

Exploration of Balanced Metrics on Symmetric Positive Definite Matrices

Yann Thanwerdas, Xavier Pennec

▶ **To cite this version:**

Yann Thanwerdas, Xavier Pennec. Exploration of Balanced Metrics on Symmetric Positive Definite Matrices. GSI 2019 - 4th conference on Geometric Science of Information, Aug 2019, Toulouse, France. pp.484–493, 10.1007/978-3-030-26980-7_50 . hal-02158525

HAL Id: hal-02158525
https://hal.science/hal-02158525

Submitted on 9 Sep 2019

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.

Exploration of Balanced Metrics on Symmetric Positive Definite Matrices

Yann Thanwerdas and Xavier Pennec

Université Côte d'Azur, Inria, Epione, France

Abstract. Symmetric Positive Definite (SPD) matrices have been used in many fields of medical data analysis. Many Riemannian metrics have been defined on this manifold but the choice of the Riemannian structure lacks a set of principles that could lead one to choose properly the metric. This drives us to introduce the principle of balanced metrics that relate the affine-invariant metric with the Euclidean and inverse-Euclidean metric, or the Bogoliubov-Kubo-Mori metric with the Euclidean and log-Euclidean metrics. We introduce two new families of balanced metrics, the mixed-power-Euclidean and the mixed-power-affine metrics and we discuss the relation between this new principle of balanced metrics and the concept of dual connections in information geometry.

1 Introduction

Symmetric Positive Definite (SPD) matrices are used in many applications: for example, they represent covariance matrices in signal or image processing [1–3] and they are diffusion tensors in diffusion tensor imaging [4–6]. Many Riemannian structures have been introduced on the manifold of SPD matrices depending on the problem and showing significantly different results from one another on statistical procedures such as the computation of barycenters or the principal component analysis. Non exhaustively, we can cite Euclidean metrics, power-Euclidean metrics [7], log-Euclidean metrics [8], which are flat; affine-invariant metrics [5, 6, 9] which are negatively curved; the Bogoliubov-Kubo-Mori metric [10] whose curvature has a quite complex expression.

Are there some relations between them? This question has practical interests. First, understanding the links between these metrics could lead to interesting formulas and allow to perform more efficient algorithms. Second, finding families of metrics that comprise these isolated metrics could allow to perform optimization on the parameters of these families to find a better adapted metric. Some relations already exist. For example, the power-Euclidean metrics [7] (resp. power-affine metrics [11]) comprise the Euclidean metric (resp. affine-invariant metric) and tend to the log-Euclidean metric when the power tends to zero.

We propose the principle of *balanced metrics* after observing two facts. The affine-invariant metric $g^A_\Sigma(X,Y) = \mathrm{tr}((\Sigma^{-1}X\Sigma^{-1})Y)$ on SPD matrices may be seen as a balanced hybridization of the Euclidean metric $g^E_\Sigma(X,Y) = \mathrm{tr}(XY)$ on one vector and of the inverse-Euclidean metric (the Euclidean metric on precision matrices) $g^I_\Sigma(X,Y) = \mathrm{tr}((\Sigma^{-1}X\Sigma^{-1})(\Sigma^{-1}Y\Sigma^{-1}))$ on the other vector.

Moreover, the definition of the Bogoliubov-Kubo-Mori metric can be rewritten as $g^{BKM}_\Sigma(X,Y) = \mathrm{tr}(\partial_X \log(\Sigma)\, Y)$ where it appears as a balance of the Euclidean metric and the log-Euclidean metric $g^{LE}_\Sigma(X,Y) = \mathrm{tr}(\partial_X \log(\Sigma)\, \partial_Y \log(\Sigma))$. These observations raise a few questions. Given two metrics, is it possible to define a balanced bilinear form in general? If yes, is it clear that this bilinear form is symmetric and positive definite? If it is a metric, are the Levi-Civita connections of the two initial metrics dual in the sense of information geometry?

In this work, we explore this principle through the affine-invariant metric, the Bogoliubov-Kubo-Mori metric and we define two new families of balanced metrics, the mixed-power-Euclidean and the mixed-power-affine metrics. In section 2, we show that if a balanced metric comes from two flat metrics, the three of them define a dually flat structure. In particular, we show that the balanced structure defined by the Euclidean and the inverse-Euclidean metrics corresponds to the dually flat structure given by the ± 1-connections of Fisher information geometry. In section 3, we enlighten the balanced structure of the BKM metric and we generalize it by defining the family of mixed-power-Euclidean metrics. In section 4, we define the family of mixed-power-affine metrics and we discuss the relation between the concepts of balanced metric and dual connections when the two initial metrics are not flat.

2 Affine-invariant metric as a balance of Euclidean and inverse-Euclidean metrics

Because the vocabulary may vary from one community to another, we shall first introduce properly the main geometric tools that we use in the article (Section 2.1). Then we examine in Section 2.2 the principle of balanced metric in the particular case of the pair Euclidean / inverse-Euclidean metrics and we formalize it in the general case of two flat metrics. In Section 2.3, we show that the ± 1-connections of the centered multivariate normal model are exactly the Levi-Civita connections of the Euclidean and inverse-Euclidean metrics.

2.1 Reminder on metrics, connections and parallel transport

On a manifold $\mathcal{M}$, we denote $\mathcal{C}^\infty(\mathcal{M})$ the ring of smooth real functions and $\mathfrak{X}(\mathcal{M})$ the $\mathcal{C}^\infty(\mathcal{M})$-module of vector fields.

Connection A *connection* is an $\mathbb{R}$-bilinear map $\nabla : \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \longrightarrow \mathfrak{X}(\mathcal{M})$ that is $\mathcal{C}^\infty(\mathcal{M})$-linear in the first variable and satisfies the Leibniz rule in the second variable. It gives notions of parallelism, parallel transport and geodesics. A vector field V is *parallel* to the curve γ if $\nabla_{\dot\gamma} V = 0$. The *parallel transport* of a vector v along a curve γ is the unique vector field $V_{\gamma(t)} = \Pi_\gamma^{0\to t} v$ that extends v and that is parallel to γ. Thus, the connection is an infinitesimal parallel transport, that is $\Pi_\gamma^{t\to 0} V_{\gamma(t)} = V_{\gamma(0)} + t\nabla_{\dot\gamma} V + o(t)$. The *geodesics* are autoparallel curves, that is curves γ satisfying $\nabla_{\dot\gamma}\dot\gamma = 0$.

Levi-Civita connection Given a metric g on a manifold $\mathcal{M}$, the *Levi-Civita connection* is the unique torsion-free connection ∇^g compatible with the metric g, that is $\nabla^g g = 0$ or more explicitly $X(g(Y,Z)) = g(\nabla^g_X Y, Z) + g(Y, \nabla^g_X Z)$ for all vector fields $X, Y, Z \in \mathfrak{X}(\mathcal{M})$. Thus a metric inherits notions of parallel transport and geodesics. Note that geodesics coincide with distance-minimizing curves with constant speed.

Dual connections Given a metric g and a connection ∇, the *dual connection* of ∇ with respect to g is the unique connection ∇^* satisfying the following equality $X(g(Y,Z)) = g(\nabla_X Y, Z) + g(Y, \nabla^*_X Z)$ for all vector fields $X, Y, Z \in \mathfrak{X}(\mathcal{M})$. It is characterized by Lemma 1 below. In this sense, the Levi-Civita connection ∇^g is the unique torsion-free self-dual connection with respect to g. We say that $(\mathcal{M}, g, \nabla, \nabla^*)$ is a *dually-flat manifold* when ∇, ∇^* are dual with respect to g and ∇ is flat (then ∇^* is automatically flat [12]).

Lemma 1 (Characterization of dual connections). *Two connections ∇, ∇' with parallel transports Π, Π' are dual with respect to a metric g if and only if the dual parallel transport preserves the metric, i.e. for all vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$ and all curve γ, $g_{\gamma(t)}(X_{\gamma(t)}, Y_{\gamma(t)}) = g_{\gamma(0)}(\Pi_\gamma^{t\to 0} X_{\gamma(t)}, (\Pi')_\gamma^{t\to 0} Y_{\gamma(t)})$.*

Proof. The direct sense is proved in [12]. Let us assume that the dual parallel transport preserves the metric and let $X, Y, Z \in \mathfrak{X}(\mathcal{M})$ be vector fields. Let $x \in \mathcal{M}$ and let γ be a curve such that $\gamma(0) = x$ and $\dot\gamma(0) = X_x$. Using the first order approximation of the parallel transport, our assumption leads to:

$$\begin{aligned} g_{\gamma(t)}(Y_{\gamma(t)}, Z_{\gamma(t)}) &= g_x(\Pi_\gamma^{t\to 0} Y_{\gamma(t)}, (\Pi')_\gamma^{t\to 0} Z_{\gamma(t)}) \\ &= g_x(Y_x + t\nabla_{\dot\gamma} Y + o(t), Z_x + t\nabla'_{\dot\gamma} Z + o(t)) \\ &= g_x(Y_x, Z_x) + t[g_x(\nabla_{\dot\gamma} Y, Z) + g_x(Y, \nabla'_{\dot\gamma} Z)] + o(t). \end{aligned}$$

So $X_x(g(Y,Z)) = g_x(\nabla_{\dot\gamma} Y, Z) + g_x(Y, \nabla'_{\dot\gamma} Z)$ and ∇, ∇' are dual w.r.t. g. □

2.2 Principle of balanced metrics

Observation We denote $\mathcal{M} = SPD_n$ the manifold of SPD matrices and $N = \dim \mathcal{M} = \frac{n(n+1)}{2}$. The (A)ffine-invariant metric g^A on SPD matrices [5, 6, 9], i.e. satisfying $g^A_{M\Sigma M^\top}(MXM^\top, MYM^\top) = g^A_\Sigma(X,Y)$ for $M \in GL_n$, is defined by:

$$g^A_\Sigma(X,Y) = \mathrm{tr}((\Sigma^{-1}X\Sigma^{-1})Y) = \mathrm{tr}(X(\Sigma^{-1}Y\Sigma^{-1})). \tag{1}$$

The (E)uclidean metric g^E on SPD matrices is the pullback metric by the embedding $id : \mathcal{M} \hookrightarrow (\mathrm{Sym}_n, \langle\cdot|\cdot\rangle_{\mathrm{Frob}})$:

$$g^E_\Sigma(X,Y) = \mathrm{tr}(XY). \tag{2}$$

The (I)nverse-Euclidean metric g^I on SPD matrices belongs to the family of power-Euclidean metrics [7] with power -1. If SPD matrices are seen as covariance matrices Σ, the inverse-Euclidean metric is the Euclidean metric on

precision matrices Σ^{-1}:

$$g^I_\Sigma(X,Y) = \mathrm{tr}(\Sigma^{-2}X\Sigma^{-2}Y) = \mathrm{tr}((\Sigma^{-1}X\Sigma^{-1})(\Sigma^{-1}Y\Sigma^{-1})). \tag{3}$$

Observing these definitions, the affine-invariant metric (1) appears as a *balance* of the Euclidean metric (2) and the inverse-Euclidean metric (3). We formalize this idea thanks to parallel transport.

Formalization The diffeomorphism $\mathrm{inv} : (\mathcal{M}, g^I) \longrightarrow (\mathcal{M}, g^E)$ is an isometry. Since these two metrics are flat, the parallel transports do not depend on the curve. On the one hand, the Euclidean parallel transport from Σ to I_n is the identity map $\Pi^E : X \in T_\Sigma\mathcal{M} \longmapsto X \in T_{I_n}\mathcal{M}$ since all tangent spaces are identified to the vector space of symmetric matrices Sym_n by the differential of the embedding $id : \mathcal{M} \hookrightarrow \mathrm{Sym}_n$. On the other hand, the isometry inv gives the inverse-Euclidean parallel transport from Σ to I_n, $\Pi^I : X \in T_\Sigma\mathcal{M} \longmapsto \Sigma^{-1}X\Sigma^{-1} \in T_{I_n}\mathcal{M}$. We generalize this situation in Definition 1. Given Lemma 1, it automatically leads to Theorem 1.

Definition 1 (Balanced bilinear form). *Let g, g^* be two flat metrics on SPD_n and Π, Π^* the associated parallel transports that do not depend on the curve. We define the balanced bilinear form $g^0_\Sigma(X,Y) = \mathrm{tr}((\Pi_{\Sigma\to I_n}X)(\Pi^*_{\Sigma\to I_n}Y))$.*

Theorem 1 (A balanced metric defines a dually flat manifold). *Let g, g^* be two flat metrics and let ∇, ∇^* be their Levi-Civita connections. If the balanced bilinear form g^0 of g, g^* is a metric, then $(\mathcal{M}, g^0, \nabla, \nabla^*)$ is a dually flat manifold.*

If two connections ∇ and ∇^* are dual connections with respect to a metric g^0, there is no reason for them to be Levi-Civita connections of some metrics. Therefore, the main advantage of the principle of balanced metrics on the concept of dual connections seems to be the metric nature of the dual connections.

Corollary 1 (Euclidean and inverse-Euclidean are dual with respect to affine-invariant). *We denote ∇^E and ∇^I the Levi-Civita connections of the Euclidean metric g^E and the inverse-Euclidean metric g^I. Then g^A is the balanced metric of g^I, g^E and $(SPD_n, g^A, \nabla^I, \nabla^E)$ is a dually flat manifold.*

2.3 Relation with Fisher information geometry

We know from [13] that the affine-invariant metric is the Fisher metric of the centered multivariate normal model. Information geometry provides a natural one-parameter family of dual connections, called α-connections [12]. In the following table, we recall the main quantities characterizing the *centered multivariate normal model* $\mathcal{P} = \{p_\Sigma : \mathbb{R}^n \longrightarrow \mathbb{R}^*_+, \Sigma \in \mathcal{M}\}$, where $\mathcal{M} = SPD_n$.

Densities	$p_\Sigma(x) = \frac{1}{\sqrt{2\pi}^n}\frac{1}{\sqrt{\det\Sigma}}\exp\left(\frac{1}{2}x^\top\Sigma^{-1}x\right)$
Log likelihood	$l_\Sigma(x) = \log p_\Sigma(x) = \frac{1}{2}\left(-n\log(2\pi) - \log\det\Sigma + x^\top\Sigma^{-1}x\right)$
Differential	$d_\Sigma l(V)(x) = -\frac{1}{2}\left[\mathrm{tr}(\Sigma^{-1}V) + x^\top\Sigma^{-1}V\Sigma^{-1}x\right]$
Fisher metric	$g_\Sigma(V,W) = \frac{1}{2}\mathrm{tr}(\Sigma^{-1}V\Sigma^{-1}W)$

We recall that the *α-connections* $\nabla^{(\alpha)}$ [12] of a family of densities $\mathcal{P}$ are defined by their Christoffel symbols $\Gamma_{ijk} = g_{lk}\Gamma^l_{ij}$ in the local basis $(\partial_i)_{1\leqslant i\leqslant N}$ at $\Sigma \in \mathcal{M}$:

$$\left(\Gamma^{(\alpha)}_{ijk}\right)_\Sigma = \mathbb{E}_\Sigma\left[\left(\partial_i\partial_j l + \frac{1-\alpha}{2}\partial_i l\partial_j l\right)\partial_k l\right]. \tag{4}$$

We give in Theorem 2 the expression of the α-connections of the centered multivariate normal model and we notice that the Euclidean and inverse-Euclidean Levi-Civita connections belong to this family.

Theorem 2 (α-connections of the centered multivariate normal model). *In the global basis of $\mathcal{M} = SPD_n$ given by the inclusion $\mathcal{M} \hookrightarrow \mathrm{Sym}(n) \simeq \mathbb{R}^N$, writing $\partial_X Y = X^i(\partial_i Y^j)\partial_j$, the α-connections of the multivariate centered normal model are given by the following formula:*

$$\nabla^{(\alpha)}_X Y = \partial_X Y - \frac{1+\alpha}{2}(X\Sigma^{-1}Y + Y\Sigma^{-1}X). \tag{5}$$

The mixture m-connection ($\alpha = -1$) is the Levi-Civita connection of the Euclidean metric $g^E_\Sigma(X,Y) = \mathrm{tr}(XY)$. The exponential e-connection ($\alpha = 1$) is the Levi-Civita connection of the inverse-Euclidean metric, i.e. the pullback of the Euclidean metric by matrix inversion, $g^I_\Sigma(X,Y) = \mathrm{tr}(\Sigma^{-2}X\Sigma^{-2}Y)$.

The formula (5) can be proved thanks to Lemma 2 which gives the results of expressions of type $\int_{\mathbb{R}^n} x^\top\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z\Sigma^{-1}x\exp\left(-\frac{1}{2}x^\top\Sigma^{-1}x\right)dx$, with $y = \Sigma^{-1/2}x$, $A = \Sigma^{-1/2}X\Sigma^{-1/2}$, $B = \Sigma^{-1/2}Y\Sigma^{-1/2}$ and $C = \Sigma^{-1/2}Z\Sigma^{-1/2}$. If one wants to avoid using the third formula of Lemma 2, one can rely on the formula (5) in the case $\alpha = 0$ which is already known from [13].

Lemma 2. *For $A, B, C \in \mathrm{Sym}_n$:*

$$\begin{aligned}
\mathbb{E}_{I_n}(y \longmapsto y^\top Ay) &= \mathrm{tr}(A),\\
\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top By) &= \mathrm{tr}(A)\mathrm{tr}(B) + 2\mathrm{tr}(AB),\\
\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top Byy^\top Cy) &= \mathrm{tr}(A)\mathrm{tr}(B)\mathrm{tr}(C) + 8\mathrm{tr}(ABC)\\
&\quad + 2(\mathrm{tr}(AB)\mathrm{tr}(C) + \mathrm{tr}(BC)\mathrm{tr}(A) + \mathrm{tr}(CA)\mathrm{tr}(B)).
\end{aligned}$$

Proof (Theorem 2). Given Lemma 2, the computation of the Christoffel symbols $\left(\Gamma^{(\alpha)}_{ijk}\right)_\Sigma$ leads to $\left(\Gamma^{(\alpha)}_{ijk}\right)_\Sigma X^iY^jZ^k = -\frac{1+\alpha}{4}\mathrm{tr}(\Sigma^{-1}[X\Sigma^{-1}Y + Y\Sigma^{-1}X]\Sigma^{-1}Z)$. On the other hand, the relation $\Gamma_{ijk} = g_{lk}\Gamma^l_{ij}$ between Christoffel symbols gives $\left(\Gamma^{(\alpha)}_{ijk}\right)_\Sigma X^iY^jZ^k = \frac{1}{2}\mathrm{tr}\left(\Sigma^{-1}\left[\left(\Gamma^l_{ij}\right)^{(\alpha)}_\Sigma X^iY^j\partial_l\right]\Sigma^{-1}Z\right)$. So we get:

$$\nabla^{(\alpha)}_X Y = \partial_X Y + \left[\left(\Gamma^l_{ij}\right)^{(\alpha)}_\Sigma X^iY^j\partial_l\right] = \partial_X Y - \frac{1+\alpha}{2}(X\Sigma^{-1}Y + Y\Sigma^{-1}X). \tag{6}$$

It is clear that the mixture connection ($\alpha = -1$) is the Euclidean connection. The inverse-Euclidean connection can be computed thanks to the Koszul formula. This calculus drives exactly to the exponential connection ($\alpha = 1$). □

In the next section, we apply the principle of balanced metrics to the pairs Euclidean / log-Euclidean (Bogoliubov-Kubo-Mori metric) and power-Euclidean / power-Euclidean (mixed-power-Euclidean metrics).

3 The family of mixed-power-Euclidean metrics

3.1 Bogoliubov-Kubo-Mori metric

The Bogoliubov-Kubo-Mori metric g^{BKM} is a metric on symmetric positive definite matrices used in quantum physics. It was introduced as $g^{BKM}_\Sigma(X,Y) = \int_0^\infty \mathrm{tr}((\Sigma+tI_n)^{-1}X(\Sigma+tI_n)^{-1}Y)dt$ and can be rewritten [10] thanks to the differential of the symmetric matrix logarithm $\log : \mathcal{M} = SPD_n \longrightarrow \mathrm{Sym}_n$ as:

$$g^{BKM}_\Sigma(X,Y) = \mathrm{tr}(\partial_X \log(\Sigma)\, Y) = \mathrm{tr}(X\, \partial_Y \log(\Sigma)). \tag{7}$$

The log-Euclidean metric g^{LE} [8] is the pullback metric of the Euclidean metric by the symmetric matrix logarithm $\log : (\mathcal{M}, g^{LE}) \longrightarrow (\mathrm{Sym}_n, g^E)$:

$$g^{LE}_\Sigma(X,Y) = \mathrm{tr}(\partial_X \log(\Sigma)\, \partial_Y \log(\Sigma)). \tag{8}$$

Therefore, the BKM metric (7) appears as the balanced metric of the Euclidean metric (2) and the log-Euclidean metric (8). As the Euclidean and log-Euclidean metrics are flat, the parallel transport does not depend on the curve and Theorem 1 ensures that they form a dually flat manifold.

Corollary 2 (Euclidean and log-Euclidean are dual with respect to BKM). *We denote ∇^E and ∇^{LE} the Levi-Civita connections of the Euclidean metric g^E and the log-Euclidean metric g^{LE}. Then g^{BKM} is the balanced metric of g^{LE}, g^E and $(SPD_n, g^{BKM}, \nabla^{LE}, \nabla^E)$ is a dually flat manifold.*

3.2 Mixed-power-Euclidean

Up to now, we observed that existing metrics (affine-invariant and BKM) were the balanced metrics of pairs of flat metrics (Euclidean / inverse-Euclidean and Euclidean / log-Euclidean). Thus, the symmetry and the positivity of the balanced bilinear forms were obvious. From now on, we build new bilinear forms thanks to the principle of balanced metrics. Therefore, it is not as obvious as before that these bilinear forms are metrics.

The family of power-Euclidean metrics $g^{E,\theta}$ [7] indexed by the power $\theta \neq 0$ is defined by pullback of the Euclidean metric by the power function $\mathrm{pow}_\theta = \exp \circ\, \theta \log : (\mathcal{M}, \theta^2 g^{E,\theta}) \longrightarrow (\mathcal{M}, g^E)$:

$$g^{E,\theta}_\Sigma(X,Y) = \frac{1}{\theta^2}\mathrm{tr}(\partial_X \mathrm{pow}_\theta(\Sigma)\, \partial_Y \mathrm{pow}_\theta(\Sigma)). \tag{9}$$

This family comprise the Euclidean metric for $\theta = 1$ and tends to the log-Euclidean metric when the power θ goes to 0. Therefore, we abusively consider that the log-Euclidean metric belongs to the family and we denote it $g^{E,0} := g^{LE}$.

We define the mixed-power-Euclidean metrics g^{E,θ_1,θ_2} as the balanced bilinear form of the power-Euclidean metrics g^{E,θ_1} and g^{E,θ_2}, where $\theta_1, \theta_2 \in \mathbb{R}$:

$$g^{E,\theta_1,\theta_2}_\Sigma(X,Y) = \frac{1}{\theta_1\theta_2}\mathrm{tr}(\partial_X \mathrm{pow}_{\theta_1}(\Sigma)\, \partial_Y \mathrm{pow}_{\theta_2}(\Sigma)). \tag{10}$$

Note that the family of mixed-power-Euclidean metrics contains the BKM metric for $(\theta_1, \theta_2) = (1, 0)$ and the θ-power-Euclidean metric for $(\theta_1, \theta_2) = (\theta, \theta)$, including the Euclidean metric for $\theta = 1$ and the log-Euclidean metric for $\theta = 0$.

At this stage, we do not know that the bilinear form g^{E,θ_1,θ_2} is a metric. This is stated by Theorem 3. As the power-Euclidean metrics are flat, Theorem 1 combined with Theorem 3 ensure that the Levi-Civita connections ∇^{E,θ_1} and ∇^{E,θ_2} of the metrics g^{E,θ_1} and g^{E,θ_2} are dual with respect to the (θ_1, θ_2)-mixed-power-Euclidean metric. This is stated by Corollary 3.

Theorem 3. *The bilinear form g^{E,θ_1,θ_2} is symmetric and positive definite so it is a metric on SPD_n. Moreover, the symmetry ensures that $g^{E,\theta_1,\theta_2} = g^{E,\theta_2,\theta_1}$.*

Corollary 3 (θ_1 and θ_2-power-Euclidean are dual with respect to (θ_1, θ_2)-mixed-power-Euclidean). *For $\theta_1, \theta_2 \in \mathbb{R}$, we denote ∇^{E,θ_1} and ∇^{E,θ_2} the Levi-Civita connections of the power-Euclidean metrics g^{E,θ_1} and g^{E,θ_2}. Then g^{E,θ_1,θ_2} is the balanced metric of $g^{E,\theta_1}, g^{E,\theta_2}$ and $(SPD_n, g^{E,\theta_1,\theta_2}, \nabla^{E,\theta_1}, \nabla^{E,\theta_2})$ is a dually flat manifold.*

To prove Theorem 3, we show that for all spectral decomposition $\Sigma = PDP^\top$ of an SPD matrix, there exists a matrix A with positive coefficients $A(i, j) > 0$ such that $g^{E,\theta_1,\theta_2}_\Sigma(X, Y) = \mathrm{tr}((A \bullet P^\top XP)(A \bullet P^\top YP))$, where $\bullet$ is the Hadamard product, i.e. $(A \bullet B)(i, j) = A(i, j)B(i, j)$, which is associative, commutative, distributive w.r.t. matrix addition and satisfies $\mathrm{tr}((A \bullet B)C) = \mathrm{tr}(B(A \bullet C))$ for symmetric matrices $A, B, C \in \mathrm{Sym}_n$. The existence of A relies on Lemma 3.

Lemma 3. *Let $\Sigma = PDP^\top$ be a spectral decomposition of $\Sigma \in \mathcal{M}$, with $P \in O(n)$ and D diagonal. For $f \in \{\exp, \log, \mathrm{pow}_\theta\}$, $\partial_V f(\Sigma) = P(\delta(f, D) \bullet P^\top VP)P^\top$ where $\delta(f, D)(i, j) = \frac{f(d_i)-f(d_j)}{d_i-d_j}$. Note that $\frac{1}{\theta}\delta(\mathrm{pow}_\theta, D)(i, j) > 0$ for all $\theta \in \mathbb{R}^*$ and $\delta(\log, D)(i, j) > 0$.*

Proof (Lemma 3). Once shown for $f = \exp$, it is easy to get for $f = \log$ by inversion and for $f = \mathrm{pow}_\theta = \exp \circ\, \theta \log$ by composition. But the case $f = \exp$ itself reduces to the case $f = \mathrm{pow}_k$ with $k \in \mathbb{N}$ by linearity, so we focus on this last case. As $\partial_V \mathrm{pow}_k(\Sigma) = \sum_{l=0}^{k-1} \Sigma^l V \Sigma^{k-1-l} = P\partial_{P^\top VP}\mathrm{pow}_k(D)P^\top$ and $\partial_{P^\top VP}\mathrm{pow}_k(D)(i, j) = \sum_{l=0}^{k-1} D^l P^\top VPD^{k-1-l}(i, j) = \frac{d_i^k - d_j^k}{d_i - d_j} P^\top VP(i, j)$, we get $\partial_V \mathrm{pow}_k(\Sigma) = P(\delta(\mathrm{pow}_k, D) \bullet P^\top VP)P^\top$.

Proof (Theorem 3). Let $\theta_1, \theta_2 \in \mathbb{R}^*$. For a spectral decomposition $\Sigma = PDP^\top$, the matrix A defined by $A(i, j) = \sqrt{\frac{1}{\theta_1}\delta(\mathrm{pow}_{\theta_1}, D)(i, j)\frac{1}{\theta_2}\delta(\mathrm{pow}_{\theta_2}, D)(i, j)} > 0$ satisfies $g^{E,\theta_1,\theta_2}_\Sigma(X, Y) = \mathrm{tr}((A \bullet P^\top XP)(A \bullet P^\top YP))$. Symmetry and non-negativity are clear since they come from the Frobenius scalar product. Finally, if $g^{E,\theta_1,\theta_2}_\Sigma(X, X) = 0$, then $A \bullet P^\top XP = 0$ so $X = 0$. So g^{E,θ_1,θ_2} is a metric. If $\theta_1 = 0$, the matrix A defined by $A(i, j) = \sqrt{\delta(\log, D)(i, j)\frac{1}{\theta_2}\delta(\mathrm{pow}_{\theta_2}, D)(i, j)} > 0$ satisfies the same property and $g^{E,0,\theta_2}$ is a metric. □

4 The family of mixed-power-affine metrics

In previous sections, we defined our balanced metric from a pair of two flat metrics and we showed that it corresponded to the duality of (Levi-Civita) connections in information geometry. In this section, we investigate the balanced metric of two non-flat metrics and we observe that the corresponding Levi-Civita connections cannot be dual with respect to this balanced metric.

The family of power-affine metrics $g^{A,\theta}$ [11] indexed by the power $\theta \neq 0$ are defined by pullback of the affine-invariant metric by the power function $\mathrm{pow}_\theta : (\mathcal{M}, \theta^2 g^{A,\theta}) \longrightarrow (\mathcal{M}, g^A)$:

$$g^{A,\theta}_\Sigma(X,Y) = \frac{1}{\theta^2}\mathrm{tr}(\Sigma^{-\theta}\,\partial_X \mathrm{pow}_\theta(\Sigma)\,\Sigma^{-\theta}\,\partial_Y \mathrm{pow}_\theta(\Sigma)) \tag{11}$$

This family comprise the affine-invariant metric for $\theta = 1$ and tends to the log-Euclidean metric when the power θ goes to 0. We consider that the log-Euclidean metric belongs to the family and we denote $g^{A,0} := g^{LE}$.

As these metrics have no cut locus because they endow the manifold with a negatively curved Riemannian symmetric structure, there exists a unique geodesic between two given points. Therefore, a canonical parallel transport can be defined along geodesics. This allows to define the balanced bilinear form of two metrics without cut locus.

Definition 2 (Balanced bilinear form). *Let g, g^* be two metrics without cut locus on SPD_n and Π, Π^* the associated geodesic parallel transports. We define the balanced bilinear form $g^0_\Sigma(X,Y) = \mathrm{tr}((\Pi_{\Sigma\to I_n} X)(\Pi^*_{\Sigma\to I_n} Y))$.*

Given that the geodesic parallel transport on the manifold $(\mathcal{M}, g^{A,\theta})$ is $\Pi_{\Sigma\to I_n} : X \in T_\Sigma\mathcal{M} \longmapsto \frac{1}{\theta}\Sigma^{-\theta/2}\partial_X \mathrm{pow}_\theta(\Sigma)\Sigma^{-\theta/2} \in T_{I_n}\mathcal{M}$, we define the mixed-power-affine metrics g^{A,θ_1,θ_2} as the balanced metric of the power-affine metrics g^{A,θ_1} and g^{A,θ_2}, where $\theta_1, \theta_2 \in \mathbb{R}$ and $\theta = (\theta_1+\theta_2)/2$:

$$g^{A,\theta_1,\theta_2}_\Sigma(X,Y) = \frac{1}{\theta_1\theta_2}\mathrm{tr}(\Sigma^{-\theta}\,\partial_X \mathrm{pow}_{\theta_1}(\Sigma)\,\Sigma^{-\theta}\,\partial_Y \mathrm{pow}_{\theta_2}(\Sigma)). \tag{12}$$

Note that the family of mixed-power-affine metrics contains the θ-power-affine metric for $(\theta_1,\theta_2) = (\theta,\theta)$, including the affine-invariant metric for $\theta = 1$ and the log-Euclidean metric for $\theta = 0$. This family has two symmetries since $g^{A,\theta_1,\theta_2} = g^{A,\pm\theta_1,\pm\theta_2}$, they come from the inverse-consistency of the affine-invariant metric. This family has a non-empty intersection with the family of mixed-power-Euclidean metrics since $g^{A,\theta_1,-\theta_1} = g^{E,\theta_1,-\theta_1} = g^{A,\theta_1}$ for all $\theta_1 \in \mathbb{R}$.

The fact that g^{A,θ_1,θ_2} is a metric can be shown exactly the same way as in the mixed-power-Euclidean case thanks to the equality $\Sigma^{-\theta/2}\partial_V \mathrm{pow}_\theta(\Sigma)\Sigma^{-\theta/2} = P(\varepsilon(\mathrm{pow}_\theta, D)\bullet P^\top V P)P^\top$ where $\varepsilon(\mathrm{pow}_\theta, D) = (d_i d_j)^{-\theta/2}\delta(\mathrm{pow}_\theta, D)$ and where $\delta(\mathrm{pow}_\theta, D)$ has been defined in Lemma 3. This is stated in Theorem 4.

Theorem 4. *The bilinear form g^{A,θ_1,θ_2} is symmetric and positive definite. Hence it is a metric on SPD_n and $g^{A,\theta_1,\theta_2} = g^{A,\theta_2,\theta_1}$.*

Power-affine metrics being non-flat, $(\mathcal{M}, g^{A,\theta_1,\theta_2}, \nabla^{A,\theta_1}, \nabla^{A,\theta_2})$, where ∇^{A,θ_1} and ∇^{A,θ_2} are Levi-Civita connections of g^{A,θ_1} and g^{A,θ_2}, cannot be a dually-flat manifold. Actually, the two connections are even not dual. It can be understood by comparison with previous sections since the duality was a consequence of the independence of the parallel transport with respect to the chosen curve, which was a consequence of the flatness of the two connections. Moreover, in the Definition 2, the vectors are parallel transported along two different curves (the geodesics relative to each connection) so it may exists a better definition for the balanced bilinear form of two metrics without cut locus or even of two general metrics.

5 Conclusion

The principle of balanced bilinear form is a procedure on SPD matrices that takes a pair of flat metrics or metrics without cut locus and builds a new metric based on the parallel transport of the initial metrics. When the two initial metrics are flat, we showed that the two Levi-Civita connections are dual with respect to the balanced metric. When the two initial metrics are not flat, the two Levi-Civita connections seem not to be dual, so the principle of balanced metrics does not reduce to the concept of dual Levi-Civita connections. A challenging objective for future works is to define properly this principle for other general pairs of metrics and to find conditions under which the balanced bilinear form is a metric.

Acknowledgements. This project has received funding from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant 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.

References

1. A. Barachant, S. Bonnet, M. Congedo, and C. Jutten. Classification of covariance matrices using a Riemannian-based kernel for BCI applications. *Elsevier Neurocomputing*, 112:172–178, 2013.
2. F. Deligianni, G. Varoquaux, B. Thirion, E. Robinson, D. Sharp, D. Edwards, and D. Rueckert. A Probabilistic Framework to Infer Brain Functional Connectivity from Anatomical Connections. *IPMI Conference*, pages 296–307, 2011.
3. G. Cheng and B. Vemuri. A novel dynamic system in the space of SPD matrices with applications to appearance tracking. *SIAM Journal on Imaging Sciences*, 6(16):592–615, 2013.
4. C. Lenglet, M. Rousson, R. Deriche, and O. Faugeras. Statistics on the Manifold of Multivariate Normal Distributions: Theory and Application to Diffusion Tensor MRI Processing. *J. of Math. Imaging and Vision*, 25(3):423–444, October 2006.

5. X. Pennec, P. Fillard, and N. Ayache. A Riemannian Framework for Tensor Computing. *International Journal of Computer Vision*, 66(1):41–66, January 2006.
6. T. Fletcher and S. Joshi. Riemannian Geometry for the Statistical Analysis of Diffusion Tensor Data. *Signal Processing*, 87:250–262, 2007.
7. I. Dryden, X. Pennec, and J.-M. Peyrat. Power Euclidean metrics for covariance matrices with application to diffusion tensor imaging. 2010.
8. V. Arsigny, P. Fillard, X. Pennec, and Ayache N. Log-Euclidean Metrics for Fast and Simple Calculus on Diffusion Tensors. *Magnetic Resonance in Medicine*, 2006.
9. C. Lenglet, M. Rousson, and R. Deriche. DTI segmentation by statistical surface evolution. *IEEE Transactions on Medical Imaging*, 25:685–700, 2006.
10. P. Michor, D. Petz, and A. Andai. The Curvature of the Bogoliubov-Kubo-Mori Scalar Product on Matrices. *Infinite Dimensional Analysis, Quantum Probability and Related Topics*, 2000.
11. Y. Thanwerdas and X. Pennec. Is affine-invariance well defined on SPD matrices? A principled continuum of metrics. *Proc. Geometric Science of Information*, 2019.
12. S. Amari and H. Nagaoka. *Methods of Information Geometry*. 2000.
13. L. Skovgaard. A Riemannian geometry of the multivariate normal model. *Scand. J. of Statistics*, 11:211–223, 1984.

6 Appendix: proofs

Let us first recall the statements of Theorem 2 and Lemma 2.

Theorem 2. **α-connections of the centered multivariate normal model** *In the global basis $(\partial_i)_{1\leqslant i\leqslant N=n(n+1)/2}$ of $\mathcal{M} = SPD_n$ given by the inclusion $\mathcal{M} \hookrightarrow \mathrm{Sym}(n) \simeq \mathbb{R}^N$, writing $\partial_X Y = X^i(\partial_i Y^j)\partial_j$, the α-connections of the multivariate centered normal model are given by the following formula (5):*

$$\nabla^{(\alpha)}_X Y = \partial_X Y - \frac{1+\alpha}{2}(X\Sigma^{-1}Y + Y\Sigma^{-1}X).$$

The mixture m-connection ($\alpha = -1$) is the Levi-Civita connection of the Euclidean metric $g^E_\Sigma(X,Y) = \mathrm{tr}(XY)$. The exponential e-connection ($\alpha = 1$) is the Levi-Civita connection of the inverse-Euclidean metric, i.e. the pullback of the Euclidean metric by matrix inversion, $g^I_\Sigma(X,Y) = \mathrm{tr}(\Sigma^{-2}X\Sigma^{-2}Y)$.

The formula (5) can be proved thanks to Lemma 2 which gives the results of expressions of type $\int_{\mathbb{R}^n} x^\top\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z\Sigma^{-1}x \exp\left(-\frac{1}{2}x^\top\Sigma^{-1}x\right)dx$, with $y = \Sigma^{-1/2}x$, $A = \Sigma^{-1/2}X\Sigma^{-1/2}$, $B = \Sigma^{-1/2}Y\Sigma^{-1/2}$ and $C = \Sigma^{-1/2}Z\Sigma^{-1/2}$. If one wants to avoid using the third formula of Lemma 2, one can rely on the formula (5) in the case $\alpha = 0$ which is already known from [13].

Lemma 2. *For $A, B, C \in \mathrm{Sym}_n$:*

$$\begin{aligned}
\mathbb{E}_{I_n}(y \longmapsto y^\top Ay) &= \mathrm{tr}(A),\\
\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top By) &= \mathrm{tr}(A)\mathrm{tr}(B) + 2\mathrm{tr}(AB),\\
\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top Byy^\top Cy) &= \mathrm{tr}(A)\mathrm{tr}(B)\mathrm{tr}(C) + 8\mathrm{tr}(ABC)\\
&\quad + 2(\mathrm{tr}(AB)\mathrm{tr}(C) + \mathrm{tr}(BC)\mathrm{tr}(A) + \mathrm{tr}(CA)\mathrm{tr}(B)).
\end{aligned}$$

6.1 Proof of Lemma 2

Let $A, B, C \in \mathrm{Sym}_n$ and let $a_1 \geqslant ... \geqslant a_n$, $b_1 \geqslant ... \geqslant b_n$ and $c_1 \geqslant ... \geqslant c_n$ be their respective eigenvalues, $A = PA'P^\top$, $B = QB'Q^\top$, $C = RC'R^\top$ with $P, Q, R \in O(n)$ and $A' = \mathrm{diag}(a_1, ..., a_n)$, $B' = \mathrm{diag}(b_1, ..., b_n)$ and $C' = \mathrm{diag}(c_1, ..., c_n)$.

We recall that if $X \sim \mathcal{N}(0,1)$, then $\mathbb{E}(X^{2k+1}) = 0$, $\mathbb{E}(X^2) = 1$, $\mathbb{E}(X^4) = 3$ and $\mathbb{E}(X^6) = 15$.

First equality The first equality is well known:

$$\begin{aligned}\mathbb{E}_{I_n}(y \longmapsto y^\top Ay) &= \int_{\mathbb{R}^n} y^\top Ay \frac{1}{\sqrt{2\pi}^n} \exp(y^\top y)dy \\ &= \int_{\mathbb{R}^n} z^\top A'z \frac{1}{\sqrt{2\pi}^n} \exp(z^\top z)dz \\ &= \sum_{i=1}^n \int_{\mathbb{R}} a_i z_i^2 \frac{1}{\sqrt{2\pi}^n} \exp(z_i^2)dz_i \\ &= \sum_{i=1}^n a_i = \mathrm{tr}(A).\end{aligned}$$

Second equality Due to the non commutativity a priori of the matrices A and B, they are not diagonalizable in the same orthonormal basis so the calculus needs a bit more attention:

$$\begin{aligned}\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top By) &= \int_{\mathbb{R}^n} y^\top Ayy^\top By \frac{1}{\sqrt{2\pi}^n} \exp(y^\top y)dy \\ &= \int_{\mathbb{R}^n} z^\top A'zz^\top U^\top B'Uz \frac{1}{\sqrt{2\pi}^n} \exp(z^\top z)dz \\ &= \int_{\mathbb{R}^n} \left(\sum_{i=1}^n a_i z_i^2\right) \left(\sum_{j=1}^n b_j (Uz)_j^2\right) \frac{1}{\sqrt{2\pi}^n} \exp(z^\top z)dz,\end{aligned}$$

where $U = Q^\top P \in O(n)$. But $(Uz)_j^2 = \sum_{k=1}^n \sum_{l=1}^n U(j,k)U(j,l)z_k z_l$ so:

$$\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top By) = \sum_{i,j,k,l} a_i b_j U(j,k)U(j,l) \int_{\mathbb{R}^n} z_i^2 z_k z_l \prod_{m=1}^n \frac{1}{\sqrt{2\pi}} \exp(z_m^2)dz_m.$$

According to the values of i, k, l, the integral takes different values:

1. if $k = l = i$, it is equal to $\mathbb{E}(X^4) = 3 = 2 + 1$,
2. if $k = l \neq i$, it is equal to $\mathbb{E}(X^2)^2 = 1$,
3. elsewhere, it is equal to 0.

$$\begin{aligned}\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top By) &= \sum_{i,j} a_i b_j [(2+1)U(j,i)U(j,i) + \sum_{k\neq i} U(j,k)U(j,k)] \\ &= \sum_{i,j} a_i b_j [2U(j,i)^2 + \sum_{k=1}^{n} U(j,k)^2] \\ &= 2\mathrm{tr}(A'U^\top B'U) + \mathrm{tr}(A')\mathrm{tr}(B') \\ &= 2\mathrm{tr}(AB) + \mathrm{tr}(A)\mathrm{tr}(B).\end{aligned}$$

Third equality The third equality is a bit more tedious to obtain because of the greater number of cases and especially because the rearrangement of the terms is more complex. Let $U = Q^\top P$ and $V = R^\top P \in O(n)$. Then:

$$\begin{aligned}&\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top Byy^\top Cy) \\ &= \int_{\mathbb{R}^n} y^\top Ayy^\top Byy^\top Cy \frac{1}{\sqrt{2\pi}^n} \exp(y^\top y)dy \\ &= \int_{\mathbb{R}^n} z^\top A'zz^\top U^\top B'Uzz^\top V^\top C'Vz \frac{1}{\sqrt{2\pi}^n} \exp(z^\top z)dz \\ &= \int_{\mathbb{R}^n} \left(\sum_{i=1}^n a_i z_i^2\right)\left(\sum_{j=1}^n b_j (Uz)_j^2\right)\left(\sum_{j=1}^n c_k (Vz)_k^2\right) \frac{1}{\sqrt{2\pi}^n} \exp(z^\top z)dz \\ &= \sum_{i,j,k,l,m,r,s} a_i b_j c_k U(j,l)U(j,m)V(k,r)V(k,s) \int_{\mathbb{R}^n} z_i^2 z_l z_m z_r z_s \prod_{t=1}^n \frac{1}{\sqrt{2\pi}} \exp(z_t^2)dz_t.\end{aligned}$$

According to the values of i, l, m, r, s, the integral takes different values:

1. if $l = m = r = s = i$, it is equal to $\mathbb{E}(X^6) = 15 = 6 \times 3 - 3$,
2. if $l = r = i \neq m = s$, it is equal to $\mathbb{E}(X^4)\mathbb{E}(X^2) = 3$,
3. if $l = s = i \neq m = r$, it is equal to $\mathbb{E}(X^4)\mathbb{E}(X^2) = 3$,
4. if $m = r = i \neq l = s$, it is equal to $\mathbb{E}(X^4)\mathbb{E}(X^2) = 3$,
5. if $m = s = i \neq l = r$, it is equal to $\mathbb{E}(X^4)\mathbb{E}(X^2) = 3$,
6. if $l = m = i \neq r = s$, it is equal to $\mathbb{E}(X^4)\mathbb{E}(X^2) = 3$,
7. if $l = m \neq i = r = s$, it is equal to $\mathbb{E}(X^4)\mathbb{E}(X^2) = 3$,
8. if $l = m = r = s \neq i$, it is equal to $\mathbb{E}(X^4)\mathbb{E}(X^2) = 3 = 1 + 1 + 1$,
9. if $l = m \neq r = s \neq i \neq l$, it is equal to $\mathbb{E}(X^2)^3 = 1$,
10. if $l = r \neq m = s \neq i \neq l$, it is equal to $\mathbb{E}(X^2)^3 = 1$,
11. if $l = s \neq m = r \neq i \neq l$, it is equal to $\mathbb{E}(X^2)^3 = 1$,
12. elsewhere, it is equal to 0.

Instead of two terms as in the second equality, we have now eleven terms to rearrange. Lines 2 to 5 lack the term of line 1 to be identified as a trace term. Line 2 completed with line 1 gives the following term:

$$3\sum_{i,j,k} a_i b_j c_k \sum_{m=1}^n U(j,i)U(j,m)V(k,i)V(k,m) = 3\mathrm{tr}(ABC)$$

and we can be convinced by symmetry of the indexes that lines 3, 4, 5 give the same result.

Lines 6 and 7 also lack the term of line 1 to be identified as a trace term. Line 6 completed with line 1 gives the following term:

$$3\sum_{i,j,k} a_i b_j c_k \sum_{r=1}^{n} U(j,i)U(j,i)V(k,r)V(k,r) = 3\mathrm{tr}(AB)\mathrm{tr}(C)$$

and we can be convinced by symmetry of the indexes that line 7 gives $3\mathrm{tr}(AC)\mathrm{tr}(B)$.

Let us write the equality with the remaining terms, after distributing line 8 to lines 9, 10, 11 and merging lines 10 and 11:

$$\begin{aligned}
&\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top Byy^\top Cy) \\
&= -3\sum_{i,j,k} a_i b_j c_k U(j,i)^2 V(k,i)^2 + 12\mathrm{tr}(ABC) + 3\mathrm{tr}(AB)\mathrm{tr}(C) + 3\mathrm{tr}(AC)\mathrm{tr}(B) \\
&+ \sum_{i,j,k} a_i b_j c_k \sum_{l\neq i} \left[\sum_{r\neq i} U(j,l)^2 V(k,r)^2 + 2\sum_{m\neq i} U(j,l)U(j,m)V(k,l)V(k,m) \right]
\end{aligned}$$

We compute the last term of the previous expression in two steps. In the first step, we fix i,j,k and we add and remove terms $r=i, m=i, l=i$. In the second step, we sum for $i,j,k \in [\![1,n]\!]$.

$$\begin{aligned}
&\sum_{l\neq i} \left[\sum_{r\neq i} U(j,l)^2 V(k,r)^2 + 2\sum_{m\neq i} U(j,l)U(j,m)V(k,l)V(k,m) \right] \\
&= \sum_{l\neq i} \left[\sum_{r} U(j,l)^2 V(k,r)^2 + 2\sum_{m} U(j,l)U(j,m)V(k,l)V(k,m) \right] \\
&\quad - \sum_{l\neq i} \left[U(j,l)^2 V(k,i)^2 + 2U(j,l)U(j,i)V(k,l)V(k,i) \right] \\
&= \sum_{l} \left[\sum_{r} U(j,l)^2 V(k,r)^2 + 2\sum_{m} U(j,l)U(j,m)V(k,l)V(k,m) \right] \\
&\quad - \sum_{l} \left[U(j,l)^2 V(k,i)^2 + 2U(j,l)U(j,i)V(k,l)V(k,i) \right] \\
&\quad - \left[\sum_{r} U(j,i)^2 V(k,r)^2 + 2\sum_{m} U(j,i)U(j,m)V(k,i)V(k,m) - 3U(j,i)^2 V(k,i)^2 \right] \\
&= 1 + 2UV^\top(j,k)^2 - V(k,i)^2 - 2UV^\top(j,k)U(j,i)V(k,i) \\
&\quad - U(j,i)^2 - 2U(j,i)V(k,i)UV^\top(j,k) + 3U(j,i)^2 V(k,i)^2
\end{aligned}$$

$$\sum_{i,j,k} a_i b_j c_k \sum_{l\neq i}\left[\sum_{r\neq i} U(j,l)^2 V(k,r)^2 + 2\sum_{m\neq i} U(j,l)U(j,m)V(k,l)V(k,m)\right]$$
$$= \mathrm{tr}(A)\mathrm{tr}(B)\mathrm{tr}(C) + 2\mathrm{tr}(A)\mathrm{tr}(BC) - (\mathrm{tr}(AB)\mathrm{tr}(C) + \mathrm{tr}(AC)\mathrm{tr}(B)) - 4\mathrm{tr}(ABC)$$
$$+ 3\sum_{i,j,k} a_i b_j c_k U(j,i)^2 V(k,i)^2$$

Then, the remaining uncomputed term cancels with the first term in $\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top Byy^\top Cy)$. Finally, we get the third formula of Lemma 2:

$$\mathbb{E}_{I_n}(y \longmapsto y^\top Ayy^\top Byy^\top Cy) = \mathrm{tr}(A)\mathrm{tr}(B)\mathrm{tr}(C) + 8\mathrm{tr}(ABC)$$
$$+ 2(\mathrm{tr}(AB)\mathrm{tr}(C) + \mathrm{tr}(BC)\mathrm{tr}(A) + \mathrm{tr}(CA)\mathrm{tr}(B)).$$

6.2 Proof of Theorem 2

Proof of formula (5) The Christoffel symbols are given by [12]:

$$\left(\Gamma^{(\alpha)}_{ijk}\right)_\Sigma = \mathbb{E}_\Sigma\left[\left(\partial_i\partial_j l + \frac{1-\alpha}{2}\partial_i l\partial_j l\right)\partial_k l\right].$$

where l is the log-likelihood of the centered multivariate normal model, namely $l_\Sigma(x) = \log p_\Sigma(x) = \frac{1}{2}\left(-n\log(2\pi) - \log\det\Sigma + x^\top\Sigma^{-1}x\right)$.

So we have to compute $\mathbb{E}_\Sigma[\partial_i\partial_j l\partial_k l]$ and $\mathbb{E}_\Sigma[\partial_i l\partial_j l\partial_k l]$. We can either use the formulas of Lemma 2 to compute each, or compute the first term and deduce the second one thanks to the formula of the Levi-Civita connection [13] corresponding to the case $\alpha = 0$.

We recall that $(Z^k\partial_k l)_{|\Sigma}(x) = d_\Sigma l(Z)(x) = -\frac{1}{2}[\mathrm{tr}(\Sigma^{-1}Z) + x^\top\Sigma^{-1}Z\Sigma^{-1}x]$ (cf. table in section 2.3). As a consequence, deriving once more in this basis, we get: $(X^iY^j\partial_i\partial_j l)_{|\Sigma}(x) = \frac{1}{2}[\mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y) + x^\top\Sigma^{-1}(X\Sigma^{-1}Y + Y\Sigma^{-1}X)\Sigma^{-1}x]$. Denoting $A = \Sigma^{-1/2}(X\Sigma^{-1}Y + Y\Sigma^{-1}X)\Sigma^{-1/2}$ and $B = \Sigma^{-1/2}Z\Sigma^{-1/2}$, we can now compute the first term in the Christoffel symbols: $X^iY^jZ^k\mathbb{E}_\Sigma[\partial_i\partial_j l\partial_k l] = \frac{1}{4}(\mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y)\mathrm{tr}(\Sigma^{-1}Z)(-1+1) + \mathrm{tr}(A)\mathrm{tr}(B) - (2\mathrm{tr}(AB) + \mathrm{tr}(A)\mathrm{tr}(B))) = -\frac{1}{2}\mathrm{tr}(AB) = -\mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z)$.

Let us compute the second term in the Christoffel symbols thanks to the two previously detailed methods.

1st method: using the third equality of Lemma 2

Thanks to the expression of $\partial_i l$ given above, and denoting $A = \Sigma^{-1/2}X\Sigma^{-1/2}$, $B = \Sigma^{-1/2}Y\Sigma^{-1/2}$ and $C = \Sigma^{-1/2}Z\Sigma^{-1/2}$, we are now able to compute $X^iY^jZ^k\mathbb{E}_\Sigma[\partial_i l\partial_j l\partial_k l] = \frac{1}{8}(-\mathrm{tr}(A))\mathrm{tr}(B)\mathrm{tr}(C)(-1+1+1+1) - \mathrm{tr}(A)(2\mathrm{tr}(BC) + \mathrm{tr}(B)\mathrm{tr}(C)) - \mathrm{tr}(B)(2\mathrm{tr}(AC) + \mathrm{tr}(A)\mathrm{tr}(C)) - \mathrm{tr}(C)(2\mathrm{tr}(AB) + \mathrm{tr}(A)\mathrm{tr}(B)) + \frac{1}{8}(8\mathrm{tr}(ABC) + 2(\mathrm{tr}(AB)\mathrm{tr}(C) + \mathrm{tr}(AC)\mathrm{tr}(B) + \mathrm{tr}(BC)\mathrm{tr}(A)) + \mathrm{tr}(A)\mathrm{tr}(B)\mathrm{tr}(C))) = \mathrm{tr}(ABC) = \mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z)$.

2^{nd} method: using the formula of the Levi-Civita connection

From [13], we know that $X^iY^jZ^k\left(\Gamma^{(0)}_{ijk}\right)_\Sigma = -\frac{1}{2}\mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z)$. On the other hand, $X^iY^jZ^k\left(\Gamma^{(0)}_{ijk}\right)_\Sigma = X^iY^jZ^k\mathbb{E}_\Sigma(\partial_i\partial_j l\partial_k l)+\frac{1}{2}X^iY^jZ^k\mathbb{E}(\partial_i l\partial_j l\partial_k l)$ and we already computed $X^iY^jZ^k\mathbb{E}_\Sigma(\partial_i\partial_j l\partial_k l) = -\mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z)$. We can deduce the second term $X^iY^jZ^k\mathbb{E}(\partial_i l\partial_j l\partial_k l) = \mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z)$.

To conclude:

$$\begin{aligned} g_\Sigma(\nabla^{(\alpha)}_X Y - \partial_X Y, Z) &= X^iY^jZ^k\left(\Gamma^{(\alpha)}_{ijk}\right)_\Sigma \\ &= -\frac{1+\alpha}{2}\mathrm{tr}(\Sigma^{-1}X\Sigma^{-1}Y\Sigma^{-1}Z) \\ &= -\frac{1+\alpha}{4}\mathrm{tr}(\Sigma^{-1}(X\Sigma^{-1}Y + Y\Sigma^{-1}X)\Sigma^{-1}Z) \end{aligned}$$

so $(\nabla^{(\alpha)}_X Y)_{|\Sigma} = (\partial_X Y)_{|\Sigma} - \frac{1+\alpha}{2}(X\Sigma^{-1}Y + Y\Sigma^{-1}X)$ and formula (5) is proved.

Mixture and exponential connections For $\alpha = -1$, there only remains the Levi-Civita connection of the Euclidean metric ∂ in the formula: this is the mixture m-connection. On the other hand, the exponential e-connection ($\alpha = 1$) is given by $\nabla^{(e)}_X Y = \partial_X Y - (X\Sigma^{-1}Y + Y\Sigma^{-1}X)$.

Let us compute the Levi-Civita connection ∇^I of the inverse-Euclidean metric given by $g^I_\Sigma(X,Y) = \mathrm{tr}(\Sigma^{-2}X\Sigma^{-2}Y)$ thanks to the Koszul formula:

$$\begin{aligned} 2g^I_\Sigma(\nabla^I_X Y, Z) &= X_\Sigma(g^I(Y,Z)) + Y_\Sigma(g^I(X,Z)) - Z_\Sigma(g^I(X,Y)) \\ &\quad + g^I([X,Y],Z) - g^I([X,Z],Y) - g^I([Y,Z],X) \\ &= -2\mathrm{tr}((\Sigma^{-2}X\Sigma^{-1} + \Sigma^{-1}X\Sigma^{-2})Y\Sigma^{-2}Z) \\ &\quad - 2\mathrm{tr}((\Sigma^{-2}Y\Sigma^{-1} + \Sigma^{-1}Y\Sigma^{-2})X\Sigma^{-2}Z) \\ &\quad + 2\mathrm{tr}((\Sigma^{-2}Z\Sigma^{-1} + \Sigma^{-1}Z\Sigma^{-2})X\Sigma^{-2}Y) \\ &\quad + 2\mathrm{tr}(\Sigma^{-2}\partial_X Y\Sigma^{-2}Z) \\ &= 2\mathrm{tr}(\Sigma^{-2}[\partial_X Y - (X\Sigma^{-1}Y + Y\Sigma^{-1}X)]\Sigma^{-2}Z) \end{aligned}$$

So $\nabla^I_X Y = \partial_X Y - (X\Sigma^{-1}Y + Y\Sigma^{-1}X)$ and $\nabla^{(e)} = \nabla^I$.