Journal of Lie Theory
Volume **32** (2022) 563–600

Curvatures of Stiefel Manifolds with Deformation Metrics

Du Nguyen

Communicated by L. San Martin

Abstract. We compute curvatures of a family of metrics on Stiefel manifolds, introduced recently by Hüper, Markina and Silva Leite. We derive the formulas from two approaches, one using curvature formulas for left-invariant metrics on homogeneous spaces, computed for the case of Cheeger/Jensen deformation metrics of a quotient space of a compact Lie group; another from a global curvature formula derived in our recent work. Allowing more than one deformation parameter, we compute Ricci curvature for a large family of diagonal metrics explicitly and obtain new Einstein metrics. We analyze the sectional curvature range and identify the parameter range where the manifold has non-negative sectional curvature. We provide the exact sectional curvature range when the number of columns in a Stiefel matrix is 2, and a conjectural range for other cases. We expect the method developed here generalizes to other homogeneous spaces.

Mathematics Subject Classification: 22E70, 53C30, 17B81, 65K10, 49Q12, 53C25, 68T05.

Key Words: Lie group, homogeneous space, optimization, Riemannian geometry, Riemannian curvature, Einstein manifold, Stiefel manifold.

1. Introduction

The purpose of this paper is to compute and analyze curvatures of a Stiefel manifold with the family of metrics defined in [13] and its generalization. It turns out that this family of metrics can be identified with metrics arising from a Cheeger deformation, one of the main tools to construct non-negative curvature metrics [7]. It was also studied by Jensen [14] in the same issue of the Journal of Differential Geometry to construct new Einstein metrics. In the case of a normal homogeneous space $\mathtt{G}/\mathtt{K}$, the quotient of two connected compact Lie groups, with $\mathtt{G}$ equipped with a bi-invariant Riemannian metric, the deformation scales the metric of a connected subgroup of $\mathtt{G}$ but maintains left-invariance and the quotient structure. Using a result of Michor [15] and the Euler-Poisson-Arnold framework [1], we derive curvature formulas from a more Lie theoretic point of view, expressing the Ricci curvature in terms of traces constructed from the ad operator and apply the result to Stiefel manifolds. We provide estimates for the sectional curvature range. For the special orthogonal group $\mathrm{SO}(n)$, we also show the Ricci curvature has a simple form when we allow a larger family of deformation metrics, generalizing previous results and providing new examples of Einstein metrics.

A contribution of the article is providing the curvature formulas for homogeneous spaces in basis-free forms, in many cases applicable to semi-Riemannian geometry. The trace form of the Ricci curvature makes computation more intuitive in the

ISSN 0949–5932 / \$2.50

instances we consider. In addition, the interest in these metrics arises from the fact that they appear in the field of Riemannian optimization, where two instances of the metrics appeared [9]. In [25], the authors show the rate of convergence of optimization algorithms on manifolds depends on sectional curvature range, thus the result here will also be helpful in applications.

In a recent paper [18], we derived global formulas to compute the curvature of a manifold $\mathcal{M}$, embedded differentiably in a Euclidean space $\mathcal{E}$, with metric defined by an operator g from $\mathcal{M}$ to the space of positive-definite operators on $\mathcal{E}$. The formulas have similar forms to the classical formula for the curvature in local coordinates, and it is straightforward conceptually to apply these formulas to Stiefel manifolds. We show this approach also gives the same curvature formulas.

Recall for two positive integers $p < n$, the real Stiefel manifold $\mathrm{St}_{p,n}$ consists of real orthogonal matrices Y of size $n\times p$. If α_1, α_0 are two positive numbers, the metric in [13] could be reparameterized so that the inner product of two tangent vectors ξ, η on $\mathrm{St}_{p,n}$, considered as elements of $\mathbb{R}^{n\times p}$ at $Y\in\mathrm{St}_{p,n}$ is given by $\alpha_0\operatorname{Tr}(\xi^{\mathsf{T}}\eta) + (\alpha_1-\alpha_0)\operatorname{Tr}(\xi^{\mathsf{T}}YY^{\mathsf{T}}\eta)$. Set $\alpha = \alpha_1/\alpha_0$, and up to scaling, we can take $\alpha_0 = 1$. This family of metrics contains both well-known metrics on Stiefel manifolds, the embedded ($\alpha = 1$, where the metric is induced from the embedding in $\mathbb{R}^{n\times p}$) and canonical metrics ($\alpha=\frac{1}{2}$) ($\mathrm{St}_{p,n}$ is normal homogeneous in this case). It will be shown in theorem 4.1 that if $\mathrm{SO}(n)$ is equipped with a Cheeger/Jensen deformation metric with deformation parameter 2α (reviewed in Theorem 2) from the right-multiplication action of $\mathrm{SO}(p)$ embedded diagonally then $\mathrm{SO}(n)/\mathrm{SO}(n-p)$ with the quotient metric could be identified with $\mathrm{St}_{p,n}$ with the metric just described. The main differential geometric concept we use is Riemannian submersion. The quotient map of a manifold by a group of isometries acting freely and properly is a Riemannian submersion. Tangent vectors on a quotient manifold could be lifted to the original manifold via *horizontal lifts*, and [19, 20] provide horizontal lift formulas for covariant derivatives and curvatures.

While Jensen [14] has provided a framework to compute curvatures for Cheeger/Jensen deformation metrics for the case of principal bundles, which applies to Stiefel manifolds, and recently [2, 4] provided explicit computation and extensions for Ricci and scalar curvature, a study of the sectional curvature range is not yet available to the best of our knowledge (note [22] is an early paper dealing with the embedded metric). We analyze the sectional curvature for the Stiefel manifold equipped with the deformation metric. We show the sectional curvature range always contains a specific interval, which is likely to be the full curvature range for metrics in the family. The ends of the interval are piecewise smooth functions described in Table 2. In particular, except for some special cases, for the embedded metric ($\alpha = 1$) on the Stiefel manifold, we show the curvature range contains the interval $[-\frac{1}{2}, 1]$, thus it could have negative curvatures, in contrast to the canonical metric ($\alpha=\frac{1}{2}$), which has range $[0, \frac{5}{4}]$. The values at the ends of each interval correspond to certain root configurations, therefore we expect that this picture generalizes to other homogeneous spaces.

Specifically, $\mathrm{St}_{2,3}$ has positive curvature for $\alpha < \frac{2}{3}$, non-negative curvature for $\alpha = \frac{2}{3}$ and both negative and positive curvature for $\alpha > \frac{2}{3}$. With $n > 3$, the Stiefel manifold $\mathrm{St}_{2,n}$ has non-negative curvature for $\alpha \leq \frac{2}{3}$, and both negative and positive curvature for $\alpha > \frac{2}{3}$, and we identify the exact sectional curvature range in this case. For $p \geq 3$, we show $\mathrm{St}_{p,n}$ has non-negative curvature for $\alpha \leq \frac{1}{2}$ and both negative and positive curvature otherwise. This agrees with [12] and we show the curvature range contains negative values in the indicated intervals.

Benefitting from earlier studies of Einstein metrics on $\mathrm{SO}(n)$ and $\mathrm{St}_{p,n}$ in [2, 4, 5, 14], we find that expressing the Ricci curvature in terms of traces give us a simple formula for Ricci curvature for certain metrics generalizing the Cheeger/Jensen metrics, and the flexibility allows us to construct new Einstein metrics. We express the Ricci curvature in eq. (9) as a sum of four traces, one is the Killing form and the others are constructed from the adjoint of the ad operator. Relative to the left-invariant metric, we have four operators. If they commute, the Ricci curvature and the Killing form could be diagonalized simultaneously, simplifying the Einstein equations. We demonstrate a family of diagonal metrics on $\mathrm{SO}(n)$ and $\mathrm{St}_{p,n}$ satisfy this condition and provide formulas for the Ricci curvature in this case, and recover a family of Einstein metrics depending on two parameters when n is even (see [24]).

For notations, if n and m are two positive integers, by $\mathbb{R}^{n\times m}$, we denote the space of $n \times m$ matrices in $\mathbb{R}$, the field of real numbers. We denote by $\mathfrak{o}(p)$ the space of antisymmetric matrices in $\mathbb{R}^{p\times p}$. The transpose of matrix is denoted by T. Working on a manifold, say $\mathcal{M}$, by $\mathrm{D}_\xi F$, we denote the directional (Lie) derivative of a scalar/vector/operator-valued function F on $\mathcal{M}$ in direction ξ (either a tangent vector defined at a point $x \in \mathcal{M}$, or a vector field on $\mathcal{M}$). The Frobenius norm is denoted by $\|\cdot\|_F$. The convention for the curvature of three vector fields $\mathtt{X}, \mathtt{Y}, \mathtt{Z}$ is

$$\mathrm{R}_{\mathtt{XY}}\,\mathtt{Z} = \nabla_{[\mathtt{X},\mathtt{Y}]}\mathtt{Z} - \nabla_{\mathtt{X}}\nabla_{\mathtt{Y}}\mathtt{Z} + \nabla_{\mathtt{Y}}\nabla_{\mathtt{X}}\mathtt{Z}$$

We would like to thank the anonymous referee for constructive suggestions and for pointing us to important references resulting in a much-improved article, the editors for the service, and our family for their support in this project.

2. Curvature formulas and deformation metrics

For a Lie group $\mathtt{G}$, with $U \in \mathtt{G}$, we will denote by $\mathcal{L}_U$ the left-multiplication map and by $d\mathcal{L}_U$ its differential. As usual, ad_A denotes the operator $X \mapsto [A, X]$ on the Lie algebra $\mathfrak{g}$ of $\mathtt{G}$ ($A, X \in \mathfrak{g}$). A left-invariant metric on $\mathtt{G}$ is defined by an inner product $\langle\rangle_{\mathrm{P}}$ on $\mathfrak{g}$, and the inner product of two tangent vectors ξ, η at $U \in \mathtt{G}$ is defined by $\langle d\mathcal{L}_{U^{-1}}\xi, d\mathcal{L}_{U^{-1}}\eta\rangle_{\mathrm{P}}$. For a quotient manifold $\mathtt{G}/\mathtt{K}$, with $\mathtt{K}$ a connected subgroup of $\mathtt{G}$ with Lie algebra $\mathfrak{k}$, a vertical vector is a tangent vector on $\mathtt{G}$ that maps to zero by the differential of the quotient map, and a horizontal vector is a tangent vector of $\mathtt{G}$ orthogonal to vertical vectors. At the identity of $\mathtt{G}$, a vertical vector corresponds to an element of $\mathfrak{k}$, while a horizontal vector corresponds to an element of $\mathfrak{m}$, the orthogonal complement of $\mathfrak{k}$ in $\mathfrak{g}$. In the following we introduce a bracket $[\ \ ]_{\mathrm{P}}$, depending on the inner product P, we also use the notation $[\ \ ]_{\mathfrak{m}}$ to denote the projection of the usual Lie bracket to the space $\mathfrak{m}$. We believe there is little risk of confusion. We recall a few results on curvatures.

Proposition 2.1. *Let $\mathtt{G}$ be a connected Lie group with Lie algebra $\mathfrak{g}$ with a left-invariant metric given by an inner product $\langle\cdot\rangle_{\mathrm{P}}$ on $\mathfrak{g}$. For $A \in \mathfrak{g}$, let $\mathrm{ad}^\dagger_A$ be the adjoint of ad_A under $\langle\cdot\rangle_{\mathrm{P}}$, that means $\mathrm{ad}^\dagger{}_A$ is a linear operator on $\mathfrak{g}$ such that $\langle [A, A_1], A_2\rangle_{\mathrm{P}} = \langle A_1, \mathrm{ad}^\dagger{}_A A_2\rangle_{\mathrm{P}}$. Define*

$$[A, B]_{\mathrm{P}} := [A, B] - \mathrm{ad}^\dagger{}_A B - \mathrm{ad}^\dagger{}_B A. \tag{1}$$

Define $\overline{\mathrm{ad}}_A B := \mathrm{ad}^{\dagger}{}_B A$ *and* $\mathrm{ad}^{\mathrm{P}}_A B := [A,B]_{\mathrm{P}}$. *Then* $\overline{\mathrm{ad}}_A$ *and* $\mathrm{ad}^{\mathrm{P}}_A$ *are antisymmetric and* $[A,B]_{\mathrm{P}} - [B,A]_{\mathrm{P}} = 2[A,B]$. *Let* $\nabla^{\mathtt{G}}$ *be the Levi-Civita connection on* $\mathtt{G}$. *For two vector fields* $\mathtt{X},\mathtt{Y}$ *on* $\mathtt{G}$, *there exists* $\mathfrak{g}$*-valued functions* $A(U), B(U)$, $U\in\mathtt{G}$ *such that* $\mathtt{X}(U) = d\mathcal{L}_U A(U), \mathtt{Y}(U) = d\mathcal{L}_U B(U)$. *We have*

$$(\nabla^{\mathtt{G}}_{\mathtt{X}}\mathtt{Y})(U) = d\mathcal{L}_U((\mathrm{D}_{\mathtt{X}} B)(U) + \frac{1}{2}[A(U),B(U)]_{\mathrm{P}}) \tag{2}$$

where $\mathrm{D}_{\mathtt{X}} B$ *is the Lie-derivative of the* $\mathfrak{g}$*-valued function* B *by the vector field* $\mathtt{X}$. *For* $\omega_1,\omega_2,\omega_3\in\mathfrak{g}$, *the curvature of* $\mathtt{G}$ *at the identity is given by*

$$\mathrm{R}^{\mathtt{G}}_{\omega_1,\omega_2}\omega_3 = \frac{1}{2}[[\omega_1,\omega_2],\omega_3]_{\mathrm{P}} - \frac{1}{4}[\omega_1[\omega_2,\omega_3]_{\mathrm{P}}]_{\mathrm{P}} + \frac{1}{4}[\omega_2[\omega_1,\omega_3]_{\mathrm{P}}]_{\mathrm{P}} \tag{3}$$

Let $\mathfrak{k}$ *be a subalgebra of* $\mathfrak{g}$ *such that* P *is* $\mathrm{ad}(\mathfrak{k})$*-invariant,* $\langle [A,K],B\rangle_{\mathrm{P}} = \langle A,[K,B]\rangle_{\mathrm{P}}$ *for* $K\in\mathfrak{k}, A,B\in\mathfrak{g}$, *and* $\mathfrak{k}$ *corresponds to a closed subgroup* $\mathtt{K}\subset\mathtt{G}$, *such that* $\mathtt{K}$ *acts freely and properly on* $\mathtt{G}$ *by isometries under right multiplication and* $\mathtt{G}/\mathtt{K}$ *is a homogeneous space. If* $\mathfrak{g} = \mathfrak{k}\oplus\mathfrak{m}$ *is an orthogonal decomposition under* $\langle\cdot\rangle_{\mathrm{P}}$, *then the horizontal lift of the curvature of* $\mathtt{M} = \mathtt{G}/\mathtt{K}$ *at* o, *the equivariant class containing the unit of* $\mathtt{G}$, *evaluated at three horizontal vectors* $\omega_1,\omega_2,\omega_3\in\mathfrak{m}$ *is*

$$\begin{aligned}\mathrm{R}^{\mathtt{M}}_{\omega_1,\omega_2}\omega_3 = (&\frac{1}{2}[[\omega_1,\omega_2],\omega_3]_{\mathrm{P}} - \frac{1}{4}[\omega_1[\omega_2,\omega_3]_{\mathrm{P}}]_{\mathrm{P}} + \frac{1}{4}[\omega_2[\omega_1,\omega_3]_{\mathrm{P}}]_{\mathrm{P}}\\ &+\frac{1}{2}\mathrm{ad}^{\dagger}{}_{\omega_3}[\omega_1,\omega_2]_{\mathfrak{k}} - \frac{1}{4}\mathrm{ad}^{\dagger}{}_{\omega_1}[\omega_2,\omega_3]_{\mathfrak{k}} + \frac{1}{4}\mathrm{ad}^{\dagger}{}_{\omega_2}[\omega_1,\omega_3]_{\mathfrak{k}})_{\mathfrak{m}}\end{aligned} \tag{4}$$

Here, $\omega_{\mathfrak{v}}$ *denotes the orthogonal projection of* ω *to* $\mathfrak{v}$ *for an element* $\omega\in\mathfrak{g}$ *and a subspace* $\mathfrak{v}$ *of* $\mathfrak{g}$. *Also, given two vector fields* $\mathtt{X},\mathtt{Y}$ *on* $\mathtt{M}$, *which lift to horizontal vector fields* $\bar{\mathtt{X}},\bar{\mathtt{Y}}$ *on* $\mathtt{G}$, *with* $\bar{\mathtt{X}}(U) = d\mathcal{L}_U A(U)$, $\bar{\mathtt{Y}} = d\mathcal{L}_U B(U)$ *for two* $\mathfrak{g}$*-valued functions* $A(U), B(U)$ *on* $\mathtt{G}$ *then the horizontal lift of* $\nabla_{\mathtt{X}}\mathtt{Y}$ *is given by*

$$\overline{\nabla_{\mathtt{X}}\mathtt{Y}(U)} = d\mathcal{L}_U((\mathrm{D}_{\bar{\mathtt{X}}} B)(U) + \frac{1}{2}[A(U),B(U)]_{\mathrm{P}})_{\mathfrak{m}} \tag{5}$$

Proof. To show $\overline{\mathrm{ad}}_A$ is antisymmetric, consider $A,B,C\in\mathfrak{g}$. We have

$$\langle\overline{\mathrm{ad}}_A B, C\rangle_{\mathrm{P}} = \langle\mathrm{ad}^{\dagger}{}_B A, C\rangle_{\mathrm{P}} = \langle A,[B,C]\rangle_{\mathrm{P}} = -\langle A,[C,B]\rangle_{\mathrm{P}} = -\langle\mathrm{ad}^{\dagger}{}_C A, B\rangle_{\mathrm{P}}$$

which is $-\langle B,\overline{\mathrm{ad}}_A C\rangle_{\mathrm{P}}$. Since $\mathrm{ad}^{\mathrm{P}}_A = \mathrm{ad}_A - \mathrm{ad}^{\dagger}{}_A - \overline{\mathrm{ad}}_A$, $\mathrm{ad}^{\mathrm{P}}_A$ is antisymmetric.

For each smooth function $F\colon\mathtt{G}\to\mathfrak{g}$ denote by $\mathcal{L}[F]$ the vector field $U\mapsto d\mathcal{L}_U F(U)$. Denote by $\langle\cdot\rangle_{\mathtt{G}}$ the left-invariant Riemannian metric induced by P on $\mathtt{G}$. For three vector fields $\mathtt{X},\mathtt{Y},\mathtt{Z}$ with $\mathtt{X} = \mathcal{L}[A], \mathtt{Y} = \mathcal{L}[B]$ and $\mathtt{Z} = \mathcal{L}[C]$ with three smooth $\mathfrak{g}$-valued functions A,B,C, we have

$$\begin{aligned}\mathrm{D}_{\mathtt{X}}\langle\mathtt{Y},\mathtt{Z}\rangle_{\mathtt{G}} &= \mathrm{D}_{\mathtt{X}}\langle B,C\rangle_{\mathrm{P}} = \langle\mathrm{D}_{\mathtt{X}} B, C\rangle_{\mathrm{P}} + \langle B,\mathrm{D}_{\mathtt{X}} C\rangle_{\mathrm{P}}\\ &= \langle\mathcal{L}[\mathrm{D}_{\mathtt{X}} B + \frac{1}{2}[A,B]_{\mathrm{P}}],\mathtt{Z}\rangle_{\mathtt{G}} + \langle\mathtt{Y},\mathcal{L}[(\mathrm{D}_{\mathtt{X}} C + \frac{1}{2}[A,C]_{\mathrm{P}}]\rangle_{\mathtt{G}}\end{aligned}$$

as the metric is left-invariant, P is constant on $\mathfrak{g}$, and $\mathrm{ad}^{\mathrm{P}}_A$ is antisymmetric. We can verify $\mathcal{L}[\mathrm{D}_{\mathtt{X}} B + \frac{1}{2}[A,B]_{\mathrm{P}}]$ satisfies the derivative rules of a connection, and we have just proved it is metric compatible. Torsion-freeness follows from $[A,B]_{\mathrm{P}} - [B,A]_{\mathrm{P}} = 2[A,B]$, thus $\mathcal{L}[\mathrm{D}_{\mathtt{X}} B + \frac{1}{2}[A,B]_{\mathrm{P}}]$ is the Levi-Civita connection.

Equation 2 is from [15], equation 3.3.2 (the author uses a right-invariant metric). It is related to the Euler-Poisson-Arnold equation (EPDiff), see equation (55) in Arnold's classical paper [1]. See also [16].

Eq. (3) now follows from the definition of the curvature $\nabla_{[\mathtt{X},\mathtt{Y}]}\mathtt{Z} - \nabla_{\mathtt{X}}\nabla_{\mathtt{Y}}\mathtt{Z} + \nabla_{\mathtt{Y}}\nabla_{\mathtt{X}}\mathtt{Z}$ applied to the invariant vector fields $\mathcal{L}[\omega_i]$, $i \in \{1,2,3\}$. Eq. 4 follows from the O'Neill equation (Theorem 2, [19]), written in $(1,3)$ tensor form. Indeed, the O'Neill tensor of two vector fields $\mathcal{L}[A], \mathcal{L}[B]$ on $\mathtt{G}$ for $\mathfrak{g}$-valued functions A and B evaluated at the coset o is $\frac{1}{2}[A,B]_{\mathfrak{k}}$ as the just proved result for covariant derivatives shows the Lie bracket $\{\mathcal{L}[A], \mathcal{L}[B]\} = \mathcal{L}[[A,B]]$, then we use Lemma 2, [19]. By properties of adjoint and projection, the right-hand side of eq. (4) is the unique vector in $\mathfrak{m}$ such that the O'Neill equation (equation 4, theorem 2, [19]) is satisfied. Equation (5) follows from the result for $\mathtt{G}$ and property of horizontal lift of a connection in Riemannian submersion, e.g. Lemma 7.45 in [20] (because of left-invariance, we can translate the projection to the identity). ■

With the notation of Proposition 2.1, we denote the projection to $\mathfrak{k}$ and $\mathfrak{m}$ by $\pi_{\mathfrak{k}}$ and $\pi_{\mathfrak{m}}$. They are both idempotent and self-adjoint. For $\mu \in \mathfrak{m}$, $\omega_1, \omega_2 \in \mathfrak{g}$, $\langle \overline{\mathrm{ad}}_\mu \pi_{\mathfrak{k}}\omega_1, \omega_2\rangle_{\mathrm{P}} = \langle \mathrm{ad}^\dagger{}_{\pi_{\mathfrak{k}}\omega_1}\mu, \omega_2\rangle_{\mathrm{P}} = \langle \mu, [\pi_{\mathfrak{k}}\omega_1, \pi_{\mathfrak{m}}\omega_2]\rangle_{\mathrm{P}} = \langle [\mu, \pi_{\mathfrak{k}}\omega_1], \pi_{\mathfrak{m}}\omega_2\rangle_{\mathrm{P}}$ by $\mathrm{ad}(\mathfrak{k})$-invariance, which gives us the equalities below and their adjoints

$$\overline{\mathrm{ad}}_\mu \pi_{\mathfrak{k}} = \pi_{\mathfrak{m}} \mathrm{ad}_\mu \pi_{\mathfrak{k}} = \mathrm{ad}_\mu \pi_{\mathfrak{k}} = \pi_{\mathfrak{m}} \overline{\mathrm{ad}}_\mu \pi_{\mathfrak{k}} \tag{6}$$

$$(\overline{\mathrm{ad}}_\mu \pi_{\mathfrak{k}})^\dagger = \pi_{\mathfrak{k}} \mathrm{ad}^\dagger{}_\mu \pi_{\mathfrak{m}} = \pi_{\mathfrak{k}} \mathrm{ad}^\dagger{}_\mu = -\pi_{\mathfrak{k}} \overline{\mathrm{ad}}_\mu \pi_{\mathfrak{m}} \tag{7}$$

We reformulate a formula for the Ricci curvature [6], chapter 7, in terms of trace, for use in section 5. The proof works for semi-Riemannian metrics.

Proposition 2.2. *For $A, B, C, D \in \mathfrak{m}$ we have*

$$\begin{aligned}\langle \mathrm{R}^{\mathtt{M}}_{AB}\, C, D\rangle_{\mathrm{P}} &= -\frac{1}{4}\langle \mathrm{ad}^\dagger{}_A C + \mathrm{ad}^\dagger{}_C A, \mathrm{ad}^\dagger{}_B D + \mathrm{ad}^\dagger{}_D B\rangle_{\mathrm{P}}\\ &+\frac{1}{4}\langle \mathrm{ad}^\dagger{}_B C + \mathrm{ad}^\dagger{}_C B, \mathrm{ad}^\dagger{}_A D + \mathrm{ad}^\dagger{}_D A\rangle_{\mathrm{P}} - \frac{1}{2}\langle [A,B]_{\mathfrak{m}}, [C,D]_{\mathfrak{m}}\rangle_{\mathrm{P}}\\ &+\frac{1}{4}\langle [B,C]_{\mathfrak{m}}, [A,D]_{\mathfrak{m}}\rangle_{\mathrm{P}} - \frac{1}{4}\langle [A,C]_{\mathfrak{m}}, [B,D]_{\mathfrak{m}}\rangle_{\mathrm{P}} + \frac{1}{4}\langle [[A,B],C], D\rangle_{\mathrm{P}}\\ &-\frac{1}{4}\langle [[A,B],D], C\rangle_{\mathrm{P}} - \frac{1}{4}\langle [A,[C,D]], B\rangle_{\mathrm{P}} - \frac{1}{4}\langle [B,[D,C]], A\rangle_{\mathrm{P}}\end{aligned} \tag{8}$$

The Ricci curvature is given by

$$\begin{aligned}\textsc{Ric}(A,A) &= \mathrm{Tr}\,\mathrm{ad}_{\mathrm{ad}^\dagger{}_A A}\pi_{\mathfrak{m}} + \frac{1}{4}\sum_{i,j=1}^{N}\langle A, [v_i, v_j]\rangle_{\mathrm{P}}\langle A, [v_i^*, v_j^*]\rangle_{\mathrm{P}}\\ &-\frac{1}{2}\sum_{i=1}^{N}\langle [A,v_i]_{\mathfrak{m}}, [A,v_i^*]_{\mathfrak{m}}\rangle_{\mathrm{P}} - \frac{1}{2}\sum_{i=1}^{N}\langle [[A,v_i]_{\mathfrak{k}}, v_i^*], A\rangle_{\mathrm{P}} - \frac{1}{2}\sum_{i=1}^{N}\langle [A,[A,v_i^*]], v_i\rangle_{\mathrm{P}}\\ &= \mathrm{Tr}\,\mathrm{ad}_{\mathrm{ad}^\dagger{}_A A}\pi_{\mathfrak{m}} - \frac{1}{4}\mathrm{Tr}\,\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}} - \frac{1}{2}\mathrm{Tr}\,\mathrm{ad}^\dagger{}_A\pi_{\mathfrak{m}}\mathrm{ad}_A\pi_{\mathfrak{m}} - \frac{1}{2}\mathrm{Tr}\,\mathrm{ad}_A\mathrm{ad}_A\end{aligned} \tag{9}$$

where $\{v_i\}_{i=1}^{\dim\mathfrak{m}}$ is a basis of $\mathfrak{m}$, and $\{v_i^*\}_{i=1}^{\dim\mathfrak{m}}$ its dual basis, thus $\langle v_i, v_j^*\rangle_{\mathrm{P}} = \delta_{ij}$ for $1 \leq i,j \leq \dim\mathfrak{m}$ with δ_{ij} is the Kronecker's delta.

Since $\textsc{Ric}(A,A)$ is the trace of $X \mapsto \mathrm{R}^{\mathtt{M}}_{A,X}A$, we can derive this result using trace manipulation only. However, we introduce eq. (8), generalizing the formulas of Püttmann [21] and Arnold [1], Jensen and Besse for $\langle \mathrm{R}^{\mathtt{M}}_{AB}A, B\rangle_{\mathrm{P}}$. This equation clearly shows $\langle \mathrm{R}^{\mathtt{M}}_{AB}C, D\rangle_{\mathrm{P}}$ satisfies the Bianchi identities.

Proof. For $A, B, C, D \in \mathfrak{g}$, set $S_{AB} = \mathrm{ad}^\dagger{}_A B + \mathrm{ad}^\dagger{}_B A$ then for $X \in \mathfrak{g}$, $\langle S_{AB}, X\rangle_{\mathrm{P}} = \langle A, [B,X]\rangle_{\mathrm{P}} + \langle B, [A,X]\rangle_{\mathrm{P}}$.

From eq. (3), since ad^{P} is antisymmetric, we have

$$
\begin{aligned}
4\langle \mathrm{R}^{\mathtt{G}}_{A,B}\, C, D\rangle_{\mathrm{P}} &= -2\langle C, [[A,B],D]_{\mathrm{P}}\rangle_{\mathrm{P}} + \langle [B,C]_{\mathrm{P}}, [A,D]_{\mathrm{P}}\rangle_{\mathrm{P}} - \langle [A,C]_{\mathrm{P}}, [B,D]_{\mathrm{P}}\rangle_{\mathrm{P}}\\
&= -2\langle C, [[A,B],D]\rangle_{\mathrm{P}} + 2\langle C, \mathrm{ad}^{\dagger}{}_{[A,B]}\, D\rangle_{\mathrm{P}} + 2\langle C, \mathrm{ad}^{\dagger}{}_{D}[A,B]\rangle_{\mathrm{P}} + \langle [B,C],[A,D]\rangle_{\mathrm{P}}\\
&\quad - \langle S_{BC}, [A,D]\rangle_{\mathrm{P}} - \langle [B,C], S_{AD}\rangle_{\mathrm{P}} + \langle S_{BC}, S_{AD}\rangle_{\mathrm{P}} - \langle [A,C],[B,D]\rangle_{\mathrm{P}}\\
&\quad + \langle S_{AC}, [B,D]\rangle_{\mathrm{P}} + \langle [A,C], S_{BD}\rangle_{\mathrm{P}} - \langle S_{AC}, S_{BD}\rangle_{\mathrm{P}}\\
&= \langle S_{BC}, S_{AD}\rangle_{\mathrm{P}} - \langle S_{AC}, S_{BD}\rangle_{\mathrm{P}} - 2\langle C, [[A,B],D]\rangle_{\mathrm{P}} + 2\langle [[A,B],C],D\rangle_{\mathrm{P}}\\
&\quad + 2\langle [D,C],[A,B]\rangle_{\mathrm{P}} + \langle [B,C],[A,D]\rangle_{\mathrm{P}} - \langle B, [C,[A,D]]\rangle_{\mathrm{P}}\\
&\quad - \langle C, [B,[A,D]]\rangle_{\mathrm{P}} - \langle [A,[B,C]],D\rangle_{\mathrm{P}} - \langle [D,[B,C]],A\rangle_{\mathrm{P}}\\
&\quad - \langle [A,C],[B,D]\rangle_{\mathrm{P}} + \langle C, [A[B,D]]\rangle_{\mathrm{P}} + \langle A, [C[B,D]]\rangle_{\mathrm{P}}\\
&\quad + \langle [B,[A,C]],D\rangle_{\mathrm{P}} + \langle [D,[A,C]],B\rangle_{\mathrm{P}}
\end{aligned}
$$

Applying the Jacobi identity we get eq. (8) for the case of $\mathtt{G}$ reducing to

$$
\begin{aligned}
-\langle [D,[B,C]],A\rangle_{\mathrm{P}} + \langle A,[C[B,D]]\rangle_{\mathrm{P}} &= -\langle A,[B[D,C]]\rangle_{\mathrm{P}}\\
-\langle B,[C,[A,D]]\rangle_{\mathrm{P}} + \langle [D,[A,C]],B\rangle_{\mathrm{P}} &= -\langle [A,[C,D]],B\rangle_{\mathrm{P}}\\
-2\langle C,[[A,B],D]\rangle_{\mathrm{P}} - \langle C,[B,[A,D]]\rangle_{\mathrm{P}} + \langle C,[A[B,D]]\rangle_{\mathrm{P}} &= -\langle C,[[A,B],D]\rangle_{\mathrm{P}}\\
2\langle [[A,B],C],D\rangle_{\mathrm{P}} - \langle [A,[B,C]],D\rangle_{\mathrm{P}} + \langle [B,[A,C]],D\rangle_{\mathrm{P}} &= \langle [[A,B],C],D\rangle_{\mathrm{P}}
\end{aligned}
$$

If $A, B, C, D \in \mathfrak{m}$, the expression for $\langle \mathrm{R}^{\mathtt{M}}_{AB}\, C, D\rangle_{\mathrm{P}}$ contains additional terms

$$
\frac{1}{2}\langle [A,B]_{\mathfrak{k}}, [C,D]_{\mathfrak{k}}\rangle_{\mathrm{P}} - \frac{1}{4}\langle [B,C]_{\mathfrak{k}}, [A,D]_{\mathfrak{k}}\rangle_{\mathrm{P}} + \frac{1}{4}\langle [A,C]_{\mathfrak{k}}, [B,D]_{\mathfrak{k}}\rangle_{\mathrm{P}}
$$

partially offsetting the corresponding terms (without $\mathfrak{k}$) giving us eq. (8).
Let $N = \dim \mathfrak{m}$. For an operator T on $\mathfrak{m}$ we have $TX = \sum_{i=1}^{N}\langle T(X), v_i\rangle_{\mathrm{P}} v_i^*$ and $\operatorname{Tr} T = \sum_{i=1}^{N}\langle Tv_i, v_i^*\rangle_{\mathrm{P}}$. For T defined on $\mathfrak{g}$, $\operatorname{Tr} T\pi_{\mathfrak{m}} = \sum_{i=1}^{N}\langle Tv_i, v_i^*\rangle_{\mathrm{P}}$, hence

$$
\begin{aligned}
\mathrm{RIC}(A,A) &= -\frac{1}{2}\langle A, [A, \sum_{i=1}^{N} \mathrm{ad}^{\dagger}{}_{v_i}\, v_i^* + \mathrm{ad}^{\dagger}{}_{v_i^*}\, v_i]\rangle_{\mathrm{P}} \qquad (10)\\
&+ \frac{1}{4}\sum_{i=1}^{N}\langle \mathrm{ad}^{\dagger}{}_{v_i}\, A + \mathrm{ad}^{\dagger}{}_{A}\, v_i, \mathrm{ad}^{\dagger}{}_{A}\, v_i^* + \mathrm{ad}^{\dagger}{}_{v_i^*}\, A\rangle_{\mathrm{P}}\\
&- \frac{1}{2}\sum_{i=1}^{N}\langle [A,v_i]_{\mathfrak{m}}, [A,v_i^*]_{\mathfrak{m}}\rangle_{\mathrm{P}} + \frac{1}{4}\sum_{i=1}^{N}\langle [v_i,A]_{\mathfrak{m}}, [A,v_i^*]_{\mathfrak{m}}\rangle_{\mathrm{P}} + \frac{1}{4}\sum_{i=1}^{N}\langle [[A,v_i],A],v_i^*\rangle_{\mathrm{P}}\\
&- \frac{1}{4}\sum_{i=1}^{N}\langle [[A,v_i],v_i^*],A\rangle_{\mathrm{P}} - \frac{1}{4}\sum_{i=1}^{N}\langle [A,[A,v_i^*]],v_i\rangle_{\mathrm{P}} - \frac{1}{4}\sum_{i=1}^{N}\langle [v_i,[v_i^*,A]],A\rangle_{\mathrm{P}}
\end{aligned}
$$

from eq. (8). Note $\langle C, \sum_{i=1}^{N}\mathrm{ad}^{\dagger}{}_{v_i}\, v_i^*\rangle_{\mathrm{P}} = \sum_{i=1}^{N}\langle [v_i, C], v_i^*]\rangle_{\mathrm{P}} = -\operatorname{Tr}\mathrm{ad}_C\,\pi_{\mathfrak{m}}$, the first term of eq. (10) reduces to $-\langle A, [A, \sum_{i=1}^{N}\mathrm{ad}^{\dagger}{}_{v_i}\, v_i^*]\rangle_{\mathrm{P}} = \operatorname{Tr}\mathrm{ad}_{\mathrm{ad}^{\dagger}{}_A A}\,\pi_{\mathfrak{m}}$. Using $(\overline{\mathrm{ad}}_A)^{\dagger} = -\overline{\mathrm{ad}}_A$ and invariance of trace, we have

$$
\operatorname{Tr}\overline{\mathrm{ad}}_A\,\mathrm{ad}^{\dagger}{}_A\,\pi_{\mathfrak{m}} = \operatorname{Tr}(\overline{\mathrm{ad}}_A\,\mathrm{ad}^{\dagger}{}_A\,\pi_{\mathfrak{m}})^{\dagger} = -\operatorname{Tr}\pi_{\mathfrak{m}}\,\mathrm{ad}_A\,\overline{\mathrm{ad}}_A = -\operatorname{Tr}\mathrm{ad}_A\,\overline{\mathrm{ad}}_A\,\pi_{\mathfrak{m}}
$$

Write $\langle \mathrm{ad}^{\dagger}{}_{v_i}\, A + \mathrm{ad}^{\dagger}{}_A\, v_i, \mathrm{ad}^{\dagger}{}_A\, v_i^* + \mathrm{ad}^{\dagger}{}_{v_i^*}\, A\rangle_{\mathrm{P}} = \langle (\overline{\mathrm{ad}}_A + \mathrm{ad}^{\dagger}{}_A)^{\dagger}(\overline{\mathrm{ad}}_A + \mathrm{ad}^{\dagger}{}_A)v_i, v_i^*\rangle_{\mathrm{P}}$, and use the just proved equality, the second term of eq. (10) is

$$\frac{1}{4}\operatorname{Tr}(-\overline{\mathrm{ad}}_A\overline{\mathrm{ad}}_A-\overline{\mathrm{ad}}_A\mathrm{ad}^{\dagger}{}_A+\mathrm{ad}_A\overline{\mathrm{ad}}_A+\mathrm{ad}_A\mathrm{ad}^{\dagger}{}_A)\pi_{\mathfrak{m}}$$
$$=\frac{1}{4}\operatorname{Tr}(-\overline{\mathrm{ad}}_A\overline{\mathrm{ad}}_A+2\,\mathrm{ad}_A\overline{\mathrm{ad}}_A+\mathrm{ad}_A\mathrm{ad}^{\dagger}{}_A)\pi_{\mathfrak{m}}$$
$$=\frac{1}{4}\operatorname{Tr}(2\,\mathrm{ad}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A-\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A+\mathrm{ad}_A\pi_{\mathfrak{m}}\mathrm{ad}^{\dagger}{}_A)\pi_{\mathfrak{m}}$$

as $2\operatorname{Tr}\mathrm{ad}_A\pi_{\mathfrak{k}}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}-\operatorname{Tr}\overline{\mathrm{ad}}_A\pi_{\mathfrak{k}}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}+\operatorname{Tr}\mathrm{ad}_A\pi_{\mathfrak{k}}\mathrm{ad}^{\dagger}{}_A\pi_{\mathfrak{m}}=0$ by eqs. (6) and (7) (converting all terms to the form $\operatorname{Tr}\mathrm{ad}_A\pi_{\mathfrak{k}}\mathrm{ad}^{\dagger}{}_A\pi_{\mathfrak{m}}$). The third and fourth term of eq. (10) is $-\frac{3}{4}\operatorname{Tr}\mathrm{ad}^{\dagger}{}_A\pi_{\mathfrak{m}}\mathrm{ad}_A\pi_{\mathfrak{m}}$, the fifth and seventh is $-\frac{1}{2}\operatorname{Tr}\mathrm{ad}_A\mathrm{ad}_A\pi_{\mathfrak{m}}$, the sixth and eighth term is $-\frac{1}{2}\operatorname{Tr}\overline{\mathrm{ad}}\,\mathrm{ad}_A\pi_{\mathfrak{m}}$. Combine,

$$\begin{aligned}\mathrm{RIC}(A,A)&=\operatorname{Tr}\mathrm{ad}_{\mathrm{ad}^{\dagger}{}_A A}\pi_{\mathfrak{m}}\\&+\frac{1}{4}(2\operatorname{Tr}\mathrm{ad}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}-\operatorname{Tr}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A+\operatorname{Tr}\mathrm{ad}_A\pi_{\mathfrak{m}}\mathrm{ad}^{\dagger}{}_A\pi_{\mathfrak{m}})\\&-\frac{3}{4}\operatorname{Tr}\mathrm{ad}^{\dagger}{}_A\pi_{\mathfrak{m}}\mathrm{ad}_A\pi_{\mathfrak{m}}-\frac{1}{2}\operatorname{Tr}\mathrm{ad}_A\mathrm{ad}_A\pi_{\mathfrak{m}}-\frac{1}{2}\operatorname{Tr}\overline{\mathrm{ad}}\,\mathrm{ad}_A\pi_{\mathfrak{m}}\\&=\operatorname{Tr}\mathrm{ad}_{\mathrm{ad}^{\dagger}{}_A A}\pi_{\mathfrak{m}}-\frac{1}{4}\operatorname{Tr}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A-\frac{1}{2}\operatorname{Tr}\mathrm{ad}^{\dagger}{}_A\pi_{\mathfrak{m}}\mathrm{ad}_A\pi_{\mathfrak{m}}\\&-\frac{1}{2}\operatorname{Tr}\mathrm{ad}_A\mathrm{ad}_A\pi_{\mathfrak{m}}-\frac{1}{2}\operatorname{Tr}\overline{\mathrm{ad}}_A\pi_{\mathfrak{k}}\mathrm{ad}_A\pi_{\mathfrak{m}}\end{aligned}$$

Since $\operatorname{Tr}\overline{\mathrm{ad}}_A\pi_{\mathfrak{k}}\mathrm{ad}_A\pi_{\mathfrak{m}}=\operatorname{Tr}\mathrm{ad}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A\pi_{\mathfrak{k}}=\operatorname{Tr}\mathrm{ad}_A\mathrm{ad}_A\pi_{\mathfrak{k}}$, the last two terms combine to $-\frac{1}{2}\operatorname{Tr}\mathrm{ad}_A\mathrm{ad}_A$, we have the trace form of eq. (9). Expand

$$\begin{aligned}\operatorname{Tr}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}&=\operatorname{Tr}(X\mapsto\overline{\mathrm{ad}}_A(\sum_i\langle\overline{\mathrm{ad}}_A X,v_i\rangle_{\mathrm{P}}v_i^*))\\&=\operatorname{Tr}(X\mapsto\sum_i\langle\overline{\mathrm{ad}}_A X,v_i\rangle_{\mathrm{P}}\overline{\mathrm{ad}}_A v_i^*)=\operatorname{Tr}(X\mapsto\sum_i\langle\mathrm{ad}^{\dagger}{}_X A,v_i\rangle_{\mathrm{P}}\mathrm{ad}^{\dagger}{}_{v_i^*}A)\\&=\operatorname{Tr}(X\mapsto\sum_i\langle A,[X,v_i]\rangle_{\mathrm{P}}\mathrm{ad}^{\dagger}{}_{v_i^*}A)=\sum_{i,j=1}^{N}\langle A,[v_j,v_i]\rangle_{\mathrm{P}}\langle\mathrm{ad}^{\dagger}{}_{v_i^*}A,v_j^*\rangle_{\mathrm{P}}\\&=-\sum_{i,j=1}^{N}\langle A,[v_i,v_j]\rangle_{\mathrm{P}}\langle A,[v_i^*,v_j^*]\rangle_{\mathrm{P}}\end{aligned}$$

and combine with the remaining traces, we get the summation form of $\mathrm{RIC}(A,A)$. ∎

Remark 2.3. On $\mathfrak{m}$, we have four symmetric bilinear forms for $A,B\in\mathfrak{m}$:

$$(A,B)\mapsto\operatorname{Tr}\mathrm{ad}_A\mathrm{ad}_B\quad\text{(Killing form)},\quad(A,B)\mapsto\frac{1}{2}(\operatorname{Tr}\mathrm{ad}_{\mathrm{ad}^{\dagger}{}_A B}\pi_{\mathfrak{m}}+\operatorname{Tr}\mathrm{ad}_{\mathrm{ad}^{\dagger}{}_B A}\pi_{\mathfrak{m}}),$$
$$(A,B)\mapsto\operatorname{Tr}\overline{\mathrm{ad}}_A\pi_{\mathfrak{m}}\overline{\mathrm{ad}}_B\pi_{\mathfrak{m}},\quad(A,B)\mapsto\operatorname{Tr}\mathrm{ad}^{\dagger}{}_A\pi_{\mathfrak{m}}\mathrm{ad}_B\pi_{\mathfrak{m}}.$$

Represent them with respect to $\langle,\rangle_{\mathrm{P}}$ by the self-adjoint operators $\mathcal{B},\mathcal{B}_1,\mathcal{B}_2,\mathcal{B}_3$.
Then $\mathrm{RIC}(A,B)$ is expressed as $\langle\mathcal{B}_1A-\frac{1}{4}\mathcal{B}_2A-\frac{1}{2}\mathcal{B}_3A-\frac{1}{2}\mathcal{B}A,B\rangle_{\mathrm{P}}$.
Since $\operatorname{Tr}\mathcal{B}_3=-\operatorname{Tr}\mathcal{B}_2=\sum_{i,j}\langle[v_i,v_j],[v_i^*,v_j^*]\rangle_{\mathrm{P}}$, if $\mathcal{Z}=-\sum_{i=1}^{N}\mathrm{ad}^{\dagger}{}_{v_i}v_i^*$ is the vector representing the functional $X\mapsto\operatorname{Tr}\mathrm{ad}_X\pi_{\mathfrak{m}}$, then $\operatorname{Tr}\mathcal{B}_1=-\|\mathcal{Z}\|_{\mathrm{P}}^2$, thus the scalar curvature formula [6], equation (7.39) becomes

$$-\|\mathcal{Z}\|_{\mathrm{P}}^2-\frac{1}{4}\operatorname{Tr}\mathcal{B}_3-\frac{1}{2}\operatorname{Tr}\mathcal{B}\tag{11}$$

For a subspace $\mathfrak{v}\subset\mathfrak{g}$, we write $\omega_{1\mathfrak{v}}$ for $(\omega_1)_{\mathfrak{v}}$, the projection of ω_1 to $\mathfrak{v}$ ($\omega_1\in\mathfrak{g}$). We write $[\omega_1,\omega_2]_{\mathfrak{v}}$, $[\omega_1,\omega_2]_{\mathrm{P}\,\mathfrak{v}}$ for the corresponding projections of brackets.
On a Lie group with a bi-invariant metric $\langle\cdot\rangle$, we now review the Cheeger/Jensen deformation metrics ([7, 12, 14, 26]) (we will call it the deformation metric for short).

The Lie algebra used in the deformation will be called $\mathfrak{a}$ here (it is often called $\mathfrak{k}$, but we use $\mathtt{K}$ for the stabilizer group. The letters $\mathfrak{a}, \mathfrak{b}$ introduced below correspond to the components A, B of a Stiefel tangent vector in Section 4). Let $\mathtt{A}$ be a connected subgroup of $\mathtt{G}$ with Lie algebra $\mathfrak{a}$. With the bi-invariant metric on $\mathtt{G}$, $\mathtt{A}$ acts via right multiplication as a group of isometries on $\mathtt{G}$. Give $\mathtt{G}\times\mathtt{A}$ a bi-invariant metric corresponding to the inner product on $\mathfrak{g}\oplus\mathfrak{a}$ evaluated as $\langle g, g\rangle + r\langle a, a\rangle$ for $(g, a)\in\mathfrak{g}\times\mathfrak{a}$ with $r > 0$, we have the submersion $\mathtt{G}\times\mathtt{A}\to\mathtt{G}$ given by $(U, Q)\mapsto UQ$ ($U\in\mathtt{G}$, $Q\in\mathtt{A}$). Let $\mathfrak{g} = \mathfrak{a}\oplus\mathfrak{n}$ be an orthogonal decomposition with respect to $\langle\cdot\rangle$. The submersion induces a new metric on $\mathtt{G}$ which is shown in [26] to be

$$\langle\omega_{\mathfrak{n}}, \omega_{\mathfrak{n}}\rangle + \frac{r}{(r+1)}\langle\omega_{\mathfrak{a}}, \omega_{\mathfrak{a}}\rangle$$

for $\omega\in\mathfrak{g}$. Denote the deformation metric P_t on $\mathfrak{g}$ by the formula $\langle\omega_{\mathfrak{n}}, \omega_{\mathfrak{n}}\rangle + t\langle\omega_{\mathfrak{a}}, \omega_{\mathfrak{a}}\rangle$ for $t > 0$. At $t = 1$, it is the original metric. For $t < 1$, the metric corresponds to the submersion above with $r = t/(1-t)$, thus $\mathtt{G}$ has non-negative curvature by O'Neill's equation. For $t > 1$, the metric on $\mathtt{G}\times\mathtt{A}$ is semi-Riemannian but the corresponding metric on $\mathtt{G}$ is Riemannian. If $\mathfrak{n}$ contains a subalgebra $\mathfrak{k}$ corresponding to a closed subgroup $\mathtt{K}$ of $\mathtt{G}$, such that $\mathfrak{k}$ commutes with $\mathfrak{a}$ then $\mathtt{G}/\mathtt{K}$ could be equipped with the quotient metric induced from P_t. Hence, we will consider the situation when $\mathfrak{k}$ is a subalgebra of an algebra $\mathfrak{h}$ commuting with $\mathfrak{a}$. Note that $\mathtt{G}/\mathtt{K}$ with the original bi-invariant metric is called a *normal homogeneous space.*

Proposition 2.4. *Assume the Lie algebra $\mathfrak{g}$ has a bi-invariant metric $\langle\cdot\rangle$. Let $\mathfrak{h}\subset\mathfrak{g}$ be a Lie subalgebra of $\mathfrak{g}$ and $\mathfrak{h}^{\perp}$ be the orthogonal complement of $\mathfrak{h}$ in $\mathfrak{g}$ under $\langle\cdot\rangle$, $\mathfrak{g} = \mathfrak{h}\oplus\mathfrak{h}^{\perp}$. Then $\mathfrak{b} := [\mathfrak{h}, \mathfrak{h}^{\perp}]\subset\mathfrak{h}^{\perp}$, or $\mathfrak{h}^{\perp}$ is a $\mathfrak{h}$-module. Let $\mathfrak{a}$ be the orthogonal complement of $\mathfrak{b}$ in $\mathfrak{h}^{\perp}$, i.e. $\mathfrak{h}^{\perp} = \mathfrak{b}\oplus\mathfrak{a}$ under $\langle\cdot\rangle$. We can characterize $\mathfrak{a}$ as the subspace $\{A\in\mathfrak{h}^{\perp}|[A, \mathfrak{h}] = 0\}$. Then*

$$\mathfrak{g} = \mathfrak{a}\oplus\mathfrak{b}\oplus\mathfrak{h} \tag{12}$$

We have $[\mathfrak{a}, \mathfrak{b}]\subset\mathfrak{b}$, $\mathfrak{a}$ is a Lie subalgebra of $\mathfrak{g}$, $[\mathfrak{a}, \mathfrak{h}] = 0$ and $\mathfrak{b}$ is both a $\mathfrak{h}$ and $\mathfrak{a}$ module. The correspondence $\mathfrak{h}\mapsto\mathfrak{a}$ is involutive on the set of all subalgebras of $\mathfrak{g}$, that means if we apply the same procedure on $\mathfrak{a}$, we recover $\mathfrak{h}$.

Proof. Let $X\in\mathfrak{h}^{\perp}$ and $A, H\in\mathfrak{h}$. Then $\langle[A, X], H\rangle = -\langle X, [A, H]\rangle = 0$ since $\mathfrak{h}$ is a subalgebra of $\mathfrak{g}$, thus $[A, X]\in\mathfrak{h}^{\perp}$. Let $\mathfrak{b} = [\mathfrak{h}, \mathfrak{h}^{\perp}]$ and $\mathfrak{h}^{\perp} = \mathfrak{b}\oplus\mathfrak{a}$, then for $A\in\mathfrak{a}$, $\langle[A, \mathfrak{h}], \mathfrak{h}^{\perp}\rangle\subset\langle A, [\mathfrak{h}, \mathfrak{h}^{\perp}]\rangle\subset\{0\}$ and $[A, \mathfrak{h}]\subset\mathfrak{h}^{\perp}$ as $\mathfrak{h}^{\perp}$ is a $\mathfrak{h}$-module. Hence, $[A, \mathfrak{h}] = 0$ as $\langle\cdot\rangle$ is non-degenerate on $\mathfrak{h}^{\perp}$. Conversely, if $A\in\mathfrak{h}^{\perp}$ and $[A, \mathfrak{h}] = 0$ then $\langle A, [\mathfrak{h}, \mathfrak{h}^{\perp}]\rangle\subset\langle[A, \mathfrak{h}], \mathfrak{h}^{\perp}]\rangle\subset\{0\}$, thus $A\in\mathfrak{a}$. We have proved $\mathfrak{a}$ is characterized as the subspace of $\mathfrak{h}^{\perp}$ such that $[A, \mathfrak{h}] = 0$ for $A\in\mathfrak{a}$.

Next, for $A\in\mathfrak{a}$, $\langle[A, \mathfrak{h}^{\perp}], \mathfrak{h}\rangle\subset\langle A, [\mathfrak{h}^{\perp}, \mathfrak{h}]\rangle\subset\{0\}$, thus $[A, \mathfrak{h}^{\perp}]\subset\mathfrak{h}^{\perp}$. Then

$$\langle[A, [\mathfrak{h}, \mathfrak{h}^{\perp}]], \mathfrak{a}\rangle\subset\langle[[A, \mathfrak{h}], \mathfrak{h}^{\perp}], \mathfrak{a}\rangle + \langle[\mathfrak{h}, [A, \mathfrak{h}^{\perp}]], \mathfrak{a}\rangle\subset\{0\}$$

as in the middle sum, the first item is zeros because $[A, \mathfrak{h}] = 0$, the second is $\langle[\mathfrak{h}, [A, \mathfrak{h}^{\perp}]], \mathfrak{a}\rangle\subset\langle[A, \mathfrak{h}^{\perp}], [\mathfrak{h}, \mathfrak{a}]\rangle\subset\{0\}$ as $[\mathfrak{h}, \mathfrak{a}] = \{0\}$. This shows $[\mathfrak{a}, \mathfrak{b}]$ is in the orthogonal complement of $\mathfrak{a}$ in $\mathfrak{h}^{\perp}$, or $[\mathfrak{a}, \mathfrak{b}]\subset\mathfrak{b}$.

Now, $\langle[\mathfrak{a}, \mathfrak{a}], \mathfrak{h}\rangle\subset\langle\mathfrak{a}, [\mathfrak{a}, \mathfrak{h}]\rangle\subset\{0\}$, thus $[\mathfrak{a}, \mathfrak{a}]\subset\mathfrak{h}^{\perp}$. But then we have

$$\langle[\mathfrak{a}, \mathfrak{a}], \mathfrak{b}\rangle\subset\langle\mathfrak{a}, [\mathfrak{a}, \mathfrak{b}]\rangle\subset\langle\mathfrak{a}, \mathfrak{b}\rangle\subset\{0\},$$

hence $[\mathfrak{a}, \mathfrak{a}]\subset\mathfrak{a}$, therefore $\mathfrak{a}$ is a subalgebra of $\mathfrak{g}$, and $\mathfrak{b}$ is an $\mathfrak{a}$-module.

Involutiveness follows from the orthogonal decomposition $\mathfrak{g} = \mathfrak{h} \oplus \mathfrak{b} \oplus \mathfrak{a}$, and the characterization of $\mathfrak{a}$ by the relation $[\mathfrak{a}, \mathfrak{h}] = 0$, which implies $\mathfrak{a}^{\perp} = \mathfrak{b} \oplus \mathfrak{h}$. ∎

Proposition 2.5. *Assume $\mathfrak{g}$ has a non-degenerate bi-invariant bilinear form $\langle \cdot \rangle$. Let P be a self-adjoint operator under $\langle \cdot \rangle$. Then under the bilinear form $\langle \cdot \rangle_{\mathrm{P}}$ defined by $\langle A_1, A_2 \rangle_{\mathrm{P}} := \langle A_1, \mathrm{P}\, A_2 \rangle$, we have $\mathrm{ad}^{\dagger}{}_A X = -\mathrm{P}^{-1}[A, \mathrm{P}\, X]$ for $X \in \mathfrak{g}$, or $\mathrm{ad}^{\dagger}{}_A = -\mathrm{P}^{-1} \circ \mathrm{ad}_A \circ \mathrm{P}$. Also $\overline{\mathrm{ad}}_A = \mathrm{P}^{-1} \circ \mathrm{ad}_{\mathrm{P}\, A}$. If $\langle \cdot \rangle_{\mathrm{P}}$ is positive definite this defines a left-invariant inner product on the corresponding Lie group.*
Let t be a positive number and $\mathfrak{a}, \mathfrak{b}, \mathfrak{h}$ as in Proposition 2.4. Let $\mathfrak{n} = \mathfrak{b} \oplus \mathfrak{h}$, thus $\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{n}$. Define the operator $\mathrm{P} = \mathrm{P}_t$ by $\mathrm{P}\, \omega = t\omega_{\mathfrak{a}} + \omega_{\mathfrak{n}}$. Then for $\omega_1, \omega_2 \in \mathfrak{g}$

$$(\mathrm{ad}^{\dagger}{}_{\omega_1} \omega_2)_{\mathfrak{a}} = -[\omega_{1\mathfrak{a}}, \omega_{2\mathfrak{a}}] - \frac{1}{t}[\omega_{1\mathfrak{n}}, \omega_{2\mathfrak{n}}]_{\mathfrak{a}} \tag{13}$$

$$(\mathrm{ad}^{\dagger}{}_{\omega_1} \omega_2)_{\mathfrak{n}} = -[\omega_{1\mathfrak{a}}, \omega_{2\mathfrak{b}}] + t[\omega_{2\mathfrak{a}}, \omega_{1\mathfrak{b}}] - [\omega_{\mathfrak{n},1}, \omega_{2\mathfrak{n}}]_{\mathfrak{n}} \tag{14}$$

$$[\omega_1, \omega_2]_{\mathrm{P}} = [\omega_1, \omega_2] + (1-t)([\omega_{1\mathfrak{a}}, \omega_{2\mathfrak{b}}] + [\omega_{2\mathfrak{a}}, \omega_{1\mathfrak{b}}]) \tag{15}$$

Let $\mathfrak{k} \subset \mathfrak{h}$ be a Lie subalgebra of $\mathfrak{h}$ and $\mathfrak{m} = \mathfrak{a} \oplus \mathfrak{b} \oplus \mathfrak{d}$ where $\mathfrak{h} = \mathfrak{k} \oplus \mathfrak{d}$ is an orthogonal decomposition, thus $\mathfrak{g} = \mathfrak{k} \oplus \mathfrak{m}$. For $\omega_3 \in \mathfrak{g}$

$$(\mathrm{ad}^{\dagger}{}_{\omega_3}[\omega_1, \omega_2]_{\mathfrak{k}})_{\mathfrak{m}} = -[\omega_{3\mathfrak{m}}[\omega_1, \omega_2]_{\mathfrak{k}}] \tag{16}$$

Proof. Let $A, Y, X \in \mathfrak{g}$. From $\langle [A, Y], \mathrm{P}\, X \rangle = \langle Y, -\mathrm{P}\,\mathrm{P}^{-1}[A, \mathrm{P}\, X] \rangle$, the first statement follows. For eq. (13) and eq. (14), we expand

$$\mathrm{ad}^{\dagger}{}_{\omega_1} \omega_2 = -\mathrm{P}^{-1}[\omega_{1\mathfrak{a}} + \omega_{1\mathfrak{n}}, t\omega_{2\mathfrak{a}} + \omega_{2\mathfrak{n}}]$$
$$= (-1/t)([t\omega_{1\mathfrak{a}}, \omega_{2\mathfrak{a}}] + [\omega_{1\mathfrak{n}}, \omega_{2\mathfrak{n}}]_{\mathfrak{a}}) - ([\omega_{1\mathfrak{a}}, \omega_{2\mathfrak{n}}] + [\omega_{1\mathfrak{n}}, t\omega_{2\mathfrak{a}}] + [\omega_{1\mathfrak{n}}, \omega_{2\mathfrak{n}}])_{\mathfrak{n}}$$

then use the fact that $[\mathfrak{a}, \mathfrak{h}] = \{0\}$. Equation 15 follows from this and the definition of $[\quad]_{\mathrm{P}}$, using anti-commutativity to cancel $1/t([\omega_{1\mathfrak{n}}, \omega_{2\mathfrak{n}}]_{\mathfrak{a}} + [\omega_{2\mathfrak{n}}, \omega_{1\mathfrak{n}}]_{\mathfrak{a}})$.
For eq. (16), let $\omega_4 \in \mathfrak{m}$, we have

$$\langle \mathrm{ad}^{\dagger}{}_{\omega_3}[\omega_1, \omega_2]_{\mathfrak{k}}, \omega_4 \rangle_{\mathrm{P}} = \langle [\omega_1, \omega_2]_{\mathfrak{k}}, \mathrm{P}[\omega_3, \omega_4] \rangle = \langle [\omega_1, \omega_2]_{\mathfrak{k}}, [\omega_3, \omega_4]_{\mathfrak{k}} \rangle$$

as when we expand $\mathrm{P}[\omega_3, \omega_4]$, only $[\omega_3, \omega_4]_{\mathfrak{k}}$ could be not orthogonal to $[\omega_1, \omega_2]_{\mathfrak{k}}$. From here we get $\langle [\omega_1, \omega_2]_{\mathfrak{k}}, [\omega_3, \omega_4]_{\mathfrak{k}} \rangle = \langle [\omega_1, \omega_2]_{\mathfrak{k}}, [\omega_3, \omega_4] \rangle = -\langle [\omega_3, [\omega_1, \omega_2]_{\mathfrak{k}}], \omega_4 \rangle$. But $[\omega_{3\mathfrak{k}}, [\omega_1, \omega_2]_{\mathfrak{k}}]$ is orthogonal to $\omega_4 \in \mathfrak{m}$, so we are left with

$$-\langle [\omega_{3\mathfrak{m}}, [\omega_1, \omega_2]_{\mathfrak{k}}], \omega_4 \rangle = -\langle [\omega_{3\mathfrak{m}}, [\omega_1, \omega_2]_{\mathfrak{k}}], \omega_4 \rangle_{\mathrm{P}}$$

as $[\omega_{3\mathfrak{m}}, [\omega_1, \omega_2]_{\mathfrak{k}}]_{\mathfrak{a}} = 0$, because $[\omega_{3\mathfrak{a}}, [\omega_1, \omega_2]_{\mathfrak{k}}] = 0$ while the remaining term is in $[\mathfrak{n}, \mathfrak{k}] \subset \mathfrak{n}$. By proposition 2.4 $[\omega_{3\mathfrak{m}}, [\omega_1, \omega_2]_{\mathfrak{k}}] \in \mathfrak{m}$ since $\mathfrak{m}$ is the orthogonal complement of $\mathfrak{k}$, this proves eq. (16). ∎

Recall o is the coset containing the identity in the homogeneous manifold $\mathtt{G}/\mathtt{K}$. The expression $\mathrm{R}^{[0]}$ in the following theorem is the curvature of a normal homogeneous manifold, probably not usually known in this format.

Proposition 2.6. *For a Lie group $\mathtt{G}$ with Lie algebra $\mathfrak{g}$ and a bi-invariant metric $\langle \cdot \rangle$, the curvature of the homogeneous manifold $\mathtt{M} = \mathtt{G}/\mathtt{K}$ under the metric P_t at o with $\mathfrak{k} \subset \mathfrak{h}$ are subalgebras of $\mathfrak{g}$, ($\mathfrak{g} = \mathfrak{a} \oplus \mathfrak{b} \oplus \mathfrak{h} = \mathfrak{a} \oplus \mathfrak{n} = \mathfrak{m} \oplus \mathfrak{k}$ as in Proposition 2.4) at $\omega_1, \omega_2, \omega_3 \in \mathfrak{m}$ is given by*

$$\mathrm{R}_{\omega_1,\omega_2} \omega_3 = \mathrm{R}^{[0]}_{\omega_1,\omega_2} \omega_3 + (1-t)\, \mathrm{R}^{[1]}_{\omega_1,\omega_2} \omega_3 + (1-t)^2\, \mathrm{R}^{[2]}_{\omega_1,\omega_2} \omega_3 \tag{17}$$

$$\mathrm{R}^{[0]}_{\omega_1,\omega_2}\,\omega_3 := \frac{1}{4}([[\omega_1,\omega_2],\omega_3]_{\mathfrak{m}} + 2[[\omega_1,\omega_2]_{\mathfrak{k}},\omega_3] - [[\omega_2,\omega_3]_{\mathfrak{k}},\omega_1] + [[\omega_1,\omega_3]_{\mathfrak{k}},\omega_2]) \quad (18)$$

$$\mathrm{R}^{[1]}_{\omega_1,\omega_2}\,\omega_3 := \frac{1}{2}([[\omega_1,\omega_2]_{\mathfrak{a}},\omega_{3\mathfrak{b}}] + [\omega_{3\mathfrak{a}},[\omega_1,\omega_2]_{\mathfrak{b}}]) \quad (19)$$
$$- \frac{1}{4}([\omega_1,[\omega_{2\mathfrak{a}},\omega_{3\mathfrak{b}}] + [\omega_{3\mathfrak{a}},\omega_{2\mathfrak{b}}]] + [\omega_{1\mathfrak{a}},[\omega_2,\omega_3]_{\mathfrak{b}}] + [[\omega_2,\omega_3]_{\mathfrak{a}},\omega_{1\mathfrak{b}}])_{\mathfrak{m}}$$
$$+ \frac{1}{4}([\omega_2,[\omega_{1\mathfrak{a}},\omega_{3\mathfrak{b}}] + [\omega_{3\mathfrak{a}},\omega_{1\mathfrak{b}}]] + [\omega_{2\mathfrak{a}},[\omega_1,\omega_3]_{\mathfrak{b}}] + [[\omega_1,\omega_3]_{\mathfrak{a}},\omega_{2\mathfrak{b}}])_{\mathfrak{m}}$$

$$\mathrm{R}^{[2]}_{\omega_1,\omega_2}\,\omega_3 := \frac{1}{4}(-[\omega_{1\mathfrak{a}},[\omega_{2\mathfrak{a}},\omega_{3\mathfrak{b}}] + [\omega_{3\mathfrak{a}},\omega_{2\mathfrak{b}}]] + [\omega_{2\mathfrak{a}},[\omega_{1\mathfrak{a}},\omega_{3\mathfrak{b}}] + [\omega_{3\mathfrak{a}},\omega_{1\mathfrak{b}}]]) \quad (20)$$

The Ricci curvature for $\omega \in \mathfrak{m}$ is

$$\mathrm{Ric}(\omega,\omega) = \frac{1}{4}\operatorname{Tr}\operatorname{ad}_\omega \pi_{\mathfrak{m}} \operatorname{ad}_\omega \pi_{\mathfrak{m}} - \frac{1}{2}\operatorname{Tr}\operatorname{ad}_\omega \operatorname{ad}_\omega + (1-t)\mathrm{Ric}^{[1]} + (1-t)^2\mathrm{Ric}^{[2]} \quad (21)$$

$$\mathrm{Ric}^{[1]}(\omega,\omega) := \frac{1}{2}\operatorname{Tr}(\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}} - \operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}) \quad (22)$$

$$\mathrm{Ric}^{[2]}(\omega,\omega) := -\frac{1}{4}\operatorname{Tr}(\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}) \quad (23)$$

Proof. We apply the formulas for $[\ \ ]_{\mathrm{P}}$, using the notation $[\ \ ]_{\mathrm{P},\mathfrak{a}}, [\ \ ]_{\mathrm{P},\mathfrak{n}}$ for the $\mathfrak{a}$ and $\mathfrak{n}$ components of $[\ \ ]_{\mathrm{P}}$

$$[[\omega_1,\omega_2],\omega_3]_{\mathrm{P},\mathfrak{a}} = [[\omega_1,\omega_2],\omega_3]_{\mathfrak{a}}$$
$$[[\omega_1,\omega_2],\omega_3]_{\mathrm{P},\mathfrak{n}} = [[\omega_1,\omega_2],\omega_3]_{\mathfrak{n}} + (1-t)([[\omega_1,\omega_2]_{\mathfrak{a}},\omega_{3,\mathfrak{b}}] + [\omega_{3,\mathfrak{a}},[\omega_1,\omega_2]_{\mathfrak{b}}]$$
$$[\omega_1[\omega_2,\omega_3]_{\mathrm{P}}]_{\mathrm{P}} = [\omega_1,[\omega_2,\omega_3]] + (1-t)[\omega_1,[\omega_{2,\mathfrak{a}},\omega_{3,\mathfrak{b}}] + [\omega_{3,\mathfrak{a}},\omega_{2,\mathfrak{b}}]]$$
$$+(1-t)([\omega_{1\mathfrak{a}},[\omega_2,\omega_3]_{\mathfrak{b}}] + [[\omega_2,\omega_3]_{\mathfrak{a}},\omega_{1,\mathfrak{b}}]) + (1-t)^2([\omega_{1\mathfrak{a}},[\omega_{2\mathfrak{a}},\omega_{3\mathfrak{b}}] + [\omega_{3\mathfrak{a}},\omega_{2\mathfrak{b}}]])$$

By the Jacobi identity, the $\mathrm{R}^{[0]}$ component of the first line of eq. (4) is

$$(\frac{1}{2}[[\omega_1,\omega_2],\omega_3] - \frac{1}{4}[\omega_1,[\omega_2,\omega_3]] + \frac{1}{4}[\omega_2,[\omega_1,\omega_3]])_{\mathfrak{m}} = \frac{1}{4}[[\omega_1,\omega_2],\omega_3]_{\mathfrak{m}}$$

while the second line has the O'Neill terms $\operatorname{ad}^\dagger_{\omega_i}[\omega_j,\omega_k]_{\mathfrak{k}}$ (i,j,k in a permutation of $\{1,2,3\}$) evaluated as $-[\omega_{i\mathfrak{m}}[\omega_j,\omega_k]_{\mathfrak{k}}]$ by eq. (16). Since we assume $\omega_i \in \mathfrak{m}$, this gives us the expression for $\mathrm{R}^{[0]}$. Only the first line of eq. (4) contributes to $\mathrm{R}^{[1]}$ and $\mathrm{R}^{[2]}$, and we get the eqs. (19) and (20) by permuting the indices and collecting terms. Some of the expressions, for example, $\mathrm{R}^{[2]}_{\omega_1\omega_2}\,\omega_3$ are already in $\mathfrak{m}$ so we do not need to apply the projection again.

From the expression for R, the Ricci curvature is a quadratic polynomial in t. The term where $t=1$ corresponds to a bi-invariant metric, thus in eq. (9) the first trace term is zero, and in this case $\operatorname{ad}^\dagger_\omega = -\operatorname{ad}_\omega, \overline{\operatorname{ad}}_\omega = \operatorname{ad}_\omega$, which gives us the first two terms in eq. (21). $\mathrm{Ric}^{[1]}$ is the trace of $X \mapsto \mathrm{R}^{[1]}_{\omega,X}\,\omega$. From eq. (19) it is

$$\operatorname{Tr}(X_{\mathfrak{a}} \mapsto -\frac{1}{4}[\omega_{\mathfrak{b}},[-\omega_{\mathfrak{b}},X_{\mathfrak{a}}]]_{\mathfrak{a}}) + \operatorname{Tr}(X_{\mathfrak{b}} \mapsto \frac{1}{2}([-\omega_{\mathfrak{b}},[\omega,X_{\mathfrak{b}}]_{\mathfrak{a}}] + [\omega_{\mathfrak{a}},[\omega,X_{\mathfrak{b}}]_{\mathfrak{b}}])$$
$$- \frac{1}{4}([\omega,[\omega_{\mathfrak{a}},X_{\mathfrak{b}}]] + [\omega_{\mathfrak{a}},[-\omega,X_{\mathfrak{b}}]_{\mathfrak{b}}] + [\omega_{\mathfrak{b}}[\omega,X_{\mathfrak{b}}]_{\mathfrak{a}}])_{\mathfrak{m}})$$
$$= \frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}} + \frac{1}{2}\operatorname{Tr}(-\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{a}}} - \operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}} + \operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{a}}} + \operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{b}}})\pi_{\mathfrak{b}}$$
$$- \frac{1}{4}\operatorname{Tr}(\operatorname{ad}_\omega\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}} - \operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}\operatorname{ad}_\omega + \operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_\omega)\pi_{\mathfrak{b}}$$

The first two terms of the second line cancel as we have

$$\operatorname{Tr}(\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}\operatorname{ad}_\omega\pi_{\mathfrak{b}}) = \operatorname{Tr}(\operatorname{ad}_\omega\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}) = \operatorname{Tr}(\operatorname{ad}_\omega\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}),$$

the first term of the first line and the last of the second line cancel as we have

$$\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega}\pi_{\mathfrak{b}} = \operatorname{Tr}\operatorname{ad}_{\omega}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}} = \operatorname{Tr}\operatorname{ad}_{\omega}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}} = \operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}},$$

the second term of the first line is zero as $\pi_{\mathfrak{a}}\operatorname{ad}_{\mathfrak{a}}\pi_{\mathfrak{b}} = 0$, the last term of the first line is zero as

$$\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}} = \operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}} = \operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{b}}} - \operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}} = \operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{b}}} = 0$$

since $\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}} = 0$. Therefore $\textsc{Ric}^{[1]} = \frac{1}{2}\operatorname{Tr}(-\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}} + \operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{a}}})\pi_{\mathfrak{b}}$. Then observe that $\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}} = \operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}$. The term $\textsc{Ric}^{[2]}$ is

$$\operatorname{Tr}(X \mapsto \mathrm{R}^{[2]}_{\omega,X}\,\omega) = \tfrac{1}{4}\operatorname{Tr}(-(\operatorname{ad}_{\omega_{\mathfrak{a}}}(-\operatorname{ad}_{\omega_{\mathfrak{b}}})\pi_{\mathfrak{a}} + \operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}})),$$

but the first term is zero as $\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\mathfrak{a} \subset \mathfrak{b}$. ∎

Remark 2.7. From the eqs. (21), (22), (23) we deduce

$$\begin{aligned}\textsc{Ric}(\omega_{\mathfrak{b}},\omega_{\mathfrak{b}}) &= \frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{m}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{m}} - \frac{1}{2}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\operatorname{ad}_{\omega_{\mathfrak{b}}} - \frac{1-t}{2}\operatorname{Tr}(\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}})\\ &= \textsc{Ric}^{\mathfrak{b}}(\omega_{\mathfrak{b}},\omega_{\mathfrak{b}}) + \frac{t}{2}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}\end{aligned}$$

where $\textsc{Ric}^{\mathfrak{b}}(\omega_{\mathfrak{b}},\omega_{\mathfrak{b}}) := \frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}} - \frac{1}{2}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\operatorname{ad}_{\omega_{\mathfrak{b}}}$ would be the Ricci curvature of the quotient space $\mathtt{B} := \mathtt{M}/\mathtt{A}$ if such quotient exists. This follows from

$$\frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{m}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{m}} - \frac{1}{2}\operatorname{Tr}(\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}})$$

$$= \frac{1}{4}(\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}} + 2\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}} + \operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}) - \frac{1}{2}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{a}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}$$

As the first term vanishes, the above reduces to $\frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{b}}$. Next,

$$\textsc{Ric}(\omega_{\mathfrak{b}},\omega_{\mathfrak{a}}) = \frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{m}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{m}} - \frac{1}{2}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\operatorname{ad}_{\omega_{\mathfrak{a}}} = -\frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}$$

since $[\mathfrak{a},\mathfrak{k}] = 0$, $\pi_{\mathfrak{m}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{m}} = \operatorname{ad}_{\omega_{\mathfrak{a}}}$, thus $\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\pi_{\mathfrak{m}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{m}} = \operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{b}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}$. Using the same observation and note $\pi_{\mathfrak{b}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}} = \operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}}$

$$\begin{aligned}\textsc{Ric}(\omega_{\mathfrak{a}},\omega_{\mathfrak{a}}) &= -\frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{a}}} + \frac{1-t}{2}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}} - \frac{(1-t)^2}{4}\operatorname{Tr}(\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}})\\ &= -\frac{1}{4}\operatorname{Tr}\operatorname{ad}_{\omega_{\mathfrak{a}}}\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{a}} - \frac{t^2}{4}(\operatorname{Tr}\omega_{\mathfrak{a}}\omega_{\mathfrak{a}} - \operatorname{Tr}\omega_{\mathfrak{a}}\omega_{\mathfrak{a}}\pi_{\mathfrak{a}})\end{aligned}$$

where we expand and collect terms, using $\operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{b}} = \operatorname{ad}_{\omega_{\mathfrak{a}}} - \operatorname{ad}_{\omega_{\mathfrak{a}}}\pi_{\mathfrak{a}}$. Thus, our result is consistent with proposition 10 in [14], if we replace t by t^2 in our formulas, and evaluate the expressions on $t^{-1}\omega_{\mathfrak{a}}$. We only need $\mathfrak{a}$ to be a subalgebra, and $\mathfrak{k} \subset \mathfrak{h}$, slightly more general assumptions than [14] (which reproved the O'Neill formula.)

3. Curvature formulas for embedded manifolds with metric operators

We will introduce some notations for embedded ambient structures that will be used subsequently, and a curvature formula that will be used in Appendix B to provide another curvature calculation for Stiefel manifolds.

Let $\mathcal{M} \subset \mathcal{E}$ be a differentiable embedding, where $\mathcal{E}$ is a Euclidean space with a given inner product $\langle\cdot\rangle_{\mathcal{E}}$, and $\mathcal{M}$ is a differentiable submanifold, and g is an operator-valued function from $\mathcal{M}$ to $End(\mathcal{E})$, such that g is positive-definite, then g induces a Riemannian metric on $\mathcal{M}$, where the inner product of two tangent vectors ξ, η at a point $x \in \mathcal{M}$ is defined by $\langle \xi, \mathsf{g}_x \eta \rangle_{\mathcal{E}}$. Here, each tangent space $T_x\mathcal{M}$ is identified with a subspace of $\mathcal{E}$ thanks to the embedding, so ξ, η are considered as elements of $\mathcal{E}$, while g_x denotes the evaluation of the operator g at x.

We call $(\mathcal{M}, \mathsf{g}, \mathcal{E})$ an embedded ambient structure. The embedding allows us to identify vector fields on $\mathcal{M}$ with $\mathcal{E}$-valued functions, thus we can take directional derivatives. Hence, for two vector fields $\mathtt{X}, \mathtt{Y}$ on $\mathcal{M}$, $\mathrm{D}_{\mathtt{X}}\, \mathtt{Y}$ makes sense with $\mathtt{Y}$ identified with an $\mathcal{E}$-valued function from $\mathcal{M}$. A Christoffel function is a function Γ from $\mathcal{M}$ with value in $Hom(\mathcal{E} \otimes \mathcal{E}, \mathcal{E})$, the space of $\mathcal{E}$-valued bilinear forms, such that the Levi-Civita connection on $\mathcal{M}$ is given by

$$\nabla_{\mathtt{X}} \mathtt{Y} = \mathrm{D}_{\mathtt{X}}\, \mathtt{Y} + \Gamma(\mathtt{X}, \mathtt{Y})$$

In [18] we proved the following curvature formula for three tangent vectors ξ, η, ϕ

$$\mathrm{R}^{\mathcal{M}}{}_{\xi,\eta}\, \phi = -(\mathrm{D}_\xi\, \Gamma)(\eta, \phi) + (\mathrm{D}_\eta\, \Gamma)(\xi, \phi) - \Gamma(\xi, \Gamma(\eta, \phi)) + \Gamma(\eta, \Gamma(\xi, \phi)) \tag{24}$$

where $\mathrm{D}_\xi\, \Gamma$ denotes the directional derivative of Γ, considered as an operator-valued function, in the direction ξ, for example. This formula has the same form as the local curvature formula, lemma 3.38 in [20].

4. Curvatures of Stiefel manifolds

In the following, $p < n$ are two positive integers. The Stiefel manifold $\mathrm{St}_{p,n}$ of orthogonal matrices in $\mathbb{R}^{n\times p}$ could be considered as the homogeneous space $\mathtt{G}/\mathtt{K} = \mathrm{SO}(n)/\mathrm{SO}(n-p)$. We can extend $Y \in \mathrm{St}_{p,n}$, considered as an orthogonal matrix in $\mathbb{R}^{n\times p}$ to a full orthogonal basis $(Y|Y_\perp)$ of $\mathbb{R}^n$, by adding an orthogonal complement $Y_\perp \in \mathbb{R}^{n\times(n-p)}$. Here, $(\cdot|\cdot)$ denotes the division of a matrix in $\mathbb{R}^{n\times n}$ to the first p (in $\mathbb{R}^{n\times p}$) and last $n-p$ (in $\mathbb{R}^{n\times(n-p)}$) column blocks. Thus, $Y_\perp Y_\perp^{\mathsf{T}} = \mathrm{I}_n - YY^{\mathsf{T}}$, $Y_\perp^{\mathsf{T}} Y_\perp = \mathrm{I}_{n-p}, Y^{\mathsf{T}} Y_\perp = 0, Y_\perp^{\mathsf{T}} Y = 0$. Any matrix $\omega \in \mathcal{E} = \mathbb{R}^{n\times p}$ could be represented in this basis as $\omega = YA + Y_\perp B$ with $A \in \mathbb{R}^{p\times p}$, $B \in \mathbb{R}^{p\times(n-p)}$ and ω is a tangent vector to $\mathrm{St}_{p,n}$ at Y if and only if A is antisymmetric, $A \in \mathfrak{o}(p)$, or equivalently $Y^{\mathsf{T}}\omega + \omega^{\mathsf{T}} Y = 0$.

In the notation of Section 2, we divide a matrix in $\mathfrak{g} = \mathfrak{o}(n)$ to blocks of the form $\begin{bmatrix} A & -B^{\mathsf{T}} \\ B & H \end{bmatrix}$, with $A \in \mathfrak{o}(p)$, $B \in \mathbb{R}^{(n-p)\times p}$ and $H \in \mathfrak{o}(n-p)$, represented more compactly by a triple $[\![A, B, H]\!]$. Take $\mathfrak{k} = \mathfrak{h} = \mathfrak{o}(n-p)$ to be the subalgebra generated by the H-blocks, identified with the bottom right $(n-p)\times(n-p)$ block of $\mathfrak{o}(n)$, then $\mathfrak{m}$ is the subspace of $\mathfrak{o}(n)$ where the H-block is zero, the subalgebra $\mathfrak{a}$ is $\mathfrak{o}(p)$ identified with the A-block, and $\mathfrak{b}$ is the subspace generated by the B and B^{T}-blocks, as below

$$\mathfrak{g} : [\![\, \mathfrak{a}, \mathfrak{b}, \mathfrak{h} \,]\!] \qquad \mathfrak{n} : [\![\, 0, \mathfrak{b}, \mathfrak{h} \,]\!] \qquad \mathfrak{m} : [\![\, \mathfrak{a}, \mathfrak{b}, 0 \,]\!]$$

The Lie bracket of $[\![\, A_1, B_1, H_1] \,]\!], [\![\, A_2, B_2, H_2 \,]\!] \in \mathfrak{o}(n)$ is given by

$$\begin{aligned} &[[\![\, A_1, B_1, H_1 \,]\!], [\![\, A_2, B_2, H_2 \,]\!]] \\ &= [\![[A_1, A_2] + B_2^{\mathsf{T}} B_1 - B_1^{\mathsf{T}} B_2, B_1 A_2 + H_1 B_2 - B_2 A_1 - H_2 B_1, [H_1, H_2] + B_2 B_1^{\mathsf{T}} - B_1 B_2^{\mathsf{T}} \,]\!] \end{aligned}$$

For $U = (Y|Y_\perp) \in \mathrm{SO}(n)$, if $\omega = (\eta|\eta_\perp)$ is a tangent vector at U to $\mathrm{SO}(n)$ then $\omega = d\mathcal{L}_U(U^{\mathsf{T}}\omega) = U\, [\![\, Y^{\mathsf{T}}\eta, Y_\perp^{\mathsf{T}}\eta, Y_\perp^{\mathsf{T}}\eta_\perp \,]\!]$.

We describe the submersion $\mathrm{SO}(n)\to \mathrm{St}_{p,n}$, identifying $\mathrm{St}_{p,n}$ with $\mathrm{SO}(n)/\mathrm{SO}(n-p)$ by the map $U\mapsto Y$, where $U=(Y|Y_{\perp})$ as just described. The map is clearly a differentiable submersion on to $\mathrm{St}_{p,n}$, the fiber over Y consists of matrices of the form $(Y|Y_{\perp}Q)$, $Q\in\mathrm{SO}(n-p)$, hence the vertical space consists of $(0|Y_{\perp}\mathrm{q})$, $\mathrm{q}\in\mathfrak{o}(n-p)$. Take the bi-invariant inner product on $\mathfrak{g}=\mathfrak{o}(n)$ to be $\langle\omega_1^{\mathsf{T}}\omega_2\rangle=\frac{1}{2}\mathrm{Tr}(\omega_1^{\mathsf{T}}\omega_2)$, $\omega_1,\omega_2\in\mathfrak{o}(n)$. Equip $\mathrm{SO}(n)$ with the metric P_t with

$$\|\,[\![A,B,H]\!]\,\|^2_{\mathrm{P}_t}=\tfrac{1}{2}(t\,\mathrm{Tr}\,AA^{\mathsf{T}}+2\,\mathrm{Tr}\,BB^{\mathsf{T}}+\mathrm{Tr}\,HH^{\mathsf{T}})$$

in Proposition 2.5. At $U=\mathrm{I}_n$, the horizontal space consists of matrices of the form $[\![A,B,0]\!]$, with $A\in\mathfrak{o}(p), B\in\mathbb{R}^{(n-p)\times p}$, and a horizontal vector at $U\in\mathrm{SO}(n)$ is in general of the form $U[\![A,B,0]\!]$. The submersion maps $\omega=(\eta|\eta_{\perp})\in T_U\mathrm{SO}(n)$ to $\eta\in T_Y\mathrm{St}_{p,n}\subset\mathbb{R}^{n\times p}$ satisfy $Y^{\mathsf{T}}\eta\in\mathfrak{o}(p)$. From eq. (15)

$$\begin{aligned}{}[[\![A_1,B_1,H_1]\!],[\![A_2,B_2,H_2]\!]]_{\mathrm{P}}=[\![&[A_1,A_2]+B_2^{\mathsf{T}}B_1-B_1^{\mathsf{T}}B_2,\\ tB_1A_2+H_1B_2+(t-2)B_2A_1-H_2B_1,&[H_1,H_2]+B_2B_1^{\mathsf{T}}-B_1B_2^{\mathsf{T}}]\!]\end{aligned}\tag{25}$$

In [13], the authors introduced a family of metrics on $\mathrm{St}_{p,n}$, which could be reparameterized using two positive real numbers α_0, α_1 with ratio $\alpha=\frac{\alpha_1}{\alpha_0}$ in [17]. In the convention of Section 3, we have $\mathcal{M}:=\mathrm{St}_{p,n}\subset\mathcal{E}:=\mathbb{R}^{n\times p}$, with the base inner product on $\mathcal{E}$ is the Frobenius inner product, thus $\langle\omega_1,\omega_2\rangle_{\mathcal{E}}=\mathrm{Tr}(\omega_1\omega_2^{\mathsf{T}})$ for $\omega_1,\omega_2\in\mathcal{E}$. Consider the metric operator

$$\mathsf{g}\omega=\mathsf{g}_Y\omega:=\alpha_0\omega+(\alpha_1-\alpha_0)YY^{\mathsf{T}}\omega\tag{26}$$

for $Y\in\mathrm{St}_{p,n}$, with inverse $\mathsf{g}^{-1}\omega=\alpha_0^{-1}\omega+(\alpha_1^{-1}-\alpha_0^{-1})YY^{\mathsf{T}}\omega$ and the inner product on $\mathcal{E}$ induced by g is $\langle\omega_1,\mathsf{g}_Y\omega_2\rangle_{\mathcal{E}}=\alpha_0\,\mathrm{Tr}\,\omega_1\omega_2^{\mathsf{T}}+(\alpha_1-\alpha_0)\,\mathrm{Tr}\,\omega_1^{\mathsf{T}}YY^{\mathsf{T}}\omega_2$, and this induces a Riemannian metric on $\mathrm{St}_{p,n}$.

Theorem 4.1. *With the above setting, the horizontal lift of a tangent vector $\eta\in T_Y\mathrm{St}_{p,n}\subset\mathbb{R}^{n\times p}$ to $\mathrm{SO}(n)$ at $U=(Y|Y_{\perp})\in\mathrm{SO}(n)$ under P_t is $\bar{\eta}=(\eta|-Y\eta^{\mathsf{T}}Y_{\perp})$ and the induced metric is*

$$\langle\eta,\eta\rangle_t=\mathrm{Tr}(\eta\eta^{\mathsf{T}}+(\frac{t}{2}-1)YY^{\mathsf{T}}\eta\eta^{\mathsf{T}})\tag{27}$$

identical to the metric given by eq. (26) for $\alpha_0=1$, $\alpha_1=\alpha=\frac{t}{2}$. The Levi-Civita connection for two vector fields $\mathtt{V},\mathtt{Z}$ on $\mathrm{St}_{p,n}$ under this metric is given by

$$\nabla_{\mathtt{V}}\mathtt{Z}=\mathrm{D}_{\mathtt{V}}\,\mathtt{Z}+\frac{1}{2}Y(\mathtt{V}^{\mathsf{T}}\mathtt{Z}+\mathtt{V}^{\mathsf{T}}\mathtt{Z})+\frac{2-t}{2}(\mathrm{I}_n-YY^{\mathsf{T}})(\mathtt{V}\mathtt{Z}^{\mathsf{T}}+\mathtt{Z}\mathtt{V}^{\mathsf{T}})Y\tag{28}$$

thus a Christoffel function for the induced metric with $t=2\alpha$ is

$$\Gamma(\omega_1,\omega_2)=\frac{1}{2}Y(\omega_1^{\mathsf{T}}\omega_2+\omega_2^{\mathsf{T}}\omega_1)+(1-\alpha)(\mathrm{I}_n-YY^{\mathsf{T}})(\omega_1\omega_2^{\mathsf{T}}+\omega_2\omega_1^{\mathsf{T}})Y\tag{29}$$

for $\omega_1,\omega_2\in\mathbb{R}^{n\times p}$. Representing three tangent vectors $\xi,\eta,\phi\in\mathbb{R}^{n\times p}$ at $Y\in\mathrm{St}_{p,n}$ in an orthogonal basis $(Y|Y_{\perp})$ of $\mathbb{R}^n$ as

$$\xi=YA_1+Y_{\perp}B_1,\quad \eta=YA_2+Y_{\perp}B_2,\quad \phi=YA_3+Y_{\perp}B_3,$$

where $A_1,A_2,A_3\in\mathfrak{o}(p)$ and $B_1,B_2,B_3\in\mathbb{R}^{(n-p)\times p}$. Then the Riemannian curvature tensor is ${\mathrm{R}^{\mathcal{M}}}_{\xi\eta}\,\phi=YA_R+Y_{\perp}B_R$ with $A_R\in\mathfrak{o}(p),B_R\in\mathbb{R}^{(n-p)\times p}$ where

$$A_R = Y^{\mathsf{T}}\,\mathrm{R}^{\mathcal{M}}{}_{\xi\eta}\,\phi = \frac{1-2\alpha}{4}(A_1B_3^{\mathsf{T}}B_2 - A_2B_3^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_3A_2 + B_2^{\mathsf{T}}B_3A_1) + \frac{1-\alpha}{2}(A_3B_1^{\mathsf{T}}B_2 - A_3B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2A_3 + B_2^{\mathsf{T}}B_1A_3) + \frac{1}{4}([[A_1,A_2],A_3] - A_1B_2^{\mathsf{T}}B_3 + A_2B_1^{\mathsf{T}}B_3 + B_3^{\mathsf{T}}B_1A_2 - B_3^{\mathsf{T}}B_2A_1) \tag{30}$$

$$B_R = Y_{\perp}^{\mathsf{T}}\,\mathrm{R}^{\mathcal{M}}{}_{\xi\eta}\,\phi = \frac{2\alpha^2-\alpha}{2}(B_1A_3A_2 - B_2A_3A_1) + (\alpha^2-\alpha)(B_3A_1A_2 - B_3A_2A_1) + (1-\alpha)(B_3B_1^{\mathsf{T}}B_2 - B_3B_2^{\mathsf{T}}B_1) + \frac{\alpha-2}{2}(B_1B_2^{\mathsf{T}}B_3 - B_2B_1^{\mathsf{T}}B_3) + \frac{\alpha}{2}(B_1A_2A_3 - B_1B_3^{\mathsf{T}}B_2 - B_2A_1A_3 + B_2B_3^{\mathsf{T}}B_1) \tag{31}$$

Christoffel functions for the cases $\alpha = 1$ and $\alpha = \frac{1}{2}$ are provided in [9]. A geodesic equation for the family was derived in [13] and later in [17]. Thus we provide here another derivation of the Christoffel function.

Proof. A matrix multiplication shows $U^{\mathsf{T}}\bar{\eta}$ is antisymmetric and could be represented as $[\![\, Y^{\mathsf{T}}\eta, Y_{\perp}^{\mathsf{T}}\eta, 0\,]\!] \in \mathfrak{m} \subset \mathfrak{o}(n)$, which is horizontal at I_n, thus $\bar{\eta}$ is horizontal and maps to η, hence it is the horizontal lift.
Using the relations $Y_{\perp}Y_{\perp}^{\mathsf{T}} + YY^{\mathsf{T}} = \mathrm{I}_n$ the induced metric is

$$\langle U^{\mathsf{T}}\bar{\eta}, U^{\mathsf{T}}\bar{\eta}\rangle_{\mathrm{P}} = \frac{1}{2}\,\mathrm{Tr}\begin{bmatrix} Y^{\mathsf{T}}\eta & -\eta^{\mathsf{T}}Y_{\perp} \\ Y_{\perp}^{\mathsf{T}}\eta & 0 \end{bmatrix}\begin{bmatrix} t\eta^{\mathsf{T}}Y & \eta^{\mathsf{T}}Y_{\perp} \\ -Y_{\perp}^{\mathsf{T}}\eta & 0 \end{bmatrix}$$
$$= \frac{1}{2}\,\mathrm{Tr}(tYY^{\mathsf{T}}\eta\eta^{\mathsf{T}} + 2Y_{\perp}Y_{\perp}^{\mathsf{T}}\eta\eta^{\mathsf{T}}) = \mathrm{Tr}(\eta\eta^{\mathsf{T}} + (\frac{t}{2}-1)YY^{\mathsf{T}}\eta\eta^{\mathsf{T}})$$

Let $\mathtt{V}$ and $\mathtt{Z}$ be two vector fields on the manifold $\mathrm{St}_{p,n}$, which lift to $\mathrm{SO}(n)$-vector fields $\bar{\mathtt{V}} = (\mathtt{V}| - Y\mathtt{V}^{\mathsf{T}}Y_{\perp})$ and $\bar{\mathtt{Z}} = (\mathtt{Z}| - Y\mathtt{Z}^{\mathsf{T}}Y_{\perp})$. Let $F = U^{\mathsf{T}}\bar{\mathtt{Z}} = [\![\, Y^{\mathsf{T}}\mathtt{Z}, Y_{\perp}^{\mathsf{T}}\mathtt{Z}, 0\,]\!]$, by eq. (5), $\nabla_{\mathtt{V}}\mathtt{Z}$ lifts to $UC_{\mathfrak{m}}$ with $C = \mathrm{D}_{\bar{\mathtt{V}}}F + \frac{1}{2}[[\![\, Y^{\mathsf{T}}\mathtt{V}, Y_{\perp}^{\mathsf{T}}\mathtt{V}, 0\,]\!], [\![\, Y^{\mathsf{T}}\mathtt{Z}, Y_{\perp}^{\mathsf{T}}\mathtt{Z}, 0\,]\!]]_{\mathrm{P}}$. Expand the Lie-derivative and the P-bracket using eq. (25)

$$C = [\![\, \mathtt{V}^{\mathsf{T}}\mathtt{Z} + Y^{\mathsf{T}}\,\mathrm{D}_{\mathtt{V}}\mathtt{Z},\ -Y_{\perp}^{\mathsf{T}}\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} + Y_{\perp}^{\mathsf{T}}\,\mathrm{D}_{\mathtt{V}}\mathtt{Z},\ 0\,]\!]$$
$$+\frac{1}{2}\,[\![[Y^{\mathsf{T}}\mathtt{V}, Y^{\mathsf{T}}\mathtt{Z}] + \mathtt{Z}^{\mathsf{T}}Y_{\perp}Y_{\perp}^{\mathsf{T}}\mathtt{V} - \mathtt{V}^{\mathsf{T}}Y_{\perp}Y_{\perp}^{\mathsf{T}}\mathtt{Z},\ tY_{\perp}^{\mathsf{T}}\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} + (t-2)Y_{\perp}^{\mathsf{T}}\mathtt{Z}Y^{\mathsf{T}}\mathtt{V},\ C_H\,]\!]$$

for $C_H \in \mathfrak{o}(n-p)$. Thus, the submersion maps $UC_{\mathfrak{m}}$ to its left p columns

$$Y(\mathtt{V}^{\mathsf{T}}\mathtt{Z} + Y^{\mathsf{T}}\,\mathrm{D}_{\mathtt{V}}\mathtt{Z} + \frac{1}{2}([Y^{\mathsf{T}}\mathtt{V}, Y^{\mathsf{T}}\mathtt{Z}] + \mathtt{Z}^{\mathsf{T}}Y_{\perp}Y_{\perp}^{\mathsf{T}}\mathtt{V} - \mathtt{V}^{\mathsf{T}}Y_{\perp}Y_{\perp}^{\mathsf{T}}\mathtt{Z}))$$
$$+ Y_{\perp}(-Y_{\perp}^{\mathsf{T}}\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} + Y_{\perp}^{\mathsf{T}}\,\mathrm{D}_{\mathtt{V}}\mathtt{Z} + \frac{1}{2}(tY_{\perp}^{\mathsf{T}}\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} + (t-2)Y_{\perp}^{\mathsf{T}}\mathtt{Z}Y^{\mathsf{T}}\mathtt{V}))$$
$$= \mathrm{D}_{\mathtt{V}}\mathtt{Z} + Y\mathtt{V}^{\mathsf{T}}\mathtt{Z} + \frac{1}{2}(YY^{\mathsf{T}}\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} - YY^{\mathsf{T}}\mathtt{Z}Y^{\mathsf{T}}\mathtt{V} + Y\mathtt{Z}^{\mathsf{T}}Y_{\perp}Y_{\perp}\mathtt{V} - Y\mathtt{V}^{\mathsf{T}}Y_{\perp}Y_{\perp}\mathtt{Z})$$
$$+ \frac{1}{2}Y_{\perp}Y_{\perp}^{\mathsf{T}}(-2\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} + t\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} + (t-2)\mathtt{Z}Y^{\mathsf{T}}\mathtt{V})$$

The last line simplifies to

$$\frac{t-2}{2}\,(\mathrm{I}_n - YY^{\mathsf{T}})(\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} + \mathtt{Z}Y^{\mathsf{T}}\mathtt{V}) = \frac{2-t}{2}(\mathrm{I}_n - YY^{\mathsf{T}})(\mathtt{V}\mathtt{Z}^{\mathsf{T}} + \mathtt{Z}\mathtt{V}^{\mathsf{T}})Y$$

while twice the remaining terms, except for $\mathrm{D}_{\mathtt{V}}\mathtt{Z}$ is

$$2Y\mathtt{V}^{\mathsf{T}}\mathtt{Z} + YY^{\mathsf{T}}\mathtt{V}Y^{\mathsf{T}}\mathtt{Z} - YY^{\mathsf{T}}\mathtt{Z}Y^{\mathsf{T}}\mathtt{V} + Y\mathtt{Z}^{\mathsf{T}}(\mathrm{I}_n - YY^{\mathsf{T}})\mathtt{V} - Y\mathtt{V}^{\mathsf{T}}(\mathrm{I}_n - YY^{\mathsf{T}})\mathtt{Z}$$
$$= Y\mathtt{V}^{\mathsf{T}}\mathtt{Z} + Y\mathtt{Z}^{\mathsf{T}}\mathtt{V} + Y(Y^{\mathsf{T}}\mathtt{V} + \mathtt{V}^{\mathsf{T}}Y)Y^{\mathsf{T}}\mathtt{Z} - Y(Y^{\mathsf{T}}\mathtt{Z} + \mathtt{Z}^{\mathsf{T}}Y)Y^{\mathsf{T}}\mathtt{V} = Y\mathtt{V}^{\mathsf{T}}\mathtt{Z} + Y\mathtt{Z}^{\mathsf{T}}\mathtt{V}.$$

Thus we have proved eq. (28).

Let us proceed to a proof of the curvature expressions. To show $f(t) = g(t/2)$ with $f(t) = f_0 + (1-t)f_1 + (1-t)^2 f_2$, where f_1, f_2, f_3 are constant matrices and g is a matrix-valued quadratic function in t, we need to show $f_0 = g(1/2)$, $-2f_1 = g'(1/2)$ and $8f_2 = g''(1/2)$. From left invariance we can take $U = \mathrm{I}_n$. Thus, we need to compute $\mathrm{R}^{[0]}, \mathrm{R}^{[1]}, \mathrm{R}^{[2]}$ and compare with values and derivatives of $g(\alpha) = [\![A_R(\alpha), B_R(\alpha), 0]\!]$ with A_R, B_R defined from eq. (30) and (31) evaluated at $\alpha = 1/2$.

Let $\xi = \omega_1, \eta = \omega_2, \phi = \omega_3$ with $\omega_i = [\![A_i, B_i, 0]\!]$ we have $[\omega_{2\mathfrak{a}}, \omega_{3\mathfrak{b}}]$ is $[\![0, -B_3A_2, 0]\!]$, $[\omega_{1\mathfrak{a}}, [\omega_{2\mathfrak{a}}, \omega_{3\mathfrak{b}}]] = [\![0, B_3A_2A_1, 0]\!]$ and permuting the indices

$$4\,\mathrm{R}^{[2]}_{\omega_1,\omega_2}\omega_3 = [\![0, -B_3A_2A_1 - B_2A_3A_1 + B_3A_1A_2 + B_1A_3A_2, 0]\!]$$

On the other hand, eq. (30) and 31 gives $A''_{R,\alpha=1/2} = 0$ and $B''_{R,\alpha=1/2}$ is

$$B''_{R,\alpha=1/2} = \frac{4}{2}(B_1A_3A_2 - B_2A_3A_1) + (2)(B_3A_1A_2 - B_3A_2A_1)$$

which confirms $8\,\mathrm{R}^{[2]}_{\omega_1,\omega_2}\omega_3 = g''(1/2)$. Next,

$$[[\omega_1,\omega_2]_{\mathfrak{a}}, \omega_{3\mathfrak{b}}] = [\![0, -B_3(([A_1,A_2] + B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2), 0]\!]$$

$$[\omega_{3\mathfrak{a}}, [\omega_1,\omega_2]_{\mathfrak{b}}]_{\mathfrak{a}} = [\![0, -(B_1A_2 - B_2A_1)A_3, 0]\!]$$

$$[\omega_1, [\omega_{2\mathfrak{a}}, \omega_{3\mathfrak{b}}]]_{\mathfrak{m}} = [[\![A_1, B_1, 0]\!], [\![0, -B_3A_2, 0]\!]]_{\mathfrak{m}} = [\![A_2B_3^{\mathsf{T}}B_1 + B_1^{\mathsf{T}}B_3A_2, B_3A_2A_1, 0]\!]$$

By permuting indices we evaluate the $\mathfrak{a}$ component of $4\,\mathrm{R}^{[1]}_{\omega_1,\omega_2}\omega_3$ from four expressions similar to $[\omega_1, [\omega_{2\mathfrak{a}}, \omega_{3\mathfrak{b}}]]_{\mathfrak{a}}$ as

$$\begin{gathered}-A_2B_3^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_3A_2 - A_3B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2A_3\\ +A_1B_3^{\mathsf{T}}B_2 + B_2^{\mathsf{T}}B_3A_1 + A_1B_2^{\mathsf{T}}B_3 + B_3^{\mathsf{T}}B_2A_1\end{gathered}$$

and evaluate the $\mathfrak{b}$ component of $4\,\mathrm{R}^{[1]}_{\omega_1,\omega_2}\omega_3$ from the remaining items as

$$\begin{gathered}2(-B_3([A_1,A_2] + B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2) - (B_1A_2 - B_2A_1)A_3)\\ -B_3A_2A_1 - B_2A_3A_1 + (B_2A_3 - B_3A_2)A_1 + B_1([A_2,A_3] + B_3^{\mathsf{T}}B_2 - B_2^{\mathsf{T}}B_3)\\ +B_3A_1A_2 + B_1A_3A_2 - (B_1A_3 - B_3A_1)A_2 - B_2([A_1,A_3] + B_3^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_3)\end{gathered}$$

Let us collect terms. Terms starting with B_3 and two A factors are

$$-B_3[A_1,A_2] - B_3A_2A_1 - B_3A_2A_1 + B_3A_1A_2 + B_3A_1A_2 = 0$$

Terms starting with B_2 and two A factors:

$$2B_2A_1A_3 - B_2A_3A_1 + B_2A_3A_1 - B_2[A_1,A_3] = B_2A_1A_3 + B_2A_3A_1$$

Terms starting with B_1 and two A factors:

$$-2B_1A_2A_3 + B_1[A_2,A_3] + B_1A_3A_2 - B_1A_3A_2 = -B_1A_2A_3 - B_1A_3A_2$$

Terms with B's only factors

$$-2B_3(B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2) + B_1(B_3^{\mathsf{T}}B_2 - B_2^{\mathsf{T}}B_3) - B_2(B_3^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_3)$$

On the other hand, we have

$$A'_{R,\alpha=1/2} = \frac{-2}{4}(A_1B_3^{\mathsf{T}}B_2 - A_2B_3^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_3A_2 + B_2^{\mathsf{T}}B_3A_1) + \frac{-1}{2}(A_3B_1^{\mathsf{T}}B_2 - A_3B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2A_3 + B_2^{\mathsf{T}}B_1A_3)$$

$$\begin{aligned}B'_{R,\alpha=1/2} = &\frac{4(1/2)-1}{2}(B_1A_3A_2 - B_2A_3A_1)\\ &+ (2(1/2)-1)(B_3A_1A_2 - B_3A_2A_1) - (B_3B_1^{\mathsf{T}}B_2 - B_3B_2^{\mathsf{T}}B_1)\\ &+ \frac{1}{2}(B_1B_2^{\mathsf{T}}B_3 - B_2B_1^{\mathsf{T}}B_3) + \frac{1}{2}(B_1A_2A_3 - B_1B_3^{\mathsf{T}}B_2 - B_2A_1A_3 + B_2B_3^{\mathsf{T}}B_1)\end{aligned}$$

and we can confirm by inspection $-2\,\mathrm{R}^{[1]}_{\omega_1,\omega_2}\omega_3 = g'(1/2)$. The constant term $\mathrm{R}^{[0]}$ is verified similarly, at $\alpha = \frac{1}{2}$, A_R and B_R reduces to

$$\begin{aligned}A_{R,\alpha=1/2} = &\frac{1}{4}(A_3B_1^{\mathsf{T}}B_2 - A_3B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2A_3 + B_2^{\mathsf{T}}B_1A_3)\\ &+ \frac{1}{4}([[A_1,A_2],A_3] - A_1B_2^{\mathsf{T}}B_3 + A_2B_1^{\mathsf{T}}B_3 + B_3^{\mathsf{T}}B_1A_2 - B_3^{\mathsf{T}}B_2A_1)\end{aligned}$$

$$\begin{aligned}B_{R,\alpha=1/2} = &-\frac{1}{4}(B_3A_1A_2 - B_3A_2A_1) + \frac{1}{2}(B_3B_1^{\mathsf{T}}B_2 - B_3B_2^{\mathsf{T}}B_1)\\ &- \frac{3}{4}(B_1B_2^{\mathsf{T}}B_3 - B_2B_1^{\mathsf{T}}B_3) + \frac{1}{4}(B_1A_2A_3 - B_1B_3^{\mathsf{T}}B_2 - B_2A_1A_3 + B_2B_3^{\mathsf{T}}B_1)\end{aligned}$$

We note that $[[\omega_1,\omega_2],\omega_3]_{\mathfrak{m}}$ is

$$\begin{aligned}&[\![[[A_1,A_2] + B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2, A_3] + B_3^{\mathsf{T}}(B_1A_2 - B_2A_1) - (B_1A_2 - B_2A_1)^{\mathsf{T}}B_3,\\ &(B_1A_2 - B_2A_1)A_3 + (B_2B_1^{\mathsf{T}} - B_1B_2^{\mathsf{T}})B_3 - B_3([A_1,A_2] + B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2), 0]\!]\end{aligned}$$

The first line verifies the formula for A_R, while B_R follows from the second line and permutations of $[[\omega_1,\omega_2]_{\mathfrak{k}},\omega_3] = [\![0, (B_2B_1^{\mathsf{T}} - B_1B_2^{\mathsf{T}})B_3, 0]\!]$. ∎

For two tangent vectors ξ and η at a point on the manifold, denote by $\langle\cdot\rangle_{\mathsf{g}}$ and $\|\cdot\|_{\mathsf{g}}$ the inner product and the norm defined by a metric operator g. We will denote the *wedge*, the *sectional curvature numerator*, and the *sectional curvature* by

$$||\xi\wedge\eta||^2_{\mathsf{g}} = ||\xi||^2_{\mathsf{g}}||\eta||^2_{\mathsf{g}} - \langle\xi,\eta\rangle^2_{\mathsf{g}},\quad \hat{\mathcal{K}}(\xi,\eta) = \langle \mathrm{R}^{\mathcal{M}}{}_{\xi,\eta}\,\xi,\eta\rangle_{\mathsf{g}},\quad \mathcal{K}(\xi,\eta) = \frac{\hat{\mathcal{K}}(\xi,\eta)}{||\xi\wedge\eta||^2_{\mathsf{g}}}. \tag{32}$$

Proposition 4.2. *With notations as in Theorem* 4.1 *the sectional curvature numerator* $\hat{\mathcal{K}}$ *is computed from one of the following*

$$\begin{aligned}\hat{\mathcal{K}} = &\operatorname{Tr}\left(\frac{2-3\alpha}{2}B_2^{\mathsf{T}}B_1B_1^{\mathsf{T}}B_2 + \frac{3\alpha-4}{2}B_2^{\mathsf{T}}B_1B_2^{\mathsf{T}}B_1 + B_2^{\mathsf{T}}B_2B_1^{\mathsf{T}}B_1 - \frac{\alpha}{4}[A_1,A_2]^2\right)\\ &+ \alpha\operatorname{Tr}((4\alpha-3)A_1A_2B_2^{\mathsf{T}}B_1 + (3-2\alpha)A_1A_2B_1^{\mathsf{T}}B_2 - \alpha A_2^2B_1^{\mathsf{T}}B_1 - \alpha A_1^2B_2^{\mathsf{T}}B_2)\end{aligned}\tag{33}$$

$$\begin{aligned}\hat{\mathcal{K}} = &\frac{\alpha}{4}\|[A_1,A_2] + (3-4\alpha)(B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2)\|_F^2\\ &+ \alpha^2\|B_1A_2 - B_2A_1\|_F^2 + \frac{1}{2}\|B_1B_2^{\mathsf{T}} - B_2B_1^{\mathsf{T}}\|_F^2 + \frac{(1-2\alpha)^3}{2}\|B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2\|_F^2\end{aligned}\tag{34}$$

In particular, if $\alpha \leq \frac{1}{2}$, *the sectional curvature is non-negative. If* ξ *and* η *are orthogonal, the sectional curvature denominator is*

$$(\alpha_1\operatorname{Tr}A_1A_1^{\mathsf{T}} + \alpha_0\operatorname{Tr}B_1B_1^{\mathsf{T}})(\alpha_1\operatorname{Tr}A_2A_2^{\mathsf{T}} + \alpha_0\operatorname{Tr}B_2B_2^{\mathsf{T}}).$$

The following expansion of eq. (34) is convenient when A_1 or A_2 is zero:

$$\begin{aligned}\hat{\mathcal{K}} = \frac{\alpha}{4}\|[A_1, A_2]\|_F^2 + \frac{\alpha(3-4\alpha)}{2}\operatorname{Tr}[A_1, A_2](B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2)^{\mathsf{T}} \\ +\frac{2-3\alpha}{4}\|B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2\|_F^2 + \alpha^2\|B_1A_2 - B_2A_1\|_F^2 + \frac{1}{2}\|B_1B_2^{\mathsf{T}} - B_2B_1^{\mathsf{T}}\|_F^2\end{aligned} \tag{35}$$

Proof. We substitute A_1, B_1 in place of A_3, B_3 in the expressions for A_R and B_R, then compute $\operatorname{Tr}(-\alpha A_RA_2 + B_RB_2^{\mathsf{T}})$

$$\begin{aligned}\hat{\mathcal{K}}(\xi, \eta) = \operatorname{Tr}(-\alpha\,(&\frac{1-2\alpha}{4}(A_1B_1^{\mathsf{T}}B_2 - A_2B_1^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_1A_2 + B_2^{\mathsf{T}}B_1A_1)\\
&+ \frac{1-\alpha}{2}(A_1B_1^{\mathsf{T}}B_2 - A_1B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2A_1 + B_2^{\mathsf{T}}B_1A_1)\\
&+ \frac{1}{4}([[A_1, A_2], A_1] - A_1B_2^{\mathsf{T}}B_1 + A_2B_1^{\mathsf{T}}B_1 + B_1^{\mathsf{T}}B_1A_2 - B_1^{\mathsf{T}}B_2A_1))A_2)\\
&+ \operatorname{Tr}((\frac{2\alpha^2-\alpha}{2}(B_1A_1A_2 - B_2A_1A_1) + (\alpha^2-\alpha)(B_1A_1A_2 - B_1A_2A_1)\\
&+ (1-\alpha)(B_1B_1^{\mathsf{T}}B_2 - B_1B_2^{\mathsf{T}}B_1) + \frac{\alpha-2}{2}(B_1B_2^{\mathsf{T}}B_1 - B_2B_1^{\mathsf{T}}B_1)\\
&+ \frac{\alpha}{2}(B_1A_2A_1 - B_1B_1^{\mathsf{T}}B_2 - B_2A_1A_1 + B_2B_1^{\mathsf{T}}B_1))B_2^{\mathsf{T}}).\end{aligned}$$

We collect terms. From $-\operatorname{Tr}([[A_1, A_2]A_1]A_2) = \operatorname{Tr}[A_1, A_2][A_1, A_2]^{\mathsf{T}}$, terms involving A_1, A_2 only are $\alpha/4\operatorname{Tr}[A_1, A_2][A_1, A_2]^{\mathsf{T}}$. Terms with both A's and B's:

$$\begin{aligned}&\operatorname{Tr}(\alpha\,(\frac{1-2\alpha}{4}(-A_1B_1^{\mathsf{T}}B_2A_2 + A_2B_1^{\mathsf{T}}B_1A_2 + B_1^{\mathsf{T}}B_1A_2^2 - B_2^{\mathsf{T}}B_1A_1A_2)\\
&\quad + \frac{1-\alpha}{2}(-A_1B_1^{\mathsf{T}}B_2A_2 + A_1B_2^{\mathsf{T}}B_1A_2 + B_1^{\mathsf{T}}B_2A_1A_2 - B_2^{\mathsf{T}}B_1A_1A_2)\\
&\quad + \frac{1}{4}(A_1B_2^{\mathsf{T}}B_1A_2 - A_2B_1^{\mathsf{T}}B_1A_2 - B_1^{\mathsf{T}}B_1A_2^2 + B_1^{\mathsf{T}}B_2A_1A_2)))\\
&\quad + \alpha\operatorname{Tr}(\frac{2\alpha-1}{2}(B_1A_1A_2B_2^{\mathsf{T}} - B_2A_1A_1B_2^{\mathsf{T}}) + (\alpha-1)(B_1A_1A_2B_2^{\mathsf{T}} - B_1A_2A_1B_2^{\mathsf{T}})\\
&\quad + \frac{1}{2}(B_1A_2A_1B_2^{\mathsf{T}} - B_2A_1A_1B_2^{\mathsf{T}}))\\
&= \alpha\operatorname{Tr}((\frac{1-\alpha}{2} + \frac{1}{4} - (\alpha-1) + \frac{1}{2})A_2A_1B_2^{\mathsf{T}}B_1 + (-\frac{1-2\alpha}{4} - \frac{1-\alpha}{2})A_2A_1B_1^{\mathsf{T}}B_2\\
&\quad + (-\frac{1-2\alpha}{4} - \frac{1-\alpha}{2} + \alpha - 1 + \frac{2\alpha-1}{2})A_1A_2B_2^{\mathsf{T}}B_1\\
&\quad + (\frac{1-\alpha}{2} + \frac{1}{4})A_1A_2B_1^{\mathsf{T}}B_2 + (2\frac{1-2\alpha}{4} - \frac{2}{4})A_2^2B_1^{\mathsf{T}}B_1 + (-\frac{2\alpha-1}{2} - \frac{1}{2})A_1^2B_2^{\mathsf{T}}B_2)\\
&= \alpha\operatorname{Tr}(\frac{9-6\alpha}{4}A_2A_1B_2^{\mathsf{T}}B_1 + \frac{4\alpha-3}{4}A_2A_1B_1^{\mathsf{T}}B_2 + \frac{12\alpha-9}{4}A_1A_2B_2^{\mathsf{T}}B_1\\
&\quad + \frac{3-2\alpha}{4}A_1A_2B_1^{\mathsf{T}}B_2 - \alpha A_2^2B_1^{\mathsf{T}}B_1 - \alpha A_1^2B_2^{\mathsf{T}}B_2)\\
&= \alpha\operatorname{Tr}((4\alpha-3)A_1A_2B_2^{\mathsf{T}}B_1 + (3-2\alpha)A_1A_2B_1^{\mathsf{T}}B_2 - \alpha A_2^2B_1^{\mathsf{T}}B_1 - \alpha A_1^2B_2^{\mathsf{T}}B_2)\end{aligned}$$

where we use $\operatorname{Tr}(A_2A_1B_2^{\mathsf{T}}B_1) = \operatorname{Tr}((A_2A_1B_2^{\mathsf{T}}B_1)^{\mathsf{T}}) = \operatorname{Tr}(A_1A_2B_1^{\mathsf{T}}B_2)$ and similarly $\operatorname{Tr}(A_2A_1B_1^{\mathsf{T}}B_2) = \operatorname{Tr}(A_1A_2B_2^{\mathsf{T}}B_1)$.

Next, we collect the terms with B_1 and B_2 only:

$$\begin{aligned}
&\operatorname{Tr}((1-\alpha)(B_1B_1^{\mathsf T}B_2B_2^{\mathsf T} - B_1B_2^{\mathsf T}B_1B_2^{\mathsf T}) + \frac{\alpha-2}{2}(B_1B_2^{\mathsf T}B_1B_2^{\mathsf T} - B_2B_1^{\mathsf T}B_1B_2^{\mathsf T})\\
&\quad + \frac{\alpha}{2}(-B_1B_1^{\mathsf T}B_2B_2^{\mathsf T} + B_2B_1^{\mathsf T}B_1B_2^{\mathsf T}))\\
&= \operatorname{Tr}((1-\frac{3\alpha}{2})B_1B_1^{\mathsf T}B_2B_2^{\mathsf T} + (\alpha-1+\frac{\alpha-2}{2})B_1B_2^{\mathsf T}B_1B_2^{\mathsf T} + (-\frac{\alpha-2}{2}+\frac{\alpha}{2})B_2B_1^{\mathsf T}B_1B_2^{\mathsf T})\\
&= \operatorname{Tr}(\frac{2-3\alpha}{2}B_1B_1^{\mathsf T}B_2B_2^{\mathsf T} + \frac{3\alpha-4}{2}B_1B_2^{\mathsf T}B_1B_2^{\mathsf T} + B_2B_1^{\mathsf T}B_1B_2^{\mathsf T})
\end{aligned}$$

This proves eq. (33). On the other hand, it is clear on the right-hand side of eq. (34), the A's only term is $\frac{\alpha}{4}\operatorname{Tr}[A_1,A_2][A_1,A_2]^{\mathsf T}$, the B's only term is:

$$\begin{aligned}
&(\frac{\alpha(3-4\alpha)^2}{4} + \frac{(1-2\alpha)^3}{2})\operatorname{Tr}(B_2^{\mathsf T}B_1 - B_1^{\mathsf T}B_2)(B_2^{\mathsf T}B_1 - B_1^{\mathsf T}B_2)^{\mathsf T}\\
&\quad + \frac{1}{2}\operatorname{Tr}(B_1B_2^{\mathsf T} - B_2B_1^{\mathsf T})(B_1B_2^{\mathsf T} - B_2B_1^{\mathsf T})^{\mathsf T}\\
&= \frac{2-3\alpha}{4}\operatorname{Tr}(B_2^{\mathsf T}B_1B_1^{\mathsf T}B_2 - B_2^{\mathsf T}B_1B_2^{\mathsf T}B_1 - B_1^{\mathsf T}B_2B_1^{\mathsf T}B_2 + B_1^{\mathsf T}B_2B_2^{\mathsf T}B_1)\\
&\quad + \frac{1}{2}\operatorname{Tr}(B_1B_2^{\mathsf T}B_2B_1^{\mathsf T} - B_1B_2^{\mathsf T}B_1B_2^{\mathsf T} - B_2B_1^{\mathsf T}B_2B_1^{\mathsf T} + B_2B_1^{\mathsf T}B_1B_2^{\mathsf T})\\
&= \operatorname{Tr}(2\frac{2-3\alpha}{4}B_1B_1^{\mathsf T}B_2B_2^{\mathsf T} + (-2\frac{2-3\alpha}{4} - 2\frac{1}{2})B_1B_2^{\mathsf T}B_1B_2^{\mathsf T} + 2\frac{1}{2}B_2B_1^{\mathsf T}B_1B_2^{\mathsf T})\\
&= \operatorname{Tr}(\frac{2-3\alpha}{2}B_1B_1^{\mathsf T}B_2B_2^{\mathsf T} + \frac{3\alpha-4}{2}B_1B_2^{\mathsf T}B_1B_2^{\mathsf T} + B_2B_1^{\mathsf T}B_1B_2^{\mathsf T}).
\end{aligned}$$

The terms with both A and B in eq. (34) are:

$$\begin{aligned}
&\alpha\operatorname{Tr}(\frac{3-4\alpha}{2}(A_2A_1 - A_1A_2)(B_2^{\mathsf T}B_1 - B_1^{\mathsf T}B_2) + \alpha(B_1A_2 - B_2A_1)^{\mathsf T}(B_1A_2 - B_2A_1))\\
&= \alpha\operatorname{Tr}\{(\frac{3-4\alpha}{2}+\alpha)A_2A_1B_2^{\mathsf T}B_1 - \frac{3-4\alpha}{2}A_2A_1B_1^{\mathsf T}B_2 - \frac{3-4\alpha}{2}A_1A_2B_2^{\mathsf T}B_1\\
&\quad + (\frac{3-4\alpha}{2}+\alpha)A_1A_2B_1^{\mathsf T}B_2 - \alpha A_2^2B_1^{\mathsf T}B_1 - \alpha A_1^2B_2^{\mathsf T}B_2\}\\
&= \alpha\operatorname{Tr}\{\frac{3-2\alpha}{2}A_2A_1B_2^{\mathsf T}B_1 - \frac{3-4\alpha}{2}A_2A_1B_1^{\mathsf T}B_2 - \frac{3-4\alpha}{2}A_1A_2B_2^{\mathsf T}B_1\\
&\quad + \frac{3-2\alpha}{2}A_1A_2B_1^{\mathsf T}B_2 - \alpha A_2^2B_1^{\mathsf T}B_1 - \alpha A_1^2B_2^{\mathsf T}B_2\}\\
&= \alpha\operatorname{Tr}\{(4\alpha-3)A_1A_2B_2^{\mathsf T}B_1 + (3-2\alpha)A_1A_2B_1^{\mathsf T}B_2 - \alpha A_2^2B_1^{\mathsf T}B_1 - \alpha A_1^2B_2^{\mathsf T}B_2\}.
\end{aligned}$$

Therefore, eq. (34) gives us the sectional curvature numerator. For the sign of the sectional curvature, in eq. (34) the terms are all positive, except for the last, which is non-negative if $\alpha \leq \frac{1}{2}$. The formula for the curvature denominator is clear. ∎

Remark 4.3. We have shown the metric in eq. (4) is P_t for $t = \frac{\alpha}{2}$. The submersion associated with the Cheeger deformation gives a sectional curvature formula for $\mathtt{G}$ with the metric P_t in proposition 2.4 of [12]. Using the O'Neill equation and eq. (16), it implies the following sectional curvature formula for $\mathtt{M} = \mathtt{G}/\mathtt{K}$ (the norm $\|\cdot\|$ corresponds to the bi-invariant inner product $\langle\cdot\rangle$)

$$\begin{aligned}
\langle \mathrm{R}^{\mathtt{M}}_{\omega_1,\omega_2}\omega_1, \mathrm{P}_t\,\omega_2\rangle &= \frac{1}{4}\|[\omega_{1\mathfrak{n}},\omega_{2\mathfrak{n}}]_{\mathfrak{n}} + t[\omega_{1\mathfrak{a}},\omega_{2\mathfrak{n}}] + t[\omega_{1\mathfrak{n}},\omega_{2\mathfrak{a}}]\|^2\\
&+ \frac{1}{4}\|[\omega_{1\mathfrak{n}},\omega_{2\mathfrak{n}}]_{\mathfrak{a}} + t^2[\omega_{1\mathfrak{a}},\omega_{2\mathfrak{a}}]\|^2 + \frac{1}{4}t(1-t)^3\|[\omega_{1\mathfrak{a}},\omega_{2\mathfrak{a}}]\|^2\\
&+ \frac{3}{4}(1-t)\|[\omega_{1\mathfrak{n}},\omega_{2\mathfrak{n}}]_{\mathfrak{a}} + t[\omega_{1\mathfrak{a}},\omega_{2\mathfrak{a}}]\|^2 + \frac{3}{4}\|[\omega_1,\omega_2]_{\mathfrak{k}}\|^2
\end{aligned} \tag{36}$$

It is a weighted sum of squares in a different format from eq. (34). It implies both the non-negativity of curvature when $t \leq 1$ and in the case $\mathfrak{a}$ is abelian, when $t \leq 4/3$.

5. Ricci curvature for diagonal homogeneous metrics

Recall an Einstein manifold is a Riemannian manifold where the Ricci curvature tensor is proportional to the metric tensor. The Einstein equation for a homogeneous space is $\mathrm{Ric}(A, A) = \lambda\langle A, A\rangle_{\mathrm{P}}$ for $A \in \mathfrak{m}$. Jensen [14] has shown we could solve for the parameter t of the deformation metrics to find Einstein manifolds. To find new Einstein metrics, we need to go beyond one parameter. From eq. (9), we expect the Ricci curvature to have a simpler form when the metric is *diagonal*, by which we mean the four trace forms could be diagonalized simultaneously. We demonstrate this for the case of $\mathrm{SO}(p)$ and Stiefel manifolds (the general construction will likely involve a root system). Let E_{ij} $1 \leq i \leq p$ be the elementary matrix in $\mathbb{R}^{p\times p}$ with the (i, j) entry is 1, and other entries 0. An element A of $\mathfrak{o}(p)$ could be written as $A = \sum_{1\leq i<j\leq p} a_{ij}(E_{ij} - E_{ji})$. The standard bi-invariant inner product is $\sum_{i<j} a_{ij}^2$.

Proposition 5.1. *Let $\{t_{ij} | 1 \leq i < j \leq p\}$ be a set of positive numbers. Consider the metric $\langle A, A\rangle_{\mathrm{P}} = \sum_{1\leq i<j\leq p} t_{ij}a_{ij}^2$ on $\mathrm{SO}(p)$, where $A = \sum_{1\leq i<j\leq p} a_{ij}(E_{ij} - E_{ji}) \in \mathfrak{o}(p)$. Its Ricci curvature is given by $\mathrm{Ric}(A, A) = \sum_{1\leq i<j\leq p} r_{ij}a_{ij}^2$ with*

$$r_{ij} = p - 2 + \frac{1}{2}\sum_{1\leq l\leq p, l\neq i, l\neq j} \frac{t_{ij}^2}{t_{il}t_{jl}} - \frac{t_{il}}{t_{jl}} - \frac{t_{jl}}{t_{il}} \tag{37}$$

Thus, the Einstein equation in this case is the system $r_{ij} = \lambda t_{ij}$ for $1 \leq i < j \leq p$ for a constant λ. The scalar curvature is given by

$$\sum_{1\leq i<j\leq p} \frac{p-2}{t_{ij}} - \frac{1}{2}\sum_{1\leq i<j\leq p}\sum_{l=1, l\neq i, l\neq j}^{p} \frac{t_{ij}}{t_{il}t_{jl}} \tag{38}$$

The simultaneous diagonalized condition is proved in the proof. We will use the items (4), (5) and (6) of Lemma A.1 to compute the traces in the Ricci curvature, using the weighted traces $\mathrm{Trw}_0, \mathrm{Trw}_1, \mathrm{Trw}_{\mathrm{ad}}$ defined there. Relative to the bi-invariant trace inner product $\langle\cdot\rangle$, any left-invariant metric is defined by a positive-definite self-adjoint operator $\mathtt{T}$ on $\mathfrak{o}(p)$ (denoted by P previously but we will use $\mathtt{T}$ in this section, considered as an operator version of the deformation parameter t). We identify $\mathtt{T}$ with an operator on $\mathbb{R}^{p\times p}$ such that $\mathtt{T}(A^{\mathsf{T}}) = \mathtt{T}(A)^{\mathsf{T}}$. Here, the operator $\mathtt{T}$ defined by $\mathtt{T}(A) = \sum_{i>j} t_{ij}a_{ij}(E_{ij} - E_{ji})$. We extend $\mathtt{T}$ to an operator on $\mathbb{R}^{p\times p}$, by defining $t_{ji} = t_{ij}$ and $t_{ii} = 0$, and set $\mathtt{T}(E_{ij}) = t_{ij}E_{ij}$, $1 \leq i, j \leq p$. On $\mathfrak{o}(p)$, $\mathtt{T}^{-1}$ operates by multiplying by t_{ij}^{-1} and we extend it to $\mathbb{R}^{p\times p}$ symmetrically, setting the diagonal entries to 0. This is an abuse of notation as $\mathtt{T}^{-1}$ as defined is only an inverse of $\mathtt{T}$ on off-diagonal entries, but this is sufficient. By $\mathrm{I}_{p\times p}$ we denote the identity operator on $\mathbb{R}^{p\times p}$, multiplying all entries by 1.

Proof. From eq. (9) we need to compute $\mathrm{Tr}\,\overline{\mathrm{ad}}_A\,\overline{\mathrm{ad}}_A$ and $\mathrm{Tr}\,\mathrm{ad}^{\dagger}{}_A\,\mathrm{ad}_A$, as $\mathrm{Tr}\,\mathrm{ad}_A\,\mathrm{ad}_A$ is well-known to be $(p-2)\,\mathrm{Tr}\,A^2 = 2(2-p)\sum_{i<j} a_{ij}^2$, (see Lemma A.1(3)) and $\mathrm{Tr}\,\mathrm{ad}_{\mathrm{ad}^{\dagger}{}_A A} = 0$. Assume $W = \mathtt{T}(A)$ with the entries $w_{ij} = t_{ij}a_{ij}$. Then we have $\overline{\mathrm{ad}}_A\,\overline{\mathrm{ad}}_A X = \mathtt{T}^{-1}[\mathtt{T}^{-1}[X, W], W]$ and $\mathrm{Tr}\,\overline{\mathrm{ad}}_A\,\overline{\mathrm{ad}}_A = -\,\mathrm{Trw}_{\mathrm{ad},(W,W,\mathtt{T}^{-1},\mathtt{T}^{-1})}$.

Using Lemma A.1(6) we obtain

$$\operatorname{Tr}\overline{\operatorname{ad}}_A\overline{\operatorname{ad}}_A = -\sum_{i<j} w_{ij}^2 \sum_{l\neq i,l\neq j}(t_{il}^{-1}t_{jl}^{-1}+t_{jl}^{-1}t_{il}^{-1}) = -2\sum_{i<j}a_{ij}^2\sum_{l\neq i,l\neq j}\frac{t_{ij}^2}{t_{il}t_{jl}}.$$

Next, $\operatorname{ad}^\dagger{}_A\operatorname{ad}_A X = -\mathtt{T}^{-1}[A,\mathtt{T}[A,X]] = -\mathtt{T}^{-1}[\mathtt{T}[X,A],A] = \mathtt{T}^{-1}[A,\mathtt{T}[X,A]]$, and

$$\operatorname{Tr}\operatorname{ad}^\dagger{}_A\operatorname{ad}_A = \operatorname{Trw}_{\operatorname{ad},(A,A,\mathtt{T},\mathtt{T}^{-1})} = \sum_{i<j}a_{ij}^2\sum_{l=1,l\neq i,l\neq j}^{p}(t_{ij}^{-1}t_{jl}+t_{jl}^{-1}t_{il}).$$

Thus, the three components give us eq. (37). The scalar curvature formula follows from here and the identity $\operatorname{Tr}\mathcal{B}_2 = -\operatorname{Tr}\mathcal{B}_3$ in remark 2.3

$$\sum_{i<j}\sum_{1\leq l\leq p,l\neq i,l\neq j}\frac{t_{il}}{t_{ij}t_{jl}}+\frac{t_{jl}}{t_{ij}t_{il}} = 2\sum_{i<j}\sum_{1\leq l\leq p,l\neq i,l\neq j}\frac{t_{ij}}{t_{il}t_{jl}} \qquad \blacksquare$$

For the Stiefel case, as $\mathfrak{k}=\mathfrak{o}(n-p)$ acts trivially on $\mathfrak{a}=\mathfrak{o}(p)$ and as p copies of the standard representation on $\mathfrak{b}=\mathbb{R}^{n-p}$, any $\operatorname{ad}(\mathfrak{k})$ invariant bilinear form on $\mathfrak{m}=\mathfrak{a}\oplus\mathfrak{b}$ could be given by an operator P represented by a pair $(\mathtt{T},\mathtt{S})$, with $\mathtt{S}=\mathtt{S}^{\mathsf{T}}$ is a positive-definite matrix in $\mathbb{R}^{p\times p}$. Assume P acts as the identity on $\mathfrak{h}$, we have

$$\mathrm{P}\,[\![\,A,B,H\,]\!] = [\![\,\mathtt{T}(A),B\mathtt{S},H\,]\!] \tag{39}$$

$$\langle[\![\,A,B,H\,]\!],[\![\,A,B,H\,]\!]\rangle_{\mathrm{P}} = \frac{1}{2}\operatorname{Tr}A^{\mathsf{T}}\mathtt{T}(A)+\operatorname{Tr}B\mathtt{S}B^{\mathsf{T}}+\frac{1}{2}\operatorname{Tr}H^{\mathsf{T}}H \tag{40}$$

With this inner product on $\mathfrak{g}$, using $\operatorname{ad}^\dagger{}_\omega = -\mathrm{P}^{-1}\operatorname{ad}_\omega\mathrm{P}$ for $\omega\in\mathfrak{g}$, we get

$$\operatorname{ad}^\dagger{}_{[\![A_1,B_1,H_1]\!]}[\![\,A_2,B_2,H_2\,]\!] = [\![\,\mathtt{T}^{-1}\{[A_1,\mathtt{T}(A_2)]+\mathtt{S}B_2^{\mathsf{T}}B_1-B_1^{\mathsf{T}}B_2\mathtt{S}\}, \tag{41}$$
$$-B_1\mathtt{T}(A_2)\mathtt{S}^{-1}-H_1B_2+B_2\mathtt{S}A_1\mathtt{S}^{-1}+H_2B_1\mathtt{S}^{-1}, -[H_1,H_2]-B_2\mathtt{S}B_1^{\mathsf{T}}+B_1\mathtt{S}B_2^{\mathsf{T}}\,]\!]$$

Let e_{ij} be the elementary matrix in $\mathbb{R}^{(n-p)\times p}$ $(1\leq i\leq n-p, 1\leq j\leq p)$ with the (i,j) entry is 1 and the other entries are zero. When $\mathtt{T}$ and $\mathtt{S}$ are both diagonal, the following proposition shows the Ricci curvature is also diagonal.

Proposition 5.2. *Representing a tangent vector at $Y\in\mathrm{St}_{p,n}$ as $YA+Y_\perp B$ for $(Y|Y_\perp)\in\mathrm{SO}(n)$ with $A=\sum_{1\leq i<j\leq p}a_{ij}(E_{ij}-E_{ji})\in\mathfrak{o}(p)$, $B=\sum_{i=1}^{n-p}\sum_{j=1}^{p}b_{ij}e_{ij}$. If the metric on $\mathrm{St}_{p,n}$ is given by $\|\,[\![\,A,B,0\,]\!]\,\|_{\mathrm{P}}^2=\sum_{1\leq i<j\leq p}a_{ij}^2t_{ij}+\sum_{i=1}^{n-p}\sum_{j=1}^{p}b_{ij}^2s_j$ then the Ricci curvature is given by $\sum_{1\leq i<j\leq p}a_{ij}^2c_{ij}+\sum_{i=1}^{n-p}\sum_{j=1}^{p}b_{ij}^2f_j$ with*

$$c_{ij} = n-2+\frac{n-p}{2}\left(\frac{t_{ij}^2}{s_is_j}-\frac{s_i}{s_j}-\frac{s_j}{s_i}\right)+\frac{1}{2}\sum_{\substack{l\neq i,\ l\neq j\\1\leq l\leq p}}\left(\frac{t_{ij}^2}{t_{il}t_{jl}}-\frac{t_{il}}{t_{jl}}-\frac{t_{jl}}{t_{il}}\right) \tag{42}$$

$$f_j = n-2+\frac{1}{2}\sum_{\substack{1\leq l\leq p\\l\neq j}}\left(\frac{s_j^2}{s_lt_{jl}}-\frac{t_{jl}}{s_l}-\frac{s_l}{t_{jl}}\right) \tag{43}$$

The Einstein equations are $c_{ij}=\lambda t_{ij}$, $f_j=\lambda s_j$. The scalar curvature is

$$\sum_{i<j}\frac{n-2}{t_{ij}}+\sum_{j=1}^{p}\frac{(n-2)(n-p)}{s_j}-\frac{1}{2}\sum_{i<j}\left\{(n-p)\left(\frac{t_{ij}}{s_is_j}+\frac{s_i}{s_jt_{ij}}+\frac{s_j}{s_it_{ij}}\right)+\sum_{\substack{l=1\\l\neq i,\ l\neq j}}^{p}\frac{t_{ij}}{t_{il}t_{jl}}\right\} \tag{44}$$

Proof. Let $\omega = [\![A, B, 0]\!]$. Then $\mathrm{Tr}\,\mathrm{ad}_\omega \mathrm{ad}_\omega = 2(2-n)(\sum_{i<j} a_{ij}^2 + \sum_{i=1}^{n-p}\sum_{j=1}^{p} b_{ij}^2)$. Using eq. (41), and the definition $\overline{\mathrm{ad}}_\omega X = \mathrm{ad}^\dagger{}_X\,\omega$, where $X = [\![X_{\mathfrak{a}}, X_{\mathfrak{b}}, 0]\!] = \hat{X}_{\mathfrak{a}} + \hat{X}_{\mathfrak{b}}$ with $\hat{X}_{\mathfrak{a}} = [\![X_{\mathfrak{a}}, 0, 0]\!]$, $\hat{X}_{\mathfrak{b}} = [\![0, X_{\mathfrak{b}}, 0]\!]$, using eq. (41) we obtain

$$\pi_{\mathfrak{m}}\,\overline{\mathrm{ad}}_{[\![A,B,0]\!]} X = \pi_{\mathfrak{m}}\,\mathrm{ad}^\dagger{}_X\,[\![A, B, 0]\!]$$
$$= [\![-\mathtt{T}^{-1}([X_{\mathfrak{a}}, \mathtt{T}(A)] + \mathtt{S}B^{\mathsf{T}}X_{\mathfrak{b}} - X_{\mathfrak{b}}^{\mathsf{T}}B\mathtt{S}), -X_{\mathfrak{b}}\mathtt{T}(A)\mathtt{S}^{-1} + B\mathtt{S}X_{\mathfrak{a}}\mathtt{S}^{-1}, 0]\!]$$
$$(\overline{\mathrm{ad}}_\omega\,\pi_{\mathfrak{m}}\,\overline{\mathrm{ad}}_\omega\,\hat{X}_{\mathfrak{a}})_{\mathfrak{a}} = -\mathtt{T}^{-1}\{[-\mathtt{T}^{-1}[X_{\mathfrak{a}}, \mathtt{T}(A)], \mathtt{T}(A)] + \mathtt{S}B^{\mathsf{T}}B\mathtt{S}X_{\mathfrak{a}}\mathtt{S}^{-1} + \mathtt{S}^{-1}X_{\mathfrak{a}}\mathtt{S}B^{\mathsf{T}}B\mathtt{S}\}$$
$$(\overline{\mathrm{ad}}_\omega\,\pi_{\mathfrak{m}}\,\overline{\mathrm{ad}}_\omega\,\hat{X}_{\mathfrak{b}})_{\mathfrak{b}} = X_{\mathfrak{b}}\mathtt{T}(A)\mathtt{S}^{-1}\mathtt{T}(A)\mathtt{S}^{-1} - B\mathtt{S}\mathtt{T}^{-1}(\mathtt{S}B^{\mathsf{T}}X_{\mathfrak{b}} - X_{\mathfrak{b}}^{\mathsf{T}}B\mathtt{S})\mathtt{S}^{-1}$$
$$(\mathrm{ad}^\dagger{}_\omega\,\pi_{\mathfrak{m}}\,\mathrm{ad}_\omega\,\hat{X}_{\mathfrak{a}})_{\mathfrak{a}} = -\mathtt{T}^{-1}[A, \mathtt{T}[A, X_{\mathfrak{a}}]] + \mathtt{T}^{-1}(\mathtt{S}X_{\mathfrak{a}}B^{\mathsf{T}}B + B^{\mathsf{T}}BX_{\mathfrak{a}}\mathtt{S})$$
$$(\mathrm{ad}^\dagger{}_\omega\,\pi_{\mathfrak{m}}\,\mathrm{ad}_\omega\,\hat{X}_{\mathfrak{b}})_{\mathfrak{b}} = B\mathtt{T}(B^{\mathsf{T}}X_{\mathfrak{b}} - X_{\mathfrak{b}}^{\mathsf{T}}B)\mathtt{S}^{-1} - X_{\mathfrak{b}}A\mathtt{S}A\mathtt{S}^{-1}$$

From here

$$\mathrm{Tr}(X_{\mathfrak{a}} \mapsto (\overline{\mathrm{ad}}_\omega\,\pi_{\mathfrak{m}}\,\overline{\mathrm{ad}}_\omega\,\hat{X}_{\mathfrak{a}})_{\mathfrak{a}}) = -\,\mathrm{Trw}_{\mathrm{ad},(\mathtt{T}(A),\mathtt{T}(A),\mathtt{T}^{-1},\mathtt{T}^{-1})} - \mathrm{Trw}_{1,(\mathtt{S}B^{\mathsf{T}}B\mathtt{S},\mathtt{S}^{-1},\mathtt{T}^{-1},\mathrm{I}_{p\times p})}$$
$$= -2\sum_{1\leq i<j\leq p}\;\sum_{l\neq i, l\neq j}\frac{t_{ij}^2}{t_{jl}t_{il}}a_{ij}^2 - \sum_{i=1}^{n-p}\sum_{j=1}^{p} b_{ij}^2 s_j^2 \sum_{l\neq j}\frac{1}{s_l t_{jl}}$$

For $\mathrm{Tr}(X_{\mathfrak{b}} \mapsto (\overline{\mathrm{ad}}_\omega\,\pi_{\mathfrak{m}}\,\overline{\mathrm{ad}}_\omega\,\hat{X}_{\mathfrak{b}})_{\mathfrak{b}})$, write the operator $X_{\mathfrak{b}} \mapsto B\mathtt{S}\mathtt{T}^{-1}(\mathtt{S}B^{\mathsf{T}}X_{\mathfrak{b}} - X_{\mathfrak{b}}^{\mathsf{T}}B\mathtt{S})\mathtt{S}^{-1}$ as $X_{\mathfrak{b}} \mapsto F_2 \circ \mathtt{T}^{-1} \circ F_1 X_{\mathfrak{b}}$ with $F_1 : X_{\mathfrak{b}} \mapsto \mathtt{S}B^{\mathsf{T}}X_{\mathfrak{b}} - X_{\mathfrak{b}}^{\mathsf{T}}B\mathtt{S}$, $F_2 : Z \mapsto B\mathtt{S}Z\mathtt{S}^{-1}$ ($Z \in \mathfrak{o}(p)$). Hence, we have $\mathrm{Tr}\, F_2 \circ \mathtt{T}^{-1} \circ F_1 = \mathrm{Tr}\,\mathtt{T}^{-1} \circ F_1 \circ F_2$ equals

$$\mathrm{Tr}(Z \mapsto \mathtt{T}^{-1}(\mathtt{S}B^{\mathsf{T}}B\mathtt{S}Z\mathtt{S}^{-1} + \mathtt{S}^{-1}Z\mathtt{S}B^{\mathsf{T}}B\mathtt{S})) = \mathrm{Trw}_{1,(\mathtt{S}B^{\mathsf{T}}B\mathtt{S},\mathtt{S}^{-1},\mathtt{T}^{-1},\mathrm{I}_{p\times p})}$$

Together with (1) of Lemma A.1 we obtain

$$\mathrm{Tr}(X_{\mathfrak{b}} \mapsto (\overline{\mathrm{ad}}_\omega\,\pi_{\mathfrak{m}}\,\overline{\mathrm{ad}}_\omega\,\hat{X}_{\mathfrak{b}})_{\mathfrak{b}}) = (n-p)\,\mathrm{Tr}(\mathtt{T}(A)\mathtt{S}^{-1}\mathtt{T}(A)\mathtt{S}^{-1}) - \mathrm{Trw}_{1,(\mathtt{S}B^{\mathsf{T}}B\mathtt{S},\mathtt{S}^{-1},\mathtt{T}^{-1},\mathrm{I}_{p\times p})}$$
$$= -2(n-p)\sum_{1\leq i<j\leq p} a_{ij}^2\frac{t_{ij}^2}{s_is_j} - \sum_{1\leq i\leq n-p}\;\sum_{1\leq j\leq p} b_{ij}^2 \sum_{1\leq l\leq p, l\neq j}\frac{s_j^2}{s_l t_{jl}}$$
$$\mathrm{Tr}(X_{\mathfrak{a}} \mapsto (\mathrm{ad}^\dagger{}_\omega\,\pi_{\mathfrak{m}}\,\mathrm{ad}_\omega\,\hat{X}_{\mathfrak{a}})_{\mathfrak{a}}) = \mathrm{Trw}_{\mathrm{ad},(A,A,\mathtt{T}^{-1},\mathtt{T})} + \mathrm{Trw}_{1,(B^{\mathsf{T}}B,\mathtt{S},\mathtt{T}^{-1},\mathrm{I}_{\mathbb{R}^{p\times p}})}$$
$$= \sum_{i<j} a_{ij}^2 \sum_{l\neq i, l\neq j}\left(\frac{t_{il}}{t_{jl}} + \frac{t_{jl}}{t_{il}}\right) + \sum_{1\leq i\leq n-p}\;\sum_{1\leq j\leq p} b_{ij}^2 \sum_{l\neq j}\frac{s_l}{t_{jl}}$$

Similarly, for $\mathrm{Tr}(X_{\mathfrak{b}} \mapsto (\mathrm{ad}^\dagger{}_\omega\,\pi_{\mathfrak{m}}\,\overline{\mathrm{ad}}_\omega\,\hat{X}_{\mathfrak{b}})_{\mathfrak{b}}$, write $X_{\mathfrak{b}} \mapsto B\mathtt{T}(B^{\mathsf{T}}X_{\mathfrak{b}} - X_{\mathfrak{b}}^{\mathsf{T}}B)S^{-1}$ as $G_2 \circ \mathtt{T} \circ G_1$ with $G_2 Z = BZ\mathtt{S}^{-1}$ ($Z \in \mathfrak{o}(p)$), $G_1 X_{\mathfrak{b}} = B^{\mathsf{T}}X_{\mathfrak{b}} - X_{\mathfrak{b}}^{\mathsf{T}}B$. Then $\mathrm{Tr}\,\mathtt{T} \circ G_1 \circ G_2 = \mathrm{Tr}(Z \mapsto \mathtt{T}(B^{\mathsf{T}}BZ\mathtt{S}^{-1} + \mathtt{S}^{-1}ZB^{\mathsf{T}}B)) = \mathrm{Trw}_{1,(B^{\mathsf{T}}B,\mathtt{S}^{-1},\mathtt{T},\mathrm{I}_{p\times p})}$, and together with (1) of Lemma A.1 we get

$$\mathrm{Tr}(X_{\mathfrak{b}} \mapsto (\mathrm{ad}^\dagger{}_\omega\,\pi_{\mathfrak{m}}\,\mathrm{ad}_\omega\,\hat{X}_{\mathfrak{b}})_{\mathfrak{b}}) = \mathrm{Trw}_{1,(B^{\mathsf{T}}B,\mathtt{S}^{-1},\mathtt{T},\mathrm{I}_{p\times p})} - (n-p)\,\mathrm{Tr}(A\mathtt{S}A\mathtt{S}^{-1})$$
$$= (n-p)\sum_{1\leq i<j\leq p} a_{ij}^2\left(\frac{s_i}{s_j} + \frac{s_j}{s_i}\right) + \sum_{1\leq i\leq n-p}\;\sum_{1\leq j\leq p} b_{ij}^2 \sum_{l\neq j}\frac{t_{jl}}{s_l}$$

Thus in the expression of the Ricci curvature in eq. (9), the coefficient of a_{ij}^2 is

$$-\frac{1}{4}\Big(-\,2\sum_{\substack{l\neq i\\ l\neq j}}\frac{t_{ij}^2}{t_{jl}t_{il}} - 2(n-p)\frac{t_{ij}^2}{s_is_j}\Big) - \frac{1}{2}\Big(\sum_{\substack{l=1\\ l\neq j}}^{p}\left(\frac{t_{il}}{t_{jl}} + \frac{t_{jl}}{t_{il}}\right) + (n-p)\left(\frac{s_i}{s_j} + \frac{s_j}{s_i}\right)\Big) - (2-n)$$

and of b_{ij}^2 is

$$-\frac{1}{4}\Big(\sum_{l=1,l\neq j}^{p}(-\frac{s_j^2}{s_lt_{jl}}-\frac{s_j^2}{s_lt_{jl}})\Big)-\frac{1}{2}\sum_{l=1,l\neq j}^{p}(\frac{s_l}{t_{jl}}+\frac{t_{jl}}{s_l})-(2-n)$$

and they rearrange to the coefficients c_{ij} and f_j in the proposition. The scalar curvature formula follows by noting there are $n-p$ terms f_j/s_j for each $j\in 1,\cdots,p$, thus the first, second terms in eq. (44) are clear. The last term follows a similar cancelation as for $\mathrm{SO}(p)$, while the third term is a simplification of

$$\frac{n-p}{2}\sum_{i<j}(\frac{t_{ij}^2}{s_is_j}-\frac{s_i}{s_jt_{ij}}-\frac{s_j}{s_it_{ij}})+\frac{n-p}{2}\sum_{j=1}^{p}\sum_{\substack{1\leq l\leq p\\ l\neq j}}(\frac{s_j}{s_lt_{jl}}-\frac{t_{jl}}{s_ls_j}-\frac{s_l}{t_{jl}s_j})$$

as $\frac{s_j}{s_lt_{jl}}-\frac{s_l}{t_{jl}s_j}$ sums to zero and $\sum_{(l,j),l\neq j}\frac{t_{jl}}{s_ls_j}=2\sum_{i<j}\frac{t_{ij}}{s_is_j}$ ■

Corollary 5.3. *For* $\mathrm{SO}(p)$, *consider a partition of* $\{1,\cdots,p\}=\cup_{I\in\mathcal{I}}I$ *to disjoint consecutive non-empty intervals. If the coefficients* t_{ij} *depends only on the partition sets, that is* $t_{i_1j_1}=t_{i_2j_2}=t_{IJ}$ *if* $i_1,i_2\in I, j_1,j_2\in J$ *for any* $I,J\in\mathcal{I}$, *then the Ricci coefficients* r_{ij} *are also only dependent on the partition sets,* $r_{i_1j_1}=r_{i_2j_2}=r_{IJ}$ *and*

$$r_{II}=\frac{|I|-2}{2}+\sum_{L\in\mathcal{I},L\neq I}\frac{|L|}{2}\frac{t_{II}^2}{t_{IL}^2}\quad \text{if } |I|>1 \tag{45}$$

$$\begin{aligned} r_{IJ}=p-2-\frac{|I|-1}{2}\frac{t_{II}}{t_{IJ}}-\frac{|J|-1}{2}\frac{t_{JJ}}{t_{IJ}}\\ +\sum_{L\in\mathcal{I},L\neq I,L\neq J}\frac{|L|}{2}(\frac{t_{IJ}^2}{t_{IL}t_{JL}}-\frac{t_{IL}}{t_{JL}}-\frac{t_{JL}}{t_{IL}})\quad \text{if } I\neq J \end{aligned}\tag{46}$$

For $\mathrm{St}_{p,n}$, *if we have further* $s_{i_1}=s_{i_2}=s_I$ *for* $i_1,i_2\in I$, $I\in\mathcal{I}$, *then the Ricci coefficients are also dependent the partition set and*

$$c_{II}=\frac{n-p}{2}\frac{t_{II}^2}{s_I^2}+\frac{|I|-2}{2}+\sum_{L\in\mathcal{I},L\neq I}\frac{|L|}{2}\frac{t_{II}^2}{t_{IL}^2}\quad \text{if } |I|>1 \tag{47}$$

$$\begin{aligned} c_{IJ}=n-2+\frac{(n-p)}{2}(\frac{t_{IJ}^2}{s_Is_J}-\frac{s_I}{s_J}-\frac{s_J}{s_I})-\frac{|I|-1}{2}\frac{t_{II}}{t_{IJ}}-\frac{|J|-1}{2}\frac{t_{JJ}}{t_{IJ}}\\ +\sum_{L\neq I,L\neq J}\frac{|L|}{2}(\frac{t_{IJ}^2}{t_{IL}t_{JL}}-\frac{t_{IL}}{t_{JL}}-\frac{t_{JL}}{t_{IL}})\quad \text{if } I\neq J \end{aligned}\tag{48}$$

$$f_J=n-2-\frac{|J|-1}{2}\frac{t_{JJ}}{s_J}+\sum_{L\neq J}\frac{|L|}{2}(\frac{s_J^2}{s_Lt_{JL}}-\frac{t_{JL}}{s_L}-\frac{s_L}{t_{JL}}) \tag{49}$$

Proof. Since t_{ii} is not defined for $i\in\{1\cdots p\}$, for r_{II} we means any value of $r_{i_1i_2}$, $i_1\neq i_2$ and $i_1,i_2\in I$, which requires $|I|>1$. For $\mathrm{SO}(p)$, in the sum $\sum_{1\leq l\leq p,l\neq i,l\neq j}\frac{t_{ij}^2}{t_{il}t_{jl}}-\frac{t_{il}}{t_{jl}}-\frac{t_{jl}}{t_{il}}$ in eq. (37), the indices $l\in I$ contribute $\frac{|I|-2}{2}(-1)$, while each $L\in\mathcal{I}$ contributes $\frac{L}{2}(\frac{t_{IJ}^2}{t_{IL}^2}-2)$. We can simplify

$$p-2-\frac{|I|-2}{2}-\sum_{L\neq I}\frac{L}{2}(2)=p-2-\frac{|I|-2}{2}-(p-|I|)=\frac{|I|-2}{2}.$$

For r_{IJ} with $I \neq J$, $l \in I$ contributes the second term and $l \in J$ contributes the third term, while $L \neq I, J$ contributes the last term. The proof for the Stiefel case follows similarly. ■

We will denote a partition of p consisting of consecutive intervals $\mathcal{I} = \{I_1, \cdots, I_q\}$ by its collection of non decreasing sizes $(k_1, \cdots, k_q)$, $\sum_{j=1}^{q} k_j = p$. Denote by $t_{[q_1 q_2]}$, $1 \leq q_1 \leq q_2 \leq q$ the coefficient $t_{I_{q_1} I_{q_2}}$, and define $s_{[q_1]} = s_{[I_{q_1}]}$.

Several calculations of the scalar and Ricci curvature for $\mathrm{SO}(p)$ and the Stiefel manifold appear in the literature [14, 2, 5] could be deducted from the results of this section. For a Stiefel manifold with deformation metric, apply the corollary for $\mathcal{I} = \{I\}$ where $I = \{1, \cdots, p\}$, $t_{ij} = t$ for $1 \leq i < j \leq p$ and $s_j = s$ for $1 \leq j \leq p$

$$c_{ij} = c_{II} = \frac{n-p}{2}\frac{t^2}{s^2} + \frac{p-2}{2} = 2(n-p)\alpha^2 + \frac{p-2}{2} \tag{50}$$

$$f_j = n - 2 - \frac{p-1}{2}\frac{t}{s} = n - 2 - (p-1)\alpha \tag{51}$$

The Einstein equation in this case gives us the Jensen metrics, if t/s satisfies the equation $(n-1)(t/s)^2 - 2(n-2)(t/s) + (p-2) = 0$. The scalar curvature is

$$(n-2)\frac{p(p-1)}{2t} + \frac{(n-2)(n-p)p}{s} - \frac{p(p-1)}{4}((n-p)\frac{t}{s^2} + (n-p)\frac{2}{t} + \frac{p-2}{t})$$
$$= \frac{p(p-1)(p-2)}{4t} + \frac{(n-2)(n-p)p}{s} - \frac{(n-p)p(p-1)t}{4s^2}$$

which reduces to the formula in section 5 of [2], where $t = 2(n-2)x_1$, $s = 2(n-2)x_{12}$, $k_1 = p$, $k_2 = n - p$ in that paper. The results of [5], equations (16), (17) could be derived from our result for $SO(n)$ by a partition $\mathcal{I} = \{I_1, I_2, I_3\}$ of $\{1, \cdots, p\}$ to three intervals of length k_1, k_2, k_3 for positive integers $k_1 + k_2 + k_3 = p$, if we set $t_{I_i, I_i} = 2(n-2)x_i$, $t_{I_i, I_j} = 2(n-2)x_{ij}$ $(i, j \in \{1, \cdots, 3\})$. The coefficients r_i, r_{ij} in that paper are $r_{I_i I_i} t^{-1}_{I_i I_i}, r_{I_i I_j} t^{-1}_{I_i I_j}$ in our notation. The results of section 6 in [2] of the Stiefel manifold $\mathrm{St}_{qk, qk+l}$ for positive integers q, k, l (we use q for s in that paper to avoid conflict of notation) corresponds to $\mathcal{I} = \{I_j | 1 \leq j \leq q\}$ obtained by dividing $\{1, \cdots, qk\}$ to q intervals I_j each of size k, where $t_{II} = 2(n-2)x$, $t_{IJ} = 2(n-2)y$ if $I \neq J \in \mathcal{I}$ and $s_J = 2(n-2)z$ for positive numbers x, y, z, $I, J \in \mathcal{I}$. We now present an explcit construction of a family of Einstein metrics for $\mathrm{SO}(p)$, likely an instant of the metrics of type 1 in [24].

Proposition 5.4. *Let $p = 2dk$ for $k \geq 3$, $d \geq 1$. Consider a partition $\mathcal{I}$ of $\{1, \cdots, p\}$ to $q = 2d$ consecutive intervals $I_1, \cdots, I_{2d}$ each of size k. Let t_1, t_2 be two positive numbers. Consider the metric on $\mathrm{SO}(p)$ with $t_{I_v I_v} = 1$, $t_{I_v I_w} = t_b$ if $1 \leq v < w \leq 2d$ and $|v - w| \equiv b \mod 2$, $b = 1, 2$. Then the Ricci curvature coefficients also takes three values r_0, r_1, r_2 with $r_{I_v I_v} = r_0$, $r_{I_v I_w} = r_b$ if $|v - w| \equiv b$ for $b = 1, 2, 1 \leq v \neq w \leq 2d$. These are the only Einstein metrics in this family:*

(a) $t_1 = t_2 = 1$, *the bi-invariant metric;*

(b) $t_1 = t_2 = ((2d+1)k - 2)/(k-2)$;

(c) $t_1 = (3dk - 2)/(dk - 2), t_2 = 1$;

(d) t_1 *is given by eq.* (52), $t_2 = xt_1$ *where x is the only positive root of eq.* (53)

$$t_1 = \frac{2(k-1)(x+1)}{x((-3d+2)kx + 2k - 4 + dk)} \tag{52}$$

$$F(x) := (3d-2)dk^2x^3 + d((-d-2)k^2+4k)x^2 \qquad (53)$$
$$+\{(3d^2+d)k^2+(-6d-2)k+4\}x+(-d^2-3d)k^2+(6d+2)k-4=0,$$

where x *is between* 0 *and* $\frac{2k-4+dk}{(3d-2)k}$ *and* $t_1 \geq t_2 \geq 0$.

For the metrics in the proposition, the q diagonal blocks have $t_{II} = 1$, blocks $t_{[12]}, t_{[23]} \cdots, t_{[2d-1,2d]}$ equal t_1 and $t_{[13]}, \cdots t_{[2d-2,2d]}$ equal t_2. We can also show $F(x)$ has only one real root and $t_1 \geq t_2 \geq 1$, the metrics b), c), d) are identical when $d = 1$ and distinct otherwise. The proof of these facts, which we will not show here, reduces to showing several long polynomial expressions are positive, by providing estimates for large k, d then analyzing the remaining range case-by-case.

Proof. From eq. (45), all the $r_{I_vI_v}$ $1 \leq v \leq 2d$ have the same value

$$r_0 = \frac{k-2}{2} + \frac{dk}{2t_1^2} + \frac{(d-1)k}{2t_2^2}$$

The first term is clear, the d terms $L = I_l$ with $|l-v| \equiv 1$ contribute $k/(2t_1^2)$ each, and $2d-d-1$ terms with $|l-v| \equiv 2$, $l \neq v$ contribute $k/(2t_2^2)$ each. Similarly from eq. (46), let $r_b = r_{I_vI_w}$ for $|v-w| \equiv b \mod 2, b = 1, 2$ for fixed v, w then

$$r_1 = p - 2 - \frac{k-1}{t_1} + \frac{q-2}{2}k(-\frac{t_2}{t_1})$$

the first term is clear, the next two terms combined to $(k-1)/t_1$, the last sum has $q-2$ components, both the odd and even case contribute $-kt_2/(2t_1)$. For r_2, the first three terms are similar, for the last sum, the case $L = I_l$ with $|l-v| \equiv |l-w| \equiv 1$ contributes d terms $k(t_2/t_1 - 2)$, while the remaining case contributes $k/2(-1)$

$$r_2 = p - 2 - \frac{k-1}{t_2} + \frac{d}{2}k(\frac{t_2}{t_1} - 2) - \frac{(d-2)k}{2}$$

We will provide the main arguments for the proof but skip the detailed expansions as the manipulations are easier done symbolically. Let $x := t_2/t_1$. We need to solve $e_1 = e_2 = 0$ for x and t_1 with $e_1 := 2x^2t_1(r_0t_1 - r_1)$, $e_2 := 2t_1x(r_0t_2 - r_2)$

$$e_1 = k(2d-2)x^3t_1 + (k-2)t_1^2x^2 + t_1(4-4kd)x^2 + (2k-2+dk)x^2 + (d-1)k$$
$$e_2 = (k-2)t_1^2x^2 - kdx^3t_1 + dkx^2 + t_1x(-kd+4-2k) + dk + k - 2$$

Let $e_3 := e_1 - e_2$, then e_3 factors to

$$e_3 = (x-1)\{t_1x(3dkx - dk - 2kx - 2k + 4) + 2(k-1)(x+1)\}$$

If $x = 1$, $e_2 = 0$ is quadratic in t_1 with two solutions $t_1 = 1$ and $t_1 = \frac{2dk+k-2}{k-2}$ giving us the solutions in a) and b). If $x \neq 1$ then we can solve for t_1 from $e_3 = 0$, giving us the relationship between t_1 and x in eq. (52). Substitute eq. (52) to $e_1 = 0$, we get an equation of degree 6 in x, which factors to $x^2((3dk-2)x + 2 - dk)F(x) = 0$. The case c) corresponds to the second factor with $x = (dk-2)/(3dk-2)$, t_1, t_2 are obtained by substitution. We are left with the case $F(x) = 0$. It is clear $F(0) < 0$ when $k \geq 3$.

Next, let $x_{asym} = (dk+2k-4)/(k(3d-2))$ be the value of x making the denominator of eq. (52) vanishes then

$$F(x_{asymp}) = \frac{8(k-2)(k-1)(dk-1)}{k(3d-2)} > 0$$

thus $F(x) = 0$ has a root between 0 and x_{asymp}, and $t_1, t_2 > 0$ at this root. ■

We can generalize (b) to $q \in \mathbb{N}$, with $t_1 = t_2 = ((q+1)k-2)/(k-2)$, it seems probable d) extends to Stiefel manifolds as in [2]. The reason c) and d) do not generalize for odd q is then $r_{[I_1I_1]} \neq r_{[I_2I_2]}$ giving us an extra equation. Observing $p = 2dk = 2d_1k_1$ with $d_1 = 1, k_1 = dk$, c) is an rearrangement of b) using (d_1, k_1). The metric in d) is not naturally reductive for the embedding $\mathrm{SO}(k)^q \subset \mathrm{SO}(p)$.

Our construction is different from the approach using an Ad-invariant metric of a subgroup of $\mathrm{SO}(p)$ but the results often turn out to be the same. In either case, several coefficients t_{ij} become identical, reducing the actual number of parameters. With our general formula, conceptually, we could have an Einstein metric on $\mathrm{SO}(p)$ with all distinct coefficients. In practice, for small p, the Einstein metrics found numerically has significantly reduced parameter count, we find no instance with $p(p-1)/2$ distinct parameters, but the parameters found are clustered and could be arranged to come from a partition $\mathcal{I}$ of $\{1, \cdots, p\}$ to $q < p$ subsets (which has the parameter counts of $d_{\mathcal{I}} = q(q+1)/2 - n_{|I|=1}$ where $n_{|I|=1}$ is the number of $I \in \mathcal{I}$ with size 1). Further, as seen in the examples above, there are additional relations between the parameters t_{II} and t_{IJ} so we may have much less than $d_{\mathcal{I}}$ parameters. The bi-invariant metric has one parameter, b) and c) of proposition 5.4 has two and d) has three (with one parameter normalized to 1).

We will not discuss the numerical methods to solve the Einstein condition. The following results could be verified independently. We find two new Einstein metrics for SO(9) with the partition $(1, 1, 3, 4)$ and the parameters

$$t_{[12]} = 11.7750044,\quad t_{[13]} = 3.49295668,\quad t_{[14]} = 11.00395784,\quad t_{[23]} = 11.42167356,$$
$$t_{[24]} = 4.85823252,\quad t_{[33]} = 1,\quad t_{[34]} = 10.63219969,\quad t_{[44]} = 2.07232593,$$

and the partition $(1, 2, 3, 3)$ with the parameters

$$t_{[12]} = 11.41086965,\quad t_{[13]} = 10.83434028,\quad t_{[14]} = 3.50463883,$$
$$t_{[22]} = 5.68181449,\quad t_{[23]} = 4.25394966,\quad t_{[24]} = 11.04723483,$$
$$t_{[33]} = 1.02563286,\quad t_{[34]} = 10.45556196,\quad t_{[44]} = 1.$$

For $\mathrm{St}_{5,7}$ we find a new Einstein metric, in addition to Jensen's and those in [4, 8], for the partition $(1, 1, 3)$, with

$$t_{[12]} = 4.15568493,\quad t_{[13]} = 3.86531511,\quad t_{[23]} = 1.78072637,\quad t_{[33]} = 0.49956303,$$
$$s_{[1]} = 1,\quad s_{[2]} = 4.09222372,\quad s_{[3]} = 3.79800548.$$

6. Sectional curvature range

We have seen the sectional curvature numerator $\hat{\mathcal{K}}$ could be expressed as a weighted sum of squares, this allows us to estimate the sectional curvature range. If $p = 1$ then the Stiefel manifold is a sphere and has constant sectional curvature. Therefore we will assume $p > 1$ below.

It is easy to establish upper and lower bounds (not tight) for the sectional curvature from eq. (34). Using the triangle inequality we can bound $\mathcal{K}$ from eq. (34) by bounding an expression of the form

$$K_1 = a\|[A_1, A_2]\|_F^2 + b\|B_1B_2^{\mathsf{T}} - B_2B_1^{\mathsf{T}}\|_F^2 + c\|B_1^{\mathsf{T}}B_2 - B_2^{\mathsf{T}}B_1\|_F^2 + d\|B_1A_2 - B_2A_1\|_F^2$$

by the curvature denominator $S := (\alpha\|A_1\|_F^2 + \|B_1\|_F^2)(\alpha\|A_2\|_F^2 + \|B_2\|_F^2)$. We use the inequality $\|[X, Z]\|_F^2 \leq \|X\|_F^2\|Z\|_F^2$, for two antisymmetric matrices in $\mathfrak{o}(n)$ if $n > 3$ ([10], Lemma 2.5 provides the explicit matrices where we have equality, see also [11], Proposition 4.2). Apply that inequality with

$$X = \frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{2\alpha}A_1 & -B_1^{\mathsf{T}}\\ B_1 & 0\end{bmatrix}, \quad \text{and} \quad Z = \frac{1}{\sqrt{2}}\begin{bmatrix}\sqrt{2\alpha}A_2 & B_2^{\mathsf{T}}\\ B_2 & 0\end{bmatrix}$$

and similar inequalities for $B_1 = B_2 = 0$, $A_1 = A_2 = 0$, we can bound each term of K_1 by S, thus getting a bound for $\mathcal{K}$.

We will attempt to provide more refined bounds. The analysis of sectional curvature range for Stiefel manifolds is more complicated than that of symmetric spaces because of the presence of both the A and B components. The manifold is homogeneous, therefore the sectional curvature range is the same at any point.

In Table 1, we show sectional curvature values of $\mathrm{St}_{p,n}$ at several sections (pairs of linearly independent tangent vectors), each defined by a quadruple (A_1, B_1, A_2, B_2).

$\mathcal{K}$	A and B	condition
0	$A_1 = A_2 = E_{12} - E_{21},\ B_1 = 2e_{13},\ B_2 = -\alpha e_{13}$	$n \geq 4,\ p \geq 3$
0	$A_1 = A_2 = 0, B_1 = e_{11},\ B_2 = e_{22}$	$n \geq 4,\ p \leq n-2$
1	$A_1 = A_2 = 0, B_1 = e_{11},\ B_2 = e_{21}$	$n \geq 4,\ p \leq n-2$
$\frac{1}{2\alpha+1}$	$A_1 = E_{12} - E_{21}, A_2 = E_{1p} - E_{p1}, B_1 = -e_{1p}, B_2 = e_{12}$	$p \geq 3$
$\frac{1}{8\alpha}$	$A_1 = E_{12} - E_{21},\ A_2 = E_{23} - E_{32},\ B_1 = B_2 = 0$	$p \geq 3$
$\frac{1}{4\alpha}$	$A_1 = E_{12} - E_{21} + E_{p-1,p} - E_{p,p-1}$	$p \geq 4$
	$A_2 = E_{1,p-1} - E_{p-1,1} - E_{2,p} + E_{p,2},\ B_1 = B_2 = 0$	
$\frac{\alpha}{2}$	$A_1 = (E_{12} - E_{21}),\ A_2 = 0, B_1 = 0,\ B_2 = e_{11}$	
$\frac{2-3\alpha}{2}$	$A_1 = 0,\ A_2 = 0,\ B_1 = e_{11},\ B_2 = e_{12}$	
$\frac{4-3\alpha}{2}$	$A_1 = A_2 = 0,\ B_1 = e_{11} + e_{22},\ B_2 = e_{12} - e_{21}$	$n \geq 4,\ p \leq n-2$
$\mathfrak{l}(\alpha)$	$A_1 = E_{12} - E_{21},\ A_2 = E_{23} - E_{31}$	
	$B_1 = \gamma_{\min}(\alpha)^{1/2}e_{11}, B_2 = \gamma_{\min}(\alpha)^{1/2}e_{13}$	$p \geq 3$, $\alpha < 7/10$

Table 1: Sectional curvature at representative sections. $\mathfrak{l}(\alpha) = \mathfrak{c}(\gamma_{\min}(\alpha))$, from eq. (55) and eq. (54).

A few of those sections come from the corresponding sections for $\mathrm{SO}(n)$, in [10] as cited. We have noted that $\mathcal{K}$ is non-negative if $\alpha \leq \frac{1}{2}$, and Table 1 shows a section with $\mathcal{K} = \frac{2-3\alpha}{2}$, thus, if $\alpha > \frac{2}{3}$, $\mathcal{K}$ always has negative values in its range. When $p = 2$, we will show $\mathcal{K}$ is non-negative if $\alpha \leq \frac{2}{3}$. When $p > 2$, $\mathcal{K}$ could be negative if $\frac{1}{2} \leq \alpha \leq \frac{2}{3}$. To see this, recall we denote E_{ij} and e_{ij} to be the elementary matrices in $\mathbb{R}^{p\times p}$ and $\mathbb{R}^{(n-p)\times p}$. Let $A_1 = E_{12} - E_{21}, B_1 = \gamma^{1/2}e_{11}, A_2 = E_{23} - E_{32}$, $B_2 = \gamma^{1/2}e_{13}$ for $\gamma \in \mathbb{R}, \gamma > 0$. Thus, $[A_1, A_2] = E_{13} - E_{31}$, $B_1A_2 = B_2A_1 = 0$,

$B_1 B_2^{\mathsf{T}} = 0$, $B_1^{\mathsf{T}} B_2 - B_2^{\mathsf{T}} B_1 = \gamma(E_{13} - E_{31})$. By eq. (35), the corresponding sectional curvature is

$$\mathfrak{c}(\gamma) = \frac{\alpha/2 + \alpha(4\alpha - 3)\gamma + (2 - 3\alpha)\gamma^2/2}{(2\alpha + \gamma)^2} \tag{54}$$

with $\frac{d}{d\gamma}\mathfrak{c}(\gamma) = \alpha((7 - 10\alpha)\gamma - 1 - 6\alpha + 8\alpha^2)/(\gamma + 2\alpha)^3$, $\mathfrak{c}$ is minimized at

$$\gamma_{\min}(\alpha) = (1 + 6\alpha - 8\alpha^2)/(7 - 10\alpha) \tag{55}$$

Substitute in, the function $\mathfrak{l}(\alpha) := \mathfrak{c}(\gamma_{\min}(\alpha))$ is slightly negative for α in the interval $(\frac{1}{2}, \frac{7}{10})$, which contains $\frac{2}{3}$. Note that $\alpha = \frac{7}{10}$ is a removable singularity of $\mathfrak{l}$, and setting $\mathfrak{l}(\frac{7}{10}) = \lim_{\gamma\to\infty} \mathfrak{c}(\gamma) = \frac{1}{2}(2 - 3 \times \frac{7}{10}) = \frac{-1}{20}$ makes it a smooth function. This function is strictly decreasing and negative in the interval $(\frac{1}{2}, \frac{7}{10})$, with $\mathfrak{l}(\frac{1}{2}) = 0$ and $\mathfrak{l}(\frac{2}{3})$ around -0.02.

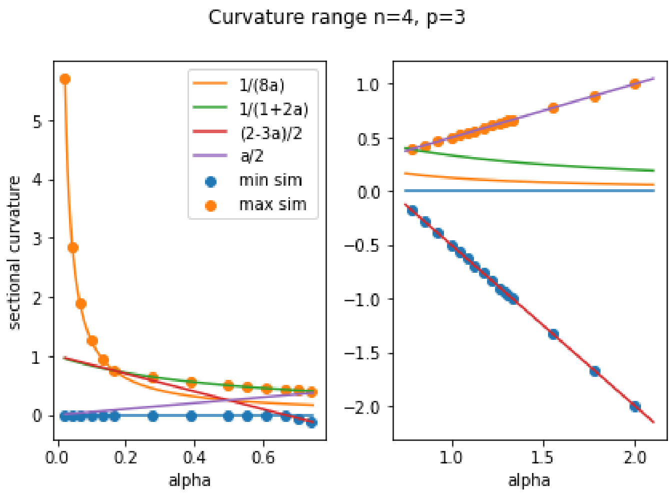

Figure 1: Numerical test for curvature range $n = 4, p = 3$. Max, min sims are curvature ranges from numerical optimization.

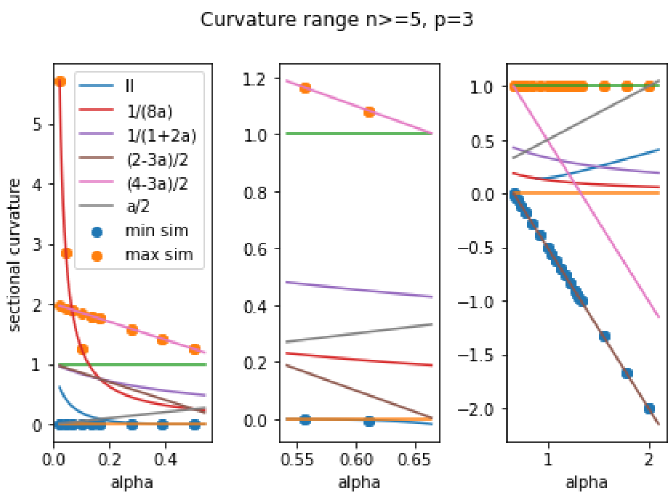

Figure 2: Numerical test for curvature range $n > 4, p = 3$

The curvature range contains the interval between the maximum and minimum of values in Table 1 if the condition in the last column of the table is satisfied. For $p = 2$, Proposition 6.1 determines the exact curvature range. For $p > 2$, numerically, the sections in that table seem to determine the range completely. For each α, the lower and upper bound of the curvature range, found numerically by optimizing $\mathcal{K}$ over the space of all sections, the Grassmann manifold of two-dimensional subspaces of $\mathbb{R}^{\dim \mathrm{St}_{p,n}}$ is within the maximal and minimal values of the sections in the table if the condition in the last column is satisfied, as shown in the Figures 1, 2, 3, 4.

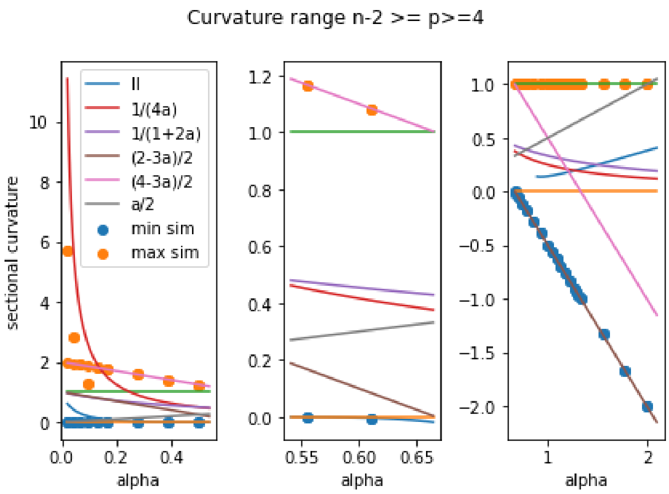

Figure 3: Numerical test for curvature range $n-2 \geq p \geq 4$. Max, min sims are curvature ranges from numerical optimization.

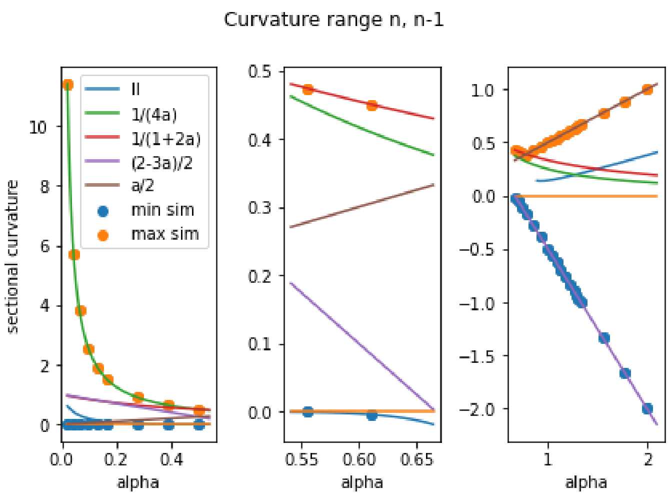

Figure 4: Numerical test for curvature range $p = n-1 \geq 4$

There, we plot the graphs of the curvatures of the list of sections as functions of α for the scenarios, and also plot the results of the numerical optimization for curvature range, for a set of 30 values of α. The optimization is done for $n = 4, p = 3$, $n = 5, p \in \{3, 4\}$, $n = 6, p \in \{3, 4, 5\}$, $n = 10, p \in \{3, 5, 10, 9\}$, $n = 100, p \in \{10, 20\}$. The curve ll in the figure is for the function $\mathfrak{l}$. The reason the numerically optimized maximum is sometimes smaller than the proposed maximum, for small α, is because the optimizer may be stuck at a local maximum.

Proposition 6.1. *If $p=2$ and $n=3$, then the sectional curvature range of* $\mathrm{St}_{p,n}$ *is* $[\frac{\alpha}{2}, \frac{2-3\alpha}{2}]$ *if* $\alpha \leq \frac{1}{2}$ *and* $[\frac{2-3\alpha}{2}, \frac{\alpha}{2}]$ *otherwise. In particular, if* $\alpha < \frac{2}{3}$, $\mathrm{St}_{2,3}$ *has strictly positive sectional curvature.*

If $p=2$ and $n>3$ then the sectional curvature range is $[0, \frac{4-3\alpha}{2}]$ *if* $\alpha \leq \frac{2}{3}$, $[\frac{2-3\alpha}{2}, 1]$ *if* $\frac{2}{3} < \alpha \leq 2$ *and* $[\frac{2-3\alpha}{2}, \frac{\alpha}{2}]$ *if* $\alpha > 2$. *Hence, when* $n > 3$, $\mathrm{St}_{2,n}$ *has non-negative curvature if* $\alpha \leq \frac{2}{3}$.

Proof. When $p=2$, $\mathfrak{o}(2)$ is one dimension so $[A_1, A_2] = 0$ and we can set $A_1 = (2\alpha)^{-1/2}c_1 J$, $A_2 = (2\alpha)^{-1/2}c_2 J$ for $J = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$, with $c_1, c_2 \in \mathbb{R}$. Further, for two orthogonal matrices in U, V of compatible dimensions, the sectional curvature is unchanged if we replace (A_1, B_1, A_2, B_2) with $(VA_1V^{\mathsf{T}}, UB_1V, VA_2V^{\mathsf{T}}, UB_2V)$. Thus, we can assume B_1 is rectangular diagonal, with diagonal entries denoted by d_i, $1 \leq i \leq \min(n-p, p)$. We denote entries of B_2 by b_{ij}, $1 \leq i \leq n-p$, $1 \leq j \leq p$. We note $B_1A_2 - B_2A_1 = (2\alpha)^{-1/2}(c_2B_1J - c_1B_2J)$, and since $JJ^{\mathsf{T}} = \mathrm{I}_2$,

$$\alpha^2 \|B_1A_1 - B_2A_2\|_F^2 = \alpha/2(c_2^2 \operatorname{Tr}(B_1B_1^{\mathsf{T}}) + c_1^2 \operatorname{Tr}(B_2B_2^{\mathsf{T}}) - 2c_1c_2 \operatorname{Tr}(B_1B_2^{\mathsf{T}})).$$

The orthogonal condition $\alpha \operatorname{Tr} A_1A_2^{\mathsf{T}} + \operatorname{Tr} B_1B_2^{\mathsf{T}} = 0$ implies

$$c_1c_2 + \operatorname{Tr} B_1B_2^{\mathsf{T}} = c_1c_2 + \sum_{i=1}^{\min(p,n-p)} d_ib_{ii} = 0,$$

or $c_1c_2 = -\operatorname{Tr} B_1B_2^{\mathsf{T}}$, so $-2c_1c_2 \operatorname{Tr} B_1B_2^{\mathsf{T}} = c_1^2c_2^2 + (\operatorname{Tr} B_1B_2)^2$. This implies

$$\alpha^2 \|B_1A_1 - B_2A_2\|_F^2 = \alpha/2(c_2^2 \operatorname{Tr}(B_1B_1^{\mathsf{T}}) + c_1^2 \operatorname{Tr}(B_2B_2^{\mathsf{T}}) + c_1^2c_2^2 + \operatorname{Tr}(B_1B_2^{\mathsf{T}})^2)$$

For the case $n=3$, from eq. (35), the curvature numerator $\hat{\mathcal{K}}$ is reduced to

$$\frac{2-3\alpha}{2} b_{12}^2 d_1^2 + \frac{\alpha}{2}(c_2^2d_1^2 + c_1^2(b_{11}^2 + b_{12}^2) + c_1^2c_2^2 + d_1^2b_{11}^2)$$

and the curvature denominator is $S = (c_1^2 + d_1^2)(c_2^2 + b_{11}^2 + b_{12}^2)$. We have

$$\hat{\mathcal{K}} - \alpha/2S = (1-2\alpha)b_{12}^2d_1^2,$$

$$\hat{\mathcal{K}} - (1-3\alpha/2)S = (2\alpha-1)(c_2^2d_1^2 + c_1^2(b_{11}^2 + b_{12}^2) + c_1^2c_2^2 + d_1^2b_{11}^2).$$

Thus, the signs of the differences are dependent on $1-2\alpha$, and $\hat{\mathcal{K}}$ is between the smaller and the larger of $\alpha/2S$ and $(1-3\alpha/2)S$. The bound is tight based on Table 1.

When $n>3$, the denominator is $S = (c_1^2 + \sum_{i=1}^2 d_i^2)(c_2^2 + \sum_{ij} b_{ij}^2)$. B_1 consists of a square diagonal block of size 2×2 and the remaining zero block of size $(n-4) \times 2$. Expand $\|B_1B_2^{\mathsf{T}} - B_2B_1^{\mathsf{T}}\|_F^2$ by dividing B_2 to a square block corresponding to indices not exceeding two, which contributes $2(b_{21}d_1 - b_{12}d_2)^2$ and the remaining blocks, which contributes $2\sum_{j=1}^2 \sum_{i>2} b_{ij}^2 d_j^2$, $\hat{\mathcal{K}}$ is

$$\frac{2-3\alpha}{2}(b_{21}d_2 - b_{12}d_1)^2 + \frac{\alpha}{2}\Big(c_2^2 \sum d_i^2 + c_1^2 \sum_{ij} b_{ij}^2 + c_1^2c_2^2 + (\sum_{i=1}^2 d_ib_{ii})^2\Big)$$
$$+ (b_{21}d_1 - b_{12}d_2)^2 + \sum_{j=1}^2 \sum_{i>2} b_{ij}^2 d_j^2.$$

The above expression appears when $\alpha \leq 2/3$, $\mathcal{K} \geq 0$. In this case, $1 \leq 2 - 3\alpha/2$, $\alpha/2 \leq 2 - 3\alpha/2$, thus $\sum_{j=1}^{2}\sum_{i>2} b_{ij}^2 d_j^2 \leq (2 - 3\alpha/2)\sum_{i=1}^{2} d_i^2 \sum_{i>2} b_{ij}^2$ and

$$\frac{\alpha}{2}\Big(c_2^2\sum d_i^2 + c_1^2\sum_{ij} b_{ij}^2 + c_1^2c_2^2\Big) \leq \frac{4-3\alpha}{2}\Big(c_2^2\sum d_i^2 + c_1^2\sum_{ij} b_{ij}^2 + c_1^2c_2^2\Big)$$

To show $\hat{\mathcal{K}} \leq (2 - 3\alpha/2)S$, we only need to show

$$\frac{2-3\alpha}{2}(b_{21}d_2 - b_{12}d_1)^2 + (b_{21}d_1 - b_{12}d_2)^2 + \frac{\alpha}{2}\Big(\sum_{i=1}^{2} d_i b_{ii}\Big)^2 \leq (2 - \frac{3\alpha}{2})\sum_{k=1}^{2} d_k^2 \sum_{i\leq 2} b_{ij}^2.$$

This follows from Cauchy-Schwarz's theorem, applying to three different combinations on the left-hand side then sum up the inequalities, as the first two terms on the left-hand side are dominated by $((2-3\alpha)/2+1)(d_1^2+d_2^2)(b_{21}^2+b_{12}^2)$, while the last one is dominated by $\alpha/2(d_1^2+d_2^2)(b_{11}^2+b_{22}^2) \leq (2-3\alpha/2)(d_1^2+d_2^2)(b_{11}^2+b_{22}^2)$.

Next, when $\alpha > 2/3$, by Cauchy-Schwarz,

$$\hat{\mathcal{K}} \geq (1 - 3\alpha/2)(b_{21}^2 + b_{12}^2)(d_1^2 + d_2^2) \geq (1 - 3\alpha/2)S,$$

as $1 - 3\alpha/2 < 0$. When $2/3 < \alpha \leq 2$, $\alpha/2 \leq 1$, thus $\mathcal{K} \leq S$, as the first term of $\hat{\mathcal{K}}$ is negative, while we can use Cauchy-Schwarz on $(\sum d_i b_{ii})^2$ and $(b_{21}d_1 - b_{12}d_2)^2$ as before. Finally, for $\alpha > 2$, $\hat{\mathcal{K}} \leq \alpha/2S$, again because the first term of $\hat{\mathcal{K}}$ is negative, while the remaining terms are dominated by the corresponding terms in $\alpha/2S$, using Cauchy-Schwarz if necessary. Again, the bounds are tight using Table 1. ■

We note $\mathrm{St}_{2,3}$ is SO(3), and could be considered as the sphere S^3 with antipodal points identified (via the quaternion representation, for example). From the formula for the metric, we see this is the projective version of the Berger sphere.

Proposition 6.2. *For $p \geq 3$, the sectional curvature range of $\mathrm{St}_{p,n}$ contains an interval $I = I(n,p,\alpha)$ as described in Table 2. The first row describes the applicable combination of (n,p), the columns labeled α_u specify the range of α where the interval formula next to it is applicable. The interval is applicable for α greater than the previous α_u (if exists) and not exceeding the current α_u.*

$(n=4,\ p=3)$		$(n,3),\ n\geq 5$		$(n,p),\ n-2\geq p\geq 4$		$(n,n-1),\ n\geq 5$	
α_u	I	α_u	I	α_u	I	α_u	I
$\frac{1}{6}$	$[0,\frac{1}{8\alpha}]$	$\frac{4-\sqrt{13}}{6}$	$[0,\frac{1}{8\alpha}]$	$\frac{4-\sqrt{10}}{6}$	$[0,\frac{1}{4\alpha}]$	$1/2$	$[0,\frac{1}{4\alpha}]$
$1/2$	$[0,\frac{1}{1+2\alpha}]$	$1/2$	$[0,\frac{4-3\alpha}{2})]$	$1/2$	$[0,\frac{4-3\alpha}{2}]$	$\frac{7}{10}$	$[\mathfrak{l}(\alpha),\frac{1}{1+2\alpha}]$
$\frac{7}{10}$	$[\mathfrak{l}(\alpha),\frac{1}{1+2\alpha}]$	$\frac{2}{3}$	$[\mathfrak{l}(\alpha),\frac{4-3\alpha}{2}]$	$\frac{2}{3}$	$[\mathfrak{l}(\alpha),\frac{4-3\alpha}{2}]$	$\frac{\sqrt{17}-1}{4}$	$[\frac{2-3\alpha}{2},\frac{1}{1+2\alpha}]$
$\frac{\sqrt{17}-1}{4}$	$[\frac{2-3\alpha}{2},\frac{1}{1+2\alpha}]$	$\frac{7}{10}$	$[\mathfrak{l}(\alpha),1]$	$\frac{7}{10}$	$[\mathfrak{l}(\alpha),1]$	∞	$[\frac{2-3\alpha}{2},\frac{\alpha}{2}]$
∞	$[\frac{2-3\alpha}{2},\frac{\alpha}{2}]$	2	$[\frac{2-3\alpha}{2},1]$	2	$[\frac{2-3\alpha}{2},1]$		
		∞	$[\frac{2-3\alpha}{2},\frac{\alpha}{2}]$	∞	$[\frac{2-3\alpha}{2},\frac{\alpha}{2}]$		

Table 2: Interval contained in the sectional curvature range of the Stiefel manifold $\mathrm{St}_{p,n}$ with metric defined by α. $\mathfrak{l}(\alpha) = \mathfrak{c}(\gamma_{\min})$ with $\mathfrak{c}$ defined in eq. (54), and $\gamma_{\min}$ in eq. (55).

For illustration: If $(n,p)=(4,3)$, for $\alpha \le \frac{1}{6}$, the sectional curvature range contains the interval $[0,\frac{1}{8\alpha}]$, for $\frac{1}{6}<\alpha\le\frac{1}{2}$, it contains the interval $[0,\frac{1}{1+2\alpha}]$, etc. In the final row, for $\alpha > \frac{\sqrt{17}-1}{4}$, it contains the interval $[\frac{2-3\alpha}{2},\frac{\alpha}{2}]$.

Proof. It is straightforward to check that for each pair (n,p) in Table 2, the values indicated correspond to a quadruple (A_1,B_1,A_2,B_2) in Table 1, which is applicable for the pair. For example, in the case $(n,p)=(4,3)$, the only applicable values from Table 1 are 0 (from the first row), $\frac{1}{2\alpha+1}$, $\frac{1}{8\alpha}$, $\mathfrak{l}(\alpha)$ and $\frac{2-3\alpha}{2}$. To show that the sectional curvature range contains I, it remains to verify the lower end of I is not greater than the upper end, which is immediate, as $\mathfrak{l}(\alpha)$ is negative between 0 and $\frac{7}{10}$, and $\frac{2-3\alpha}{2}$ is negative for $\alpha>\frac{7}{10}>\frac{2}{3}$.

The graphs in Figures 1, 2, 3, 4 display the relative values of these functions. As all the functions involved are simple algebraic functions, except for $\mathfrak{l}$, if we can assess the contribution of $\mathfrak{l}$, it will be easy to check that the lower end of I corresponds to the smallest value among the applicable values, and the upper to the largest of the applicable values.

The function $\gamma_{\min}$ from eq. (55) has a root at $\alpha_s=\frac{3+\sqrt{17}}{8}$ at around 0.89, and is negative in the interval $(\frac{7}{10},\alpha_s)$, hence $\sqrt{\gamma_{\min}}$ and B_1, B_2 for this section are not defined, so $\mathfrak{l}(\alpha)$ cannot be an extremum for $\alpha\in(\frac{7}{10},\alpha_s)$. In the interval $[\alpha_s,2]$, $\mathfrak{l}$ has the approximate range of $[0.14,0.38]$, less than 1, and in the interval $[\alpha_s,\frac{\sqrt{17}-1}{4}]$ it is less than $\frac{1}{1+2\alpha}$. For large α, $\gamma_{\min}$ is approximated by 0.8α, thus $\mathfrak{l}(\alpha)$ has an asymptote with slope $\frac{4\times0.8-3\times0.8^2/2}{2.8^2}\approx 0.286$, smaller than the slope of $\frac{\alpha}{2}$. It is also easy to graph $\mathfrak{l}$ in the interim to show beyond the contribution to the lower bound in $[1/2,\frac{7}{10}]$, $\mathfrak{l}$ has no other effect on the curvature range.

With that analysis, for the case $(n,p)=(4,3)$, the only applicable values from Table 1 are 0 (from the first row), $\frac{1}{2\alpha+1}$, $\frac{1}{8\alpha}$, $\mathfrak{l}(\alpha)$ and $\frac{2-3\alpha}{2}$.

If $\alpha<1/2$, all these functions are non-negative, and thus 0 is the smallest value among them.

When $\frac{1}{2}<\alpha<\frac{7}{10}$, $\mathfrak{l}(\alpha)$ is negative, and in the interval $[\frac{2}{3},\frac{7}{10}]$, $\frac{2-3\alpha}{2}$ is also negative, but $\mathfrak{l}(\alpha)$ is the lesser of the two, while we have discussed $\mathfrak{l}(\alpha)$ has no effect for $\alpha>\frac{7}{10}$. Thus, for $\alpha>\frac{7}{10}$ the upper end of I is $\max(\frac{1}{1+2\alpha},\frac{\alpha}{2})$, with the break-even point $\frac{\sqrt{17}-1}{4}$. In general, consider the upper or lower ends of I as functions of α, the values in column α_u corresponds to nonsmooth points of these functions or infinity.

For the case $n\ge5$, $p=n-1$, $(0,\frac{1}{4\alpha},\frac{1}{8\alpha},\frac{1}{2\alpha+1},\mathfrak{l}(\alpha),\frac{2-3\alpha}{2},\frac{\alpha}{2})$ are the applicable curvature values. Again, with $\mathfrak{l}$ having only an effect in $[\frac{2}{3},\frac{7}{10}]$, it is straightforward to verify the piecewise smooth function $\max(0,\frac{1}{4\alpha},\frac{1}{8\alpha},\frac{1}{2\alpha+1},\mathfrak{l}(\alpha),\frac{2-3\alpha}{2},\frac{\alpha}{2})$ has the form corresponding to the upper end of I, and the lower end corresponding to the minimum of those functions, for $\alpha>\frac{7}{10}$. We address the case $p\le n-2$ similarly. ■

For $\alpha=\frac{1}{2}$, when $p=n-1, n\ge4$, the range contains $[0,\frac{1}{2}]$, and it could be shown to be exactly $[0,\frac{1}{2}]$ as the manifold is isometric to $\mathrm{SO}(n)$ with a bi-invariant metric. If $2\le p\le n-2, n\ge5$, the range contains $[0,2-3\alpha/2]=[0,5/4]$, which is proved to be the exact range in [23]. For $\alpha=1$, the interval is $[-1/2,1]$. From the numerical evidence mentioned, this seems to be tight. We note for $p\ge3$, both when α is large or α is small, the curvature range becomes large.

7. Discussion

In this paper, we have obtained explicit formulas for curvatures of real Stiefel manifolds with deformation metrics and obtained several results related to Ricci curvature and sectional curvature range, including parameter values corresponding to non-negative sectional curvatures. We expect similar results could be obtained for complex and quaternionic Stiefel manifolds. We hope the availability of explicit curvature formulas for a family of metrics on an important class of manifolds will be helpful in both theory and applications. The curvature formula for diagonal homogeneous metrics has good potential for generalization. Besides the family of Einstein metrics obtained here, the approach pioneered in [2, 4, 5, 3] could be applied to yield more low parameter count families of metrics. It is also of interest to apply the symbolic approach for diagonal metrics on a low dimension group. The set of solutions to the Einstein equation on $\mathrm{SO}(p)$ of the equation has a structure based on partitions of p, which could be extended to other Lie groups, and we hope future studies may reveal additional interesting properties.

A. A few trace formulas

We collect a few results on the trace of common operators that will be useful in the computation of the Ricci curvature for matrix spaces. Some are known, item 3 gives us the Killing form of $\mathfrak{o}(p)$, but we do not have the exact references.

Lemma A.1. (1) *Let X be a matrix in $\mathbb{R}^{m\times n}$. The trace of the operator $X \mapsto AXB$ where $A \in \mathbb{R}^{m\times m}$ and $B \in \mathbb{R}^{n\times n}$ is $\mathrm{Tr}(A)\,\mathrm{Tr}(B)$. In particular, the trace of $X \mapsto AX$ is $n\,\mathrm{Tr}(A)$, the trace of $X \mapsto XB$ is $m\,\mathrm{Tr}(B)$. The trace of the operator $X \mapsto AX^{\mathsf{T}}B$ where A and B are matrices of size $m\times n$ is $\mathrm{Tr}(AB^{\mathsf{T}})$.*

(2) *The trace of the operator $X \mapsto AXB + B^{\mathsf{T}}XA^{\mathsf{T}}$ from the space Sym_p to itself is $\mathrm{Tr}(A)\,\mathrm{Tr}(B) + \mathrm{Tr}(AB^{\mathsf{T}})$. In particular, the trace of the operator $X \mapsto AX + XA^{\mathsf{T}}$ is $(p+1)\,\mathrm{Tr}(A)$. The trace of the operator $X \mapsto \mathrm{Tr}(AX)B$, with B is a symmetric matrix and A is a $p\times p$ matrix is $\mathrm{Tr}(\frac{1}{2}(A + A^{\mathsf{T}})B)$.*

(3) *The trace of the operator $X \mapsto AXB + B^{\mathsf{T}}XA^{\mathsf{T}}$, from the space $\mathfrak{o}(p)$ to itself, where A and B are $p\times p$ matrices, is $\mathrm{Tr}(A)\,\mathrm{Tr}(B) - \mathrm{Tr}(AB^{\mathsf{T}})$. In particular, if A and B are antisymmetric matrices then the trace of $X \mapsto [[AX]B]$ is $(2-p)\,\mathrm{Tr}(AB)$.*

(4) *Let P and Q be two diagonal operators on $\mathbb{R}^{p\times p}$, that is for $X \in \mathbb{R}^{p\times p} = (X_{ij})_{1\leq i,j\leq 1}$, $P(X)_{ij} = p_{ij}X_{ij}$, $Q(X)_{ij} = q_{ij}X_{ij}$ for two matrices $(p_{ij})_{1\leq i,j\leq p}$, $(q_{ij})_{1\leq i,j\leq p}$. Assume further that they are symmetric, $p_{ij} = p_{ji}$, $q_{ij} = q_{ji}$. Let $A = (a_{ij})_{1\leq i,j\leq p}$ and $B = (b_{ij})_{1\leq i,j\leq p}$ be two matrices in $\mathbb{R}^{p\times p}$. Then the operator $X \mapsto Q\{AP(BX) + P(XB^{\mathsf{T}})A^{\mathsf{T}}\}, X \in \mathfrak{o}(p)$ operates on $\mathfrak{o}(p)$. We have*

$$\begin{aligned}\mathrm{Trw}_{0,(A,B,Q,P)} &:= \mathrm{Tr}(X \mapsto Q\{AP(BX) + P(XB^{\mathsf{T}})A^{\mathsf{T}}\}) \\ &= \sum_{i=1}^{p}\sum_{j=1}^{p} a_{ji}b_{ij}(\sum_{l=1,l\neq j}^{p} q_{jl}p_{il})\end{aligned} \tag{56}$$

In particular, $\mathrm{Tr}(X \mapsto Q\{AP(X) + P(X)A\})$ is $\sum_{i=1}^{p} a_{ii}\sum_{l=1,l\neq j} q_{il}p_{il}$. If A and B are antisymmetric, $\mathrm{Tr}(X \mapsto Q(AP(BX) + P(XB)A))$ is

$$-\sum_{i=1}^{p}\sum_{j=i+1}^{p} a_{ij}b_{ij}\{q_{ij}(p_{ii} + p_{jj}) + \sum_{l=1,l\neq i,l\neq j}^{p} (q_{il}p_{jl} + q_{jl}p_{il})\} \tag{57}$$

(5) *With the same assumption as above, the operator*

$$X \mapsto Q(AP(XB) + P(B^{\mathsf{T}}X)A^{\mathsf{T}})$$

operates on $\mathfrak{o}(p)$*, with trace*

$$\mathrm{Trw}_{1,(A,B,Q,P)} := \sum_{i<j} q_{ij}(p_{ij}a_{ii}b_{jj} - p_{jj}a_{ij}b_{ij} + p_{ij}b_{ii}a_{jj} - p_{ii}b_{ji}a_{ji}) \tag{58}$$

In particular, if B *is diagonal,* Trw_1 *is* $\sum_{i=1}^{p}\sum_{j=1,j\neq i}^{p} q_{ij}p_{ij}a_{ii}b_{jj}$*, if* A *and* B *are antisymmetric, the trace is* $\sum_{i<j} -q_{ij}a_{ij}b_{ij}(p_{ii}+p_{jj})$*.*

(6) *If* A *and* B *are antisymmetric then* $\mathrm{Tr}(X \mapsto Q[A, P[X, B]])$ *is* $\mathrm{Trw}_1 - \mathrm{Trw}_0$*, or*

$$\mathrm{Trw}_{\mathrm{ad},(A,B,Q,P)} := \sum_{i<j} a_{ij}b_{ij} \sum_{l=1,l\neq i,l\neq j}^{p} (q_{il}p_{jl} + q_{jl}p_{il}) \tag{59}$$

Proof. Let E_{ij} be the matrix with the ij-entry equal to 1, and other entries equal to 0 and of the same size as X. All the statements are proved similarly, by summing the coefficients of the operators on an appropriate base based on E_{ij}. Let entries of A be a_{ij} and entries of B be b_{ij}.

(1): $(AE_{ij}B)_{ij} = a_{ii}b_{jj}$, so the trace of $X \mapsto AXB$ is $\sum_{ij} a_{ii}b_{jj} = \mathrm{Tr}(A)\,\mathrm{Tr}(B)$. Since $(AE_{ij}^{\mathsf{T}}B)_{ij} = a_{ij}b_{ij}$, $\mathrm{Tr}(X \mapsto AX^{\mathsf{T}}B)$ is $\sum_{ij} a_{ij}b_{ij} = \mathrm{Tr}(AB^{\mathsf{T}})$.

(2): A basis of Sym_p consists of matrices E_{ii} $(i = 1, \cdots, p)$ and $E_{ij} + E_{ji}$ for $i < j$. We now compute the trace of $X \mapsto AXB + B^{\mathsf{T}}XA^{\mathsf{T}}$ with respect to this basis. For E_{ii}, $(AE_{ii}B + B^{\mathsf{T}}E_{ii}A^{\mathsf{T}})_{ii} = 2a_{ii}b_{ii}$, for $E_{ij} + E_{ji}$, the coefficient is

$$(A(E_{ij} + E_{ji})B + B^{\mathsf{T}}(E_{ij} + E_{ji})A^{\mathsf{T}})_{ij} = a_{ii}b_{jj} + a_{ij}b_{ji} + b_{ii}a_{jj} + b_{ij}a_{ij}$$

Hence the trace is

$$\sum_i 2a_{ii}b_{ii} + \sum_{i<j}(a_{ii}b_{jj} + a_{ij}b_{ij} + b_{ii}a_{jj} + b_{ij}a_{ij}) = \sum_i a_{ii}\sum_j b_{jj} + \sum_{ij} a_{ij}b_{ij}$$

which is $\mathrm{Tr}(A)\,\mathrm{Tr}(B) + \mathrm{Tr}(AB^{\mathsf{T}})$, as $\sum_i a_{ii}b_{ii} + \sum_{i<j}(a_{ii}b_{jj} + b_{ii}a_{jj})$ rearranges to the first sum, and the sum of remaining terms is $\mathrm{Tr}(AB^{\mathsf{T}})$.

With $B = \mathrm{I}_p$ we have the trace of $X \mapsto AX + XA^{\mathsf{T}}$ is $(p+1)\,\mathrm{Tr}(A)$.

For the operator $X \mapsto \mathrm{Tr}(AX)B$, the coefficient corresponding to E_{ii} is $a_{ii}b_{ii}$, corresponding to $E_{ij} + E_{ji}$ is $(a_{ij} + a_{ji})b_{ij}$. The trace is

$$\sum_i a_{ii}b_{ii} + \sum_{i<j}(a_{ij} + a_{ji})b_{ij} = \frac{1}{2}\,\mathrm{Tr}((A + A^{\mathsf{T}})B)$$

(3): A basis of $\mathfrak{o}(p)$ consists of matrices $E_{ij} - E_{ji}$ for $i < j$. The coefficient corresponds to $E_{ij} - E_{ji}$ is

$$(A(E_{ij} - E_{ji})B + B^{\mathsf{T}}(E_{ij} - E_{ji})A^{\mathsf{T}})_{ij} = a_{ii}b_{jj} - a_{ij}b_{ij} + b_{ii}a_{jj} - b_{ji}a_{ji}$$

The trace is

$$\sum_{i<j} a_{ii}b_{jj} - a_{ij}b_{ij} + b_{ii}a_{jj} - b_{ji}a_{ji} = \sum_{ij} a_{ii}b_{jj} - \sum_{ij} a_{ij}b_{ij},$$

which is $\mathrm{Tr}(A)\,\mathrm{Tr}(B) - \mathrm{Tr}(AB^{\mathsf{T}})$.

For the trace of

$$X \mapsto [[AX]B] = (AX - XA)B - B(AX - XA) = AXB + BXA - BAX - XAB,$$

we have:

$$\mathrm{Tr}(X \mapsto AXB + BXA) = -\mathrm{Tr}(AB^{\mathsf{T}}) = \mathrm{Tr}(AB)$$
$$\mathrm{Tr}(X \mapsto BAX + XAB) = \mathrm{Tr}(\mathrm{I}_p)\,\mathrm{Tr}(BA) - \mathrm{Tr}(BA) = (p-1)\,\mathrm{Tr}(BA)$$

from here we get $\mathrm{Tr}(X \mapsto [[AX]B]) = (2-p)\,\mathrm{Tr}(AB)$.
(4): We note the symmetric assumption means $P(E_{ij}^{\mathsf{T}}) = p_{ji}E_{ji} = P(E_{ij})^{\mathsf{T}}$, so $P(X^{\mathsf{T}}) = P(X)^{\mathsf{T}}$ and similarly $Q(X^{\mathsf{T}}) = Q(X)^{\mathsf{T}}$ for $X \in \mathbb{R}^{p\times p}$, so it is clear $X \mapsto Q\{AP(BX) + P(XB^{\mathsf{T}})A^{\mathsf{T}}\}$ maps $\mathfrak{o}(p)$ to itself. We have

$$(AP(B(E_{ij} - E_{ji})) + P((E_{ij} - E_{ji})B^{\mathsf{T}})A^{\mathsf{T}})_{ij} = \sum_l p_{lj}a_{il}b_{li} + p_{il}b_{lj}a_{jl}$$

We can see this by first noting the terms ending or starting with E_{ji} does not contribute, then enumerate for each remaining term on the left, the possible three steps paths from i to j, starting or ending with E_{ij} and count the contribution applying the corresponding factor p. Applying Q contribute a factor q_{ij}, thus the trace is $\sum_{i<j}\sum_{l=1}^{p}\sum_l q_{ij}p_{lj}a_{il}b_{li} + q_{ij}p_{il}b_{lj}a_{jl}$. For each pair of integer $1 \leq r, s \leq p$, the coefficient of $a_{sr}b_{rs}$ from the first term of the sum is $\sum_{j>s} q_{sj}p_{rj}$, in the second term is $\sum_{i<s} q_{is}p_{ir}$. Combining, the trace is $\sum_{rs}\sum_{l\neq s} a_{sr}b_{rs}\sum_{l\neq s} q_{sl}p_{rl}$ as stated. For the case $B = \mathrm{I}_p$ we get the next statement. Flipping indices when $r > s$, the antisymmetric case follows from

$$\sum_{rs}(-a_{rs}b_{rs}(\sum_{l\neq s} q_{sl}p_{rl}) = -\sum_{r<s} a_{rs}b_{rs}(q_{rs}p_{rr} + q_{rs}p_{ss} + \sum_{l\neq s, l\neq r}(q_{sl}p_{rl} + q_{rl}p_{sl}))$$

(5): We have

$$(AP((E_{ij} - E_{ji})B) + P(B^{\mathsf{T}}(E_{ij} - E_{ji}))A^{\mathsf{T}})_{ij} = p_{ij}a_{ii}b_{jj} - p_{jj}a_{ij}b_{ij} + p_{ij}b_{ii}a_{jj} - p_{ii}b_{ji}a_{ji}$$

(each term on the left contributes a 3-step path going from i to j, bridged by E_{ij} and E_{ji}). Applying Q would add a factor of q_{ij}, thus the trace is

$$\sum_{i<j} q_{ij}(p_{ij}a_{ii}b_{jj} - p_{jj}a_{ij}b_{ij} + p_{ij}b_{ii}a_{jj} - p_{ii}b_{ji}a_{ji})$$

The special cases of (5) are clear, and (6) is also clear. ■

B. Direct calculation of the curvature of a Stiefel manifold

We use eq. (24) to prove the curvature formulas in theorem 4.1 directly. First, with Γ given by eq. (29)

$$\begin{aligned}\mathrm{D}_\xi\,\Gamma(\eta,\phi) &= \frac{1}{2}\xi(\eta^{\mathsf{T}}\phi + \phi^{\mathsf{T}}\eta)\\ &\quad + (1-\alpha)\{(\mathrm{I}_n - YY^{\mathsf{T}})(\eta\phi^{\mathsf{T}} + \phi\eta^{\mathsf{T}})\xi - (\xi Y^{\mathsf{T}} + Y\xi^{\mathsf{T}})(\eta\phi^{\mathsf{T}} + \phi\eta^{\mathsf{T}})Y\}\end{aligned}$$

Expanding ξ, η, ϕ we obtain

$$Y_\perp^\mathsf{T}(\mathrm{D}_\xi \Gamma)(\eta, \phi) = \frac{1}{2}B_1(-A_2A_3 - A_3A_2 + B_2^\mathsf{T}B_3 + B_3^\mathsf{T}B_2)$$
$$+ (1-\alpha)\{B_2(-A_3Y^\mathsf{T} + B_3^\mathsf{T}Y_\perp) + B_3(-A_2Y^\mathsf{T} + B_2^\mathsf{T}Y_\perp)\}(YA_1 + Y_\perp B_1)$$
$$- (1-\alpha)(B_1Y^\mathsf{T}(-YA_2A_3 - YA_3A_2)$$
$$= \frac{B_1B_2^\mathsf{T}B_3}{2} + \frac{B_1B_3^\mathsf{T}B_2}{2} + (\frac{1}{2} - \alpha)(B_1A_2A_3 + B_1A_3A_2)$$
$$+ (1-\alpha)(-B_2A_3A_1 + B_2B_3^\mathsf{T}B_1 - B_3A_2A_1 + B_3B_2^\mathsf{T}B_1).$$

Simplify as follows: $Y^\mathsf{T}(\xi Y^\mathsf{T} + Y\xi^\mathsf{T}) = A_1Y^\mathsf{T} - A_1Y^\mathsf{T} + B_1^\mathsf{T}Y_\perp^\mathsf{T} = B_1^\mathsf{T}Y_\perp^\mathsf{T}$. Then

$$Y^\mathsf{T}(\mathrm{D}_\xi \Gamma)(\eta, \phi) = \frac{1}{2}A_1(-A_2A_3 + B_2^\mathsf{T}B_3 - A_3A_2 + B_3^\mathsf{T}B_2)$$
$$- (1-\alpha)(B_1^\mathsf{T}Y_\perp^\mathsf{T})(-Y_\perp B_2A_3Y^\mathsf{T} - Y_\perp B_3A_2Y^\mathsf{T})Y$$
$$= (1-\alpha)(B_1^\mathsf{T}B_2A_3 + B_1^\mathsf{T}B_3A_2) + \frac{1}{2}(-A_1A_2A_3 - A_1A_3A_2 + A_1B_2^\mathsf{T}B_3 + A_1B_3^\mathsf{T}B_2)$$

As noted, any $\omega \in \mathbb{R}^{n\times p}$ could be expressed as $\omega = YA + Y_\perp B$, however A may not be antisymmetric. By direct substitution we have

$$(\mathrm{I}_n - YY^\mathsf{T})(\eta\omega^\mathsf{T} + \omega\eta^\mathsf{T})Y = Y_\perp(B_2A^\mathsf{T} - BA_2)$$

and $$\Gamma(\eta, \omega) = \frac{1}{2}Y(-A_2A + A^\mathsf{T}A_2 + B^\mathsf{T}B_2 + B_2^\mathsf{T}B) + (1-\alpha)Y_\perp(B_2A^\mathsf{T} - BA_2)$$

In particular we have

$$Y^\mathsf{T}\Gamma(\eta, \omega) = \frac{1}{2}(-A_2A + A^\mathsf{T}A_2 + B^\mathsf{T}B_2 + B_2^\mathsf{T}B),$$

and $$Y_\perp^\mathsf{T}\Gamma(\eta, \omega) = (1-\alpha)(B_2A^\mathsf{T} - BA_2).$$

Use the formula for $\Gamma(\xi, \omega)$ with $\omega = \Gamma(\eta, \phi)$

$$Y^\mathsf{T}\Gamma(\xi, \Gamma(\eta, \phi)) = \frac{1}{2}(-A_1(\frac{1}{2}(-A_2A_3 - A_3A_2 + B_3^\mathsf{T}B_2 + B_2^\mathsf{T}B_3))$$
$$+ (\frac{1}{2}(-A_2A_3 - A_3A_2 + B_3^\mathsf{T}B_2 + B_2^\mathsf{T}B_3))^\mathsf{T}A_1 + B_1^\mathsf{T}((1-\alpha)(-B_2A_3 - B_3A_2))$$
$$+ ((1-\alpha)(-B_2A_3 - B_3A_2))^\mathsf{T}B_1)$$
$$= \frac{1-\alpha}{2}(A_2B_3^\mathsf{T}B_1 + A_3B_2^\mathsf{T}B_1 - B_1^\mathsf{T}B_2A_3 - B_1^\mathsf{T}B_3A_2) + \frac{1}{4}(A_1A_2A_3$$
$$+ A_1A_3A_2 - A_1B_2^\mathsf{T}B_3 - A_1B_3^\mathsf{T}B_2 - A_2A_3A_1 - A_3A_2A_1 + B_2^\mathsf{T}B_3A_1 + B_3^\mathsf{T}B_2A_1)$$
$$Y_\perp(\Gamma(\xi, \Gamma(\eta, \phi)) = (1-\alpha)\{B_1(\frac{1}{2}(-A_2A_3 - A_3A_2 + B_3^\mathsf{T}B_2 + B_2^\mathsf{T}B_3)^\mathsf{T}$$
$$+ ((1-\alpha)(-B_2A_3 - B_3A_2))A_1)\}$$
$$= (\alpha-1)^2(B_2A_3A_1 + B_3A_2A_1) + \frac{\alpha-1}{2}(B_1A_2A_3 + B_1A_3A_2 - B_1B_2^\mathsf{T}B_3 - B_1B_3^\mathsf{T}B_2)$$

Therefore:

$$
\begin{aligned}
Y^{\mathsf{T}}\,\mathrm{R}^{\mathcal{M}}{}_{\xi\eta}\,\phi = &-\Big\{(1-\alpha)(B_1^{\mathsf{T}}B_2A_3 + B_1^{\mathsf{T}}B_3A_2)\\
&\quad+\frac{1}{2}(-A_1A_2A_3 - A_1A_3A_2 + A_1B_2^{\mathsf{T}}B_3 + A_1B_3^{\mathsf{T}}B_2)\Big\}\\
&+\{(1-\alpha)(B_2^{\mathsf{T}}B_1A_3+B_2^{\mathsf{T}}B_3A_1)+\frac{1}{2}(-A_2A_1A_3-A_2A_3A_1+A_2B_1^{\mathsf{T}}B_3+A_2B_3^{\mathsf{T}}B_1)\}\\
&-\{\frac{1-\alpha}{2}(A_2B_3^{\mathsf{T}}B_1 + A_3B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2A_3 - B_1^{\mathsf{T}}B_3A_2) + \frac{1}{4}(A_1A_2A_3\\
&+A_1A_3A_2 - A_1B_2^{\mathsf{T}}B_3 - A_1B_3^{\mathsf{T}}B_2 - A_2A_3A_1 - A_3A_2A_1 + B_2^{\mathsf{T}}B_3A_1 + B_3^{\mathsf{T}}B_2A_1)\}\\
&+\{\frac{1-\alpha}{2}(A_1B_3^{\mathsf{T}}B_2 + A_3B_1^{\mathsf{T}}B_2 - B_2^{\mathsf{T}}B_1A_3 - B_2^{\mathsf{T}}B_3A_1) + \frac{1}{4}(A_2A_1A_3\\
&+A_2A_3A_1 - A_2B_1^{\mathsf{T}}B_3 - A_2B_3^{\mathsf{T}}B_1 - A_1A_3A_2 - A_3A_1A_2 + B_1^{\mathsf{T}}B_3A_2 + B_3^{\mathsf{T}}B_1A_2)\}\\
=&\ \frac{1-2\alpha}{4}\,(A_1B_3^{\mathsf{T}}B_2 - A_2B_3^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_3A_2 + B_2^{\mathsf{T}}B_3A_1)\\
&+\frac{1-\alpha}{2}(A_3B_1^{\mathsf{T}}B_2 - A_3B_2^{\mathsf{T}}B_1 - B_1^{\mathsf{T}}B_2A_3 + B_2^{\mathsf{T}}B_1A_3) + \frac{1}{4}(A_1A_2A_3 - A_1B_2^{\mathsf{T}}B_3\\
&-A_2A_1A_3 + A_2B_1^{\mathsf{T}}B_3 - A_3A_1A_2 + A_3A_2A_1 + B_3^{\mathsf{T}}B_1A_2 - B_3^{\mathsf{T}}B_2A_1).
\end{aligned}
$$

The last expression follows from a term by term collection, for example, the coefficient of $A_1A_2A_3$ is $-(-1/2)-1/4 = 1/4$, and similarly for all terms with coefficient $1/4$. The coefficient for $A_1B_3^{\mathsf{T}}B_2$ is $-1/2+1/4+(1-\alpha)/2 = (1-2\alpha/4)$, and similar to all the terms with that coefficient.

$$
\begin{aligned}
Y_{\perp}^{\mathsf{T}}\,\mathrm{R}^{\mathcal{M}}{}_{\xi\eta}\,\phi = &-(\frac{B_1B_2^{\mathsf{T}}B_3}{2} + \frac{B_1B_3^{\mathsf{T}}B_2}{2} + (\frac{1}{2}-\alpha)(B_1A_2A_3 + B_1A_3A_2)\\
&+(1-\alpha)(-B_2A_3A_1 + B_2B_3^{\mathsf{T}}B_1 - B_3A_2A_1 + B_3B_2^{\mathsf{T}}B_1))\\
&+(\frac{B_2B_1^{\mathsf{T}}B_3}{2} + \frac{B_2B_3^{\mathsf{T}}B_1}{2} + (\frac{1}{2}-\alpha)(B_2A_1A_3 + B_2A_3A_1)\\
&+(1-\alpha)(-B_1A_3A_2 + B_1B_3^{\mathsf{T}}B_2 - B_3A_1A_2 + B_3B_1^{\mathsf{T}}B_2))\\
&-(\alpha-1)^2(B_2A_3A_1+B_3A_2A_1) - \frac{\alpha-1}{2}(B_1A_2A_3+B_1A_3A_2-B_1B_2^{\mathsf{T}}B_3-B_1B_3^{\mathsf{T}}B_2)\\
&+(\alpha-1)^2(B_1A_3A_2+B_3A_1A_2) + \frac{\alpha-1}{2}(B_2A_1A_3+B_2A_3A_1-B_2B_1^{\mathsf{T}}B_3-B_2B_3^{\mathsf{T}}B_1).
\end{aligned}
$$

Again, we collect term by term, (we do use a symbolic calculation program helper). The coefficient for $B_1B_2^{\mathsf{T}}B_3$ is $-1/2+(\alpha-1)/2 = (\alpha-2)/2$, and similar for $B_2B_1^{\mathsf{T}}B3$. The coefficient for $B_1B_3^{\mathsf{T}}B_2$ is $-1/2+(1-\alpha)+(\alpha-1)/2 = -\alpha/2$, and similar for $B_2B_3^{\mathsf{T}}B_1$. The coefficient for $B_1A_2A_3$ is $-(1/2-\alpha)-(\alpha-1)/2 = \alpha/2$, and similar for $B_2A_1A_3$. The coefficient for $B_1A_3A_2$ is

$$
-(\tfrac{1}{2}-\alpha)-(1-\alpha)-\tfrac{\alpha-1}{2}+(\alpha-1)^2 = \alpha^2-\tfrac{\alpha}{2} = \tfrac{2\alpha^2-\alpha}{2}
$$

and similar for $B_2A_3A_1$. The coefficient for $B_3A_2A_1$ is $(1-\alpha)-(\alpha-1)^2 = \alpha-\alpha^2$, and $B_3A_1A_2$ follows by permutation. The coefficient for $B_3B_2^{\mathsf{T}}B_1$ is $-(1-\alpha)$, and similar for $B_3B_1^{\mathsf{T}}B_2$. Finally we have

$$
\begin{aligned}
Y_{\perp}^{\mathsf{T}}\,\mathrm{R}^{\mathcal{M}}{}_{\xi\eta}\,\phi = &\ \frac{2\alpha^2-\alpha}{2}(B_1A_3A_2 - B_2A_3A_1) + (\alpha^2-\alpha)(B_3A_1A_2 - B_3A_2A_1)\\
&+(1-\alpha)(B_3B_1^{\mathsf{T}}B_2 - B_3B_2^{\mathsf{T}}B_1) + \frac{\alpha-2}{2}(B_1B_2^{\mathsf{T}}B_3 - B_2B_1^{\mathsf{T}}B_3)\\
&+\frac{\alpha}{2}(B_1A_2A_3 - B_1B_3^{\mathsf{T}}B_2 - B_2A_1A_3 + B_2B_3^{\mathsf{T}}B_1).
\end{aligned}
$$

■

References

[1] V. Arnold: *Sur la géométrie différentielle des groupes de Lie de dimension infinie et ses applications à l'hydrodynamique des fluides parfaits*, Ann. l'Inst. Fourier 16/1 (1966) 319–361.

[2] A. Arvanitoyeorgos, V. V. Dzhepko, Yu. G. Nikonorov: *Invariant Einstein metrics on some homogeneous spaces of classical Lie groups*, Canadian J. Math. 61/6 (2009) 1201–1213.

[3] A. Arvanitoyeorgos, K. Mori, Y. Sakane: *Einstein metrics on compact Lie groups which are not naturally reductive*, Geom. Dedicata 160/1 (2012) 261–285.

[4] A. Arvanitoyeorgos, Y. Sakane, M. Statha: *New homogeneous Einstein metrics on Stiefel manifolds*, Diff. Geometry Appl. 35 (2013) 2–18.

[5] A. Arvanitoyeorgos, Y. Sakane, M. Statha: *New Einstein metrics on the Lie group SO(n) which are not naturally reductive*, Geom. Imaging Comp. 2/2 (2015) 77–108.

[6] A. L. Besse: *Einstein Manifolds*, Classics in Mathematics, Springer, Berlin (2007).

[7] J. Cheeger: *Some examples of manifolds of nonnegative curvature*, J. Diff. Geometry 8/4 (1973) 623–628.

[8] J. E. D'Atri, W. Ziller: *Naturally Reductive Metrics and Einstein Metrics on Compact Lie Groups*, Memoirs of the American Mathematical Society 215, American Mathematical Society, Providence (1979).

[9] A. Edelman, T. A. Arias, S. T. Smith: *The geometry of algorithms with orthogonality constraints*, SIAM J. Matrix Analysis Appl. 202 (1999) 303–353.

[10] J. Ge: *DDVV-type inequality for skew-symmetric matrices and Simons-type inequality for Riemannian submersions*, Advances Math. 251 (2014) 62–86.

[11] K. Grove, H. Karcher, E. Ruh: *Group actions and curvature*, Invent. Math. 23 (1974) 31–48.

[12] K. Grove, W. Ziller: *Curvature and symmetry of Milnor spheres*, Annals Math. 152 (2000) 331–367.

[13] K. Hüper, I. Markina, F. Silva Leite: *A Lagrangian approach to extremal curves on Stiefel manifolds*, J. Geometric Mech. 13/1 (2021) 55–72.

[14] G. R. Jensen: *Einstein metrics on principal fibre bundles*, J. Diff. Geometry 8/4 (1973) 599–614.

[15] P. W. Michor: *Some geometric evolution equations arising as geodesic equations on groups of diffeomorphisms including the Hamiltonian approach*, in: *Phase Space Analysis of Partial Differential Equations*, A. Bove, F. Colombini, D. Del Santo (eds.), Birkhäuser, Boston (2007) 133–215.

[16] J. Milnor: *Curvatures of left invariant metrics on Lie groups*, Advances Math. 21/3 (1976) 293–329.

[17] D. Nguyen: *Operator-valued formulas for Riemannian gradient and Hessian and families of tractable metrics in optimization and machine learning*, arXiv: 2009.10159 (2020).

[18] D. Nguyen: *Riemannian geometry with differentiable ambient space and metric operator*, arXiv: 2105.01583 (2021).

[19] B. O'Neill: *The fundamental equations of a submersion*, Michigan Math. J. 13/4 (1966) 459–469.

[20] B. O'Neill: *Semi-Riemannian Geometry with Applications to Relativity*, Pure and Applied Mathematics 103, Academic Press, New York (1983).

[21] T. Püttmann: *Optimal pinching constants of odd dimensional homogeneous spaces*, Invent. Math. 138/3 (1999) 631–684.

[22] T. Rapcsak: *Sectional curvatures in nonlinear optimization*, J. Global Optimization 40 (2008) 375–388.

[23] Q. Rentmeesters: *Algorithms for Data Fitting on Some Common Homogeneous Spaces*, Université Catholique de Louvain, Louvain (2013).

[24] Z. Yan, S. Deng: *Einstein metrics on compact simple Lie groups attached to standard triples*, Trans. Amer. Math. Soc. 369 (2016) 8587–8605.

[25] H. Zhang, S. Sra: *First-order methods for geodesically convex optimization*, in: Proceedings of Machine Learning Research 49 (2016) 1617–1638.

[26] W. Ziller: *Examples of Riemannian manifolds with non-negative sectional curvature*, in: *Metric and Comparison Geometry*, K. Grove, J. Cheeger (eds.), Surveys in Differential Geometry 11, International Press, Boston (2007) 63–102.

Du Nguyen, 1 Gardiner Street, Darien, CT 06820, U.S.A.; nguyendu@post.harvard.edu

Received May 19, 2021
and in final form January 28, 2022