



**HAL**  
open science

# Theoretically and computationally convenient geometries on full-rank correlation matrices

Yann Thanwerdas, Xavier Pennec

► **To cite this version:**

Yann Thanwerdas, Xavier Pennec. Theoretically and computationally convenient geometries on full-rank correlation matrices. *SIAM Journal on Matrix Analysis and Applications*, In press. hal-03527072v1

**HAL Id: hal-03527072**

**<https://hal.science/hal-03527072v1>**

Submitted on 14 Jan 2022 (v1), last revised 15 Nov 2022 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# THEORETICALLY AND COMPUTATIONALLY CONVENIENT GEOMETRIES ON FULL-RANK CORRELATION MATRICES \*

YANN THANWERDAS<sup>†</sup> AND XAVIER PENNEC<sup>‡</sup>

**Abstract.** In contrast to SPD matrices, few tools exist to perform Riemannian statistics on the open ellipsope of full-rank correlation matrices. The quotient-affine metric was recently built as the quotient of the affine-invariant metric by the congruence action of positive diagonal matrices. The space of SPD matrices had always been thought of as a Riemannian homogeneous space. In contrast, we view in this work SPD matrices as a Lie group and the affine-invariant metric as a left-invariant metric. This unexpected new viewpoint allows us to generalize the construction of the quotient-affine metric and to show that the main Riemannian operations can be computed numerically. However, the uniqueness of the Riemannian logarithm or the Fréchet mean are not ensured, which is bad for computing on the ellipsope. Hence, we define three new families of Riemannian metrics on full-rank correlation matrices which provide Hadamard structures, including two flat. Thus the Riemannian logarithm and the Fréchet mean are unique. We also define a nilpotent group structure for which the affine logarithm and the group mean are unique. We provide the main Riemannian/group operations of these four structures in closed form.

**Key words.** SPD matrices; correlation matrices; Lie group; Lie group actions; quotient-affine metric; Lie-Cholesky metrics; poly-hyperbolic-Cholesky metrics; Euclidean-Cholesky metrics; log-Euclidean-Cholesky metrics

**AMS subject classifications.** 15B48, 15B99, 53-08, 53B21, 15A63, 53C22, 62H20, 58D17.

**1. Introduction.** The (open) ellipsope is the set of full-rank correlation matrices, it is open in the affine space of symmetric matrices with unit diagonal. Its geometry has been much less studied than the one of the cone of Symmetric Positive Definite (SPD) matrices. Nevertheless, several applications could benefit from well suited geometries on this manifold: graphical networks, brain connectomes [24], finance [20, 16] or phylogenetic trees [9]. A few operations have been proposed such as sampling by projecting samples on spheres [20, 11], computing distances via the Hilbert geometry of convex sets [17] or via a recent parametrization by  $\mathbb{R}^{n(n-1)/2}$  [1].

The open ellipsope was also recently characterized as the quotient manifold of SPD matrices by the smooth, proper and free congruence action of positive diagonal matrices [7, 6]. Indeed, a correlation matrix is obtained from the covariance matrix by dividing by the standard deviation of each variable. This is equivalent to multiply the covariance matrix on left and right by the diagonal matrix of the inverse square roots of the variances. Hence all covariance matrices that are congruent up to a diagonal matrix represent the same correlation matrix. This structure allowed to quotient the well known affine-invariant metrics on SPD matrices to the so called quotient-affine metrics on full-rank correlation matrices [7, 6, 22]. They only differ by a scaling factor so they are often referred as *the* quotient-affine metric. It offers promising perspectives since it is geodesically complete with closed form expressions for the exponential map, the Levi-Civita connection and the sectional curvature. However, it was difficult to determine the sign and potential bounds of the curvature of the quotient-affine metric. In this work, we show that the sectional curvature can take both positive and negative values, that it is bounded from below and unbounded from above. Therefore, the ellipsope of full-rank correlation matrices endowed with

---

\*Submitted to the editors January 14th, 2022.

<sup>†</sup>Université Côte d'Azur and Inria, Epione Project Team ([yann.thanwerdas@inria.fr](mailto:yann.thanwerdas@inria.fr)).

<sup>‡</sup>Université Côte d'Azur and Inria, Epione Project Team ([xavier.pennec@inria.fr](mailto:xavier.pennec@inria.fr)).

the quotient-affine metric is neither a Hadamard space nor a  $\text{CAT}(k)$  space for any  $k \in \mathbb{R}$ .

This calls for new Riemannian metrics on the ellipsope. Indeed, we could expect to find suitable metrics with non-positive or even null curvature since the ellipsope is an open set of an affine space. The structure of the space is an important element of modeling because simple structures often bring good theoretical properties and better computability of the geometric operations. For example, in a Hadamard space, the Fréchet mean is unique. Recall that in a metric space  $(\mathcal{M}, d)$ , a Fréchet mean of points  $x_1, \dots, x_k \in \mathcal{M}$  is a point  $x \in \mathcal{M}$  which minimizes the function  $x \in \mathcal{M} \mapsto \sum_{i=1}^k d(x, x_i)^2 \geq 0$ . This is the case of the affine-invariant metric on SPD matrices [21] or the Fisher metric of beta and Dirichlet distributions [12]. Moreover, if one has a Euclideanization of the space, i.e. a smooth diffeomorphism to a Euclidean space, then all operations become trivial. This is the case of the log-Euclidean [2] or the log-Cholesky [15] metrics on SPD matrices.

Another strong element of modeling is the invariance of the geometry under a given group action. Let us give some examples. First, on SPD matrices, one could require the invariance of the geometry under all affine transformations of the feature vector. Indeed in EEG, if we assume that there is an affine transformation from one brain to another at first order, then electromagnetic fields are transformed similarly since they satisfy linear equations. The affine-invariant metric significantly improved the results of classification in BCI [3]. Second, one could want the analysis to be invariant from the scale of each variable. As explained previously, this corresponds to the invariance under the congruence action of positive diagonal matrices and it suggests to focus more on correlation matrices than on covariance matrices. Third, all the operations mentioned above on correlation matrices are invariant under permutations. It means that the statistical analyses are invariant under any joint permutation of the rows and columns of the correlation matrices in the dataset. It is an advantage if the way the variables (or channels) are ordered is arbitrary. It is a drawback if the order is chosen for a certain reason. For instance, in the auto-correlation matrix of a signal, the variables represent different times, which are not exchangeable. One could want a structure on correlation matrices that respects the time structure instead of being invariant under permutations. In this work, we focus on non-permutation-invariant geometric structures.

Moreover, an invariance may also lead to simple structures with good properties. For example, a Riemannian homogeneous space is geodesically complete, which is well suited for interpolation and extrapolation. The geodesics of naturally reductive homogeneous spaces are the orbits of the one-parameter subgroups so they are known in closed form. In a Riemannian symmetric space, the parallel transport is obtained in closed form by composition of two symmetries. So it is another good reason to study Lie group actions on our space.

**1.1. Overview of the results.** In this work, we define new non-permutation-invariant geometric structures on full-rank correlation matrices following two directions. Firstly, we define a non-permutation-invariant generalization of quotient-affine metrics. Our approach consists in studying the congruence action of several matrix Lie groups on SPD matrices. Our main result is a characterization of affine-invariant metrics, i.e. metrics that are invariant under the congruence action of the general linear group  $\text{GL}(n)$ , by the joint invariance of a pair of subgroups. These are the group of permutation matrices  $\mathfrak{S}(n)$  and the group of lower triangular matrices with positive diagonal  $\text{LT}^+(n)$ . In other words, the affine-invariant metrics are the

unique  $(\mathfrak{S}(n) \times \text{LT}^+(n))$ -invariant metrics on SPD matrices. Therefore, the family of  $\text{LT}^+(n)$ -invariant metrics appears as a natural non-permutation-invariant generalization of affine-invariant metrics. Moreover, we show that such metrics are exactly the pullback metrics of left-invariant metrics on the Lie group  $\text{LT}^+(n)$  by the Cholesky map, so we call them Lie-Cholesky metrics. In particular, there exists a Lie group structure on SPD matrices (namely, the product of the Cholesky factors) such that affine-invariant metrics are left-invariant metrics on that Lie group. Finally, since the Lie group  $\text{LT}^+(n)$  contains positive diagonal matrices, Lie-Cholesky metrics descend to quotient metrics on full-rank correlation matrices by the same procedure as affine-invariant metrics. We show that several Riemannian operations are numerically computable such as the Riemannian metric, the exponential map, the logarithm map or the Riemannian distance. However, despite the nice theoretical results and the computability of the main geometric operations, quotient-Lie-Cholesky metrics don't have obvious nice geometric properties. For example, it is not clear whether the Riemannian logarithm and the Fréchet mean are unique or not.

Hence, secondly, to solve the drawbacks of quotient-affine and quotient-Lie-Cholesky metrics, we propose a series of new geometries on full-rank correlation matrices for which the mean of finite samples is unique. They are built in a different way than the Lie-Cholesky metrics, the only common point being the use of the Cholesky map to convey the structures from triangular matrices to correlation matrices. Hence they are not permutation-invariant either for  $n \geq 3$ . We define the poly-hyperbolic-Cholesky metrics as the pullbacks of weighted products of hyperbolic spaces  $\mathbb{H}^1 \times \dots \times \mathbb{H}^{n-1}$ . The metrics of this family provide a structure of Riemannian symmetric space with non-positive bounded curvature, thus Hadamard. All operations are known in closed form. We define the Euclidean-Cholesky and the log-Euclidean-Cholesky metrics as the pullbacks by two maps derived from the Cholesky map of Euclidean structures on the vector space  $\text{LT}^0(n)$  of strictly lower triangular matrices. The metrics in these two families are flat and geodesically complete so all Riemannian operations are known in closed form, the former being less costly than the latter. Finally, we define a Lie group structure as the pullback of the Lie group (for matrix multiplication)  $\text{LT}^1(n)$  of lower triangular matrices with unit diagonal. Here, we consider the geometric structure induced by the canonical Cartan-Schouten connection rather than by a Riemannian metric. Indeed, the (group) exponential map allows to define a group mean, which is unique here because the Lie algebra of the Lie group is nilpotent [5]. Finally, in dimension 2, we show that on the one hand, the quotient-affine and the poly-hyperbolic-Cholesky metrics coincide, and on the other hand, the Euclidean-Cholesky and the log-Euclidean-Cholesky metrics coincide and their geodesics coincide with the group geodesics. In particular, these geodesics provide a new interpolation of the correlation coefficient, different than the one proposed in [22].

In the remainder of this section, we introduce the matrix notations, the notions of covariance and correlation matrices and the Cholesky map. In Section 2, we recall the definition of the quotient-affine metrics and we prove that its sectional curvature takes both negative and positive values and is unbounded from above. In Section 3, we give a characterization of affine-invariant metrics in function of the congruence action of other groups. This allows us to introduce Lie-Cholesky metrics on SPD matrices and quotient-Lie-Cholesky metrics on full-rank correlation matrices. In Section 4, we introduce four new non-permutation-invariant structures on full-rank correlation matrices for which the mean is unique. We conclude in Section 5.

## 1.2. Concepts and notations.

**1.2.1. Matrix notations.** Our main matrix space notations are given in Table

1.1. In the paper, we also use the following linear maps:

- $(A, B) \in \text{Mat}(n) \times \text{Mat}(n) \mapsto A \bullet B = [A_{ij}B_{ij}]_{1 \leq i, j \leq n} \in \text{Mat}(n)$  is the Hadamard product of matrices,
- $\text{Diag} : M \in \text{Mat}(n) \mapsto [\delta_{ij}M_{ij}]_{1 \leq i, j \leq n} \in \text{Diag}(n)$  selects the diagonal terms,
- $\text{off} = \text{Id}_{\text{Mat}(n)} - \text{Diag} : \text{Mat}(n) \rightarrow \ker \text{Diag}$  selects the off-diagonal terms,
- $\text{low} : M \in \text{Mat}(n) \mapsto [\delta_{i \geq j}M_{ij}]_{1 \leq i, j \leq n} \in \text{LT}(n)$  selects the lower triangular terms, *including* the diagonal terms,
- $\text{low}_0 = \text{low} - \text{Diag} : \text{Mat}(n) \rightarrow \text{LT}^0(n)$  selects the strictly lower triangular terms, *excluding* the diagonal terms,
- $\text{low}_S = \text{low}_0 + \frac{1}{2}\text{Diag} : \text{Sym}(n) \rightarrow \text{LT}(n)$  selects the strictly lower triangular terms and half of the diagonal terms, so that when a symmetric matrix writes  $M = L + L^\top \in \text{Sym}(n)$  with  $L \in \text{LT}(n)$ , then  $L = \text{low}_S(M)$ ,
- $\text{sum} : M \in \text{Mat}(n) \mapsto \sum_{i, j} M_{ij} = \mathbf{1}^\top M \mathbf{1}$  sums the terms of the matrix,
- $\text{sum} : x \in \mathbb{R}^n \mapsto \sum_i x_i = \mathbf{1}^\top x$  sums the terms of the vector,

where  $\mathbf{1} = (1, \dots, 1)^\top \in \mathbb{R}^n$ .

Matrix vector spaces		Matrix manifolds	
$\text{Mat}(n)$	Squared	$\text{GL}(n)$	General linear group
$\text{Skew}(n)$	Skew-symmetric	$\text{O}(n)$	Orthogonal group
		$\text{SO}(n)$	Special orthogonal group
$\text{Sym}(n)$	Symmetric	$\text{Sym}^+(n)$	Sym Positive Definite cone
$\text{Hol}(n)$	Symmetric hollow	$\text{Cor}^+(n)$	Full-rank correlation ellipptope
$\text{LT}(n)$	Lower Triangular	$\text{LT}^+(n)$	LT with positive diagonal
$\text{LT}^0(n)$	LT with null diagonal	$\text{LT}^1(n)$	LT with unit diagonal
$\text{Diag}(n)$	Diagonal	$\text{Diag}^+(n)$	Positive diagonal group

TABLE 1.1

*Matrix space notations*

The possibly uncommon notations in Table 1.1 are:

- $\text{Sym}^+(n) = \{\Sigma \in \text{Sym}(n) | \Sigma > 0\}$  where  $>$  is the Loewner order,
- $\text{Cor}^+(n) = \{\Sigma \in \text{Sym}^+(n) | \text{Diag}(\Sigma) = I_n\}$ ,
- $\text{Hol}(n) = \{M \in \text{Sym}(n) | \text{Diag}(M) = 0\}$  where “hollow” means vanishing diagonal,
- $\text{LT}(n) = \{\text{low}(M) | M \in \text{Mat}(n)\}$ ,
- $\text{LT}^+(n) = \{L \in \text{LT}(n) | \text{Diag}(L) \in \text{Diag}^+(n)\}$ ,
- $\text{LT}^0(n) = \{L \in \text{LT}(n) | \text{Diag}(L) = 0\}$ ,
- $\text{LT}^1(n) = \{L \in \text{LT}(n) | \text{Diag}(L) = I_n\} = I_n + \text{LT}^0(n)$ .

We denote  $\mathfrak{S}(n)$  the permutation group of order  $n$ .

**1.2.2. Covariance and correlation matrices.** Given an invertible covariance matrix  $\Sigma = (\text{Cov}(X_i, X_j))_{1 \leq i, j \leq n} \in \text{Sym}^+(n)$  of a random vector  $X$ , the corresponding correlation matrix is defined by  $C = \text{cor}(\Sigma) = (\text{Cor}(X_i, X_j))_{1 \leq i, j \leq n}$  where:

$$(1.1) \quad \text{Cov}(X_i, X_j) = \mathbb{E}(X_i X_j) - \mathbb{E}(X_i)\mathbb{E}(X_j),$$

$$(1.2) \quad \text{Cor}(X_i, X_j) = \frac{\text{Cov}(X_i, X_j)}{\sqrt{\text{Cov}(X_i, X_i)}\sqrt{\text{Cov}(X_j, X_j)}} = \frac{\Sigma_{ij}}{\sqrt{\Sigma_{ii}}\sqrt{\Sigma_{jj}}}$$

$$(1.3) \quad = [\text{Diag}(\Sigma)^{-1/2} \Sigma \text{Diag}(\Sigma)^{-1/2}]_{ij},$$

with  $\text{Diag}(\Sigma) = \text{diag}(\Sigma_{11}, \dots, \Sigma_{nn}) \in \text{Diag}^+(n)$ . Moreover, if  $D \in \text{Diag}^+(n)$ , then the correlation matrix associated to  $D\Sigma D$  is again  $\text{cor}(\Sigma)$ . It is known as the invariance of the correlation matrix under the scaling of each component of the random vector. In other words, the surjective map  $\text{cor} : \Sigma \in \text{Sym}^+(n) \mapsto \text{cor}(\Sigma) \in \text{Cor}^+(n)$  is invariant under the group action  $(D, \Sigma) \in \text{Diag}^+(n) \times \text{Sym}^+(n) \mapsto D\Sigma D \in \text{Sym}^+(n)$  and the orbit space  $\text{Sym}^+(n)/\text{Diag}^+(n)$  can be identified with  $\text{Cor}^+(n)$ . Note that, via this identification, the induced topology  $\text{Cor}^+(n) \hookrightarrow \text{Sym}(n)$  coincides with the quotient topology of  $\text{Sym}^+(n)/\text{Diag}^+(n)$ . The action being proper [7], the quotient manifold theorem [13] states that the orbit space  $\text{Sym}^+(n)/\text{Diag}^+(n)$  has a unique smooth manifold structure such that the canonical surjection  $\text{Sym}^+(n) \rightarrow \text{Sym}^+(n)/\text{Diag}^+(n)$  is a smooth submersion. Since  $\text{cor} : \text{Sym}^+(n) \rightarrow \text{Cor}^+(n)$  is a smooth submersion and since  $\text{Cor}^+(n)$  is  $\text{Sym}^+(n)/\text{Diag}^+(n)$  as a topological space, this smooth structure coincides with the induced smooth structure of  $\text{Cor}^+(n)$ .

**1.2.3. The Cholesky map.** A Cholesky decomposition of a symmetric positive semi-definite matrix  $\Sigma$  is a factorization of the form  $\Sigma = LL^\top$  where  $L \in \text{LT}(n)$  is a lower triangular matrix. If  $\Sigma$  is positive definite, then there exists a unique triangular matrix with positive diagonal  $L \in \text{LT}^+(n)$  such that  $\Sigma = LL^\top$ . This allows to define the Cholesky bijective map:

$$(1.4) \quad \text{Chol} : \Sigma \in \text{Sym}^+(n) \mapsto L \in \text{LT}^+(n),$$

whose inverse is the smooth map  $\phi : L \in \text{LT}^+(n) \mapsto LL^\top \in \text{Sym}^+(n)$ . Moreover, the Cholesky map is smooth. Indeed, it is the product of two smooth maps  $\text{Chol}(\Sigma) = L(\Sigma)\sqrt{D(\Sigma)}$  where  $L \equiv L(\Sigma) \in \text{LT}^1(n)$  and  $D \equiv D(\Sigma) \in \text{Diag}^+(n)$  are recursively defined for all  $(i, j) \in \{1, \dots, n\}^2$  with  $i > j$  by:

$$(1.5) \quad D_{ii} = \Sigma_{ii} - \sum_{j=1}^{i-1} L_{ij}^2 D_{jj} > 0,$$

$$(1.6) \quad L_{ij} = \frac{1}{D_{jj}} \left( \Sigma_{ij} - \sum_{k=1}^{j-1} L_{ik} L_{jk} D_{kk} \right),$$

the order of the computations being  $\{D_{11}\} \rightarrow \dots \rightarrow \{L_{i1} \rightarrow L_{i2} \rightarrow \dots \rightarrow L_{i,i-1} \rightarrow D_{ii}\} \rightarrow \dots \rightarrow \{L_{n1} \rightarrow \dots \rightarrow L_{n,n-1} \rightarrow D_{nn}\}$ .

We denote:

$$\mathcal{L} = \text{Chol}(\text{Cor}^+(n)) = \left\{ L = \begin{pmatrix} L_1 \\ \vdots \\ L_n \end{pmatrix} \in \text{LT}^+(n) \mid \forall i \in \{1, \dots, n\}, \|L_i\|^2 = L_i L_i^\top = 1 \right\},$$

which is diffeomorphic to  $\text{Cor}^+(n)$  via the Cholesky map. We also define a Cholesky-based diffeomorphism between  $\text{Cor}^+(n)$  and  $\text{LT}^1(n)$ :

$$(1.7) \quad \Theta : C \in \text{Cor}^+(n) \mapsto \Gamma = \text{Diag}(\text{Chol}(C))^{-1} \text{Chol}(C) \in \text{LT}^1(n),$$

$$(1.8) \quad \Phi : \Gamma \in \text{LT}^1(n) \mapsto C = \text{Diag}(\Gamma\Gamma^\top)^{-1/2} \Gamma\Gamma^\top \text{Diag}(\Gamma\Gamma^\top)^{-1/2} \in \text{Cor}^+(n).$$

We clearly have the relations  $\Phi = \Theta^{-1} = \text{cor} \circ \phi$ .

We compute the differentials of  $\phi : \text{LT}^+(n) \rightarrow \text{Sym}^+(n)$  and  $\text{Chol} : \text{Sym}^+(n) \rightarrow$

$\text{LT}^+(n)$ . For all  $Z \in T_L \text{LT}^+(n) \simeq \text{LT}(n)$ :

$$(1.9) \quad \begin{aligned} V &:= d_L \phi(Z) = ZL^\top + LZ^\top, \\ L^{-1}VL^{-\top} &= L^{-1}Z + (L^{-1}Z)^\top, \\ \text{low}_S(L^{-1}VL^{-\top}) &= L^{-1}Z, \end{aligned}$$

$$(1.10) \quad Z = d_\Sigma \text{Chol}(V) = L \text{low}_S(L^{-1}VL^{-\top}).$$

In particular,  $d_{I_n} \text{Chol} = \text{low}_S$ .

**2. Quotient-affine metrics.** In this section, we briefly recall how to build quotient metrics to fix the notations (Section 2.1), then we recall the definition of quotient-affine metrics (Section 2.2) and finally we show that the sectional curvature takes both negative and positive curvature and is unbounded from above (Section 2.3).

**2.1. Quotient metrics.** Given smooth manifolds  $\mathcal{M}, \mathcal{M}'$  and a smooth submersion  $\pi : \mathcal{M} \rightarrow \mathcal{M}'$ , we can define the vertical space  $\mathcal{V}_x = \ker d_x \pi \subset T_x \mathcal{M}$ , where  $d_x \pi : T_x \mathcal{M} \rightarrow T_{\pi(x)} \mathcal{M}'$  is the differential of the map  $\pi$  at point  $x \in \mathcal{M}$ . The horizontal space  $\mathcal{H}_x$  can be any supplementary vector space, i.e. such that  $\mathcal{V}_x \oplus \mathcal{H}_x = T_x \mathcal{M}$ . Given a horizontal distribution  $x \mapsto \mathcal{H}_x$ , there exist vertical and horizontal projections  $\text{ver}_x : T_x \mathcal{M} \rightarrow \mathcal{V}_x$  and  $\text{hor}_x : T_x \mathcal{M} \rightarrow \mathcal{H}_x$ . Moreover, the linear map  $(d_x \pi)|_{\mathcal{H}_x} : \mathcal{H}_x \rightarrow T_{\pi(x)} \mathcal{M}'$  is an isomorphism. Its inverse isomorphism is called the horizontal lift and denoted  $\#_x : X \in T_{\pi(x)} \mathcal{M}' \rightarrow X_x^\# \in \mathcal{H}_x$ . When  $\mathcal{M}$  is endowed with a Riemannian metric, there is a canonical choice of horizontal space which is the orthogonal  $\mathcal{H}_x = \mathcal{V}_x^\perp$ . In this case, the projections are orthogonal.

Given a smooth manifold  $\mathcal{M}$  on which a Lie group  $G$  acts smoothly, properly and freely, the quotient space  $\mathcal{M}/G$  admits a unique smooth manifold structure that turns the canonical projection  $\pi : \mathcal{M} \rightarrow \mathcal{M}/G$  into a smooth submersion [13, Theorem 21.10]. The vertical distribution is  $G$ -equivariant, i.e.  $\mathcal{V}_{a \cdot x} = a \cdot \mathcal{V}_x$  for all  $a \in G$  (where  $\cdot$  is the group action on  $\mathcal{M}$  and  $T\mathcal{M}$ ). Given a  $G$ -invariant Riemannian metric  $g$  on  $\mathcal{M}$ , the horizontal distribution is  $G$ -equivariant and the metric descends to a metric  $g'$  on  $\mathcal{M}/G$  defined by  $g'_{\pi(x)}(X, X) = g_x(X_x^\#, X_x^\#)$ .

**2.2. Definition of quotient-affine metrics.** Applying this to  $\mathcal{M} = \text{Sym}^+(n)$ ,  $G = \text{Diag}^+(n)$  and  $\mathcal{M}/G \simeq \text{Cor}^+(n)$ , the submersion  $\text{cor} : \Sigma \in \text{Sym}^+(n) \mapsto \text{Diag}(\Sigma)^{-1/2} \Sigma \text{Diag}(\Sigma)^{-1/2} \in \text{Cor}^+(n)$  [7, 6] allows to descend any  $\text{Diag}^+(n)$ -invariant Riemannian metric on  $\text{Sym}^+(n)$  to a Riemannian metric on  $\text{Cor}^+(n)$  [18].

A natural example of  $\text{Diag}^+(n)$ -invariant Riemannian metric on SPD matrices is provided by the affine-invariant metric defined for all  $\Sigma \in \text{Sym}^+(n)$  and  $V \in T_\Sigma \text{Sym}^+(n) \simeq \text{Sym}(n)$  by:

$$(2.1) \quad g_\Sigma^{\text{AI}(\alpha, \beta)}(V, V) = \alpha \text{tr}(\Sigma^{-1}V\Sigma^{-1}V) + \beta \text{tr}(\Sigma^{-1}V)^2$$

where  $\alpha > 0$  and  $\beta > -\frac{\alpha}{n}$ . Its sectional curvature is for  $V, W \in \text{Sym}(n)$  [21, 23]:

$$(2.2) \quad \kappa_\Sigma^{\text{AI}(\alpha, \beta)}(V, W) = \frac{1}{4\alpha} \text{tr}((\Sigma^{-1}V\Sigma^{-1}W - \Sigma^{-1}W\Sigma^{-1}V)^2) \in \left[-\frac{1}{2\alpha}; 0\right]$$

The quotient-affine metric is the quotient of the affine-invariant metric via the submersion  $\text{cor} : \text{Sym}^+(n) \rightarrow \text{Cor}^+(n)$ . It does not depend on  $\beta$  and it writes  $\alpha g^{\text{QA}}$  with, for all  $C \in \text{Cor}^+(n)$  and  $X \in T_C \text{Cor}^+(n) \simeq \text{Hol}(n)$ :

$$(2.3) \quad g_C^{\text{QA}}(X, X) = \text{tr}((C^{-1}X)^2) - 2\mathbf{1}^\top \text{Diag}(C^{-1}X)(I_n + C \bullet C^{-1})^{-1} \text{Diag}(C^{-1}X)\mathbf{1}.$$



Note that it is invariant under permutations.

The vertical/horizontal distributions/projections and the horizontal lift are for all  $\Sigma \in \text{Sym}^+(n)$ ,  $V \in T_\Sigma \text{Sym}^+(n) \simeq \text{Sym}(n)$  and  $X \in T_{\text{cor}(\Sigma)} \text{Cor}^+(n) \simeq \text{Hol}(n)$  [22]:

- $\mathcal{V}_\Sigma = \{D\Sigma + \Sigma D \mid D \in \text{Diag}(n)\}$ ,
- $\mathcal{H}_\Sigma = \{V \in \text{Sym}(n) \mid \Sigma^{-1}V + V\Sigma^{-1} \in \text{Hol}(n)\} = \mathcal{S}_{\Sigma^{-1}}(\text{Hol}(n))$ ,
- $\text{ver}_\Sigma(V) = D\Sigma + \Sigma D \in \mathcal{V}_\Sigma$  with  $D = \text{diag}((I_n + \Sigma \bullet \Sigma^{-1})^{-1} \text{Diag}(\Sigma^{-1}V)\mathbf{1})$ ,
- $\text{hor}_\Sigma(V) = V - \text{ver}_\Sigma(V) \in \mathcal{H}_\Sigma$ ,
- $X_\Sigma^\# = \text{hor}_\Sigma(\text{Diag}(\Sigma)^{1/2} X \text{Diag}(\Sigma)^{1/2}) \in \mathcal{H}_\Sigma$ ,

where  $\mathcal{S}_A(X)$  is the unique solution of the Sylvester equation  $A\mathcal{S}_A(X) + \mathcal{S}_A(X)A = X$ .

These operations allow to write the sectional curvature for all  $C \in \text{Cor}^+(n)$  and  $X, Y \in T_C \text{Cor}^+(n) \simeq \text{Hol}(n)$  [22]:

$$(2.4) \quad \kappa_C^{\text{QA}}(X, Y) = \underbrace{\kappa_C^{\text{AI}}(X_C^\#, Y_C^\#)}_{\in [-\frac{1}{2}; 0]} + \underbrace{\frac{3}{8} \frac{\mu^\top (I_n + C \bullet C^{-1})^{-1} \mu}{g_C^{\text{QA}}(X, X)g_C^{\text{QA}}(Y, Y) - g_C^{\text{QA}}(X, Y)^2}}_{\geq 0},$$

with  $\mu \in \mathbb{R}^n$  defined by  $\mu = [D(X, Y) - D(Y, X)]\mathbf{1} \in \mathbb{R}^n$  and  $D(X, Y) \in \text{Diag}(n)$  defined by  $D(X, Y) = \text{Diag}([C^{-1} \text{Diag}(X_C^\#)C, C^{-1}Y_C^\#])$ , where  $[A, B] = AB - BA$  is the commutator of squared matrices.

### 2.3. Complement: bounds of curvature.

**THEOREM 2.1.** *The sectional curvature of the quotient-affine metric takes positive and negative values. It is bounded from below and unbounded from above.*

*Proof.* First of all,  $\kappa_C(X, Y) \geq \kappa_C^{\text{AI}}(X, Y) \geq -\frac{1}{2}$  so the curvature is bounded from below. Second, at  $C = I_n$ ,  $X^\# = X$  and  $Y^\# = Y$  so  $\text{Diag}(X^\#) = \text{Diag}(Y^\#) = 0$  and  $\mu = 0$ . Hence,  $\kappa_{I_n}(X, Y) = \kappa_{I_n}^{\text{AI}}(X, Y) \leq 0$  and for example  $\kappa_{I_n}(E_{ij}, E_{ik}) = -\frac{1}{8} < 0$  with  $i \neq j \neq k \neq i \in \{1, \dots, n\}$  [23]. So the curvature takes negative values. Third, let  $X = \mathbf{1}\mathbf{1}^\top - I_n$  and  $Y = \mu\mathbf{1}^\top + \mathbf{1}\mu^\top - 2\text{diag}(\mu)$  with  $\text{sum}(\mu) = \mathbf{1}^\top \mu = 0$  where  $\mu \in \mathbb{R}^n$ . Let  $C = (1 - \rho)I_n + \rho\mathbf{1}\mathbf{1}^\top \in \text{Cor}^+(n)$  for  $\rho \in (-\frac{1}{n-1}, 1)$ . We show in the supplementary material that  $\kappa_C(X, Y)$  tends to  $+\infty$  when  $\rho \rightarrow -\frac{1}{n-1}$ , which proves that the curvature takes positive values and it is not bounded from above.  $\square$

Hence, the Riemannian manifold of full-rank correlation matrices endowed with the quotient-affine metric is geodesically complete but it is not a  $\text{CAT}(k)$  space for any  $k \in \mathbb{R}$ . In particular, it is not a Hadamard space as one could have hoped. It may be a problem for many algorithms because the Riemannian logarithm and the Fréchet mean are not ensured to be unique.

In Section 3, we relax the invariance under permutations that might be unsuitable in some contexts. In Section 4, we define non-permutation-invariant metrics which in addition bring uniqueness of the mean.

**3. Generalization of quotient-affine metrics.** In this section, we define a family of quotient metrics which generalize quotient-affine metrics and which are not invariant by permutations. In Section 3.1, we start by giving an overview of the congruence action of several subgroups of the general linear group  $\text{GL}(n)$  on SPD matrices. In particular, we show that affine-invariant metrics are exactly  $(\mathfrak{S}(n) \times \text{LT}^+(n))$ -invariant metrics so the family of  $\text{LT}^+(n)$ -invariant metrics appears as a natural generalization of affine-invariant metrics. In Section 3.2, we show that  $\text{LT}^+(n)$ -invariant metrics are exactly pullbacks by the Cholesky map of left-invariant metrics on the Lie group  $\text{LT}^+(n)$ . We call them Lie-Cholesky metrics. In particular, we define a



Lie group structure on SPD matrices such that the affine-invariant metric is a left-invariant metric on that Lie group. This is an unexpected result as the SPD cone endowed with an affine-invariant metric is always seen as a Riemannian homogeneous (symmetric) space. In Section 3.3, we define quotient-Lie-Cholesky metrics and we show that we can compute numerically the exponential map, the logarithm map and the Riemannian distance.

**3.1. Congruence actions of matrix Lie groups on SPD matrices.** The action of congruence of the real general linear group  $\text{GL}(n)$  on SPD matrices is:

$$(3.1) \quad \begin{cases} \text{GL}(n) \times \text{Sym}^+(n) & \longrightarrow \text{Sym}^+(n) \\ (A, \Sigma) & \longmapsto A\Sigma A^\top \end{cases} .$$

In the following theorem, we focus on the subactions given by the following subgroups.

1. The group of matrices with positive determinant  $\text{GL}^+(n)$ .
2. The special linear group  $\text{SL}(n) = \{A \in \text{GL}(n) | \det(A) = 1\} \subset \text{GL}^+(n)$ , which is interesting for the invariance of covariance matrices under volume-preserving linear transformations of the feature vector.
3. The orthogonal group  $\text{O}(n) = \{R \in \text{GL}(n) | RR^\top = I_n\}$  and the special orthogonal group  $\text{SO}(n) = \text{O}(n) \cap \text{SL}(n)$ , which are interesting for the invariance of covariance matrices under rotations and symmetries.
4. The group of lower triangular matrices with positive diagonal  $\text{LT}^+(n)$ , which is interesting because an  $\text{LT}^+(n)$ -invariant metric on SPD matrices is a left-invariant metric on a Lie group (cf. Section 3.2).
5. The positive diagonal group  $\text{Diag}^+(n) \subset \text{LT}^+(n)$ , which is interesting for the invariance of covariance matrices under scalings on each variable.
6. The group of positive real numbers  $\mathbb{R}^+$  which injects itself into the general linear group via the map  $\lambda \in \mathbb{R}^+ \mapsto \lambda I_n \in \text{GL}^+(n)$ . It is interesting for the invariance of covariance matrices under global scaling.
7. The permutation group  $\mathfrak{S}(n)$  which injects itself into the orthogonal group via the map  $\sigma \in \mathfrak{S}(n) \mapsto P_\sigma \in \text{O}(n)$  defined by  $[P_\sigma]_{ij} = \delta_{i, \sigma(j)}$  for all  $i, j \in \{1, \dots, n\}$ . It is interesting for the invariance of covariance matrices under permutation of the axes.

In the proof, we also use the following notations:  $\mathbb{R}^*$  is the group of invertible scalar matrices,  $\text{Diag}^*(n)$  is the group of invertible diagonal matrices,  $\text{LT}^*(n)$  is the group of invertible lower triangular matrices,  $\text{UT}^+(n)$  is the group of upper triangular matrices with positive diagonal.

**THEOREM 3.1** (Characterization of affine-invariant metrics). *Let  $g$  be a Riemannian metric on SPD matrices. The following statements are equivalent:*

1.  $g$  is  $\text{GL}(n)$ -invariant,
2.  $g$  is  $\text{GL}^+(n)$ -invariant,
3.  $g$  is  $\text{SL}(n)$ -invariant and  $\mathbb{R}^+$ -invariant,
4.  $g$  is  $\text{SO}(n)$ -invariant and  $\text{Diag}^+(n)$ -invariant,
5.  $g$  is  $\mathfrak{S}(n)$ -invariant and  $\text{LT}^+(n)$ -invariant.

*Proof.* The first statement clearly implies the others. To prove  $2 \implies 1$ , we need to take a general  $\text{GL}^+(n)$ -invariant metric on SPD matrices and prove that it is  $\text{GL}(n)$ -invariant. It amounts to prove that  $\text{SO}(n)$ -invariant inner products on symmetric matrices are  $\text{O}(n)$ -invariant, which is well known so  $2 \implies 1$ . To prove  $3 \implies 2$ ;  $4 \implies 2$  and  $5 \implies 1$ , it suffices to prove that the pairs of groups respectively generate  $\text{GL}^+(n)$ ,  $\text{GL}^+(n)$  and  $\text{GL}(n)$ .

The group generated by  $\mathrm{SL}(n)$  and  $\mathbb{R}^+$  is  $\mathrm{GL}^+(n)$  so  $3 \implies 2$ . The group generated by  $\mathrm{SO}(n)$  and  $\mathrm{Diag}^+(n)$  is also  $\mathrm{GL}^+(n)$  (it is clearly included in  $\mathrm{GL}^+(n)$  and conversely it contains  $\mathrm{Sym}^+(n)$  by the spectral theorem and  $\mathrm{GL}^+(n)$  by polar decomposition) so  $4 \implies 2$ .

To prove that  $5 \implies 1$ , let us show that the group generated by  $\mathfrak{S}(n)$  and  $\mathrm{LT}^+(n)$  is  $\mathrm{GL}(n)$ . First, the LU decomposition exists for any square matrix modulo a permutation. More precisely, for all  $A \in \mathrm{GL}(n)$ , there exists a permutation  $\sigma \in \mathfrak{S}(n)$ , a lower triangular matrix  $L \in \mathrm{LT}^*(n)$  and an upper triangular matrix  $U \in \mathrm{UT}^+(n)$  with  $\mathrm{Diag}(U) = I_n$  such that  $A = P_\sigma LU$ . If we permute the rows and columns of  $U$  according to the permutation  $\sigma_0 : k \mapsto n+1-k$ , then  $P_{\sigma_0} U P_{\sigma_0}^\top$  is lower triangular with ones on the diagonal so it is in  $\mathrm{LT}^+(n)$ . Hence the group generated by  $\mathfrak{S}(n)$  and  $\mathrm{LT}^*(n)$  is  $\mathrm{GL}(n)$ . Since  $\mathrm{LT}^*(n)$  is generated by  $\mathrm{LT}^+(n)$  and  $\mathrm{Diag}^*(n)$ , it suffices to prove that  $\mathrm{Diag}^*(n)$  is generated by  $\mathfrak{S}(n)$  and  $\mathrm{LT}^+(n)$ . Since  $\mathrm{Diag}^*(n)$  is generated by  $\mathrm{Diag}^+(n)$  and matrices of the form  $\mathrm{diag}(\pm 1, \dots, \pm 1)$  and since  $\mathrm{Diag}^+(n) \subset \mathrm{LT}^+(n)$ , it suffices to prove that matrices of the form  $\mathrm{diag}(\pm 1, \dots, \pm 1)$  are generated by  $\mathfrak{S}(n)$  and  $\mathrm{LT}^+(n)$ . By matrix product and permutations, it suffices to prove that  $\mathrm{diag}(-1, 1, \dots, 1)$  is generated by  $\mathfrak{S}(n)$  and  $\mathrm{LT}^+(n)$ . The following product of matrices:

$$(3.2) \quad \mathfrak{S}(2) \ni \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} \begin{array}{c} \nearrow \\ \downarrow \\ \nearrow \\ \downarrow \\ \nearrow \\ \downarrow = \end{array} \begin{array}{c} \overbrace{\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}}^{\in \mathrm{LT}^+(2)} \\ \begin{pmatrix} -1 & 1 \\ 1 & 0 \end{pmatrix} \end{array} \begin{array}{c} \nearrow \\ \downarrow \\ \nearrow \\ \downarrow \\ \nearrow \\ \downarrow = \end{array} \begin{array}{c} \overbrace{\begin{pmatrix} 1 & 1 \\ 0 & 1 \end{pmatrix}}^{\in \mathrm{UT}^+(2)} \\ \begin{pmatrix} -1 & 0 \\ 1 & 1 \end{pmatrix} \end{array} \begin{array}{c} \nearrow \\ \downarrow \\ \nearrow \\ \downarrow = \end{array} \begin{array}{c} \overbrace{\begin{pmatrix} 1 & 0 \\ -1 & 1 \end{pmatrix}}^{\in \mathrm{LT}^+(2)} \\ \begin{pmatrix} -1 & 0 \\ 0 & 1 \end{pmatrix} \end{array}$$

can be generalized in dimension  $n$  by adding a diagonal block  $I_{n-2}$ . Hence, the matrix  $\mathrm{diag}(-1, 1, \dots, 1)$  is generated by  $\mathfrak{S}(n)$  and  $\mathrm{LT}^+(n)$  so  $\mathrm{LT}^*(n)$  as well and finally  $\mathrm{GL}(n)$  entirely.  $\square$

One interpretation of this result is that these pairs of invariance ( $\mathrm{SL}(n)$  and  $\mathbb{R}^+$ ;  $\mathrm{SO}(n)$  and  $\mathrm{Diag}^+(n)$ ;  $\mathfrak{S}(n)$  and  $\mathrm{LT}^+(n)$ ) are incompatible except in the affine-invariant metrics. In other words, the family of  $\mathrm{LT}^+(n)$ -invariant metrics is a natural non-permutation-invariant extension of the family of affine-invariant metrics. In the next section, we characterize them as pullback metrics of left-invariant metrics on the Lie group  $\mathrm{LT}^+(n)$  by the Cholesky map.

### 3.2. Lie-Cholesky metrics.

**3.2.1.  $\mathrm{LT}^+(n)$ -invariant metrics are Lie-Cholesky metrics.** The manifold  $\mathrm{LT}^+(n)$  is an open set of the vector space  $\mathrm{LT}(n)$ . It is a Lie group where the internal law is the matrix multiplication.

**DEFINITION 3.2** (Lie-Cholesky metrics on  $\mathrm{Sym}^+(n)$ ). *A Lie-Cholesky metric on  $\mathrm{Sym}^+(n)$  is the pullback of a left-invariant metric on the Lie group  $(\mathrm{LT}^+(n), \times)$  via the Cholesky diffeomorphism. It is geodesically complete.*

**THEOREM 3.3** (Lie-Cholesky are  $\mathrm{LT}^+(n)$ -invariant metrics). *A Riemannian metric on  $\mathrm{Sym}^+(n)$  is a Lie-Cholesky metric if and only if it is  $\mathrm{LT}^+(n)$ -invariant.*

*Proof.* Let  $g$  be a metric on  $\mathrm{Sym}^+(n)$ . Let  $\Sigma \in \mathrm{Sym}^+(n)$ ,  $V \in T_\Sigma \mathrm{Sym}^+(n)$ ,  $L = \mathrm{Chol}(\Sigma)$  and  $Z = d_\Sigma \mathrm{Chol}(X)$ . Note that Equation (1.10) rewrites  $L^{-1}Z = d_{I_n} \mathrm{Chol}(L^{-1}VL^{-\top})$ . Thus, by definition of the pushforward, we have the following

equalities:

$$(3.3) \quad (\text{Chol}_*g)_L(Z, Z) = g_\Sigma(V, V)$$

$$(3.4) \quad (\text{Chol}_*g)_{I_n}(L^{-1}Z, L^{-1}Z) = g_{I_n}(L^{-1}VL^{-\top}, L^{-1}VL^{-\top})$$

Hence,  $g$  is a Lie-Cholesky metric if and only if  $\text{Chol}_*g$  is left-invariant if and only if the left terms are equal if and only if the right terms are equal if and only if  $g$  is invariant under the action of  $\text{LT}^+(n)$ .  $\square$

### 3.2.2. Consequences.

**COROLLARY 3.4.** *The Riemannian metrics on  $\text{Sym}^+(n)$  which are  $\text{LT}^+(n)$ -invariant are geodesically complete.*

**COROLLARY 3.5.** *Permutation-invariant Lie-Cholesky metrics are affine-invariant metrics.*

**COROLLARY 3.6** (Affine-invariant metrics on SPD matrices are left-invariant metrics on a Lie group). *The manifold  $\text{Sym}^+(n)$  has a structure of Lie group given by the matrix multiplication of the Cholesky factors, i.e. for all  $\Sigma, \Sigma' \in \text{Sym}^+(n)$ , denoting  $L = \text{Chol}(\Sigma)$  and  $L' = \text{Chol}(\Sigma')$ , the internal law is:*

$$(3.5) \quad \Sigma \star \Sigma' = (LL')(LL')^\top = L\Sigma'L^\top.$$

*The affine-invariant metrics  $g_\Sigma(X, X) = \alpha \text{tr}(\Sigma^{-1}X\Sigma^{-1}X) + \beta \text{tr}(\Sigma^{-1}X)^2$  are left-invariant metrics for this Lie group structure. They are pullbacks of the left-invariant metrics on the Lie group  $\text{LT}^+(n)$  characterized by the following inner products at  $I_n \in \text{LT}^+(n)$ , where  $Z \in T_{I_n}\text{LT}^+(n) \simeq \text{LT}(n)$ :*

$$(3.6) \quad \begin{aligned} \langle Z|Z \rangle_{\alpha, \beta} &= g_{I_n}(Z + Z^\top, Z + Z^\top) \\ &= \alpha \text{tr}((Z + Z^\top)(Z + Z^\top)) + \beta \text{tr}(Z + Z^\top)^2 \\ &= 2\alpha \text{tr}(ZZ^\top) + 2\alpha \text{tr}(\text{Diag}(Z)^2) + 4\beta \text{tr}(Z)^2. \end{aligned}$$

*We denote  $\langle \cdot | \cdot \rangle = \langle \cdot | \cdot \rangle_{1/2, 0}$ , i.e.  $\langle Z|Z' \rangle = \text{tr}(ZZ^\top) + \text{tr}(\text{Diag}(Z)^2)$  for  $Z \in \text{LT}(n)$ .*

The left-invariant metrics on  $\text{LT}^+(n)$  are parametrized by the inner products at  $I_n$ . Hence, they can be parameterized by self-adjoint positive definite linear maps  $f : \text{LT}(n) \rightarrow \text{LT}(n)$  as follows, for all  $Z, Z' \in \text{LT}(n)$ :

$$(3.7) \quad Z \cdot Z' = \langle f(Z)|Z' \rangle = \langle Z|f(Z') \rangle.$$

In other words, if  $\text{vec} : \text{LT}(n) \rightarrow \mathbb{R}^{n(n+1)/2}$  denotes the linear map defined by  $\text{vec}(Z) = (Z_{11}, Z_{21}, Z_{22}, \dots, Z_{n, n-1}, Z_{nn})^\top$  for all  $Z \in \text{LT}(n)$ , then there exists an SPD matrix  $A \in \text{Sym}^+(\frac{n(n+1)}{2})$  such that  $Z \cdot Z' = \text{vec}(Z)^\top A \text{vec}(Z')$ . Thus the Lie-Cholesky metrics are parametrized by self-adjoint positive definite linear maps  $f : \text{LT}(n) \rightarrow \text{LT}(n)$  or SPD matrices  $A \in \text{Sym}^+(\frac{n(n+1)}{2})$ . They are denoted  $\text{LC}(f)$  or  $\text{LC}(A)$ .

Therefore, we can see the family of Lie-Cholesky metrics as a natural non-permutation-invariant extension of affine-invariant metrics. Since they are left-invariant metrics on a Lie group, the numerical computation of the geodesics (exponential map, logarithm map) and the parallel transport is simpler and much more stable and precise than for other metrics [10]. The curvature is known in closed form modulo the computation of the operator  $\text{ad}^* : \text{LT}(n) \times \text{LT}(n) \rightarrow \text{LT}(n)$  defined for all  $X, Y, Z \in \text{LT}(n)$  by  $\langle \text{ad}^*(X)(Y)|Z \rangle = \langle Y|[X, Z] \rangle$  [4, Theorem 7.30]. The operator  $\text{ad}^*$  for Lie-Cholesky metrics is given in the following lemma.

LEMMA 3.7 (Operator  $\text{ad}^*$  for Lie-Cholesky metrics).

Let  $\text{LC}(A)$  be a Lie-Cholesky metric characterized by  $A \in \text{Sym}^+(\frac{n(n+1)}{2})$ . Then for  $X, Y \in \text{LT}(n)$ ,  $\text{vec}(\text{ad}^*(X)(Y)) = A(I_n \otimes X^\top - X \otimes I_n)A^{-1}\text{vec}(Y)$ .

*Proof.* In the matrix Lie group  $\text{LT}(n)$ , the Lie bracket is the commutator  $[X, Z] = XZ - ZX$ . Therefore:

$$\begin{aligned} \text{ad}^*(X)(Y) \cdot Z &= Y \cdot [X, Z] = \text{vec}(Y)^\top A \text{vec}(XZ - ZX) \\ &= \text{vec}(Y)^\top A(I_n \otimes X - X^\top \otimes I_n) \text{vec}(Z) \\ &= [\text{vec}(Y)^\top A(I_n \otimes X - X^\top \otimes I_n)A^{-1}] A \text{vec}(Z) \\ \text{vec}(\text{ad}^*(X)(Y)) &= A^{-1}(I_n \otimes X^\top - X \otimes I_n)A \text{vec}(Y). \quad \square \end{aligned}$$

**3.3. Quotient-Lie-Cholesky metrics.** Lie-Cholesky metrics being  $\text{LT}^+(n)$ -invariant, they are in particular  $\text{Diag}^+(n)$ -invariant as the affine-invariant metric. As recalled earlier, the smooth map  $\text{cor} : \text{Sym}^+(n) \rightarrow \text{Cor}^+(n)$  is a  $\text{Diag}^+(n)$ -invariant submersion so Lie-Cholesky metrics descend to quotient-Lie-Cholesky metrics on full-rank correlation matrices. Quotient-Lie-Cholesky metrics naturally extend quotient-affine metrics. To use them, one needs to compute the vertical and horizontal projections. After the definition of quotient-Lie-Cholesky metrics, the following lemma gives the expression of the vertical and horizontal distributions.

DEFINITION 3.8 (Quotient-Lie-Cholesky metric). A quotient-Lie-Cholesky metric is the quotient metric of a Lie-Cholesky metric by the submersion  $\text{cor} : \text{Sym}^+(n) \rightarrow \text{Cor}^+(n)$ . Equivalently, it is the pushforward metric by the diffeomorphism  $\Phi : L \in \text{LT}^1(n) \mapsto \text{Diag}(LL^\top)^{-1/2}LL^\top\text{Diag}(LL^\top)^{-1/2} \in \text{Cor}^+(n)$  of a quotient metric on  $\text{LT}^1(n) \simeq \text{Diag}(n) \setminus \text{LT}^+(n)$  defined by quotient via the submersion  $\pi : L \in \text{LT}^+(n) \mapsto \text{Diag}(L)^{-1}L \in \text{LT}^1(n)$ .

In this section, we express all the Riemannian operations of quotient-Lie-Cholesky metrics  $\text{LT}^1(n)$ . It suffices to push them forward by  $\Phi$  to get them on  $\text{Cor}^+(n)$ .

LEMMA 3.9 (Vertical distribution, horizontal distribution). The vertical and horizontal distributions associated to the quotient-Lie-Cholesky associated to  $f$  are  $\mathcal{V}_L = \text{Diag}(n)L$  and  $\mathcal{H}_L^{\text{LC}(f)} = Lf^{-1}(L^{-1}d_{LL^\top}\text{Chol}(\mathcal{S}_{(LL^\top)^{-1}}(\text{Hol}(n))))$ .

*Proof.* The vertical space is the tangent space of the fiber  $\text{Diag}^+(n)L$  or the kernel of the differential of the submersion  $d_L\pi(Z) = \text{Diag}(L)^{-1}(Z - \text{Diag}(L^{-1}Z)L)$ . Hence  $\mathcal{V}_L = \text{Diag}(n)L$ . Moreover, we know that when  $f = \text{Id}_{\text{LT}(n)}$ , the horizontal space is  $\mathcal{H}_L^{\text{AI}} = d_{LL^\top}\text{Chol}(\mathcal{S}_{(LL^\top)^{-1}}(\text{Hol}(n)))$ . Indeed, it is the pushforward by the Cholesky map of the horizontal space on SPD matrices (cf. Section 2.2). In other words, for all  $Z \in \mathcal{H}_L^{\text{AI}}$ , for all  $Z' \in \mathcal{V}_L$ , we have  $\langle L^{-1}Z | L^{-1}Z' \rangle = 0$ . Therefore, the horizontal space of the metric  $\text{LC}(f)$  is:

$$\begin{aligned} \mathcal{H}_L^{\text{LC}(f)} &= \{Z \in \text{LT}(n) | \forall Z' \in \mathcal{V}_L, \langle f(L^{-1}Z) | L^{-1}Z' \rangle = 0\} \\ &= \{Z \in \text{LT}(n) | \forall Z' \in \mathcal{V}_L, \langle L^{-1}Lf(L^{-1}Z) | L^{-1}Z' \rangle = 0\} \\ &= \{Z \in \text{LT}(n) | Lf(L^{-1}Z) \in \mathcal{H}_L^{\text{AI}}\} \\ &= Lf^{-1}(L^{-1}\mathcal{H}_L^{\text{AI}}). \quad \square \end{aligned}$$

Then, to compute the vertical and horizontal projections, it suffices to take bases of the vertical space  $\mathcal{V}_L$  and the horizontal space  $\mathcal{H}_L$ , to orthonormalize them by

Gram-Schmidt process and to project onto these orthonormal bases. Then the horizontal lift can be computed as follows, which allows to compute the metric and the exponential map.

LEMMA 3.10 (Horizontal lift, Riemannian metric, exponential map). *Given the horizontal projection  $\text{hor}_L : T_L \text{LT}^+(n) \rightarrow \mathcal{H}_L$  and the exponential map of the Lie-Cholesky metric  $\text{Exp}^{\text{LC}} : \text{LT}(n) \rightarrow \text{LT}^+(n)$ , we have for all  $L \in \text{LT}^+(n)$ , for all  $\Gamma = \text{Diag}(L)^{-1}L \in \text{LT}^1(n)$ , for all  $\xi \in T_\Gamma \text{LT}^1(n) \simeq \text{LT}^0(n)$ :*

- (Horizontal lift)  $\xi_L^\# = \text{hor}_L(\text{Diag}(L)\xi)$ ,
- (Riemannian metric)  $g_\Gamma^{\text{QLC}}(\xi, \xi) = \langle f(\Gamma^{-1}\text{hor}_\Gamma(\xi)) | \Gamma^{-1}\text{hor}_\Gamma(\xi) \rangle$ ,
- (Exponential map)  $\text{Exp}_\Gamma^{\text{QLC}}(t\xi) = \text{Diag}(\text{Exp}_\Gamma^{\text{LC}}(t \text{hor}(\xi)))^{-1} \text{Exp}_\Gamma^{\text{LC}}(t \text{hor}(\xi))$ .

*In particular, quotient-Lie-Cholesky metrics are geodesically complete.*

*Proof.* For all  $\xi \in \text{LT}^0(n)$ , we have  $d_L \pi(\text{Diag}(L)\xi) = \text{Diag}(L)^{-1}(\text{Diag}(L)\xi - \text{Diag}(\xi)L) = \xi$ . Hence  $\xi_L^\# = \text{hor}_L(\text{Diag}(L)\xi)$ . In particular,  $\xi_\Gamma^\# = \text{hor}_\Gamma(\xi)$ . Then the metric and the exponential map simply write  $g_\Gamma^{\text{QLC}}(\xi, \xi) = g_\Gamma^{\text{LC}}(\xi_\Gamma^\#, \xi_\Gamma^\#)$  and  $\text{Exp}_\Gamma^{\text{QLC}}(t\xi) = \pi(\text{Exp}_\Gamma^{\text{LC}}(t\xi_\Gamma^\#))$ .  $\square$

The Riemannian logarithm and the Riemannian distance can then be computed by minimizing the Lie-Cholesky distance along a fiber. The Lie-Cholesky distance is itself computed numerically with efficient tools on Lie groups [10].

In this section, we gave an overview of the congruence action of several matrix Lie groups on SPD matrices. It allowed to understand that the natural extension of affine-invariant metrics to non-permutation-invariant metrics is the family of  $\text{LT}^+(n)$ -invariant metrics. We showed that such metrics are pullbacks of left-invariant metrics on the Lie group  $\text{LT}^+(n)$  by the Cholesky diffeomorphism. Hence, the space of lower triangular matrices and the Cholesky map naturally appear when one wants to get rid of the invariance under permutations on SPD matrices. Following the same construction as for the quotient-affine metric, we built quotient-Lie-Cholesky metrics and we showed that most of the interesting Riemannian operations can be computed numerically. However, this geometry is not completely satisfying since we have no formula in closed form and no obvious nice geometric properties except geodesic completeness. Moreover, the numerical computation of the Riemannian logarithm and the Riemannian distance are done with the generic methods in quotient manifolds which can be heavy and unstable. Furthermore, the quotient-affine metric has unbounded curvature as shown in Section 2 and it is difficult to say something about the curvature of general quotient-Lie-Cholesky metrics.

Hence in the following section, we continue to rely on lower triangular matrices and the Cholesky map to define Riemannian metrics with simpler geometries than the general quotient geometry of quotient-Lie-Cholesky metrics.

**4. New geometric structures with unique mean.** In this section, we define new families of metrics on the open ellipptope of full-rank correlation matrices. They are based on simple geometries of subspaces of lower triangular matrices and transported to the ellipptope via the Cholesky map or the derived map  $\Theta$ . In Section 4.1, we introduce the poly-hyperbolic-Cholesky metrics which provide Riemannian symmetric structures. Then we introduce the Euclidean-Cholesky metrics (Section 4.2) and the log-Euclidean-Cholesky metrics (Section 4.3) that provide Euclidean structures. Equipped with a metric of one of these three families, the open ellipptope is Hadamard so the Riemannian logarithm and Fréchet mean are unique. In Section 4.4, we introduce a Lie group structure that allows to define Cartan-Schouten affine connections

and left-invariant metrics. The group mean of the canonical Cartan-Schouten connection is unique. In Section 4.5, we give the geodesics in dimension 2, which correspond to interpolations of one correlation coefficient.

**4.1. Riemannian symmetric space structure: poly-hyperbolic-Cholesky metrics.** In this section, we use the diffeomorphism  $\text{Chol}|_{\text{Cor}^+(n)} : \text{Cor}^+(n) \rightarrow \mathcal{L}$  where  $\mathcal{L}$  is the set of lower triangular matrices with positive diagonal such that each row is unit normed for the canonical Euclidean norm. Thus, the  $k$ -th row of  $L = \text{Chol}(C) \in \mathcal{L}$  writes  $(L_{k1}, \dots, L_{k,k-1}, L_{kk}, 0, \dots, 0)$  with  $L_{kk} > 0$ . It belongs to the open hemisphere  $\text{HS}^{k-1} = \{x \in \mathbb{R}^k \mid \|x\| = 1 \text{ and } x_k > 0\}$ . So each  $L \in \mathcal{L}$  is a point in  $\text{HS}^0 \times \dots \times \text{HS}^{n-1}$ . This construction is clearly bijective and diffeomorphic. Note that since  $L_{11} = 1$  and  $\text{HS}^0 = \{1\}$ , we can remove it from the Cartesian product. Hence, we can define the diffeomorphism  $\Psi : \mathcal{L} \rightarrow \text{HS}^1 \times \dots \times \text{HS}^{n-1}$ . Moreover, an open hemisphere is one of the avatars of the hyperbolic space, which is the Riemannian manifold of negative constant curvature.

Before introducing the Riemannian metric induced by the diffeomorphism  $\Psi$ , we recall the definition and Riemannian operations of the hyperbolic space in the model of the hyperboloid.

**THEOREM 4.1** (Hyperbolic geometry of the hyperboloid). *The vector space  $\mathbb{R}^{k+1}$  is endowed with the non-degenerate quadratic form  $Q(x) = \sum_{i=1}^k x_i^2 - x_{k+1}^2$ . We denote  $(x, y) \mapsto Q(x, y)$  the associated symmetric bilinear form. The hyperboloid  $\mathbb{H}^k$  is the Riemannian manifold defined by  $\mathbb{H}^k = \{x \in \mathbb{R}^{k+1} \mid Q(x) = -1\}$  endowed with the induced pseudo-metric, which is a Riemannian metric on  $\mathbb{H}^k$ . The tangent space writes  $T_x \mathbb{H}^k = \{v \in \mathbb{R}^{k+1} \mid Q(x, v) = 0\}$ . For all  $x, y \in \mathbb{H}^k$ ,  $v \in T_x \mathbb{H}^k$  and  $w \in T_x \mathbb{H}^k$  non colinear to  $v$ , the Riemannian operations are:*

- (Riemannian metric)  $g_x^{\mathbb{H}}(v, v) = Q(v)$ ,
- (Riemannian norm)  $\|v\|_{\mathbb{H}} = \sqrt{Q(v)}$ ,
- (Riemannian distance)  $d_{\mathbb{H}}(x, y) = \arccos(-Q(x, y))$ ,
- (Exponential map)  $\text{Exp}_x^{\mathbb{H}}(v) = \cosh(\|v\|_{\mathbb{H}})x + \sinh(\|v\|_{\mathbb{H}}) \frac{v}{\|v\|_{\mathbb{H}}}$ ,
- (Logarithm map)  $\text{Log}_x^{\mathbb{H}}(y) = d^{\mathbb{H}}(x, y) \frac{y + Q(x, y)x}{\|y + Q(x, y)x\|_{\mathbb{H}}}$ ,
- (Sectional curvature)  $\kappa_x^{\mathbb{H}}(v, w) = -1$ .

The formulae on the other models of the hyperbolic space can be obtained by pull-back via the appropriate diffeomorphism. In particular, the diffeomorphism between the open hemisphere  $\text{HS}^k = \{(x_1, \dots, x_{k+1}) \in \mathbb{R}^k \times \mathbb{R}^+ \mid \sum_{i=1}^{k+1} x_i^2 = 1\}$  and the hyperbolic space  $\mathbb{H}^k$  is:

$$(4.1) \quad \varphi^{\text{SH}} : (x_1, \dots, x_{k+1}) \in \text{HS}^k \mapsto \frac{1}{x_{k+1}}(x_1, \dots, x_k, 1) \in \mathbb{H}^k.$$

Thus the natural metric on the open hemisphere is the pullback metric  $g^{\text{HS}} = (\varphi^{\text{SH}})^* g^{\mathbb{H}}$ .

**DEFINITION 4.2** (Poly-hyperbolic-Cholesky metrics). *Let  $\alpha_1, \dots, \alpha_{n-1} > 0$  be positive coefficients. A poly-hyperbolic-Cholesky metric on  $\text{Cor}^+(n)$  is the pullback metric  $g^{\text{PHC}} = (\Psi \circ \text{Chol})^*(\alpha_1 g^{\text{HS}^1} \oplus \dots \oplus \alpha_{n-1} g^{\text{HS}^{n-1}})$  by the map  $\text{Chol} \circ \Psi$  of a weighted product metric on the product of hyperbolic spaces  $\text{HS}^1 \times \dots \times \text{HS}^{n-1}$ . The PHC metric with all weights equal to 1 is called the canonical PHC metric.*

**THEOREM 4.3** (Symmetric space structure). *The manifold of full-rank correlation matrices  $\text{Cor}^+(n)$  equipped with a poly-hyperbolic-Cholesky metric is a Riemannian symmetric space of non-positive sectional curvature bounded by  $[a, 0]$  with  $a = -\frac{1}{\min_{i \geq 2} \alpha_i}$ . For  $n \geq 3$ , it is not of constant curvature. The canonical PHC metric*

writes for all  $C \in \text{Cor}^+(n)$  and  $X \in T_C \text{Cor}^+(n) \simeq \text{Hol}(n)$ :

$$(4.2) \quad g_C^{\text{CPHC}}(X, X) = \|\text{Diag}(L)^{-1} L \text{low}_S(L^{-1} X L^{-\top})\|^2,$$

where  $L = \text{Chol}(C) \in \mathcal{L}$ . The square distance between  $C$  and  $C' = \phi(L')$  writes:

$$(4.3) \quad d_{\text{CPHC}}(C, C')^2 = \sum_{i=2}^n \arccos(-Q(L_i^\top, L'_i{}^\top))^2,$$

where  $L_i, L'_i$  are the  $i$ -th rows of  $L, L'$  respectively.

*Proof.* The product of Riemannian symmetric spaces is a Riemannian symmetric space. The product of manifolds with sectional curvature bounded by  $[a, b]$  with  $a \leq 0 \leq b$  has its sectional curvature bounded by  $[a, b]$ . (This is also valid for  $(-\infty, b]$  and  $[a, +\infty)$ .) Let  $k \geq 2$  such that  $\alpha_k = \min_{i \geq 2} \alpha_i$ . The values  $a = -\frac{1}{\alpha_k}$  and  $b = 0$  are clearly reached for  $n \geq 3$  by bivectors  $(X, Y)$  and  $(X, Z)$  respectively, where  $X = (0, \dots, 0, X_k, 0, \dots, 0)$ ,  $Y = (0, \dots, 0, Y_k, 0, \dots, 0)$  and  $Z = (Z_1, 0, \dots, 0) \in T(\text{HS}^1 \times \dots \times \text{HS}^{n-1})$ . To express the Riemannian metric, we take the pullback of the Riemannian metric on the hyperboloid  $\mathbb{H}^n$  defined by  $g_x(v, v) = \sum_{k=1}^n v_k^2 - v_{n+1}^2$  by the diffeomorphism  $\varphi^{\text{SH}} : (x_1, \dots, x_{n+1}) \in \text{HS}^n \mapsto \frac{1}{x_{n+1}}(x_1, \dots, x_n, 1) \in \mathbb{H}^n$ . We compute for all  $x \in \text{HS}^n$  and all  $v \in T_x \text{HS}^n$ , using  $\sum_{k=1}^{n+1} x_k^2 = 1$  and  $\sum_{k=1}^{n+1} x_k v_k = 0$ :

$$\begin{aligned} d_x \varphi^{\text{SH}}(v) &= \frac{1}{x_{n+1}} \left( v_1 - x_1 \frac{v_{n+1}}{x_{n+1}}, \dots, v_n - x_n \frac{v_{n+1}}{x_{n+1}}, -\frac{v_{n+1}}{x_{n+1}} \right), \\ g_x^{\text{HS}}(v, v) &= g_{\varphi(x)}^{\mathbb{H}}(d_x \varphi^{\text{SH}}(v), d_x \varphi^{\text{SH}}(v)) \\ &= \frac{1}{x_{n+1}^2} \left( \sum_{k=1}^n \left( v_k - x_k \frac{v_{n+1}}{x_{n+1}} \right)^2 - \frac{v_{n+1}^2}{x_{n+1}^2} \right) \\ &= \frac{1}{x_{n+1}^2} \left( \sum_{k=1}^n \left( v_k^2 - 2x_k v_k \frac{v_{n+1}}{x_{n+1}} + x_k^2 \frac{v_{n+1}^2}{x_{n+1}^2} \right) - \frac{v_{n+1}^2}{x_{n+1}^2} \right) \\ &= \frac{1}{x_{n+1}^2} \left( \sum_{k=1}^n v_k^2 + 2v_{n+1}^2 + (1 - x_{n+1}^2) \frac{v_{n+1}^2}{x_{n+1}^2} - \frac{v_{n+1}^2}{x_{n+1}^2} \right) \\ &= \frac{\|v\|^2}{x_{n+1}^2}. \end{aligned}$$

Hence, for all  $C \in \text{Cor}^+(n)$ ,  $X \in \text{Hol}(n)$ ,  $L = \text{Chol}(C) \in \mathcal{L}$  and  $Y = d_C \text{Chol}(X) = L \text{low}_S(L^{-1} X L^{-\top})$  with  $Y_{11} = \frac{1}{2} L_{11} [L^{-1} X L^{-\top}]_{11} = \frac{X_{11}}{2L_{11}} = 0$ , we have:

$$\begin{aligned} g_C^{\text{PHC}}(X, X) &= g_L^{\text{HS}^1 \times \dots \times \text{HS}^{n-1}}(Y, Y) \\ &= \sum_{i=2}^n \frac{\|Y_{i\bullet}\|^2}{L_{ii}^2} = \sum_{i=1}^n \|\text{Diag}(L)_{ii}^{-1} Y_{i\bullet}\|^2 \\ &= \|\text{Diag}(L)^{-1} Y\|^2 = \|\text{Diag}(L)^{-1} L \text{low}_S(L^{-1} X L^{-\top})\|^2. \end{aligned}$$

Note that the general PHC metric writes  $g_C^{\text{PHC}}(X, X) = \sum_{i=2}^n \|\text{Diag}(L)_{ii}^{-1} \alpha_i Y_{i\bullet}\|^2$  and the general PHC distance writes  $d(C, C')^2 = \sum_{i=2}^n \alpha_i \arccos(-Q(L_i^\top, L'_i{}^\top))^2$ .  $\square$



**4.2. Vector space structure: Euclidean-Cholesky metrics.** The Cholesky map was already used on SPD matrices [25]. The Euclidean-Cholesky metric is defined as the Euclidean metric on the Cholesky factor belonging to  $\text{LT}^+(n)$ . However, it is not complete because  $\text{LT}^+(n)$  is open in  $\text{LT}(n)$ . A solution to this problem is to take the logarithm of the diagonal before taking the Euclidean metric [19, 15]. It amounts to define the product metric of a Euclidean metric on the strictly lower part and a log-Euclidean metric on the diagonal part.

For correlation matrices, we can use the diffeomorphism  $\Theta : \text{Cor}^+(n) \rightarrow \text{LT}^1(n)$ . Since the diagonal is  $I_n$ , the two mentioned metrics reduce to the same metric on the open ellipsope. We call it the Euclidean-Cholesky metric.

**DEFINITION 4.4** (Euclidean-Cholesky metrics). *The Euclidean-Cholesky metrics on full-rank correlation matrices are pullback metrics by  $\Theta : \text{Cor}^+(n) \rightarrow \text{LT}^1(n)$  of inner products on  $\text{LT}^1(n) = I_n + \text{LT}^0(n)$ .*

**THEOREM 4.5** (Riemannian operations). *Let  $\|\cdot\|$  be a Euclidean norm on  $\text{LT}^0(n)$ . The Riemannian operations of the Euclidean-Cholesky metric associated to this norm are, for all  $C, C', C_i \in \text{Cor}^+(n)$ ,  $X \in T_C\text{Cor}^+(n) \simeq \text{Hol}(n)$ ,  $t \in \mathbb{R}$ :*

- (Exponential map)  $\text{Exp}_C(tX) = \Theta^{-1}(\Theta(C) + t d_C\Theta(X))$ ,
- (Logarithm map)  $\text{Log}_C(C') = (d_C\Theta)^{-1}(\Theta(C') - \Theta(C))$ ,
- (Geodesic)  $\gamma_{C \rightarrow C'}(t) = \Theta^{-1}((1-t)\Theta(C) + t\Theta(C'))$ ,
- (Distance)  $d(C, C') = \|\Theta(C') - \Theta(C)\|$ ,
- (Parallel transport)  $\Pi_{C \rightarrow C'}X = (d_{C'}\Theta)^{-1}(d_C\Theta(X))$ ,
- (Curvature) Null,
- (Fréchet mean)  $\bar{C} = \Theta^{-1}(\frac{1}{n} \sum_{i=1}^n \Theta(C_i))$ ,

where  $d_C\Theta(X) = \Theta(C)\text{low}_S(L^{-1}XL^{-\top}) - \frac{1}{2}\text{Diag}(L^{-1}XL^{-\top})\Theta(C)$  and  $L = \text{Chol}(C)$ . The Euclidean-Cholesky metrics are geodesically complete.

These metrics are flat, geodesically complete and the Riemannian operations are trivial. Since they reduce to (the pullback of) an inner product on a *vector space*, we prefer not to use the term *Lie group* for them, contrarily to the terminology of [14, 15]. We prefer to reserve it for Lie groups that are not vector spaces, such as the natural Lie group structure of  $\text{LT}^+(n)$  (with matrix multiplication) underlying Lie-Cholesky metrics.

**4.3. Vector space structure: log-Euclidean-Cholesky metrics.** Another map was used to Euclideanize the manifold  $\text{LT}^+(n)$ : the matrix logarithm [14]. Indeed, the matrix exponential is a smooth diffeomorphism from  $\text{LT}(n)$  to  $\text{LT}^+(n)$  [8]. We can use the same idea for correlation matrices since the matrix exponential is a smooth diffeomorphism from  $\text{LT}^0(n)$  to  $\text{LT}^1(n)$ . Moreover it has a particularly simple expression:

$$(4.4) \quad \exp(\xi) = \sum_{k=0}^{n-1} \frac{1}{k!} \xi^k,$$

because  $\text{LT}^0(n)$  is a nilpotent algebra. Then the logarithm  $\log : \text{LT}^1(n) \rightarrow \text{LT}^0(n)$  is simply:

$$(4.5) \quad \log(Z) = \sum_{k=1}^{n-1} \frac{(-1)^{k-1}}{k} (Z - I_n)^k.$$

Therefore, the differential of the logarithm writes:

$$(4.6) \quad d_Z \log(\xi) = \sum_{k=1}^{n-1} \frac{(-1)^k}{k} [(Z - I_n)^{k-1} \xi + (Z - I_n)^{k-2} \xi (Z - I_n) + \dots + \xi (Z - I_n)^{k-1}].$$

**DEFINITION 4.6** (Log-Euclidean-Cholesky metrics). *The log-Euclidean-Cholesky metrics on full-rank correlation matrices are pullback metrics by  $\log \circ \Theta : \text{Cor}^+(n) \rightarrow \text{LT}^0(n)$  of inner products on  $\text{LT}^0(n)$ .*

**THEOREM 4.7** (Riemannian operations). *Let  $\|\cdot\|$  be a Euclidean norm on  $\text{LT}^0(n)$ . The Riemannian operations of log-Euclidean-Cholesky metrics associated to this norm are, for all  $C, C', C_i \in \text{Cor}^+(n)$ ,  $X \in T_C \text{Cor}^+(n) \simeq \text{Hol}(n)$ ,  $t \in \mathbb{R}$ :*

- (Exponential map)  $\text{Exp}_C(tX) = \Theta^{-1} \circ \exp(\log(\Theta(C)) + t d_C(\log \circ \Theta)(X))$ ,
- (Logarithm map)  $\text{Log}_C(C') = (d_C(\log \circ \Theta))^{-1}(\log(\Theta(C')) - \log(\Theta(C)))$ ,
- (Geodesic)  $\gamma_{C \rightarrow C'}(t) = \Theta^{-1} \circ \exp((1-t) \log(\Theta(C)) + t \log(\Theta(C')))$ ,
- (Distance)  $d(C, C') = \|\log(\Theta(C')) - \log(\Theta(C))\|$ ,
- (Parallel transport)  $\Pi_{C \rightarrow C'} X = (d_{C'}(\log \circ \Theta))^{-1}(d_C \log \circ \Theta(X))$ ,
- (Curvature) Null,
- (Fréchet mean)  $\bar{C} = \Theta^{-1} \circ \exp(\frac{1}{n} \sum_{i=1}^n \log(\Theta(C_i)))$ ,

where  $d_C(\log \circ \Theta)(X) = d_{\Theta(C)} \log(d_C \Theta(X))$ . *The log-Euclidean-Cholesky metrics are geodesically complete.*

The computation of the Riemannian operations of the Euclidean-Cholesky metrics is more straightforward than those of the log-Euclidean-Cholesky because one has to compute the differential of the triangular matrix logarithm for the latter.

**4.4. Nilpotent Lie group structure.** Another interesting structure is given by the natural Lie group structure of  $\text{LT}^1(n)$  for the matrix multiplication. This equips full-rank correlation matrices with a Lie group structure via the diffeomorphism  $\Theta : \text{Cor}^+(n) \rightarrow \text{LT}^1(n)$ . Hence, left-invariant metrics can be defined. In analogy with  $\text{Sym}^+(n) \simeq \text{LT}^+(n)$ , they can also be called Lie-Cholesky metrics. Then all Riemannian operations can be computed numerically [10] and the space is ensured to be geodesically complete. However, this doesn't give information on the sign of the curvature.

More interestingly, one can rely on the canonical Cartan-Schouten connection to define the group exponential and the notion of group mean. We can also name them after Lie-Cholesky.

**THEOREM 4.8** (Group operations). *The group operations associated to the Lie-Cholesky group structure on full-rank correlation matrices are, for all  $C, C', C_i \in \text{Cor}^+(n)$ ,  $X \in T_C \text{Cor}^+(n) \simeq \text{Hol}(n)$ ,  $t \in \mathbb{R}$ :*

- (Exponential map)  $\text{Exp}_C(tX) = \Theta^{-1}(\Theta(C) \exp(t \Theta(C)^{-1} d_C \Theta(X)))$ ,
- (Logarithm map)  $\text{Log}_C(C') = (d_C \Theta)^{-1}(\Theta(C) \log(\Theta(C)^{-1} \Theta(C')))$ ,
- (Geodesic)  $\gamma_{C \rightarrow C'}(t) = \Theta^{-1}(\Theta(C) (\Theta(C)^{-1} \Theta(C'))^t)$ ,
- (Group mean) *Unique, characterized by  $\sum_{i=1}^k \log(\Theta(\bar{C})^{-1} \Theta(C_i)) = 0$ .*

*Proof.* The exponential map, logarithm map and geodesics are pullbacks by  $\Theta$  of corresponding operations in  $\text{LT}^1(n)$ , which are for all  $\Gamma, \Gamma' \in \text{LT}^+(n)$ ,  $\xi \in T_\Gamma \text{LT}^+(n) \simeq \text{LT}(n)$  and  $t \in \mathbb{R}$ :

- (Exponential map)  $\text{Exp}_\Gamma(t\xi) = \Gamma \exp(t \Gamma^{-1} \xi)$ ,
- (Logarithm map)  $\text{Log}_\Gamma(\Gamma') = \Gamma \log(\Gamma^{-1} \Gamma')$ ,
- (Geodesic)  $\gamma_{\Gamma \rightarrow \Gamma'}(t) = \text{Exp}_\Gamma(t \Gamma^{-1} \text{Log}_\Gamma(\Gamma')) = \Gamma(\Gamma^{-1} \Gamma')^t$ .

Since the Lie algebra  $\text{LT}^0(n)$  is nilpotent, the group mean  $\bar{\Gamma}$  of the finite sample  $\Gamma_1, \dots, \Gamma_k \in \text{LT}^1(n)$  is unique [5, Example 8.1.8]. It is characterized by  $0 = \sum_{i=1}^k \text{Log}_{\bar{\Gamma}}(\Gamma_i) = \bar{\Gamma} \sum_{i=1}^k \log(\bar{\Gamma}^{-1}\Gamma_i)$ , which is equivalent to  $\sum_{i=1}^k \log(\bar{\Gamma}^{-1}\Gamma_i) = 0$ .  $\square$

**4.5. Explicit geodesics in dimension 2.** In dimension 2, the ellipsope is reduced to one parameter. All full-rank correlation matrices write  $C = C(\rho) = \begin{pmatrix} 1 & \rho \\ \rho & 1 \end{pmatrix}$  with  $\rho \in (-1, 1)$ . Therefore, the quotient-affine metric and the metrics defined in Section 4 only depend on one scaling parameter. They actually split in two groups and the geodesics can be computed in closed forms. The two formulae in the following result provide two different interpolations of the correlation coefficient. The proof is in the supplementary material.

**THEOREM 4.9** (Geodesics in dimension 2). *Let  $C_1 = C(\rho_1), C_2 = C(\rho_2) \in \text{Cor}^+(2)$  with  $\rho_1, \rho_2 \in (-1, 1)$ .*

1. *Quotient-affine metrics and poly-hyperbolic-Cholesky metrics coincide (up to a scaling factor). The geodesic between  $C_1$  and  $C_2$  is  $C(\rho(t))$  for  $t \in \mathbb{R}$  where:*

$$(4.7) \quad \rho(t) = \frac{\rho_1 \cosh(\lambda t) + \sinh(\lambda t)}{\rho_1 \sinh(\lambda t) + \cosh(\lambda t)},$$

where  $\lambda = \log \sqrt{\frac{1+\rho_2}{1-\rho_2}} - \log \sqrt{\frac{1+\rho_1}{1-\rho_1}}$  is known as the difference of the Fisher transformation of the correlation coefficients  $\rho_1$  and  $\rho_2$ .

2. *Euclidean-Cholesky and log-Euclidean-Cholesky metrics coincide. The geodesic between  $C_1$  and  $C_2$  is  $C(\rho(t))$  for  $t \in \mathbb{R}$  where:*

$$(4.8) \quad \rho(t) = \frac{F(t)}{\sqrt{1 + F(t)^2}},$$

where  $F(t) = (1 - t) \frac{\rho_1}{\sqrt{1-\rho_1^2}} + t \frac{\rho_2}{\sqrt{1-\rho_2^2}}$ . This geodesic also coincides with the Lie-Cholesky group geodesic of Section 4.4.

**5. Conclusion.** In this work, we proposed new Riemannian metrics on the open ellipsope of full-rank correlation matrices that are not invariant under permutations. To the best of our knowledge, all the existing geometric structures were invariant under permutations. Thus the geometries we propose significantly departs from the classical structures. This can be a good assumption in some applications and an irrelevant characteristic in some others. We generalized the recently introduced quotient-affine metrics by studying the congruence action of several matrix Lie groups on SPD matrices. We showed that the family of  $\text{LT}^+(n)$ -invariant metrics is a natural non-permutation-invariant generalization of affine-invariant metrics. Moreover, they are pullbacks of left-invariant metrics on the Lie group  $\text{LT}^+(n)$  by the Cholesky map. They are invariant under the congruence action of positive diagonal matrices so they descend to the ellipsope. We explained that the main Riemannian operations can be computed numerically for these quotient-Lie-Cholesky metrics. However, we also showed that the curvature of quotient-affine metrics is unbounded and we can conjecture that the situation is not better for quotient-Lie-Cholesky metrics. In addition, the Riemannian operations are not computable in closed form a priori.

That is why we introduced new Riemannian metrics on the ellipsope in a different way. We kept the Cholesky map which seems to be a good alternative to the invariance under permutations since the space of lower triangular matrices is not stable by permutations. Thus we defined the poly-hyperbolic-Cholesky (PHC) metrics

which provide non-positively curved Riemannian symmetric space structures. We also defined two kinds of vector space structures that are flat, geodesically complete and for which all operations are known in closed form. Thus, these three families of metrics provide a Hadamard structure, in particular the Riemannian logarithm and the Fréchet mean are unique. We also put forward a nilpotent Lie group structure for which the group mean is unique. Finally, we proved that in dimension 2, the PHC geodesics are the quotient-affine geodesics and the geodesics of the three last structures coincide. This provides a new interpolation of the correlation coefficient.

It would be nice to test these new metrics on different kinds of data in future works. Moreover, all metrics on correlation matrices provide new product metrics on covariance matrices by decoupling the scales of the variables and the correlations between them. This approach seems promising since in many problems, the correlation gives more information than the covariance on the strength of the relations between the variables, although the scales can remain interesting. Thus, one question could be to adjust the weights between the two components and also between the scales of the variables. The possibilities are multiplied now we have many metrics on correlation matrices. Another direction of research is to investigate permutation-invariant Riemannian metrics on correlation matrices with a simpler geometry than the one of the quotient-affine metrics, for example Hadamard or even flat.

**Acknowledgments.** This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation program (grant G-Statistics agreement No 786854). This work has been supported by the French government, through the UCAJEDI Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-15-IDEX-01 and through the 3IA Côte d'Azur Investments in the Future project managed by the National Research Agency (ANR) with the reference number ANR-19-P3IA-0002. The authors warmly thank Nicolas Guigui for insightful discussions on lower triangular matrices.

**A. Proof appendix of Theorem 2.1.** Let  $X = \mathbf{1}\mathbf{1}^\top - I_n$  and  $Y = \mu\mathbf{1}^\top + \mathbf{1}\mu^\top - 2\text{diag}(\mu)$  with  $\text{sum}(\mu) = \mathbf{1}^\top \mu = 0$  where  $\mu \in \mathbb{R}^n$ . Let  $C = (1-\rho)I_n + \rho\mathbf{1}\mathbf{1}^\top \in \text{Cor}^+(n)$  for  $\rho \in (-\frac{1}{n-1}, 1)$ . Let us show that  $\kappa_C(X, Y)$  tends to  $+\infty$  when  $\rho \rightarrow -\frac{1}{n-1}$ , which proves that the curvature is not bounded from above.

The symmetric matrix  $\mathbf{1}\mathbf{1}^\top$  has two eigenvalues: 0 with multiplicity  $n-1$  and  $n$  with multiplicity 1. Since  $(\mathbf{1}\mathbf{1}^\top)^2 = n\mathbf{1}\mathbf{1}^\top$ , the minimal polynomial is  $P_{\mathbf{1}\mathbf{1}^\top}(x) = x(x-n)$  for  $x \in \mathbb{R}$ . For all  $\alpha, \beta \in \mathbb{R}$ , the symmetric matrix  $\Sigma = \alpha I_n + \beta\mathbf{1}\mathbf{1}^\top$  has minimal polynomial  $P_\Sigma(x) = (x-\alpha)(x-(\alpha+n\beta))$  for  $x \in \mathbb{R}$ , which is of degree 2. Hence  $\Sigma$  is positive definite if and only if  $\alpha > 0$  and  $\alpha+n\beta > 0$ . In this case, its inverse is a polynomial in  $\mathbf{1}\mathbf{1}^\top$  of degree 1. More precisely,  $\Sigma^{-1} = \alpha' I_n + \beta'\mathbf{1}\mathbf{1}^\top$  with  $\alpha' = \frac{1}{\alpha}$  and  $\alpha' + n\beta' = \frac{1}{\alpha+n\beta}$ , i.e.  $\beta' = -\frac{\beta}{\alpha+n\beta}$ . Note that  $\alpha\beta' + \beta\alpha' + n\beta\beta' = 0$ .

Moreover, for all  $i \neq j \in \{1, \dots, n\}$ ,  $[\Sigma \bullet \Sigma^{-1}]_{ii} = (\alpha+\beta)(\alpha'+\beta')$  and  $[\Sigma \bullet \Sigma^{-1}]_{ij} = \beta\beta'$ . Therefore,  $I_n + \Sigma \bullet \Sigma^{-1} = AI_n + B\mathbf{1}\mathbf{1}^\top$  with  $A = 1 + (\alpha+\beta)(\alpha'+\beta') - \beta\beta' = 2 + \alpha\beta' + \beta\alpha' = \frac{2\alpha(\alpha+n\beta)+n\beta^2}{\alpha(\alpha+n\beta)}$  and  $B = \beta\beta'$ . Note that  $A+nB = 2$ . And  $(I_n + \Sigma \bullet \Sigma^{-1})^{-1} = A'I_n + B'\mathbf{1}\mathbf{1}^\top$  with  $A' = \frac{1}{A}$  and  $B' = -\frac{B}{2A}$ .

We compute  $\kappa_C(X, Y)$  where  $C = \alpha I_n + \beta\mathbf{1}\mathbf{1}^\top \in \text{Cor}^+(n)$  with  $\alpha + \beta = 1$ .

$$\begin{aligned} C^{-1}X &= (\alpha' I_n + \beta'\mathbf{1}\mathbf{1}^\top)(\mathbf{1}\mathbf{1}^\top - I_n) \\ &= -\alpha' I_n + (\alpha' + (n-1)\beta')\mathbf{1}\mathbf{1}^\top \end{aligned}$$

$$\begin{aligned}
 &= -\frac{1}{\alpha}I_n + \frac{1}{\alpha(\alpha+n\beta)}\mathbf{1}\mathbf{1}^\top, \\
 \text{Diag}(C^{-1}X) &= (n-1)\beta'I_n, \\
 (I_n + C \bullet C^{-1})^{-1}\text{Diag}(C^{-1}X)\mathbf{1} &= (n-1)\beta'(A'I_n + B'\mathbf{1}\mathbf{1}^\top)\mathbf{1} \\
 &= \frac{n-1}{2}\beta'\mathbf{1}, \\
 X^\# &= X - (n-1)\beta'C, \\
 \text{Diag}(X^\#) &= -(n-1)\beta'I_n, \\
 D(X, Y) &= \text{Diag}(C^{-1}\text{Diag}(X^\#)Y^\# - C^{-1}Y^\#C^{-1}\text{Diag}(X^\#)C) \\
 &= 0, \\
 C^{-1}Y &= (\alpha'I_n + \beta'\mathbf{1}\mathbf{1}^\top)(\mu\mathbf{1}^\top + \mathbf{1}\mu^\top - 2\text{diag}(\mu)) \\
 &= \alpha'\mu\mathbf{1}^\top + (\alpha' + (n-2)\beta')\mathbf{1}\mu^\top - 2\alpha'\text{diag}(\mu), \\
 \text{Diag}(C^{-1}Y) &= (n-2)\beta'\text{diag}(\mu), \\
 (I_n + C \bullet C^{-1})^{-1}\text{Diag}(C^{-1}Y)\mathbf{1} &= (n-2)\beta'(A'I_n + B'\mathbf{1}\mathbf{1}^\top)\mu \\
 &= (n-2)\beta'A'\mu, \\
 Y^\# &= Y - (n-2)\beta'A'(\text{diag}(\mu)C + C\text{diag}(\mu)), \\
 \text{Diag}(Y^\#) &= -2(n-2)\beta'A'\text{diag}(\mu), \\
 C^{-1}\text{Diag}(Y^\#)X^\# &= -2(n-2)\beta'A'(\alpha'I_n + \beta'\mathbf{1}\mathbf{1}^\top)\text{diag}(\mu)(\mathbf{1}\mathbf{1}^\top - I_n) \\
 &= -2(n-2)\beta'A'(-\alpha'\text{diag}(\mu) + \alpha'\mu\mathbf{1}^\top - \beta'\mathbf{1}\mu^\top),
 \end{aligned}$$

$$\begin{aligned}
 &C^{-1}X^\#C^{-1}\text{Diag}(Y^\#)C \\
 &= -2(n-2)\beta'A'(-\alpha'I_n + (\alpha' + (n-1)\beta')\mathbf{1}\mathbf{1}^\top)(\alpha'I_n + \beta'\mathbf{1}\mathbf{1}^\top)\text{diag}(\mu)(\alpha I_n + \beta\mathbf{1}\mathbf{1}^\top) \\
 &= -2(n-2)\beta'A'(-\alpha'I_n + (\alpha' + (n-1)\beta')\mathbf{1}\mathbf{1}^\top)(\text{diag}(\mu) + \alpha'\beta\mu\mathbf{1}^\top + \alpha\beta'\mathbf{1}\mu^\top) \\
 &= -2(n-2)\beta'A'[-\alpha'(\text{diag}(\mu) + \alpha'\beta\mu\mathbf{1}^\top + \alpha\beta'\mathbf{1}\mu^\top) + (\alpha' + (n-1)\beta')(1 + n\alpha\beta')\mathbf{1}\mu^\top],
 \end{aligned}$$

$$\begin{aligned}
 &C^{-1}\text{Diag}(Y^\#)X^\# - C^{-1}X^\#C^{-1}\text{Diag}(Y^\#)C \\
 &= -2(n-2)\beta'A'[\alpha'(1 + \alpha'\beta)\mu\mathbf{1}^\top - (\alpha' + (n-1)\beta')(1 + n\alpha\beta')\mathbf{1}\mu^\top],
 \end{aligned}$$

$$\begin{aligned}
 \text{(A.1)} \quad D(Y, X) &= -2(n-2)\beta'A' \left( \frac{1}{\alpha^2} - \frac{1 + (n-1)\alpha\beta'}{\alpha + n\beta} \right) \text{diag}(\mu) \\
 &= -2(n-2)\beta'A' \left( \frac{1}{\alpha^2} - \frac{1}{(\alpha + n\beta)^2} \right) \text{diag}(\mu) \\
 &= 2n(n-2) \frac{\beta^2}{2\alpha(\alpha + n\beta) + n\beta^2} \frac{2\alpha + n\beta}{\alpha^2(\alpha + n\beta)^2} \text{diag}(\mu),
 \end{aligned}$$

$$\begin{aligned}
 \mathbf{1}^\top D(I_n + C \bullet C^{-1})^{-1} D\mathbf{1} &= A' \left( 2n(n-2) \frac{\beta^2}{2\alpha(\alpha + n\beta) + n\beta^2} \frac{2\alpha + n\beta}{\alpha^2(\alpha + n\beta)^2} \right)^2 \|\mu\|^2 \\
 &= 4n^2(n-2)^2 \frac{\beta^4(2\alpha + n\beta)^2}{[\alpha(\alpha + n\beta)(2\alpha(\alpha + n\beta) + n\beta^2)]^3} \|\mu\|^2,
 \end{aligned}$$

where we used  $\alpha + \beta = 1$  from Equation (A.1).

Now, we compute  $g_C^{\text{QA}}(X, X)$ ,  $g_C^{\text{QA}}(Y, Y)$  and  $g_C^{\text{QA}}(X, Y)$ .

$$\begin{aligned}
g_C(X, X) &= \text{tr}((C^{-1}X)^2) - 2\mathbf{1}^\top \text{Diag}(C^{-1}X)(I_n + C \bullet C^{-1})^{-1} \text{Diag}(C^{-1}X)\mathbf{1} \\
&= \text{tr}\left((- \alpha' I_n + (\alpha' + (n-1)\beta')\mathbf{1}\mathbf{1}^\top\right)^2 - 2(n-1)^2\beta'^2 \text{sum}(A' I_n + B'\mathbf{1}\mathbf{1}^\top) \\
&= \text{tr}(\alpha'^2 I_n + (n(\alpha' + (n-1)\beta')^2 - 2\alpha'(\alpha' + (n-1)\beta'))\mathbf{1}\mathbf{1}^\top) \\
&\quad - 2(n-1)^2\beta'^2 n(A' + nB') \\
&= n((n-1)\alpha'^2 + 2(n-1)^2\alpha'\beta' + n(n-1)^2\beta'^2) - n(n-1)^2\beta'^2 \\
&= \frac{n(n-1)}{\alpha^2(\alpha+n\beta)^2}((\alpha+n\beta)^2 - 2(n-1)\beta(\alpha+n\beta) + (n-1)^2\beta^2) \\
&= \frac{n(n-1)}{\alpha^2(\alpha+n\beta)^2}, \\
g_C(Y, Y) &= \text{tr}((C^{-1}Y)^2) - 2\mathbf{1}^\top \text{Diag}(C^{-1}Y)(I_n + C \bullet C^{-1})^{-1} \text{Diag}(C^{-1}Y)\mathbf{1} \\
&= \text{tr}\left((\alpha'\mu\mathbf{1}^\top + (\alpha' + (n-2)\beta')\mathbf{1}\mu^\top - 2\alpha'\text{diag}(\mu))^2\right) \\
&\quad - 2(n-2)^2\beta'^2\mu^\top(A' I_n + B'\mathbf{1}\mathbf{1}^\top)\mu \\
&= \text{tr}(4\alpha'^2\text{diag}(\mu)^2 + \alpha'(\alpha' + (n-2)\beta')(n\mu\mu^\top + \|\mu\|^2\mathbf{1}\mathbf{1}^\top)) \\
&\quad - 2\alpha'\text{tr}((2\alpha' + (n-2)\beta')\mu\mu^\top + \alpha'(\mu \bullet \mu)\mathbf{1}^\top + (\alpha' + (n-2)\beta')\mathbf{1}(\mu \bullet \mu)^\top)) \\
&\quad - 2(n-2)^2\beta'^2 A'\|\mu\|^2 \\
&= \|\mu\|^2(4\alpha'^2 + 2n\alpha'(\alpha' + (n-2)\beta') - 4\alpha'(2\alpha' + (n-2)\beta') - 2(n-2)^2\beta'^2 A') \\
&= \|\mu\|^2(2(n-2)\underbrace{\alpha'(\alpha' + (n-2)\beta')}_{\frac{1+\beta}{\alpha^2(\alpha+n\beta)}} - 2(n-2)^2 \underbrace{\beta'^2 A'}_{\frac{\beta^2}{\alpha(\alpha+n\beta)(2\alpha(\alpha+n\beta)+n\beta^2)}}) \\
&= \frac{2(n-2)\|\mu\|^2}{\alpha^2(\alpha+n\beta)(2\alpha(\alpha+n\beta)+n\beta^2)} \underbrace{((1+\beta)(2\alpha(\alpha+n\beta)+n\beta^2) - (n-2)\alpha\beta^2)}_{2\alpha(1+\beta)(\alpha+n\beta)+2\beta^2(\alpha+n\beta)=2(\alpha+n\beta)} \\
&= \frac{4(n-2)\|\mu\|^2}{\alpha^2(2\alpha(\alpha+n\beta)+n\beta^2)},
\end{aligned}$$

$$\begin{aligned}
g_C(X, Y) &= \text{tr}(C^{-1}XC^{-1}Y) - 2\mathbf{1}^\top \text{Diag}(C^{-1}X)(I_n + C \bullet C^{-1})^{-1} \text{Diag}(C^{-1}Y)\mathbf{1} \\
&= \text{tr}((- \alpha' I_n + (\alpha' + (n-1)\beta')\mathbf{1}\mathbf{1}^\top)(\alpha'\mu\mathbf{1}^\top + (\alpha' + (n-2)\beta')\mathbf{1}\mu^\top - 2\alpha'\text{diag}(\mu))) \\
&\quad - 2(n-1)(n-2)\beta'^2\mathbf{1}^\top(A' I_n + \beta'\mathbf{1}\mathbf{1}^\top)\mu \\
&= \text{constant} \times \text{sum}(\mu) = 0.
\end{aligned}$$

Finally:

$$\frac{3}{8} \frac{\mathbf{1}^\top D(I_n + C \bullet C^{-1})^{-1} D\mathbf{1}}{g_C(X, X)g_C(Y, Y) - g_C(X, Y)^2} = \frac{3n(n-2)}{8(n-1)} \frac{\alpha\beta^4}{\alpha+n\beta} \left( \frac{2\alpha+n\beta}{2\alpha(\alpha+n\beta)+n\beta^2} \right)^2.$$

When  $\beta \rightarrow -\frac{1}{n-1}$  (and  $\alpha = 1 - \beta \rightarrow \frac{n}{n-1}$ ), we have  $\alpha + n\beta \rightarrow 0$  and we have  $\alpha\beta^4 \left( \frac{2\alpha+n\beta}{2\alpha(\alpha+n\beta)+n\beta^2} \right)^2 \rightarrow \frac{n}{(n-1)^3}$  so this quantity tends to  $+\infty$ . Finally with  $C = (1 - \rho)I_n + \rho\mathbf{1}\mathbf{1}^\top \in \text{Cor}^+(n)$  with  $\rho \in (-\frac{1}{n-1}; 1)$ ,  $X = I_n - \mathbf{1}\mathbf{1}^\top \in T_C\text{Cor}^+(n)$

and  $Y = \mu \mathbf{1}^\top + \mathbf{1} \mu^\top - 2 \text{diag}(\mu) \in T_C \text{Cor}^+(n)$ , we have:

$$(A.2) \quad \kappa_C(X, Y) \geq -\frac{1}{2} + \frac{3}{8} \frac{\mathbf{1}^\top D(I_n + C \bullet C^{-1})^{-1} D \mathbf{1}}{g_C(X, X)g_C(Y, Y) - g_C(X, Y)^2} \xrightarrow{\rho \rightarrow -\frac{1}{n-1}} +\infty.$$

This proves that the quotient-affine sectional curvature is not bounded from above.

### B. Proof of Theorem 4.9.

1. The formula of the geodesic is known for the quotient-affine metrics in dimension 2. Hence it suffices to show that the quotient-affine metrics and the poly-hyperbolic-Cholesky metrics coincide up to a scaling factor. Let  $C = C(\rho)$  and  $X = \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} \in T_C \text{Cor}^+(2)$ . We compute the quotient-affine metric  $g_C^{\text{QA}}(X, X) = \text{tr}(C^{-1}XC^{-1}X) - 2 \text{sum}(D(I_n + C \bullet C^{-1})^{-1}D)$  and the canonical PHC metric  $g_C^{\text{CPHC}}(X, X) = \|\text{Diag}(L)^{-1}L \text{low}_S(L^{-1}XL^{-\top})\|^2$  where  $D = \text{Diag}(C^{-1}X)$  and  $L = \text{Chol}(C)$ .

$$\begin{aligned} C^{-1} &= \frac{1}{1-\rho^2} \begin{pmatrix} 1 & -\rho \\ -\rho & 1 \end{pmatrix}, \\ C^{-1}X &= \frac{x}{1-\rho^2} \begin{pmatrix} -\rho & 1 \\ 1 & -\rho \end{pmatrix}, \\ C^{-1}XC^{-1}X &= \frac{x^2}{(1-\rho^2)^2} \begin{pmatrix} 1+\rho^2 & -2\rho \\ -2\rho & 1+\rho^2 \end{pmatrix}, \\ \text{tr}(C^{-1}XC^{-1}X) &= \frac{2(1+\rho^2)}{(1-\rho^2)^2} x^2, \\ I_n + C \bullet C^{-1} &= \frac{1}{1-\rho^2} \begin{pmatrix} 2-\rho^2 & -\rho^2 \\ -\rho^2 & 2-\rho^2 \end{pmatrix}, \\ (I_n + C \bullet C^{-1})^{-1} &= \frac{1}{4} \begin{pmatrix} 2-\rho^2 & \rho^2 \\ \rho^2 & 2-\rho^2 \end{pmatrix}, \\ D = \text{Diag}(C^{-1}X) &= -\frac{\rho x}{1-\rho^2} I_2, \\ \text{sum}(D(I_n + C \bullet C^{-1})^{-1}D) &= \frac{\rho^2 x^2}{(1-\rho^2)^2}, \\ g_C(X, X) &= \frac{2x^2}{(1-\rho^2)^2}, \end{aligned}$$

$$\begin{aligned} L = \text{Chol}(C) &= \begin{pmatrix} 1 & 0 \\ \rho & \sqrt{1-\rho^2} \end{pmatrix}, \\ L^{-1}XL^{-\top} &= \begin{pmatrix} 1 & 0 \\ -\frac{\rho}{\sqrt{1-\rho^2}} & \frac{1}{\sqrt{1-\rho^2}} \end{pmatrix} \begin{pmatrix} 0 & x \\ x & 0 \end{pmatrix} \begin{pmatrix} 1 & -\frac{\rho}{\sqrt{1-\rho^2}} \\ 0 & \frac{1}{\sqrt{1-\rho^2}} \end{pmatrix} \\ &= \begin{pmatrix} 0 & x \\ \frac{x}{\sqrt{1-\rho^2}} & -\frac{\rho x}{\sqrt{1-\rho^2}} \end{pmatrix} \begin{pmatrix} 1 & -\frac{\rho}{\sqrt{1-\rho^2}} \\ 0 & \frac{1}{\sqrt{1-\rho^2}} \end{pmatrix} \\ &= \begin{pmatrix} 0 & \frac{x}{\sqrt{1-\rho^2}} \\ \frac{x}{\sqrt{1-\rho^2}} & -\frac{2\rho x}{1-\rho^2} \end{pmatrix}, \end{aligned}$$



$$\begin{aligned}
\text{Diag}(L)^{-1}L \text{low}_S(L^{-1}XL^{-\top}) &= \begin{pmatrix} 1 & 0 \\ \frac{\rho}{\sqrt{1-\rho^2}} & 1 \end{pmatrix} \begin{pmatrix} 0 & 0 \\ \frac{x}{\sqrt{1-\rho^2}} & -\frac{\rho x}{1-\rho^2} \end{pmatrix} \\
&= \begin{pmatrix} 0 & 0 \\ \frac{x}{\sqrt{1-\rho^2}} & -\frac{\rho x}{1-\rho^2} \end{pmatrix}, \\
g_C^{\text{CPHC}}(X, X) &= \left( \frac{1}{1-\rho^2} + \frac{\rho^2}{(1-\rho^2)^2} \right) x^2 \\
&= \frac{x^2}{(1-\rho^2)^2}.
\end{aligned}$$

2. Euclidean-Cholesky and log-Euclidean-Cholesky metrics coincide in dimension 2 because  $\exp(\xi) = I_n + \xi$  and  $\log(\Gamma) = \Gamma - I_2$  for  $\xi \in \text{LT}^0(2)$  and  $\Gamma \in \text{LT}^1(2)$ . Thus their common Riemannian exponential and logarithm in  $\text{LT}^1(2)$  are simply  $\text{Exp}_\Gamma(\xi) = \Gamma + \xi$  and  $\text{Log}_\Gamma(\Gamma') = \Gamma' - \Gamma$ . On the other hand, the group exponential in  $\text{LT}^1(2)$  is  $\text{Exp}_\Gamma^{\text{LT}^1(2)}(\xi) = \Gamma \exp(\Gamma^{-1}\xi) = \Gamma(I_n + \Gamma^{-1}\xi) = \Gamma + \xi = \text{Exp}_\Gamma(\xi)$ . Hence, the group geodesics coincide with the (log-)Euclidean-Cholesky geodesics. Let us compute them.

$$\begin{aligned}
\Gamma_1 &:= \text{Diag}(\text{Chol}(C_1))^{-1}\text{Chol}(C_1) = \begin{pmatrix} 1 & 0 \\ \frac{\rho_1}{\sqrt{1-\rho_1^2}} & 1 \end{pmatrix}, \\
\Gamma_2 &:= \text{Diag}(\text{Chol}(C_2))^{-1}\text{Chol}(C_2) = \begin{pmatrix} 1 & 0 \\ \frac{\rho_2}{\sqrt{1-\rho_2^2}} & 1 \end{pmatrix}, \\
C(t) &= \Theta^{-1}((1-t)\Gamma_1 + t\Gamma_2) = \Theta^{-1} \begin{pmatrix} 1 & 0 \\ F(t) & 1 \end{pmatrix} \\
&= \text{cor} \begin{pmatrix} 1 & F(t) \\ F(t) & 1 + F(t)^2 \end{pmatrix} = \begin{pmatrix} 1 & \frac{F(t)}{\sqrt{1+F(t)^2}} \\ \frac{F(t)}{\sqrt{1+F(t)^2}} & 1 \end{pmatrix}.
\end{aligned}$$

We can also compute the (log-)Euclidean-Cholesky metric in dimension 2.

$$\begin{aligned}
g_C(X, X) &= \|d_C\Theta(X)\|^2 = \|(\Theta \circ C)'(\rho)\|^2 \\
&= f'(\rho)^2 x^2,
\end{aligned}$$

where  $f(\rho) = \frac{\rho}{\sqrt{1-\rho^2}}$ . So  $f'(\rho) = \frac{\sqrt{1-\rho^2} + \frac{\rho^2}{\sqrt{1-\rho^2}}}{1-\rho^2} = \frac{1}{(1-\rho^2)^{3/2}}$  and:

$$g_C(X, X) = \frac{x^2}{(1-\rho^2)^3}.$$

#### REFERENCES

- [1] I. ARCHAKOV AND P. R. HANSEN, *A New Parametrization of Correlation Matrices*, *Econometrica*, 89 (2021), pp. 1699–1715.
- [2] V. ARSIGNY, P. FILLARD, X. PENNEC, AND N. AYACHE, *Log-Euclidean metrics for fast and simple calculus on diffusion tensors*, *Magnetic Resonance in Medicine*, 56 (2006), pp. 411–421.
- [3] A. BARACHANT, S. BONNET, M. CONGEDO, AND C. JUTTEN, *Classification of covariance matrices using a Riemannian-based kernel for BCI applications*, *Neurocomputing*, 112 (2013), pp. 172–178.

- [4] A. L. BESSE, *Einstein Manifolds*, Ergebnisse der Mathematik und ihrer Grenzgebiete, Springer, Berlin, Heidelberg, 1987.
- [5] P. BUSER AND H. KARCHER, *Gromov's almost flat manifolds*, Société Mathématique de France, 1981.
- [6] P. DAVID, *A Riemannian Quotient Structure for Correlation Matrices with Applications to Data Science*, PhD thesis, Institute of Mathematical Sciences, Claremont Graduate University, 2019.
- [7] P. DAVID AND W. GU, *A Riemannian structure for correlation matrices*, Operators and Matrices, 13 (2019), pp. 607–627.
- [8] J. GALLIER, *Logarithms and Square Roots of Real Matrices*, 2008. ArXiv e-prints.
- [9] M. K. GARBA, T. M. W. NYE, J. LUEG, AND S. F. HUCKEMANN, *Information geometry for phylogenetic trees*, Journal of Mathematical Biology, 82 (2021), p. 19.
- [10] N. GUIGUI AND X. PENNEC, *A Reduced Parallel Transport Equation on Lie Groups with a Left-Invariant Metric*, in Geometric Science of Information, F. Nielsen and F. Barbaresco, eds., Lecture Notes in Computer Science, Cham, 2021, Springer International Publishing, pp. 119–126.
- [11] A. N. KERCHEVAL, *On Rebonato and Jäckel's parametrization method for finding nearest correlation matrices*, International Journal of Pure and Applied Mathematics, 45 (2008), pp. 383–390.
- [12] A. LE BRIGANT, S. C. PRESTON, AND S. PUECHMOREL, *Fisher-Rao geometry of Dirichlet distributions*, Differential Geometry and its Applications, 74 (2021), p. 101702.
- [13] J. M. LEE, *Introduction to Smooth Manifolds*, Springer, New York, NY, 2nd ed., 2012.
- [14] P. LI, Q. WANG, H. ZENG, AND L. ZHANG, *Local log-euclidean multivariate gaussian descriptor and its application to image classification*, IEEE Transactions on Pattern Analysis and Machine Intelligence, 39 (2017), pp. 803–817.
- [15] Z. LIN, *Riemannian Geometry of Symmetric Positive Definite Matrices via Cholesky Decomposition*, SIAM Journal on Matrix Analysis and Applications, 40 (2019), pp. 1353–1370.
- [16] G. MARTI, V. GOUBET, AND F. NIELSEN, *cCorrGAN: Conditional Correlation GAN for Learning Empirical Conditional Distributions in the Elliptope*, in Geometric Science of Information, F. Nielsen and F. Barbaresco, eds., Lecture Notes in Computer Science, Cham, 2021, Springer International Publishing, pp. 613–620.
- [17] F. NIELSEN AND K. SUN, *Clustering in Hilbert's Projective Geometry: The Case Studies of the Probability Simplex and the Elliptope of Correlation Matrices*, in Geometric Structures of Information, F. Nielsen, ed., Springer International Publishing, Cham, 2019, pp. 297–331.
- [18] B. O'NEILL, *The fundamental equations of a submersion.*, Michigan Mathematical Journal, 13 (1966), pp. 459–469.
- [19] J. C. PINHEIRO AND D. M. BATES, *Unconstrained parametrizations for variance-covariance matrices*, Statistics and Computing, 6 (1996), pp. 289–296.
- [20] R. REBONATO AND P. JAECKEL, *The most general methodology to create a valid correlation matrix for risk management and option pricing purposes*, Journal of Risk, 2 (2001), pp. 17–27.
- [21] L. T. SKOVGAARD, *A Riemannian Geometry of the Multivariate Normal Model*, Scandinavian Journal of Statistics, 11 (1984), pp. 211–223.
- [22] Y. THANWERDAS AND X. PENNEC, *Geodesics and Curvature of the Quotient-Affine Metrics on Full-Rank Correlation Matrices*, in Proceedings of GSI 2021 - 5th conference on Geometric Science of Information, vol. 12829 of Lecture Notes in Computer Science, Paris, France, July 2021, Springer International Publishing, pp. 93–102.
- [23] Y. THANWERDAS AND X. PENNEC, *O(n)-invariant Riemannian metrics on SPD matrices*, Sept. 2021. ArXiv e-prints.
- [24] G. VAROQUAUX, F. BARONNET, A. KLEINSCHMIDT, P. FILLARD, AND B. THIRION, *Detection of brain functional-connectivity difference in post-stroke patients using group-level covariance modeling*, in Medical Image Computing and Computer Added Intervention, D. Shen, A. Frangi, and G. Szekely, eds., Lecture Notes in Computer Science, Beijing, China, Sept. 2010, Tianzi Jiang, Springer, pp. 200–208.
- [25] Z. WANG, B. C. VEMURI, Y. CHEN, AND T. H. MARECI, *A constrained variational principle for direct estimation and smoothing of the diffusion tensor field from complex dwi*, IEEE Transactions on Medical Imaging, 23 (2004), pp. 930–939.