22] considers a mixed flat/curved parameter space in a model that explains atmospheric sulfur dioxide pollution incidents. The contribution [31] presented an ARMA predictor for widely linear systems.

The present paper aims at providing a general method for generating random paths on real-valued matrix Riemannian manifolds, on the basis of their differential-geometric properties. The presented method is thought of to be sufficiently general and sufficiently practical and easy to customize to specific cases of interest. It is based on the generalization of the classical ARMA generative model to matrix Riemannian manifolds. The method for generating random paths on real-valued Riemannian matrix manifolds is presented in two versions: A version is based on parallel translation, while a version is based on a numerically-convenient approximation of parallel translation known as ‘vector transport’.

In order to measure quantitatively the statistical features of a matrix sequence generated by the devised ARMA model, an extension of the classical autocorrelation function is suggested. Such notion also allows recovering the standard notions of uncorrelatedness of a sequence on a real-valued matrix Riemannian manifold.

The paper is organized as follows. Section 2 explains the general devised method to accomplish the generation of matrix-valued random paths on real-valued Riemannian manifolds as well as the proposed notion of auto-correlation function on Riemannian matrix manifolds; Section 2 also suggests possible applications of the devised model and mentions a possible extension of the notion of partial auto-correlation function to manifold-valued temporal sequences. Section 3 examines specific cases of matrix manifolds of interest in the scientific literature. Section 4 illustrates the devised theory through several numerical examples. Section 5 concludes the paper.

2. Random paths on matrix Riemannian manifolds and their autocorrelation function. Subsection 2.1 recalls notions of differential geometry that are instrumental in the development of a geometrically-sound method to generate random paths on Riemannian manifolds. Subsections 2.2 explains the devised method and Subsection 2.3 presents the suggested extension of the classical autocorrelation function to matrix Riemannian manifolds. Subsection 2.4 mentions an extension of the notion of partial auto-correlation function. Section 2.5 mentions possible applications of the devised ARMA model.
2.1. Notions of differential geometry. Let \( M \) denote a \( p \)-dimensional Riemannian manifold, with \( p \) finite. Associated with each point \( x \) in a \( p \)-dimensional differentiable manifold \( M \) is a tangent space, denoted by \( T_x M \). This is a \( p \)-dimensional vector space whose elements can be thought of as equivalence classes of curves passing through the point \( x \). Symbol \( TM \overset{\text{def}}{=} \{(x, v) \mid x \in M, v \in T_x M\} \) denotes the tangent bundle associated with the manifold \( M \). Symbol \( \mathfrak{X}(M) \) denotes the set of vector fields on \( M \). A vector field \( \mathfrak{F} \in \mathfrak{X}(M) \) is a map \( \mathfrak{F} : x \in M \mapsto \mathfrak{F}(x) \in T_x M \).

A Riemannian manifold \( M \ni x \) is endowed with an inner product \( \langle \cdot, \cdot \rangle : T_x M \times T_x M \rightarrow \mathbb{R} \). An inner product is a positive-definite, smooth, symmetric, bilinear map that assigns a real number to pairs of tangent vectors at each tangent space of the manifold. An inner product induces a local norm, namely \( \|v\|_x \overset{\text{def}}{=} \sqrt{\langle v, v \rangle_x} \).

A fundamental notion in differential geometry is parallel translation that allows, e.g., comparing vectors belonging to different tangent spaces. On a smooth Riemannian manifold \( M \), fix a smooth curve \( \gamma : [-a, a] \rightarrow M \) \((a > 0)\) and a vector field \( \mathfrak{F} \in \mathfrak{X}(M) \). The parallel translation map \( \Gamma^s_t(\gamma) : T_{\gamma(s)} M \rightarrow T_{\gamma(t)} M \) associated with the curve \( \gamma \) is a linear isomorphism for every \( s, t \in [-a, a] \). The parallel translation map depends smoothly on its arguments and is such that \( \Gamma^s_t(\gamma) = \Gamma^t_s(\gamma) \) for every \( s, a, t \in [-a, a] \).

The notion of geodesic arc generalizes the notion of straight line of Euclidean spaces. A distinguishing feature of a straight line of an Euclidean space is that it translates parallel to itself, namely, it is self-parallel. The notion of ‘straight line’ on a curved space inherits such a distinguishing feature: A geodesic arc on a manifold \( M \) is a curve \( \gamma \) such that \( \dot{\gamma} \) is parallel translated along \( \gamma \) itself, namely, for every \( s, t \in [-a, a] \), it holds that \( \Gamma^s_t(\gamma) \dot{\gamma}(s) = \dot{\gamma}(t) \). A geodesic arc may be parametrized in terms of two parameters, such as the initial values \( \gamma(0) = x \in M \) and \( \gamma'(0) = v \in T_x M \), in which case the geodesic arc will be denoted as \( \gamma_{x,v}(t) \).

A geodesic arc \( \gamma_{x,v} : [0, 1] \rightarrow M \) is associated with the function \( \exp : TM \rightarrow M \) defined as:

\[
\exp_{x,v}(t) \overset{\text{def}}{=} \gamma_{x,v}(t)
\]

for any \((x, v) \in TM\). The function (1) is termed exponential map with pole \( x \in M \).

Given two points \( x, y \in M \) and a geodesic arc \( \gamma : [0, 1] \rightarrow M \) that joins them, namely, such that \( \gamma(0) = x \) and \( \gamma(1) = y \), the parallel translation operator \( \Gamma^y_x(\delta) \) may be simply represented as \( \Gamma^y_x : T_x M \rightarrow T_y M \), namely \( w = \Gamma^y_x(v) \), \( v \in T_x M \), \( w \in T_y M \), with the convention that the operator \( \Gamma^y_x \) coincides with the identity map in \( T_x M \).

For further references on differential geometry, readers may want to consult, e.g., the series of books [36] and the more recent book [9].

An important problem related to the implementation on a computation platform is that parallel translation may be difficult to express in closed form. In such cases, it may be replaced by the notion of vector transport, which appears as an attractive numerical alternative. In order to define the notion of vector transport, the \( p \)-dimensional manifold \( M \) is supposed to be embedded onto a large linear space \( E \) of appropriate dimension\(^1\) whose tangent bundle \( TE \) contains the bundle \( TM \). Upon endowing the tangent bundle \( TE \) by a Euclidean inner product, it is possible to define an orthogonal projection operator \( P_z : TE \rightarrow T_z M \), \( z \in M \). Define the

\(^1\)The existence of such embedding is ensured by classical results of differential geometry, which also state the best dimension of the linear space \( E \) on the basis of the structural properties of the embedding, such as the strong Whitney’s embedding theorem and the Nash isometric embedding theorem for Riemannian manifolds.
Whitney sum $TM \oplus TM = \{(x, (v, w)) | x \in M, v \in T_x M, w \in T_x M\}$. Then, the vector transport operator $vt : (x, (v, u)) \in TM \oplus TM \mapsto w \in T_{\exp_x(v)} M$ associated with an exponential map $\exp : TM \to M$ is defined as:

$$v_{t x, v}(u) \defeq P_{\exp_x(v)}(u).$$

In practice, instead of parallel-translating the tangent vector $u \in T_x M$ along a geodesic arc emanating from the point $x \in M$ in the direction $v \in T_x M$ as $\Gamma^1_{x,v}(u)$, vector transport moves rigidly the vector $u$ within $TE$ up to the point $y$ and then orthogonally-projects the vector $u$ over the vector space $T_y M$ where $y = \exp_x(v)$. With a slight abuse of the exact terminology, one may refer to vector transport of a vector $u \in T_x M$ to a tangent space $T_y M$ by $P_y(u)$ for some $y \in M$. An example of interest in the present paper concerns finite-dimensional matrix manifolds. On a Riemannian manifold $M$ of real $n \times p$ matrices (naturally embedded onto the space $\mathbb{R}^{n \times p}$), endowed with the inner product $\langle \cdot, \cdot \rangle$, the orthogonal projection operator $P_Y : \mathbb{R}^{n \times p} \to T_Y M$ at a matrix-point $Y \in M$ is defined by the two conditions:

- $P_Y(U) \in T_Y M$ for any $Y \in M$ and $U \in \mathbb{R}^{n \times p}$,
- $\langle V, P_Y(U) \rangle_Y = \text{tr}(U^T V)$ for all $V \in T_Y M$.

It is worth remarking that vector transport should be taken as a computationally-convenient shortcut that does not enjoy the structural properties of parallel translation. For instance, while parallel translation preserves the angles between translated vectors and the length of tangent vectors (namely, it represents an isometry), vector transport does not preserve such intrinsic quantities, in general (see Subsection 3.1 for a specific example).

2.2. Generation of a random path on a matrix Riemannian manifold.

The fields of differential geometry and of ARMA discrete-time dynamical system theory have touched each other in several research endeavors. A noteworthy example is offered by the paper [1], that studies the geometry of the manifold of a parametric family of ARMA$(r, q)$ systems on $\mathbb{R}$, parametrized by their coefficients, where the geometric structure is induced by the structural properties of the systems, such as invertibility. A similar approach, related to a specific implementation of bounded-input/bounded-output ARMA discrete-time filters, was described in the contribution [16]. In the paper [40], ARMA modeling on the space $\mathbb{R}^p$ of video sequences is proposed, and it is assumed that the parameters of the ARMA model lay on differentiable manifolds, hence, an inference algorithm is designed to estimate the parameters of the model on a Graßmann manifold. More recently, a study of the differential-geometrical structure of the parameter space of linear models of multiple neural spike train sequences was proposed in [35]. The aim of the present contribution is to devise an ARMA model that produces a pseudo-random temporal sequence of matrices that form a path on a matrix-type differentiable manifold. Hence, not only the parameters of the ARMA model are specially structured but the input and the output of the ARMA model are structured matrix sequences that belong to a differentiable manifold $M$ rather than to the space $\mathbb{R}$ (or $\mathbb{R}^p$).

An ARMA model on a manifold is described in terms of a pair of sequences $(x_k, v_k) \in TM$, where $x_k$ denotes the output signal of the model, $v_k$ denotes the state of the model and $k \in \mathbb{Z}$ denotes a discrete-time index. In the conventional case $M = \mathbb{R}^p$, choosing the standard identification $T_x \mathbb{R}^p \cong \mathbb{R}^p$, for any $x \in \mathbb{R}^p$,
an ARMA\((r, q)\) model\(^2\) on the tangent bundle \(T\mathbb{R}^p\) is made of a state-evolution equation and of a state-to-output transform equation. The equation that governs the evolution of the internal state \(v_k \in T_{x_k} \mathbb{R}^p\) reads:

\[
v_k = \sum_{i=1}^{r} A_i v_{k-i} + \sum_{i=0}^{q} B_i w_{k-i}, \tag{3}
\]

where transition matrices \(A_i \in \mathbb{R}^{p \times p}\) and \(B_i \in \mathbb{R}^{p \times p}\) denote the parameters of the auto-regressive part and of the moving-average part, respectively, and the sequence \(w_k \in T_{x_k} \mathbb{R}^p\) is a white-noise sequence (also termed random innovation term). The equation that describes how the internal state transforms into the output signal \(x_k \in M\) reads:

\[
x_{k+1} = x_k + v_k + c, \tag{4}
\]

with \(c \in \mathbb{R}^p\) being a constant bias (or drift term). It is worth remarking that, since the tangent bundle \(T\mathbb{R}^p\) is trivial (in fact, \(T\mathbb{R}^p\) is isomorphic to \(\mathbb{R}^p \times \mathbb{R}^p\)), the equation (3) is based on vector addition. On a general Riemannian manifold \(M\), however, vector addition in the tangent bundle \(TM\) (that is, in the AR part and in the MA part) is not defined without the concept of parallel translation. Likewise, equation (4) is based on vector addition on \(\mathbb{R}^p\), which is not defined on a general (curved) Riemannian manifold without the concept of exponential map.

The proposed extension of the above ARMA model to a Riemannian manifold \(M\) is based on the following operations:

1. At each step \(k\), it is necessary to compute a linear combination of several tangent vectors in \(T_{x_k} M\) via transition operators \(A_i : (x, v) \in TM \mapsto w \in T_x M\) for the auto-regressive part and \(B_i : (x, v) \in TM \mapsto w \in T_x M\) for the moving-average part, where operators \(A_i\) and \(B_i\) are linear in the second argument. Note that \(A_i, B_i \in \mathfrak{X}(M)\), namely, they are vector fields depending on a variable.

2. In order to compute the linear combination of previous and current values of the state variable, at each step \(k\), it is necessary to parallel translate each tangent vector \(v_{k-1}, \ldots, v_{k-r}, w_{k-1}, \ldots, w_{k-q}\) to the tangent space \(T_{x_k} M\) from the tangent spaces that they belong to.

3. In order to propagate the constant bias at each step \(k\), it is necessary to parallel translate the constant term to each tangent space \(T_{x_k} M\).

On the basis of the above considerations, an ARMA\((r, q)\) model on the tangent bundle \(TM\) of the real-valued finite-dimensional matrix manifold \(M\) may be written as:

\[
\begin{aligned}
V_k &= \sum_{i=1}^{r} A_i \left( X_k, \Gamma_{X_{k-1}}^{X_k}(V_{k-i}) \right) + \sum_{i=0}^{q} B_i \left( X_k, \Gamma_{X_{k-i}}^{X_k}(W_{k-i}) \right), \\
X_{k+1} &= \exp_{X_k}(V_k + \Gamma_{X_0}^{X_k}(C)),
\end{aligned} \tag{5}
\]

for \(k = 0, 1, 2, \ldots\). To start the generation of a random path, namely, for \(k = 0\), it is necessary to select the initial seed \(X_0 \in M\) as well as the constant bias \(C \in T_{X_0} M\), and to put into effect an extension of standard zero-padding:

\[
\begin{aligned}
V_{-1} &= \ldots = V_{-r} = 0, \\
W_{-1} &= W_{-2} = \ldots = W_{-q} = 0, \\
X_{-1} &= X_{-2} = \ldots = X_{-\beta} = X_0, \tag{6}
\end{aligned}
\]

\(^2\)In order to distinguish the scalar ARMA models from the vector-type ARMA models, the latter may be denoted as VARMA models [37] or, in the auto-regressive instance only, as VAR models [41].
Algorithm 1 Pseudocode to implement the ARMA($r, q$) random structured matrix process generation rules (5) and (7) on manifold $M$ with $K$ time-steps.

1: Choose a metric for the manifold $M$ and compute the exponential map $\exp : TM \to M$ and the parallel-transport operator $\Gamma^X_X : T_X M \to T_Y M$ or the vector-transport operator $P_X$, with $X, Y \in M$ (see Section 3 for examples on specific manifolds)
2: Choose the initial seed $X_0 \in M$ and the initial drift $C \in T_{X_0} M$
3: Choose $r$ operators $A_i : (X, V) \in TM \mapsto W \in T_X M$ and $q$ operators $B_i : (X, V) \in TM \mapsto W \in T_X M$ (see Section 4 for examples on specific manifolds)
4: Set $\beta \defeq \max\{r, q\}$, $V_{-1} = \ldots = V_{-r} = 0$, $W_{-1} = W_{-2} = \ldots = W_{-q} = 0$, $X_{-1} = X_{-2} = \ldots = X_{-\beta} = X_0$
5: for $k = 0$ to $K$ do
6: Generate a random tangent matrix $W_k \in T_{X_k} M$ by a random generator on a linear space (see Section 3 for examples on specific manifolds)
7: Compute the matrices $V_{k,i} = \Gamma^{X_k}_{X_{k-i}}(V_{k-i})$ and $W_{k,j} = \Gamma^{X_k}_{X_{k-j}}(W_{k-j})$, or $V_{k,i} = P_{X_k}(V_{k-i})$ and $W_{k,j} = P_{X_k}(W_{k-j})$, for $i = 1, \ldots, r$ and $j = 0, \ldots, q$
8: Compute the matrix $V_k = \sum_{i=1}^r A_i(X_k, V_{k,i}) + \sum_{j=0}^q B_j(X_k, W_{k,j})$ and the actualized drift term $D_k = \Gamma^{X_k}_{X_k}(C)$ or $D_k = P_{X_k}(C)$
9: Compute the next sample $X_{k+1} = \exp_{X_k}(V_k + D_k)$
10: end for

with $\beta = \max\{r, q\}$. The tangent vectors $V_k$ are computed as the output of an autoregressive moving-average model on the tangent bundle $TM$. In an ARMA($r, q$) model with state-variable $V_k \in T_{X_k} M$ and input variable $W_k \in T_{X_k} M$, the current state variable value $V_k$ depends linearly of the past state-variable values $V_{k-1}, \ldots, V_{k-r}$ and of the current and past input-variable values $W_k, W_{k-1}, \ldots, W_{k-q}$. At each step $k$, it is necessary to generate only a new random matrix $W_k \in T_{X_k} M$. The additive term $\Gamma^{X_k}_{X_k}(C)$ in the second equation of the system (5) represents a drift term that slightly moves the current state even when the scalar velocity $\|V_k\|_{X_k}$ is close to zero.

The random path generation rule (5), where parallel translation is replaced by vector transport, reads:

$$\left\{ \begin{array}{l} V_k = \sum_{i=1}^r A_i(X_k, P_{X_k}(V_{k-i})) + \sum_{j=0}^q B_j(X_k, P_{X_k}(W_{k-i})), \\ X_{k+1} = \exp_{X_k}(V_k + P_{X_k}(C)), \end{array} \right. \quad (7)$$

for $k = 0, 1, 2, \ldots$. In some circumstances, even if a parallel-translation formula is available, vector transport is to be preferred due to easier implementation and lighter computational burden.

The steps to implement the ARMA($r, q$) random process generation rules (5) and (7) are explained in Algorithm 1.

By adding extra terms to the right-hand side of the first equation of the system (5) or (7), the discussed ARMA model on manifold may be extended to an ARMAX (ARMA with exogenous inputs). Such extension is not pursued in the present paper.

2.3. An auto-correlation function of a matrix-sequence on a real-valued Riemannian manifold. Autocorrelation refers to the correlation of a time series with its own past and future values [6]. Large positive autocorrelation denounces
'persistence', the tendency of a system to remain in the same state from one observation to the next. The set of autocorrelation coefficients arranged as a function of separation in time (or lag) is the sample autocorrelation function.

In the conventional case of the manifold \( M = \mathbb{R}^p \) (endowed with Euclidean geometry), the sample autocorrelation function \( R^v \) associated with a sequence \( v_k \in \mathbb{R}^p, k = 0, \ldots, K, \) is defined as:

\[
R^v(m, n) = \frac{1}{K - \max\{m, n\}} \sum_{k=1}^{K - \max\{m, n\}} v_{k+n}v_{k+m}, \quad m, n = 0, \ldots, L, \tag{8}
\]

where \( m \) and \( n \) denote the lags at which the correlation is computed. The constant integers \( K \) and \( L \) denote, respectively, the number of samples of the sequence \( v_k \) and the maximum considered lag.

In the case that the manifold \( M \) is a matrix Riemannian manifold endowed with the scalar product \( \langle \cdot, \cdot \rangle \), the above definition of sample autocorrelation function may be extended to the sequence \( (X_k, V_k) \in TM, k = 0, \ldots, K, \) as:

\[
R^V(m, n) = \frac{1}{K - \max\{m, n\}} \sum_{k=1}^{K - \max\{m, n\}} \langle \Gamma^X_{X_{k+n}}(V_{k+n}), \Gamma^X_{X_{k+m}}(V_{k+m}) \rangle x_k, \tag{9}
\]

where \( m, n = 0, \ldots, L \). The function \( R^V \) is symmetric in its integer arguments, namely \( R^V(m, n) = R^V(n, m) \), because the inner product is a symmetric operator.

The above definition of autocorrelation function for a discrete-time sequence on a tangent bundle \( TM \) of a real-valued Riemannian manifold inherits several properties of the autocorrelation functions on the space \( \mathbb{R}^p \).

If the correlation coefficients \( R^V(m, n) \) depend only on the difference \( m - n \) (namely, the matrix with entries \( R^V(m, n) \) is 'banded'), the sequence \( (X_k, V_k) \) is termed second-order stationary. In such case, the only independent entries are those on the row \( R^V(m, 0) \) (or, equivalently, on the column \( R^V(0, n) \)), hence the correlation function may be denoted as \( R^V(m) \). The one-parameter autocorrelation function is defined, explicitly, as:

\[
R^V(m) = \frac{1}{K - m} \sum_{k=1}^{K - m} \langle V_k, \Gamma^X_{X_{k+m}}(V_{k+m}) \rangle x_k, \tag{10}
\]

where \( m = 0, \ldots, L \). A sequence \( (X_k, V_k) \) is said to be uncorrelated (or white noise) if \( R^V(m) = 0 \) for \( m \geq 1 \).

The diagonal entries \( R^V(m, m) \) of an autocorrelation function possess a particular meaning. Such terms may be written explicitly as:

\[
R^V(m, m) = \frac{1}{K - m} \sum_{k=1}^{K - m} \langle \Gamma^X_{X_{k+m}}(V_{k+m}), \Gamma^X_{X_{k+m}}(V_{k+m}) \rangle x_k \\
= \frac{1}{K - m} \sum_{k=1}^{K - m} \| \Gamma^X_{X_{k+m}}(V_{k+m}) \|^2_{X_k}, \tag{11}
\]

Since parallel translation along a geodesic line is an isometry with respect to its own metrics, it holds that \( \| \Gamma^X_{X_{k+m}}(V_{k+m}) \|^2_{X_k} = \| V_{k+m} \|^2_{X_{k+m}} \), therefore:

\[
R^V(m, m) = \frac{1}{K - m} \sum_{k=1}^{K - m} \| V_{k+m} \|^2_{X_{k+m}} = \frac{1}{K - m} \sum_{k=m}^{K} \| V_k \|^2_{X_k}. \tag{12}
\]
Under the hypothesis of stationarity, in particular, all the coefficients \( R^V(m, m) \) coincide with the one-parameter autocorrelation function at zero lag, namely, with:

\[
R^V(0) = \frac{1}{K} \sum_{k=1}^{K} \| V_k \|^2_{X_k},
\]

which coincides with the Riemannian variance of a random distribution on the tangent bundle \( TM \). Clearly \( R^V(0) \geq 0 \). By using the Cauchy-Schwarz inequality for inner products and the fact that parallel translation is an isometry, it is straightforward to show that \( R^V(m) \leq \frac{1}{K-m} R^V(0) \) for every lag \( m \).

The above definition of autocorrelation function may be extended to include the drift term that appears in the second equation of the system (5) as follows:

\[
R^V(m) \overset{\text{def}}{=} \frac{1}{K - m} \sum_{k=1}^{K-m} \langle V_k + \Gamma^{X_k}_{X_0}(C), \Gamma^{X_{k+m}}_{X_0}(V_{k+m} + \Gamma^{X_{k+m}}_{X_0}(C)) \rangle_{X_k}, \ m = 0, \ldots, L.
\]

(14)

The presence of the drift term in the definition (14) may be shown to cause approximately a lift of the autocorrelation function (10) for small values of the lag \( m \) compared to the length \( K \) of the sequence. In fact, by the properties of the parallel translation operator \( \Gamma \) and by the bilinearity of the inner product \( \langle \cdot, \cdot \rangle \) it follows that:

\[
\begin{align*}
\langle V_k + \Gamma^{X_k}_{X_0}(C), \Gamma^{X_{k+m}}_{X_0}(V_{k+m} + \Gamma^{X_{k+m}}_{X_0}(C)) \rangle_{X_k} = \\
\langle V_k + \Gamma^{X_k}_{X_0}(C), \Gamma^{X_{k+m}}_{X_0}(V_{k+m}) + \Gamma^{X_{k+m}}_{X_k+m} \circ \Gamma^{X_{k+m}}_{X_0}(C) \rangle_{X_k} = \\
\langle V_k, \Gamma^{X_{k+m}}_{X_k+m}(V_{k+m}) \rangle_{X_k} + \langle V_k, \Gamma^{X_{k+m}}_{X_k+m} \circ \Gamma^{X_{k+m}}_{X_0}(C) \rangle_{X_k} + \\
\langle \Gamma^{X_k}_{X_0}(C), \Gamma^{X_{k+m}}_{X_k+m}(V_{k+m}) \rangle_{X_k} + \langle \Gamma^{X_k}_{X_0}(C), \Gamma^{X_{k+m}}_{X_k+m} \circ \Gamma^{X_{k+m}}_{X_0}(C) \rangle_{X_k}.
\end{align*}
\]

In the hypothesis that \( \Gamma^{X_k}_{X_k+m} \circ \Gamma^{X_{k+m}}_{X_0}(C) \approx \Gamma^{X_k}_{X_0}(C) \), it is convenient to define the quantity:

\[
Z^V(m) \overset{\text{def}}{=} \frac{1}{K - m} \sum_{k=1}^{K-m} \langle V_{k+m}, \Gamma^{X_{k+m}}_{X_0}(C) \rangle_{X_{k+m}}.
\]

(15)

For \( m \ll K \), the autocorrelation function (14) may be approximated as:

\[
R^V(m) \approx R^V(0) + 2Z^V(0) + \langle C, C \rangle_{X_0}.
\]

(16)

It is worth underlining that, whenever vector transport is used instead of parallel translation in the discrete-time dynamical system that implements the suggested ARMA model on manifolds, the only geometrical notion retained from the general theory of Riemannian manifolds is that of inner product, which is necessary to define the autocorrelation function (10). In this case, the autocorrelation function associated with the sequence \((X_k, V_k) \in TM, k = 0, \ldots, K\), reads:

\[
R^V(m) \overset{\text{def}}{=} \frac{1}{K - m} \sum_{k=1}^{K-m} \langle V_k, P_{X_k}(V_{k+m}) \rangle_{X_k}, \ m = 0, \ldots, L.
\]

(17)

2.4. Partial auto-correlation function. In a similar manner, it would be possible to construct a partial autocorrelation function, which is typically used to infer the order of an autoregressive model [6] (interested readers might also want to consult the classical paper by F.L. Ramsey on the characterization of the partial autocorrelation function [34]).
The Partial Auto-Correlation Function is an important tool, often used in association with the Auto-Correlation Function, in the theory of scalar (linear) ARMA models. An important application is the estimation of the nature of a model underlying a given temporal series (namely, AR, MA or mixed) and the estimation of the order of a model, thanks to its prominent cutoff property. The Partial Auto-Correlation Function at a given lag represents the correlation between a time series and its lagged version once the correlation between these and the intermediate-lagged sequences has been removed. In other terms, the value of the Partial Auto-Correlation Function at a given lag accounts for the degree of autocorrelation not explained by the values of the Partial Auto-Correlation Functions corresponding to intervening lags.

The definition of Partial Auto-Correlation Function and its properties are essentially related with the linear structure of the model underlying the data. For non-linear models, the properties of the Partial Auto-Correlation Functions differ (see, for instance, the study [24], which shows that, for non-linear ARMA or VARMA models, it is necessary to introduce other quantities, such as the ‘R’ function based on mutual entropy, to detect the order of a non-linear model). Since the models (5) and (7) are inherently non-linear because they are formulated on curved spaces, it is unknown, at this stage, whether it will be possible to recover the useful properties of the Partial Auto-Correlation Functions of signals generated/explained by linear models.

2.5. Impact on practical applications. Although the present research topic is still at the seminal stage, some possible applications of the general ARMA\((r, q)\) model on Riemannian manifolds may be suggested on the basis of existing works, published in the scientific literature, that make use of simple models tailored to specific spaces.

An example of possible application is offered by the contribution [38] that deals with robust state estimation algorithms on manifolds. With a growing geometric awareness among state estimation practitioners, geometrically sound algorithms tailored for particular applications are emerging. A common application is the estimation of the orientations of physical objects, which involve the manifold of three-dimensional rotations on the ordinary space, denoted as SO(3). Such application utilizes an ARMA\((1, 0)\) system to model a sequence of points on the manifold SO(3) and employs a particle filtering technique to estimate the state of such model. The application of a random path generator on a manifold might, therefore, be used to produce temporal sequences of data, generated with a known model, to be used in the testing/tuning stage of such state estimation algorithm.

A further example of possible application is offered by the paper [4]. Visual ego-motion estimation refers to the process of estimating the three-dimensional camera pose based on bi-dimensional image sequences captured by a camera. Visual ego-motion estimation plays an important role in computer vision and robotics applications such as visual simultaneous localization and augmented reality. A camera pose can be represented by a matrix belonging to the special Euclidean group SE(3). The state equation utilized in the contribution [4] to represent camera pose during motion is an ARMA\((1, 0)\) model on the manifold SE(3). A possible application of the general ARMA model (5) is to provide a more accurate modeling of a camera pose during motion by making use of an ARMA model with a wider memory span.
A potentially far-reaching application is suggested again by the manuscript [4] that introduces a Particle Swarm Optimization (PSO) method suitable for differentiable manifolds. A close examination of the equations that constitute the PSO algorithm on manifold reveals that such algorithm is formulated as a distributed system of optimization sub-algorithms which locally are ARMA systems insisting on the same manifold. The structural properties of such optimization algorithm, such as its convergence, are likely to emerge from the study of the structural properties of ARMA systems on manifolds. Therefore, an application of the general theory being developed is to study the structural properties of centralized recursive algorithms and distributed recursive sub-algorithms.

The actual realization of the above-mentioned applications still need some further theoretical advancements in the understanding of ARMA models on manifolds, such as the study of structural properties (such as asymptotic stability) and the study of deep relationships between the order and type (AR, MA or mixed) of a model and its auto-correlation/partial autocorrelation functions. Such topics are currently under investigation.

3. Matrix manifolds of interest. The present Section recalls the differential-geometric features of some matrix manifolds of interest in applications in view of performing numerical simulations in Section 4. In particular, the present Section examines the cases of compact Stiefel manifold and of the group of symmetric positive-definite matrices as well as related manifolds of interest in the scientific literature.

3.1. Compact Stiefel manifold and related manifolds. The Compact Stiefel manifold is defined as:

$$\text{St}(n, p) = \{ X \in \mathbb{R}^{n \times p} | X^T X = I_p, \ p \leq n \},$$

(18)

where superscript $^T$ denotes matrix transpose and symbol $I_p$ denotes a $p \times p$ identity matrix\(^3\). The tangent space to the manifold $\text{St}(n, p)$ in a point $X \in \text{St}(n, p)$ has structure:

$$T_X \text{St}(n, p) = \{ V \in \mathbb{R}^{n \times p} | V^T X + X^T V = 0_p \},$$

(19)

where symbol $0_p$ denotes an all-zero $p \times p$ matrix. Exemplary applications where the orthogonality constraint plays a prominent role are blind source separation upon signal pre-whitening and independent component analysis [7], non-negative matrix factorization [47], best basis search/selection [2], electronic structures computation within local density approximation [13] and factor analysis in psychometrics [14].

A metric that the Stiefel manifold may be endowed with is the Euclidean metric:

$$\langle V, W \rangle_X \overset{\text{def}}{=} \text{tr}(V^T W), \ V, W \in T_X \text{St}(n, p),$$

(20)

where $(X, V) \in T\text{St}(n, p)$ and symbol $\text{tr}(\cdot)$ denotes matrix trace. The metric defined in (20) may be said to be Euclidean because it does not depend on the point $X$ (namely, it is uniform). However, this does not mean that the space $\text{St}(n, p)$ is an Euclidean (flat) manifold. In fact, $\text{St}(n, p)$ is a curved Riemannian manifold and parallel translation of tangent vectors (namely, of matrices of the type (19)) along

\(^3\)In the present manuscript, with a slight abuse of terminology, the name ‘Stiefel manifold’ is used to denote the compact Stiefel manifold $\text{St}(n, p, \mathbb{R})$ of real-valued matrices, as opposed to the compact Stiefel manifold $\text{St}(n, p, \mathbb{C})$ of complex-valued matrices and the non-compact Stiefel manifolds, where symbol $\mathbb{C}$ denotes the field of complex numbers.
geodesic lines is not trivial. The structure of a geodesic arc in the Euclidean metric reads [13]:

$$\gamma_{X, V}(t) = [X \ V] \text{Exp} \left( t \left[ \begin{array}{cc} X^T V & -V^T V \\ I_p & X^T V \end{array} \right] \right) I_{2p, p} \text{Exp}(-tX^T V),$$

(21)

where symbol $\text{Exp}(\cdot)$ denotes matrix exponential and symbol $I_{r, s}$ denotes a $r \times s$ pseudo-identity matrix. The exponential of a matrix $X \in \mathbb{R}^{p \times p}$ is defined by the series:

$$\text{Exp}(X) \overset{\text{def}}{=} I_p + \sum_{r=1}^{\infty} \frac{X^r}{r!}.$$  

(22)

In numerical linear algebra, there is a large body of numerical recipes to compute the exponential of a matrix (see, for example, [25]). The orthogonal projection operator $P_Y : \mathbb{R}^{n \times p} \rightarrow T_Y \text{St}(n, p)$, for $Y \in \text{St}(n, p)$, associated with the Euclidean metrics on the compact Stiefel manifold, reads:

$$P_Y(U) = \left( I_n - \frac{1}{2} YY^T \right) U - \frac{1}{2} YU^T Y,$$

(23)

with $U \in \mathbb{R}^{n \times p}$. It can be used to compute vector transport in the equations (7).

The compact Stiefel manifold may be endowed with a second kind of metric, termed canonical metric. The associated inner product reads:

$$\langle V, W \rangle_X \overset{\text{def}}{=} \text{tr} \left( V^T \left( I_n - \frac{1}{2} XX^T \right) W \right), \quad V, W \in T_X \text{St}(n, p),$$

(24)

which, unlike the Euclidean metric (20), is not uniform over the Stiefel manifold. The structure of a geodesic arc in the canonical metric reads [13]:

$$\gamma_{X, V}(t) = [X \ Q] \text{Exp} \left( t \left[ \begin{array}{cc} X^T V & -R^T \\ R & 0_p \\ 0_p & 0_p \end{array} \right] \right) \left[ \begin{array}{c} I_p \\ 0_p \end{array} \right],$$

(25)

where $(X, V) \in T \text{St}(n, p)$ and $Q$ and $R$ denote the factors of the compact QR decomposition of the matrix $(I_n - XX^T)V$. The orthogonal projection operator $P_Y : \mathbb{R}^{n \times p} \rightarrow T_Y \text{St}(n, p)$, for $Y \in \text{St}(n, p)$, associated with the canonical metrics on the compact Stiefel manifold:

$$P_Y(U) = U - YU^T Y,$$

(26)

with $U \in \mathbb{R}^{n \times p}$. For the convenience of the readers, a derivation of the expression (26) is reported in the Appendix A, along with a short discussion on the differences between parallel translation and vector transport based on the projector (26).

The unit hypersphere $S^{n-1} \overset{\text{def}}{=} \{ x \in \mathbb{R}^n : x^T x = 1 \}$ coincides with the Stiefel manifold $\text{St}(n, p)$ with $p = 1$. Moreover, on the Stiefel manifold $\text{St}(n, 1)$, the Euclidean metric and the canonical metric coincide, namely, they collapse into the uniform inner product:

$$\langle v, w \rangle_x = v^T w, \quad \forall v, w \in T_x S^{n-1},$$

(27)

where:

$$T_x S^{n-1} = \{ v \in \mathbb{R}^n : v^T x = 0 \}.$$  

(28)

There is a number of applications that deal with the manifold $S^{n-1}$, as, for instance, blind deconvolution algorithms [15, 17], algorithms for data classification by linear discrimination based on non-gaussianity discovery [33] and algorithms for adaptive pattern recognition [48]. Many important algorithms developed in robotics and related areas require sampling over spheres [45]. For example, the paradigm of sampling-based motion planning has led to algorithms that can solve challenging
problems by combining collision detection, search algorithms and sampling strategies over the configuration space. General sampling over spheres arises in many forms of planning and optimization in which some number of directions are locally explored. The norm associated with the inner product \( \langle x, y \rangle \) is the standard vector 2-norm \( \| \cdot \| \) and the associated geodesic \( \gamma_{x,v}(t) \), with \( (x, v) \in TS^{n-1} \) and \( t \in [0, 1] \), may be given a closed-form expression, for \( v \neq 0 \), that is:

\[
\gamma_{x,v}(t) = x \cos\left(\|v\|t\right) + v\|v\|^{-1} \sin\left(\|v\|t\right),
\]

(29)

which represents a great circle on the hypersphere, while \( \gamma_{x,0}(t) = x \) for any value of the parameter \( t \). Given a geodesic arc \( \gamma_{x,v} : [0, 1] \to S^{n-1} \) with \( (x, v) \in TS^{n-1} \) and a tangent vector \( u \in T_xS^{n-1} \), the expression of the parallel-translation operator that translates the vector \( u \) along the geodesic arc \( \gamma_{x,v}(t) \) toward the endpoint \( t = 1 \) is known:

\[
\Gamma^1_0(\gamma_{x,v})u = u - \frac{v^T u}{\|v\|} \left( x \sin(\|v\|) + v \frac{1 - \cos(\|v\|)}{\|v\|} \right).
\]

(30)

For a reference, readers might want to consult [19]. Starting from this, some mathematical work leads to the following result: Given two points \( x, y \in S^{n-1} \) and a vector \( u \in T_xS^{n-1} \), the parallel translation operator \( \Gamma^y_x : T_xS^{n-1} \to T_yS^{n-1} \) has the following structure:

\[
\Gamma^y_x(u) = \left[ I_n - \frac{(I_n - xx^T)y y^T}{1 + x^T y} - xy^T \right] u,
\]

(31)

provided that \( x^T y \neq -1 \) (namely, the points \( x \) and \( y \) are not antipodal one to another). Note that setting \( y = x \) in the equation (31) yields the expected result \( \Gamma^x_x(u) = \left[ I_n - (I_n - xx^T)x x^T / 2 - xx^T \right] u = u \), for any \( u \in T_xS^{n-1} \). Moreover, it is straightforward to show that \( \Gamma^y_x(u) \in T_yS^{n-1} \) because \( y^T \Gamma^y_x(u) = 0 \) for any \( x, y \in S^{n-1} \) and \( u \in T_xS^{n-1} \).

A manifold related to the hypersphere is the oblique manifold \([39]\), defined as:

\[
\text{OB}(n) = \{ X \in \mathbb{R}^{n \times n} | \text{diag}(X^TX) = I_n \},
\]

(32)

where \( \text{diag}(\cdot) \) returns the zero matrix except for the main diagonal that is copied from the main diagonal of its argument. The following identification holds true:

\[
\text{OB}(n) \cong S^{n-1} \times S^{n-1} \times \cdots \times S^{n-1},
\]

(33)

where \( n \) times

hence, each column of a \( \text{OB}(n) \)-matrix may be treated as a \( S^{n-1} \)-vector.

The hyper-ellipsoid \( L_{n-1}^{n-1} \) is defined as:

\[
L_{n-1}^{n-1} = \{ x \in \mathbb{R}^n | \|DRx\|^2 = 1 \},
\]

(34)

with \( D \in \mathbb{R}^{n \times n} \) being diagonal and positive-definite and \( R \in \mathbb{R}^{n \times n} \) being such that \( R^T R = I_n \) and \( \text{det}(R) = 1 \). The hyper-ellipsoid \( L_{n-1}^{n-1} \) is related to the hypersphere \( S^{n-1} \) and its geometry may be studied on the basis of the geometry of the hypersphere. The hyper-ellipsoid \( L_{n-1}^{n-1} \) is used, e.g., in the calibration of magnetometers [42].

A generalization of the Stiefel manifold, recently introduced in the context of principal subspace tracking in presence of a colored noise, has been studied in the contribution [46]. Such generalized Stiefel manifold is defined as:

\[
\text{St}_B(n,p) = \{ X \in \mathbb{R}^{n \times p} | X^T B X = I_p \},
\]

(35)
where $B$ denotes any symmetric, positive-definite $n \times n$ matrix. The contribution [46] studied the structure of the tangent bundle of the generalized Stiefel manifold and suggests a computationally-convenient retraction map on it (that is, a numerically-convenient approximation of the exponential map [26]).

3.2. Group of symmetric positive-definite matrices and related manifolds.
Symmetric positive-definite matrices find a wide range of applications. For instance, symmetric positive-definite matrices are applied in automatic and intelligent control [8], in computational biology [10], in cognitive computation [18] as well as in psychophysics [27].

The manifold of symmetric positive-definite matrices is defined as:

$$S^+(n) = \{X \in \mathbb{R}^{n \times n} | X^T - X = 0_n, \ X > 0_n\}. \quad (36)$$

The tangent spaces have structure:

$$T_X S^+(n) = \{V \in \mathbb{R}^{n \times n} | V^T - V = 0_n\}, \quad (37)$$

namely, every tangent space $T_X S^+(n)$ coincides with the set of $n \times n$ symmetric matrices. The canonical metric adopted for the manifold of symmetric positive-definite matrices is:

$$\langle V, W \rangle_X = \text{tr}(V X^{-1} W X^{-1}), \quad (38)$$

whose associated geodesic curve reads:

$$\gamma_{X,V}(t) = \sqrt{X} \text{Exp} \left( t \sqrt{X} V \sqrt{X}^{-1} \right) \sqrt{X}, \quad (39)$$

where the symbol $\sqrt{X}$ denotes a symmetric square root. The above expression requires the evaluation of square root of a symmetric positive-definite matrix which exists surely. By recalling that $X^{-1} \text{Exp}(V) X = \text{Exp}(X^{-1} V X)$, the above expression simplifies into

$$\gamma_{X,V}(t) = X \text{Exp}(t X^{-1} V) = \text{Exp}(t V X^{-1}) X. \quad (40)$$

Given a geodesic arc $\gamma_{X,V} : [0, 1] \to S^+(n)$ and a tangent vector $U \in T_X S^+(n)$, the expression of the parallel-translation operator that translates the tangent vector $U$ along the geodesic arc $\gamma_{X,V}(t)$ toward the endpoint $t = 1$ reads:

$$\Gamma^1_{0}(\gamma_{X,V}) U = \text{Exp} \left( V X^{-1} \over 2 \right) \text{Exp} \left( X^{-1} V \over 2 \right). \quad (41)$$

Some mathematical work leads to the following result: Given two points $X, Y \in S^+(n)$ and a tangent vector $U \in T_X S^+(n)$, the parallel translation operator $\Gamma^Y_X : T_X S^+(n) \to T_Y S^+(n)$ has the following structure:

$$\Gamma^Y_X(U) = \sqrt{Y X^{-1}} U \sqrt{X^{-1} Y}. \quad (42)$$

Note that setting $Y = X$ in equation (42) immediately yields the expected result $\Gamma^X_X(U) = U$, for any $U \in T_X S^+(n)$. Moreover, it may be verified that $\Gamma^Y_X(U) \in T_Y S^+(n)$ for any $X, Y \in S^+(n)$ and $U \in T_X S^+(n)$.

A related problem arises about the manifold $S^+(n, p)$ of symmetric fixed-rank positive-semidefinite matrices, that may be defined as:

$$S^+(n, p) \overset{\text{def}}{=} \{U R^2 U^T | U \in \text{St}(n, p), \ R^2 \in S^+(p), \ p < n\}. \quad (43)$$

The interest in symmetric fixed-rank positive-semidefinite matrices stems from the observation that, with the growing use of low-rank approximations of matrices as a way to retain tractability in large-scale applications, there is a need to extend the calculus of positive-definite matrices to their low-rank counterparts. The manifold
of symmetric fixed-rank positive-semidefinite matrices admits the quotient representation:

\[ S^+(n, p) \cong (\text{St}(n, p) \times S^+(p))/O(p), \]  

(44)

that expresses the fact that any matrix \( X \in S^+(n, p) \) may be represented as \( X = U R^2 U^T \), with \((U, R^2) \in \text{St}(n, p) \times S^+(p)\), up to the equivalence relation \((U, R^2) \sim (UP, P^T R^2 P)\) for any \( P \in O(p)\). On the basis of the quotient representation (44), paper [29] proposes a Riemannian structure for the space \( S^+(n, p) \), which is based on a metric that is defined as the linear combination of the Euclidean metric in \( \text{St}(n, p) \) (which do not correspond to Riemannian distances, though).

4. Numerical examples. The present Section illustrates the features of the discussed method to generate random paths on real-valued Riemannian matrix manifolds by a number of numerical examples. The general ARMA\((r, q)\) model on manifolds developed in Subsection 2.2 and the notion of autocorrelation function on manifold developed in Subsection 2.3 are applied to the unit hypersphere, to the manifold of symmetric positive-definite matrices and to the compact Stiefel manifold. In particular, examples about how to choose the auto-regressive and the moving-average operators \( A_i \) and \( B_i \) are given and practical details about the implementation of the relative ARMA models are explained as well.

4.1. Generation of random paths on the unit hypersphere. In order to generate a random vector \( v \in T_2^{n-1} \), given a point \( x \in S^{n-1} \), the following scheme may be put into effect:

- Generate a random vector \( \tilde{v} \in \mathbb{R}^n \) by a random scalar generator;
- Project the vector \( \tilde{v} \) into the tangent space \( T_2^{n-1} \) by the rule \( v = \tilde{v} - xx^T \) \( x \).

The generation of a random path over the manifold \( S^{n-1} \) requires to choose the parameters that take part in the ARMA\((r, q)\) model (5). In particular, a way to select the transition operators \( A_i : (x, v) \in TS^n \mapsto w \in T_2^{n-1} \) and \( B_i : (x, v) \in T_2^{n-1} \mapsto w \in T_2^{n-1} \) is:

\[
\begin{cases}
A_i(x, v) = (I_n - xx^T)A_i v, & i \geq 1, \\
B_i(x, v) = (I_n - xx^T)B_i v, & i \geq 0,
\end{cases}
\]  

(45)

where matrices \( A_i, B_i \in \mathbb{R}^{n \times n} \) may be chosen arbitrarily.

In the present numerical example, it was chosen \( n = 3, r = 2 \) and \( q = 4 \). Moreover, the implementation based on parallel translation was opted. The resulting random-path generation rule ARMA\((2, 4)\) on the manifold \( S^2 \) may be particularized from the general expression (5) as follows:

\[
v_k = (I_3 - x_kx_k^T)A_1 \Gamma^{x_k}_{x_{k-1}}(v_{k-1}) + (I_3 - x_kx_k^T)A_2 \Gamma^{x_k}_{x_{k-2}}(v_{k-2}) + (I_3 - x_kx_k^T)B_0 w_k + (I_3 - x_kx_k^T)B_1 \Gamma^{x_k}_{x_{k-1}}(w_{k-1}) + (I_3 - x_kx_k^T)B_2 \Gamma^{x_k}_{x_{k-2}}(w_{k-2}) + (I_3 - x_kx_k^T)B_3 \Gamma^{x_k}_{x_{k-3}}(w_{k-3}) + (I_3 - x_kx_k^T)B_4 \Gamma^{x_k}_{x_{k-4}}(w_{k-4}),
\]

(46)

\[
x_{k+1} = \exp_{x_k}(v_k + \Gamma^{x_k}_{x_0}(c)),
\]

(47)

where the exponential map on the hypersphere is implemented as:

\[
\exp_{x}(v) = x \cos(\|v\|) + v\|v\|^{-1} \sin(\|v\|),
\]

(48)
with \( x_k \in S^2 \), \( w_k, v_k \in T_{x_k}S^2 \) for \( k = 0, 1, 2, \ldots \) and \( c \in T_{x_0}S^2 \) generated randomly by projection over the tangent space \( T_{x_0}S^2 \).

An example of random path generated over the bundle \( TS^2 \) by an ARMA(2,4) model is illustrated in the Figure 1. The initial seed was chosen as the vector \( x_0 = \frac{1}{\sqrt{3}}[-1 -1 1]^T \). A total of \( K = 5,000 \) temporal samples were generated.

The Figure 2 illustrates the autocorrelation function \( R^v(m, n) \) associated with the random path \((x_k, v_k) \in TS^2\). The maximum lag was chosen as \( L = 10 \). The autocorrelation function \( R^v(m, n) \) is clearly symmetric and banded. Such numerical results confirms that, for small lags \( m, n \), the informative part of the autocorrelation function \( R^v(m, n) \) may be retained as \( R^v(m, 0) \).

The sample auto-correlation function \( R^v(m) \) associated with the sequence \((x_k, v_k) \in TS^2, k = 0, \ldots , K\), reads:

\[
R^v(m) = \frac{1}{K-m} \sum_{k=1}^{K-m} v_k^T \Gamma x_{k+m} (v_{k+m}), \; m = 0, \ldots , L.
\] (49)

The Figure 3 illustrates the autocorrelation function \( R^w(m) \) associated with the random path \((x_k, w_k) \in TS^2\) and the autocorrelation function \( R^w(m) \) associated with the random path \((x_k, w_k) \in TS^2\). The maximum lag was chosen as \( L = 20 \). The

**Figure 1.** Example of random path generated over the sphere \( S^2 \) by an ARMA(2,4) model \((5)\). The initial seed is denoted by a diamond and the generated points are denoted by dots. For clarity, the temporal consequence of two samples is denoted by the geodesic arc that links them.
autocorrelation function $R^w(m)$ shows the typical decay shape, while the shape of the autocorrelation function $R^w(m)$ suggests that the sequence $w_k$ is uncorrelated. The moving-average component of the model (46)-(47) 'colors' the white noise $w_k$.

4.2. Generation of random paths on the manifold of symmetric, positive-definite matrices. To generate a random tangent matrix $V \in T_X S^+(n)$, given a point $X \in S^+(n)$, the following scheme may be put into effect:

- Generate a random matrix $\tilde{V} \in \mathbb{R}^{n \times n}$ by a random scalar generator;
- Project the matrix $\tilde{V}$ into the tangent space $T_X S^+(n)$ by the rule $V = \frac{1}{2}(\tilde{V} + \tilde{V}^T)$.

The generation of a random path over the manifold $S^+(n)$ requires to choose the parameters that take part in the ARMA model (5). In particular, a suitable way to select the transition operators $A_i : (X, V) \in TS^+(n) \mapsto W \in T_X S^+(n)$ and $B_i : (X, V) \in TS^+(n) \mapsto W \in T_X S^+(n)$ is:

\[
\begin{align*}
A_i(X, V) &= a_i(XV + VX), \quad i \geq 1, \\
B_i(X, V) &= b_i(XV + VX), \quad i \geq 0,
\end{align*}
\]

where $a_i$ and $b_i$ denote real-valued constants chosen arbitrarily.

In the present numerical example, it the dimension of the manifold and the depths of the ARMA($r, q$) discrete-time system were chosen as $n = 2$, $r = 2$ and $q = 4$. The implementation based on parallel translation was opted. An example of random process generated over the manifold $S^+(2)$ by an ARMA(2, 4) model is illustrated in the Figure 4. A total of $K = 5,000$ temporal samples were generated. The initial seed was chosen as $X_0 = I_2$. 
Figure 3. Example of random path generated over the unit-hypersphere $S^2$ by an ARMA(2,4) model. Top panel: Autocorrelation function $R^v(m)$ associated with the random path $(x_k, v_k) \in TS^2$. Bottom panel: Autocorrelation function $R^w(m)$ associated with the random path $(x_k, w_k) \in TS^2$.

The sample auto-correlation function associated with the sequence $(X_k, V_k) \in TS^+(2)$, $k = 0, \ldots, K$, reads:

$$R^V(m) = \frac{1}{K - m} \sum_{k=1}^{K-m} \text{tr}(\Gamma_{X_k}^{X_{k+m}}(V_{k+m})X_k^{-1}V_kX_k^{-1}), \ m = 0, \ldots, L. \quad (51)$$

The Figure 5 illustrates the autocorrelation function $R^V(m)$ associated with the random path $(X_k, V_k) \in TS^+(2)$ and the autocorrelation function $R^W(m)$ associated with the random path $(X_k, W_k) \in TS^+(2)$. The maximum lag was chosen as $L = 20$. The autocorrelation function $R^V(m)$ shows the typical decay shape, while the shape of the autocorrelation function $R^W(m)$ suggests that the sequence $W_k$ is uncorrelated.

4.3. Generation of random paths on the compact Stiefel manifold. An inspection of the equations that express geodesic arcs and vector transport corresponding to the Euclidean metric and the canonical metric for the Stiefel manifold reported in Subsection 3.1 reveals that the quantities that correspond to the canonical metric are easier to implement on a computational platform. Hence, the canonical metric is the metric of choice to perform numerical examples in the present Subsection.

With the aim to generate a random tangent matrix $V \in T_X \text{St}(n, p)$, given a point $X \in \text{St}(n, p)$, the following scheme may be taken advantage of:

- Generate a random matrix $\tilde{V} \in \mathbb{R}^{n \times p}$ by a random scalar generator;
Figure 4. Example of random path generated over the manifold $S^+(2)$ by an ARMA(2,4) model \((5)\). The initial seed is denoted by a diamond and the generated points are denoted by dots. The temporal consequence of two samples is denoted by the geodesic arc that links them.

- Project the matrix $\tilde{V}$ into the tangent space $T_XSt(n,p)$ by the rule $V = \tilde{V} - XV^TX$.

The generation of a random path over the manifold $St(n,p)$ requires to choose the transition operators $\mathcal{A}_i : (X,V) \in TSt(n,p) \mapsto W \in T_XSt(n,p)$ and $\mathcal{B}_i : (X,V) \in TSt(n,p) \mapsto W \in T_XSt(n,p)$. A possible way is:

\[
\begin{align*}
\mathcal{A}_i(X,V) &= a_i V, \\
\mathcal{B}_i(X,V) &= b_i XV^TX,
\end{align*}
\]

where $a_i$ and $b_i$ are arbitrarily chosen real-valued constants.

In the present numerical example, it was chosen $n = 5$, $p = 2$, $r = 2$ and $q = 4$. The initial seed was chosen randomly in $St(5,2)$ and no constant drift term was included, in this simulation. The random path generation rule \((7)\) particularizes, in this case, to:

\[
\begin{align*}
V_k &= P_{X_k} \left( a_1 V_{k-1} + a_2 V_{k-2} + X_k^T (\sum_{i=0}^{4} b_i W_{k-i}) X_k \right), \\
X_{k+1} &= \exp_{X_k}(V_k),
\end{align*}
\]

for $k = 0, 1, 2, \ldots$ The exponential is implemented as:

\[
\exp_X(V) = [X \ Q] \text{Exp} \left( \begin{bmatrix} X^TV & -R^T \\ R & 0_2 \end{bmatrix} \right) \begin{bmatrix} I_2 \\ 0_2 \end{bmatrix},
\]

where $Q$ and $R$ denote the factors of the compact QR decomposition of the matrix $(I_5 - XX^T)V$. An example of random path generated over the manifold $St(5,2)$ by an ARMA(2,4) model is illustrated in the Figure 6. Again, a total of $K = 5,000$ temporal samples were generated.
The sample auto-correlation function associated with the sequence $(X_k, V_k) \in T\text{St}(5, 2)$, $k = 0, \ldots, K$, may be expressed as:

$$ R^V(m) = \frac{1}{K - m} \sum_{k=1}^{K-m} \text{tr} \left( P^T_{X_k}(V_{k+m}) \left( I_5 - \frac{1}{2} X_k X_k^T \right) V_k \right), \quad m = 0, \ldots, L. \quad (55) $$

The Figure 7 illustrates the autocorrelation function $R^V(m)$ associated with the random path $(X_k, V_k) \in T\text{St}(5, 2)$ and the autocorrelation function $R^W(m)$ associated with the random path $(X_k, W_k) \in T\text{St}(5, 2)$. The maximum lag was chosen as $L = 20$.

5. Conclusion. Although ad hoc methods may be envisaged to generate random matrices of specific forms, it appears that a sufficiently general method to simulate a random path on a matrix manifold was still lacking from the scientific literature. The present paper suggests a unifying framework that shifts the problem of generating a random path on a matrix manifold into the problem of generating a random path on the tangent bundle associated with the manifold by extending the classical notion of (scalar, linear) ARMA model to real-valued Riemannian matrix manifolds. In addition, the present contribution suggests a way to measure the temporal dependency between the samples generated by an ARMA model by extending the notion of sample autocorrelation function to real-valued Riemannian matrix manifolds. The outcome of such research endeavor is illustrated through several numerical example on matrix manifolds of interest in the scientific literature.
(a) The panel in the $i^{th}$ row and $j^{th}$ column represents the course of the entry $x_{ij,k}$ of the generated random process $X_k$.

(b) Scalar velocity at each time step.

Figure 6. Example of random path generated over the manifold $S(5, 2)$ by an ARMA(2, 4) model. The sub-figure (a) shows the actual random process entry-wise, while the sub-figure (b) shows the the norm of the corresponding random velocity.

The present manuscript covers the state of advancement of the present research topic. Crucial studies to be pursued in the future will be the BIBO (bounded input, bounded output) stability analysis as well as the asymptotic stability analysis of the devised ARMA models on matrix manifolds. Although ad hoc methods to study the stability of specific non-linear ARMA models on the space $\mathbb{R}^p$ are available in the scientific literature (see, e.g., the method to analyze switched ARMA systems presented in [20]), such methods do not appear immediately extendable.
Figure 7. Example of random path generated over the manifold \( \text{St}(5, 2) \) by an ARMA(2, 4) model (only the first 500 time steps are actually shown). Top panel: Autocorrelation function \( R^V(m) \) associated with the random path \((X_k, V_k) \in \text{TSt}(5, 2)\). Bottom panel: Autocorrelation function \( R^W(m) \) associated with the random path \((X_k, W_k) \in \text{TSt}(5, 2)\).

to the analysis of ARMA models on manifolds. Future research efforts will be directed along the lines of extending the ARMA model on manifolds to different models currently used, e.g., in econometrics, such as the ARCH model family [5, 32]. In addition, current ARMA models on manifolds could be extended to include further non-linear dependencies on the basis of previous research studies on non-linear ARMA models on flat spaces [44].

Acknowledgments. The content of this paper was presented during an informal talk given at Prof. Amari’s Lab (RIKEN, Wako-shi, Japan) in June 2013. I wish to thank Prof. Amari and the laboratory members for fruitful discussions on the subject. I would also like to express my gratitude to the reviewers for their stimulating comments, which contributed to improving the quality of presentation. The present research was partially funded under the grant Ricerca Scientifica di Ateneo UNIVPM 2012: “Reti neurali artificiali per l’elaborazione di segnali su varietà differenziali”.

Appendix A. Derivation of the orthogonal projector (26) and discussion about the features of vector transport based on such projector. According to the definition of orthogonal projection on a tangent space of a manifold given in Subsection 2.1, given a point \( Y \in \text{St}(n, p) \) and a matrix \( U \in \mathbb{R}^{n \times p} \), the orthogonal projection \( P_Y(U) \) must satisfy the:

- Tangency condition, namely: \( P_Y(U) \in T_Y\text{St}(n, p) \), and the
• Metric compatibility condition, namely: \( \langle V, P_Y(U) \rangle_Y = \text{tr}(V^TU) \) for all \( V \in T_Y \text{St}(n, p) \).

Such pair of conditions determine univocally the value of the projection \( P_Y(U) \). The orthogonal projection (26), in particular, is referred to the canonical metric \( \langle V, W \rangle_X \overset{\text{def}}{=} \text{tr} \left( V^T \left( I_n - \frac{1}{2} XX^T \right) W \right) \), with \( V, W \in T_X \text{St}(n, p) \). The compatibility condition with the canonical metric reads:

\[
\text{tr} \left( V^T \left( P_Y(U) - \frac{1}{2} YY^T P_Y(U) - U \right) \right) = 0 \text{ for all } V \in T_Y \text{St}(n, p). 
\] (56)

Such equation implies that the matrix \( P_Y(U) - \frac{1}{2} YY^T P_Y(U) - U \) be orthogonal to all tangent vectors in \( T_Y \text{St}(n, p) \), namely, it must be a normal vector. It is known (see, e.g., [13]), that the normal space at a point \( Y \) of a Stiefel manifold is formed by matrices of the type \( YS \), with \( S \) being \( p \times p \) symmetric. Therefore, there should exists some symmetric matrix \( S \) such that:

\[
P_Y(U) - \frac{1}{2} YY^T P_Y(U) - U = YS. 
\] (57)

Pre-multiplying both sides of the latter equation by \( Y^T \) yields:

\[
\frac{1}{2} Y^T P_Y(U) - Y^T U = S. 
\] (58)

Transposing both sides of the above equation yields:

\[
\frac{1}{2} P_Y^T(U)Y - U^T Y = S. 
\] (59)

The tangency condition implies that \( P_Y^T(U)Y = -Y^T P_Y(U) \). Summing up, term by term, the equations (58) and (59) and plugging in the tangency condition yields, for the matrix \( S \), the value:

\[
S = -\frac{1}{2}(Y^T U + U^T Y). 
\] (60)

Replacing such value of the matrix \( S \) within the relationship (58) gives:

\[
Y^T P_Y(U) = Y^T U - U^T Y = Y^T(U - YU^T Y). 
\] (61)

Since the matrix \( U - YU^T Y \) lies on the tangent space \( T_Y \text{St}(n, p) \) (as it may be verified directly), the above equation implies that \( P_Y(U) = U - YU^T Y \).

Now, take two points \( X, Y \in \text{St}(n, p) \) and take \( U \in T_X \text{St}(n, p) \subset \mathbb{R}^{n \times p} \). Vector transport of \( U \) to the tangent space \( T_Y \text{St}(n, p) \) is defined as \( P_Y(U) \). Such vector transport differs from the actual parallel translation. As a first observation, it is immediate to verify that vector transport is not an isometry, namely, it does not preserves the length of the transported vector:

\[
\langle U, U \rangle_X \neq \langle P_Y(U), P_Y(U) \rangle_Y. 
\] (62)

Moreover, such vector transport does not preserves angles between transported vectors, namely, given another tangent vector \( V \in T_X \text{St}(n, p) \), it can be immediately verified that:

\[
\frac{\langle U, V \rangle_X}{\sqrt{\langle U, U \rangle_X \langle V, V \rangle_X}} \neq \frac{\langle P_Y(U), P_Y(V) \rangle_Y}{\sqrt{\langle P_Y(U), P_Y(U) \rangle_Y \langle P_Y(V), P_Y(V) \rangle_Y}}, 
\] (63)

where the expression on the left-hand side represents the cosine of the angle between the tangent vectors \( U, V \) and the the expression on the right-hand side represents the cosine of the angle between the tangent vectors \( P_Y(U), P_Y(V) \).
REFERENCES

[1] S.-i. Amari, Differential geometry of a parametric family of invertible linear systems – Riemannian metric, dual affine connections, and divergence, Mathematical Systems Theory, 20 (1987), 53–82.
[2] S.-i. Amari, Natural gradient learning for over- and under-complete bases in ICA, Neural Computation, 11 (1989), 1875–1883.
[3] T. W. Anderson, I. Olkin and L. G. Underhill, Generation of random orthogonal matrices, SIAM Journal on Scientific and Statistical Computing, 8 (1987), 625–629.
[4] Y. K. Baik, J. Kwon, H. S. Lee and K. M. Lee, Geometric particle swarm optimization for robust visual ego-motion estimation via particle filtering, Image and Vision Computing, 31 (2013), 565–579.
[5] T. Bollerslev, Generalized autoregressive conditional heteroskedasticity, Journal of Econometrics, 31 (1986), 307–327.
[6] P. Brockwell and R. Davis, Time Series: Theory and Methods, 2nd Edition, Springer, 2009.
[7] E. Celledoni and S. Fiori, Neural learning by geometric integration of reduced ‘rigid-body’ equations, Journal of Computational and Applied Mathematics, (JCAM), 172 (2004), 247–269.
[8] Y. Chen and J. E. McInroy, Estimation of symmetric positive-definite matrices from imperfect measurements, IEEE Transactions on Automatic Control, 47 (2002), 1721–1725.
[9] G. S. Chirikjian, Stochastic Models, Information Theory, and Lie Groups, Volume 1: Classical Results and Geometric Methods, Birkhäuser, 2009.
[10] P. Chossat and O. Faugeras, Hyperbolic planforms in relation to visual edges and textures perception, PLoS Computational Biology, 5 (2009), e1000625.
[11] P. I. Davies and N. J. Higham, Numerically stable generation of correlation matrices and their factors, BIT, 40 (2000), 640–651.
[12] P. Diaconis and M. Shahshahani, The subgroup algorithm for generating uniform random variables, Probability in Engineering and Informational Sciences, 1 (1987), 15–32.
[13] A. Edelman, T. A. Arias and S. T. Smith, The geometry of algorithms with orthogonality constraints, SIAM Journal on Matrix Analysis Applications, 20 (1998), 303–353.
[14] L. Eldén and H. Park, A Procrustes problem on the Stiefel manifold, Numerical Mathematics, 82 (1999), 599–619.
[15] S. Fiori, A fast fixed-point neural blind deconvolution algorithm, IEEE Transactions on Neural Networks, 15 (2004), 455–459.
[16] S. Fiori, Blind adaptation of stable discrete-time IIR filters in state-space form, IEEE Transactions on Signal Processing, 54 (2006), 2596–2605.
[17] S. Fiori, Geodesic-based and projection-based neural blind deconvolution algorithms, Signal Processing, 88, (2008), 521–538.
[18] S. Fiori, Learning the Fréchet mean over the manifold of symmetric positive-definite matrices, Cognitive Computation, 1 (2009), 279–291.
[19] S. Fiori, Blind deconvolution by a Newton method on the non-unitary hypersphere, International Journal of Adaptive Control and Signal Processing, 27 (2013), 488–518.
[20] G. Fonseca, On the Stability of Nonlinear ARMA Mmodels, Technical report 2005/03, Facoltà di Economia, Università dell’Insubria, 2005.
[21] A. Genz, Methods for Generating Random Orthogonal Matrices, in Monte Carlo and Quasi-Monte Carlo Methods 1998, (Eds. H. Niederreiter and J. Spanier), Springer-Verlag, Berlin, 1999, 199–213.
[22] W. Gonzales-Manteiga, G. Henry and D. Rodriguez, Partly linear models on Riemannian manifolds, Journal of Applied Statistics, 39 (2012), 1797–1809.
[23] R. Harman and V. Lacko, On decompositional algorithms for uniform sampling from n-spheres and n-balls, Journal of Multivariate Analysis, 101 (2010), 2297–2304.
[24] J. L. Harvill and B. K. Ray, An investigation of lag identification tools for vector nonlinear time series, Communications in Statistics - Theory and Methods, 29 (2000), 1677–1701.
[25] N. J. Higham, Functions of Matrices: Theory and Computation, SIAM, 2008.
[26] T. Kaneko, S. Fiori and T. Tanaka, Empirical arithmetic averaging over the compact Stiefel manifold, IEEE Transactions on Signal Processing, 61 (2013), 883–894.
[27] R. Lenz, R. Mochizuki and J. Chao, Iwasawa decomposition and computational Riemannian geometry, in Proceedings of the 20th International Conference on Pattern Recognition (ICPR), 2010, 4472–4475.
N. L. Lim, Clustering and Sampling Algorithms on Shape Manifolds, Technical report of the Bio-Informatics Institute (Biopolis, Singapore), 2012.

G. Meyer, Regression on fixed-rank positive semidefinite matrices: A Riemannian approach, Journal of Machine Learning Research, 12 (2011), 593–625.

F. Mezzadri, How to generate random matrices from the classical compact groups, Notices of the AMS, 54 (2007), 592–604.

J. Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm, IEEE Transactions on Signal Processing, 56 (2008), 3061–3068.

D. B. Nelson, Conditional heteroskedasticity in asset returns: A new approach, Econometrica, 59 (1991), 347–370.

G. Meyer, Regression on fixed-rank positive semidefinite matrices: A Riemannian approach, Journal of Machine Learning Research, 12 (2011), 593–625.

F. Mezzadri, How to generate random matrices from the classical compact groups, Notices of the AMS, 54 (2007), 592–604.

J. Navarro-Moreno, ARMA prediction of widely linear systems by using the innovations algorithm, IEEE Transactions on Signal Processing, 56 (2008), 3061–3068.

D. B. Nelson, Conditional heteroskedasticity in asset returns: A new approach, Econometrica, 59 (1991), 347–370.

P. Pajunen and M. Girolami, Implementing decisions in binary decision trees using independent component analysis, in Proceedings of the International Workshop on Independent Component Analysis and Blind Signal Separation (Helsinki, Finland, 2000), 2000, 477–481.

F. L. Ramsey, Characterization of the partial autocorrelation function, The Annals of Statistics, 2 (1974), 1296–1301.

H. Shimazaki, S.-i. Amari, E. N. Brown and S. Grün, State-space analysis of time-varying higher-order spike correlation for multiple neural spike train data, PLoS Computational Biology, 8 (2012), e1002385.

M. Spivak, A Comprehensive Introduction to Differential Geometry, 1, 2nd Edition, Berkeley, CA: Publish or Perish Press, 1979.

G. C. Tiao and R. S. Tsay, Model specification in multivariate time series, Journal of Royal Statistical Society, Series B (Methodological), 51 (1989), 157–213.

H. Tidéfelt and T. B. Schön, Robust point-mass filters on manifold, in Proceedings of the 15th IFAC Symposium on System Identification (SYSID, Saint-Malo, France, 2009), 2009, 1–6.

N. T. Trendafilov and R. A. Lippert, The multimode Procrustes problem, Linear Algebra and its Applications, 349 (2002), 245–264.

P. Turaga, A. Veeraraghavan and R. Chellappa, Statistical analysis on Stiefel and Grassmann manifolds with applications in computer vision, in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR 2008, Anchorage (AK), USA, 2008), 1–8.

P. A. Valdes-Sosa, Spatio-temporal autoregressive models defined over brain manifolds, Neuroinformatics, 2 (2004), 239–250.

J. F. Vasconcelos, G. Elkaim, C. Silvestre, P. Oliveira and B. Cardeira, Geometric approach to strapdown magnetometer calibration in sensor frame, IEEE Transactions on Aerospace and Electronic Systems, 47 (2011), 1293–1306.

A. Veeraraghavan, A. K. Roy-Chowdhury and R. Chellappa, Matching shape sequences in video with applications in human movement analysis, IEEE Transactions on Pattern Analysis and Machine Intelligence, 27 (2005), 1896–1909.

N. Wessel, H. Malberg, R. Bauernschmitt, A. Schirdewan and J. Kurths, Nonlinear additive autoregressive model-based analysis of short-term heart rate variability, Medical and Biological Engineering and Computing, 44 (2006), 321–330.

A. Yershova and S. M. LaValle, Deterministic sampling methods for spheres and SO(3), in Proceedings of the IEEE International Conference on Robotics and Automation (ICRA 2004, New Orleans (LA, USA), 2004), 4 (2004), 3974–3980.

F. Yger, M. Berar, G. Gasso and A. Rakotomamonjy, Oblique principal subspace tracking on manifold, in Proceedings of the 2012 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP, Kyoto (Japan)), 2012, 2429–2432.

J. Yoo and S. Choi, Orthogonal nonnegative matrix factorization: multiplicative updates on Stiefel manifolds, in Proceedings of the Intelligent Data Engineering and Automated Learning (IDEAL 2008), Springer Berlin/Heidelberg, 2008, 140–147.

Y. Zhu, W. Mio and X. Liu, Optimal dimension reduction for image retrieval with correlation metrics, in Proceedings of the International Conference on Neural Networks (IJCNN 2009, Atlanta (GA, USA)), 2009, 3565–3570.

Received December 2013; revised April 2014.

E-mail address: s.fiori@univpm.it