Building Neural Networks on Matrix Manifolds: A Gyrovector Space Approach

Xuan Son Nguyen [1]   Shuo Yang [1]

Abstract

Matrix manifolds, such as manifolds of Symmetric Positive Definite (SPD) matrices and Grassmann manifolds, appear in many applications. Recently, by applying the theory of gyrogroups and gyrovector spaces that is a powerful framework for studying hyperbolic geometry, some works have attempted to build principled generalizations of Euclidean neural networks on matrix manifolds. However, due to the lack of many concepts in gyrovector spaces for the considered manifolds, e.g., the inner product and gyroangles, techniques and mathematical tools provided by these works are still limited compared to those developed for studying hyperbolic geometry. In this paper, we generalize some notions in gyrovector spaces for SPD and Grassmann manifolds, and propose new models and layers for building neural networks on these manifolds. We show the effectiveness of our approach in two applications, i.e., human action recognition and knowledge graph completion.

1. Introduction

Deep neural networks (DNNs) usually assume Euclidean geometry in their computations. However, in many applications, data exhibit a strongly non-Euclidean latent structure such as those lying on Riemannian manifolds (Bronstein et al., 2017). Therefore, a lot of effort has been put into building DNNs on Riemannian manifolds in recent years. The most common representation spaces are Riemannian manifolds of constant non-zero curvature, e.g., spherical and hyperbolic spaces (Ganea et al., 2018; Skopek et al., 2020). Beside having closed-form expressions for the distance function, exponential and logarithmic maps, and parallel transport that ease the task of learning parametric models, such spaces also have the nice algebraic structure of gyrovector

spaces (Ungar, 2014) that enables principled generalizations of DNNs to the manifold setting (Ganea et al., 2018). Another type of Riemannian manifolds referred to as matrix manifolds (Absil et al., 2007), where elements can be represented in the form of matrix arrays, are also popular in representation learning. Typical examples are SPD and Grassmann manifolds. Unlike the works in Ganea et al. (2018); Skopek et al. (2020), the works in (Huang & Gool, 2017; Dong et al., 2017; Huang et al., 2018; Nguyen et al., 2019a;b; 2020; Nguyen, 2021; Wang et al., 2021) approach the problem of generalizing DNNs to the considered manifolds in an unprincipled way. This makes it hard for them to generalize a broad class of DNNs to these manifolds. Another line of research (Chakraborty et al., 2018; Chakraborty et al., 2020; Weiler et al., 2021; Banerjee et al., 2022; Xu et al., 2022) that proposes analogs of convolutional neural networks (CNNs) on Riemannian manifolds relies on the notion of equivariant neural networks (Bronstein et al., 2017). However, these works mainly focus on building CNNs where many essential building blocks of DNNs for solving a wide range of problems are missing.

Recently, some works (Kim, 2020; Nguyen, 2022b) have attempted to explore analogies that SPD and Grassmann manifolds share with Euclidean and hyperbolic spaces. Although these works show how to construct some basic operations, e.g., the binary operation and scalar multiplication from the Riemannian geometry of the considered manifolds, it might be difficult to obtain closed-form expressions for such operations even if closed-form expressions for the exponential and logarithmic maps, and the parallel transport exist. This is the case of Grassmann manifolds from the ONB (orthonormal basis) perspective (Edelman et al., 1998; Bendokat et al., 2020), where the expressions of the exponential map and parallel transport are based on singular value decomposition (SVD). In some applications, it is more advantageous (Bendokat et al., 2020) to represent points on Grassmann manifolds with orthogonal matrices (the ONB perspective) than with projection matrices. Furthermore, due to the lack of some concepts in gyrovector spaces for the considered matrix manifolds, e.g., the inner product and gyroangles (Ungar, 2014), it is not trivial for these works to generalize many traditional machine learning models, e.g., multinomial logistic regression (MLR) to the considered manifolds.

[1]ETIS, UMR 8051, CY Cergy Paris Université, ENSEA, CNRS, Cergy, France. Correspondence to: Xuan Son Nguyen <xuanson.nguyen@ensea.fr>.

In this paper, we propose a new method for constructing the basic operations and gyroautomorphism of Grassmann manifolds from the ONB perspective. We also improve existing works by generalizing some notions in gyrovector spaces for SPD and Grassmann manifolds. This leads to the development of MLR on SPD manifolds and isometric models on SPD and Grassmann manifolds, which we refer to as SPD and Grassmann gyroisometries. These are the counterparts of Euclidean isometries on SPD and Grassmann manifolds. Our motivation for studying such isometries is that they can be seen as transformations that deform the manifold without affecting its local structure. In the context of geometric deep learning (GDL) (Bronstein et al., 2017) on manifolds, one aims to construct functions acting on signals defined on a manifold that are invariant to isometries. This invariance property, referred to as geometric stability, is one of the geometric principles in GDL for learning stable representations of high-dimensional data. Therefore, the characterization of SPD and Grassmann gyroisometries is one of the first steps to build such functions (e.g., neural networks) on the considered manifolds.

2. Proposed Approach

2.1. Notations

We adopt the notations used in Nguyen (2022b). Let $\mathcal{M}$ be a homogeneous Riemannian manifold, $T_{\mathbf{P}}\mathcal{M}$ be the tangent space of $\mathcal{M}$ at $\mathbf{P} \in \mathcal{M}$. Denote by $\exp(\mathbf{P})$ and $\log(\mathbf{P})$ the usual matrix exponential and logarithm of $\mathbf{P}$, $\mathrm{Exp}_{\mathbf{P}}(\mathbf{W})$ the exponential map at $\mathbf{P}$ that associates to a tangent vector $\mathbf{W} \in T_{\mathbf{P}}\mathcal{M}$ a point of $\mathcal{M}$, $\mathrm{Log}_{\mathbf{P}}(\mathbf{Q})$ the logarithmic map of $\mathbf{Q} \in \mathcal{M}$ at $\mathbf{P}$, $\mathcal{T}_{\mathbf{P}\rightarrow\mathbf{Q}}(\mathbf{W})$ the parallel transport of $\mathbf{W}$ from $\mathbf{P}$ to $\mathbf{Q}$ along geodesics connecting $\mathbf{P}$ and $\mathbf{Q}$, $D\phi_{\mathbf{P}}(\mathbf{W})$ the directional derivative of map ϕ at point $\mathbf{P}$ along direction $\mathbf{W}$. Denote by $\mathrm{M}_{n,m}$ the space of $n \times m$ matrices, Sym_n^+ the space of $n \times n$ SPD matrices, Sym_n the space of $n \times n$ symmetric matrices, $\mathrm{Gr}_{n,p}$ the p-dimensional subspaces of $\mathbb{R}^n$ from the projector perspective (Bendokat et al., 2020). For clarity of presentation, let $\widetilde{\mathrm{Gr}}_{n,p}$ be the p-dimensional subspaces of $\mathbb{R}^n$ from the ONB perspective. We will use superscripts for the exponential and logarithmic maps, and the parallel transport to indicate their associated Riemannian metric (in the case of SPD manifolds) or the considered manifold (in the case of Grassmann manifolds). Other notations will be introduced in appropriate paragraphs of the paper.

2.2. Gyrovector Spaces Induced by Isometries

In this section, we study the connections of the basic operations and gyroautomorphisms of two homogeneous Riemannian manifolds that are related by an isometry. A review of gyrogroups and gyrovector spaces is given in Appendix B.

Let M and N be two homogeneous Riemannian manifolds. Assuming that there exists a bijective isometry between the two manifolds

$$\phi : M \rightarrow N.$$

Assuming in addition that one can construct the binary operations, scalar multiplications, and gyroautomorphisms for the two manifolds that verify the axioms of gyrovector spaces from the following equations (Nguyen, 2022b):

$$\mathbf{P} \oplus \mathbf{Q} = \mathrm{Exp}_{\mathbf{P}}(\mathcal{T}_{\mathbf{I}\rightarrow\mathbf{P}}(\mathrm{Log}_{\mathbf{I}}(\mathbf{Q}))), \quad (1)$$

$$t \otimes \mathbf{P} = \mathrm{Exp}_{\mathbf{I}}(t\,\mathrm{Log}_{\mathbf{I}}(\mathbf{P})), \quad (2)$$

$$\mathrm{gyr}[\mathbf{P}, \mathbf{Q}]\mathbf{R} = \big(\ominus(\mathbf{P} \oplus \mathbf{Q})\big) \oplus \big(\mathbf{P} \oplus (\mathbf{Q} \oplus \mathbf{R})\big), \quad (3)$$

where $\mathbf{P}$, $\mathbf{Q}$, and $\mathbf{R}$ are three points on the considered manifold, $\mathbf{I}$ is the identity element of the manifold, $t \in \mathbb{R}$, $\oplus$, $\otimes$, and $\mathrm{gyr}[.,.]$ denote respectively the binary operation, scalar multiplication, and gyroautomorphism of the manifold, $\ominus\mathbf{Y}$ denotes the left inverse of any point $\mathbf{Y}$ on the manifold such that $\ominus\mathbf{Y} \oplus \mathbf{Y} = e$, e is the left identity of the corresponding gyrovector space.

With a slight abuse of terminology, we will refer to manifolds M and N as gyrovector spaces. Finally, assuming that $\bar{\mathbf{I}}$ and $\phi(\bar{\mathbf{I}})$ are respectively the identity elements of manifolds M and N and that $\bar{\mathbf{I}}$ and $\phi(\bar{\mathbf{I}})$ are respectively the left identities of gyrovector spaces M and N. We will study the connections between the basic operations and gyroautomorphisms of the two gyrovector spaces. Lemma 2.1 gives such a connection for the binary operations.

Lemma 2.1. *Let $\mathbf{P}, \mathbf{Q} \in M$. Denote by $\oplus_m$ and $\oplus_n$ the binary operations of gyrovector spaces M and N, respectively. Then*

$$\mathbf{P} \oplus_m \mathbf{Q} = \phi^{-1}(\phi(\mathbf{P}) \oplus_n \phi(\mathbf{Q})). \quad (4)$$

Proof See Appendix E.

Lemma 2.1 states that the binary operation $\oplus_m$ can be performed by first mapping its operands to gyrovector space N via mapping $\phi(.)$, then computing the result of the binary operation $\oplus_n$ with the two resulting points in gyrovector space N, and finally returning back to the original gyrovector space M via inverse mapping $\phi^{-1}(.)$ of $\phi(.)$. Similarly, Lemmas 2.2 and 2.3 give the connections for the scalar multiplications and gyroautomorphisms.

Lemma 2.2. *Let $\mathbf{P} \in M$ and $t \in \mathbb{R}$. Denote by $\otimes_m$ and $\otimes_n$ the scalar multiplications of gyrovector spaces M and N, respectively. Then*

$$t \otimes_m \mathbf{P} = \phi^{-1}(t \otimes_n \phi(\mathbf{P})). \quad (5)$$

Proof See Appendix F.

Lemma 2.3. *Let* $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in M$. *Denote by* $\mathrm{gyr}_m[.,.]$ *and* $\mathrm{gyr}_n[.,.]$ *the gyroautomorphisms of gyrovector spaces* M *and* N, *respectively. Then*

$$\mathrm{gyr}_m[\mathbf{P}, \mathbf{Q}]\mathbf{R} = \phi^{-1}(\mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\mathbf{R})). \quad (6)$$

Proof See Appendix G.

The results from Lemmas 2.1, 2.2, 2.3 suggest an effective method for deriving closed-form expressions of the basic operations and gyroautomorphisms for certain matrix manifolds (see Section 2.3.1 and Appendix C). This method is supported by Theorems 2.4 and 2.5.

Theorem 2.4. *Let* $(G_n, \oplus_n, \otimes_n)$ *be a gyrovector space. Let* $\oplus_m$, $\otimes_m$, *and* $\mathrm{gyr}_m[.,.]$ *be respectively the binary operation, scalar multiplication, and gyroautomorphism defined by Eqs. (4), (5), and (6) where* $\phi(.)$ *is a bijective isometry. Then* $(G_m, \oplus_m, \otimes_m)$ *forms a gyrovector space.*

Proof See Appendix H.

A direct consequence of Theorem 2.4 follows.

Theorem 2.5. *Let* $(G, \oplus_n)$ *be a gyrocommutative and gyrononreductive gyrogroup. Let* $\oplus_m$ *and* $\mathrm{gyr}_m[.,.]$ *be respectively the binary operation and gyroautomorphism defined by Eqs. (4) and (6) where* $\phi(.)$ *is a bijective isometry. Then* $(G, \oplus_m)$ *forms a gyrocommutative and gyrononreductive gyrogroup.*

2.3. Grassmann Manifolds

We show how to construct the basic operations and gyroautomorphism for Grassmann manifolds from the ONB perspective in Section 2.3.1. In Section 2.3.2, we study some isometries of Grassmann manifolds with respect to the canonical metric (Edelman et al., 1998).

2.3.1. GRASSMANN GYROCOMMUTATIVE AND GYRONONREDUCTIVE GYROGROUPS: THE ONB PERSPECTIVE

In Nguyen (2022b), closed-form expressions of the basic operations and gyroautomorphism for $\mathrm{Gr}_{n,p}$ have been derived. These can be obtained from Eqs. (1), (2), and (3) as the exponential and logarithmic maps, and the parallel transport appear in closed-forms. However, the same method cannot be applied to Grassmann manifolds from the ONB perspective. This is because the exponential map and parallel transport in this case are all based on SVD operations.

We tackle the above problem using the following diffeomorphism (Helmke & Moore, 1994) between $\widetilde{\mathrm{Gr}}_{n,p}$ and $\mathrm{Gr}_{n,p}$:

$$\tau : \widetilde{\mathrm{Gr}}_{n,p} \to \mathrm{Gr}_{n,p}, \ \mathbf{U} \mapsto \mathbf{U}\mathbf{U}^T,$$

where $\mathbf{U} \in \widetilde{\mathrm{Gr}}_{n,p}$. This leads to the following definitions.

Definition 2.6. For $\mathbf{U}, \mathbf{V} \in \widetilde{\mathrm{Gr}}_{n,p}$, assuming that $\mathbf{I}_{n,p}$ and $\mathbf{U}\mathbf{U}^T$ are not in each other's cut locus, then the binary operation $\mathbf{U}\widetilde{\oplus}_{gr}\mathbf{V}$ can be defined as

$$\mathbf{U}\widetilde{\oplus}_{gr}\mathbf{V} = \exp([\overline{\mathbf{P}}, \mathbf{I}_{n,p}])\mathbf{V}, \quad (7)$$

where $\mathbf{I}_{n,p} = \begin{bmatrix} \mathbf{I}_p & 0 \\ 0 & 0 \end{bmatrix} \in M_{n,n}$ is the identity element of $\mathrm{Gr}_{n,p}$, $[.,.]$ denotes the matrix commutator, and $\overline{\mathbf{P}} = \mathrm{Log}_{\mathbf{I}_{n,p}}^{gr}(\mathbf{U}\mathbf{U}^T)$ is the logarithmic map of $\mathbf{U}\mathbf{U}^T \in \mathrm{Gr}_{n,p}$ at $\mathbf{I}_{n,p}$.

Definition 2.7. For $\mathbf{U} \in \widetilde{\mathrm{Gr}}_{n,p}$ and $t \in \mathbb{R}$, assuming that $\mathbf{I}_{n,p}$ and $\mathbf{U}\mathbf{U}^T$ are not in each other's cut locus, then the scalar multiplication $t\widetilde{\otimes}_{gr}\mathbf{U}$ can be defined as

$$t\widetilde{\otimes}_{gr}\mathbf{U} = \exp([t\overline{\mathbf{P}}, \mathbf{I}_{n,p}])\widetilde{\mathbf{I}}_{n,p}, \quad (8)$$

where $\widetilde{\mathbf{I}}_{n,p} = \begin{bmatrix} \mathbf{I}_p \\ 0 \end{bmatrix} \in M_{n,p}$, and $\overline{\mathbf{P}} = \mathrm{Log}_{\mathbf{I}_{n,p}}^{gr}(\mathbf{U}\mathbf{U}^T)$.

Definition 2.8. Define the binary operation $\widetilde{\oplus}_{gr}$ and the scalar multiplication $\widetilde{\otimes}_{gr}$ by Eqs. (7) and (8), respectively. For $\mathbf{U}, \mathbf{V}, \mathbf{W} \in \widetilde{\mathrm{Gr}}_{n,p}$, assuming that $\mathbf{I}_{n,p}$ and $\mathbf{U}\mathbf{U}^T$ are not in each other's cut locus, $\mathbf{I}_{n,p}$ and $\mathbf{V}\mathbf{V}^T$ are not in each other's cut locus, $\mathbf{I}_{n,p}$ and $\mathbf{U}\mathbf{U}^T \oplus_{gr} \mathbf{V}\mathbf{V}^T$ are not in each other's cut locus where $\oplus_{gr}$ is the binary operation (Nguyen, 2022b) on $\mathrm{Gr}_{n,p}$, then the gyroautomorphism generated by $\mathbf{U}$ and $\mathbf{V}$ can be defined as

$$\widetilde{\mathrm{gyr}}_{gr}[\mathbf{U}, \mathbf{V}]\mathbf{W} = \widetilde{F}_{gr}(\mathbf{U}, \mathbf{V})\mathbf{W},$$

where $\widetilde{F}_{gr}(\mathbf{U}, \mathbf{V})$ is given by

$$\widetilde{F}_{gr}(\mathbf{U}, \mathbf{V}) = \exp(-[\overline{\mathbf{P} \oplus_{gr} \mathbf{Q}}, \mathbf{I}_{n,p}]) \exp([\overline{\mathbf{P}}, \mathbf{I}_{n,p}])$$
$$\exp([\overline{\mathbf{Q}}, \mathbf{I}_{n,p}]),$$

where $\overline{\mathbf{P}} = \mathrm{Log}_{\mathbf{I}_{n,p}}^{gr}(\mathbf{U}\mathbf{U}^T)$, $\overline{\mathbf{Q}} = \mathrm{Log}_{\mathbf{I}_{n,p}}^{gr}(\mathbf{V}\mathbf{V}^T)$, and $\overline{\mathbf{P} \oplus_{gr} \mathbf{Q}} = \mathrm{Log}_{\mathbf{I}_{n,p}}^{gr}(\mathbf{U}\mathbf{U}^T \oplus_{gr} \mathbf{V}\mathbf{V}^T)$.

2.3.2. GRASSMANN GYROISOMETRIES - THE ISOMETRIES OF GRASSMANN MANIFOLDS

Let $\ominus_{gr}$ and $\mathrm{gyr}_{gr}[.,.]$ be the inverse operation and gyroautomorphism of $\mathrm{Gr}_{n,p}$. Guided by analogies with the Euclidean and hyperbolic geometries, we investigate in this section some isometries of Grassmann manifolds. First, we need to define the inner product on these manifolds.

Definition 2.9 (The Grassmann Inner Product). Let $\mathbf{P}, \mathbf{Q} \in \mathrm{Gr}_{n,p}$. Then the Grassmann inner product of $\mathbf{P}$ and $\mathbf{Q}$ is defined as

$$\langle \mathbf{P}, \mathbf{Q} \rangle = \langle \mathrm{Log}_{\mathbf{I}_{n,p}}^{gr}(\mathbf{P}), \mathrm{Log}_{\mathbf{I}_{n,p}}^{gr}(\mathbf{Q}) \rangle_{\mathbf{I}_{n,p}},$$

where $\langle ., . \rangle_{\mathbf{I}_{n,p}}$ denotes the inner product at $\mathbf{I}_{n,p}$ given by the canonical metric of $\mathrm{Gr}_{n,p}$. Note that we use the notation

$\langle .,. \rangle$ without subscript to denote the inner product that is defined directly on Grassmann manifolds, and the notation $\langle .,. \rangle$ with subscript to denote the inner product on tangent spaces of Grassmann manifolds.

The counterpart of the Euclidean distance function on $\mathrm{Gr}_{n,p}$ is defined below.

Definition 2.10 (The Grassmann Gyrodistance Function). Let $\mathbf{P}, \mathbf{Q} \in \mathrm{Gr}_{n,p}$. Then the Grassmann gyrodistance function $d(\mathbf{P}, \mathbf{Q})$ is defined as

$$d(\mathbf{P}, \mathbf{Q}) = \| \ominus_{gr} \mathbf{P} \oplus_{gr} \mathbf{Q} \|,$$

where $\|.\|$ denotes the Grassmann norm induced by the Grassmann inner product given in Definition 2.9.

Grassmann gyroisometries now can be defined as follows.

Definition 2.11 (Grassmann Gyroisometries). Let $\mathbf{P}, \mathbf{Q} \in \mathrm{Gr}_{n,p}$. Then a map $\omega : \mathrm{Gr}_{n,p} \to \mathrm{Gr}_{n,p}$ is a Grassmann gyroisometry if it preserves the Grassmann gyrodistance between $\mathbf{P}$ and $\mathbf{Q}$, i.e.,

$$d(\omega(\mathbf{P}), \omega(\mathbf{Q})) = d(\mathbf{P}, \mathbf{Q}).$$

The definitions of the Grassmann gyrodistance function and Grassmann gyroisometries agree with those of the hyperbolic gyrodistance function and hyperbolic isometries (Ungar, 2014). Theorems 2.12, 2.13, and 2.14 characterize some Grassmann gyroisometries.

Theorem 2.12. *For any $\mathbf{P} \in \mathrm{Gr}_{n,p}$, a left Grassmann gyrotranslation by $\mathbf{P}$ is the map $\psi_{\mathbf{P}} : \mathrm{Gr}_{n,p} \to \mathrm{Gr}_{n,p}$ given by*

$$\psi_{\mathbf{P}}(\mathbf{Q}) = \mathbf{P} \oplus_{gr} \mathbf{Q},$$

where $\mathbf{Q} \in \mathrm{Gr}_{n,p}$. Then left Grassmann gyrotranslations are Grassmann gyroisometries.

Proof See Appendix I.

Theorem 2.13. *Gyroautomorphisms $\mathrm{gyr}_{gr}[.,.]$ are Grassmann gyroisometries.*

Proof See Appendix J.

Theorem 2.14. *A Grassmann inverse map is the map $\lambda : \mathrm{Gr}_{n,p} \to \mathrm{Gr}_{n,p}$ given by*

$$\lambda(\mathbf{P}) = \ominus_{gr}\mathbf{P},$$

where $\mathbf{P} \in \mathrm{Gr}_{n,p}$. Then Grassmann inverse maps are Grassmann gyroisometries.

Proof See Appendix K.

We note that the isometries of Grassmann manifolds with respect to different metrics have been investigated in Botelho et al. (2013); Gehér & Semrl (2016; 2018); Qian et al.

(2021). However, these works only show the general forms of these isometries, while our work gives specific expressions of some Grassmann gyroisometries with respect to the Grassmann gyrodistance function, thank to the closed-form expressions of left Grassmann gyrotranslations, gyroautomorphisms, and Grassmann inverse maps. To the best of our knowledge, these expressions of Grassmann gyroisometries have not appeared in previous works.

2.4. SPD Manifolds

In this section, we examine the similar concepts in Section 2.3.2 for SPD manifolds. Section 2.4.1 presents some isometries of SPD manifolds with Log-Euclidean (Arsigny et al., 2005), Log-Cholesky (Lin, 2019), and Affine-Invariant (Pennec et al., 2004) metrics. In Section 2.4.2, we define hyperplanes on SPD manifolds, and introduce the notion of SPD pseudo-gyrodistance from a SPD matrix or a set of SPD matrices to a hyperplane on SPD manifolds. These notations allow us to generalize MLR on SPD manifolds.

2.4.1. SPD GYROISOMETRIES - THE ISOMETRIES OF SPD MANIFOLDS

In Nguyen (2022a;b), the author has shown that SPD manifolds with Log-Euclidean, Log-Cholesky, and Affine-Invariant metrics form gyrovector spaces referred to as LE, LC, and AI gyrovector spaces, respectively. We adopt the notations in these works and consider the case where $r = 1$ (see Nguyen (2022b), Definition 3.1). Let $\oplus_{le}, \oplus_{lc}$, and $\oplus_{ai}$ be the binary operations in LE, LC, and AI gyrovector spaces, respectively. Let $\otimes_{le}, \otimes_{lc}$, and $\otimes_{ai}$ be the scalar multiplications in LE, LC, and AI gyrovector spaces, respectively. Let $\mathrm{gyr}_{le}[.,.], \mathrm{gyr}_{lc}[.,.]$, and $\mathrm{gyr}_{ai}[.,.]$ be the gyroautomorphisms in LE, LC, and AI gyrovector spaces, respectively. For convenience of presentation, we use the letter g in the subscripts and superscripts of notations to indicate the Riemannian metric of the considered SPD manifold where $g \in \{le, lc, ai\}$, unless otherwise stated. Denote by $\mathbf{I}_n$ the $n \times n$ identity matrix. We repeat the approach used in Section 2.3.2 for SPD manifolds. The inner product on these manifolds is given below.

Definition 2.15 (The SPD Inner Product). Let $\mathbf{P}, \mathbf{Q} \in \mathrm{Sym}_n^+$. Then the SPD inner product of $\mathbf{P}$ and $\mathbf{Q}$ is defined as

$$\langle \mathbf{P}, \mathbf{Q} \rangle = \langle \mathrm{Log}_{\mathbf{I}_n}^g(\mathbf{P}), \mathrm{Log}_{\mathbf{I}_n}^g(\mathbf{Q}) \rangle_{\mathbf{I}_n},$$

where $\langle \cdot, \cdot \rangle_{\mathbf{I}_n}$ denotes the inner product at $\mathbf{I}_n$ given by the Riemannian metric of the considered manifold.

The SPD norm, SPD gyrodistance function, SPD gyroisometries, left SPD gyrotranslations, and SPD inverse maps are defined in the same way[1] as those on Grassmann manifolds.

---

[1] For simplicity, we use the same notations for the SPD inner

Theorems 2.16 and 2.17 characterize some SPD gyroisometries of LE, LC, and AI gyrovector spaces that are fully analogous with Grassmann gyroisometries.

Theorem 2.16. *Left SPD gyrotranslations are SPD gyroisometries.*

Proof See Appendix L.

Theorem 2.17. *Gyroautomorphisms* $\mathrm{gyr}_g[.,.]$ *are SPD gyroisometries.*

Proof See Appendix M.

Theorem 2.18. *SPD inverse maps are SPD gyroisometries.*

Proof See Appendix N.

The SPD gyroisometries given in Theorems 2.16, 2.17, and 2.18 belong to a family of isometries of SPD manifolds discussed in Molnár (2015); Molnár & Szokol (2015). The difference between these works and ours is that our SPD gyroisometries are obtained from the gyrovector space perspective. Furthermore, our method can be applied to any metric on SPD manifolds as long as the basic operations and gyroautomorphism associated with that metric verify the axioms of gyrovector spaces considered in Nguyen (2022b).

2.4.2. MULTICLASS LOGISTIC REGRESSION ON SPD MANIFOLDS

Inspired by the works in Lebanon & Lafferty (2004); Ganea et al. (2018) that generalize MLR to multinomial and hyperbolic geometries, here we aim to generalize MLR to SPD manifolds.

Given K classes, MLR computes the probability of each of the output classes as

$$p(y = k|x) = \frac{\exp(w_k^T x + b_k)}{\sum_{i=1}^{K} \exp(w_i^T x + b_i)} \propto \exp(w_k^T x + b_k), \tag{9}$$

where x is an input sample, $b_k \in \mathbb{R}$, $x, w_k \in \mathbb{R}^n$, $k = 1, \ldots, K$.

As shown in Lebanon & Lafferty (2004); Ganea et al. (2018), Eq. (9) can be rewritten as

$$p(y = k|x) \propto \exp(\mathrm{sign}(w_k^T x + b_k)\|w_k\| d(x, \mathcal{H}_{w_k, b_k})),$$

where $d(x, \mathcal{H}_{w_k, b_k})$ is the margin distance from point x to a hyperplane $\mathcal{H}_{w_k, b_k}$.

The generalization of MLR to SPD manifolds thus requires the definitions of hyperplanes and margin distances in such manifolds. Guided by analogies with hyperbolic geometry (Ungar, 2014; Ganea et al., 2018), hyperplanes on SPD manifolds can be defined as follows.

---

product, SPD norm, and SPD gyrodistance function as those on Grassmann manifolds since they should be clear from the context.
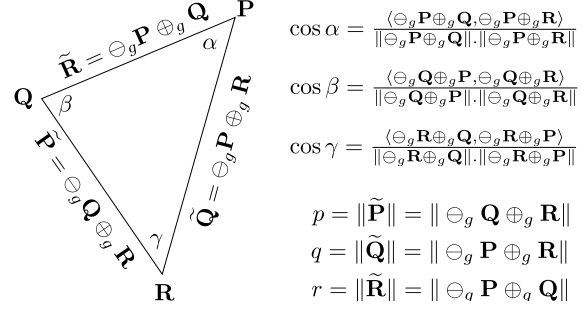

$$\cos \alpha = \frac{\langle \ominus_g \mathbf{P} \oplus_g \mathbf{Q}, \ominus_g \mathbf{P} \oplus_g \mathbf{R} \rangle}{\|\ominus_g \mathbf{P} \oplus_g \mathbf{Q}\|.\|\ominus_g \mathbf{P} \oplus_g \mathbf{R}\|}$$

$$\cos \beta = \frac{\langle \ominus_g \mathbf{Q} \oplus_g \mathbf{P}, \ominus_g \mathbf{Q} \oplus_g \mathbf{R} \rangle}{\|\ominus_g \mathbf{Q} \oplus_g \mathbf{P}\|.\|\ominus_g \mathbf{Q} \oplus_g \mathbf{R}\|}$$

$$\cos \gamma = \frac{\langle \ominus_g \mathbf{R} \oplus_g \mathbf{Q}, \ominus_g \mathbf{R} \oplus_g \mathbf{P} \rangle}{\|\ominus_g \mathbf{R} \oplus_g \mathbf{Q}\|.\|\ominus_g \mathbf{R} \oplus_g \mathbf{P}\|}$$

$$p = \|\widetilde{\mathbf{P}}\| = \|\ominus_g \mathbf{Q} \oplus_g \mathbf{R}\|$$
$$q = \|\widetilde{\mathbf{Q}}\| = \|\ominus_g \mathbf{P} \oplus_g \mathbf{R}\|$$
$$r = \|\widetilde{\mathbf{R}}\| = \|\ominus_g \mathbf{P} \oplus_g \mathbf{Q}\|$$

Figure 1. Illustration of a SPD gyrotriangle, SPD gyroangles, and SPD gyrosides in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$.

Definition 2.19 (**SPD Hypergyroplanes**)**.** For $\mathbf{P} \in \mathrm{Sym}_n^+$, $\mathbf{W} \in \mathcal{T}_\mathbf{P} \mathrm{Sym}_n^+$, SPD hypergyroplanes are defined as

$$\mathcal{H}_{\mathbf{W},\mathbf{P}} = \{\mathbf{Q} \in \mathrm{Sym}_n^+ : \langle \mathrm{Log}_\mathbf{P}^g(\mathbf{Q}), \mathbf{W} \rangle_\mathbf{P} = 0\}. \tag{10}$$

In order to define the margin distance from a SPD matrix to a SPD hypergyroplane, we need to generalize the notion of gyroangles on SPD manifolds, given below.

Definition 2.20 (**The SPD Gyrocosine Function and SPD Gyroangles**)**.** Let $\mathbf{P}, \mathbf{Q}$, and $\mathbf{R}$ be three distinct SPD gyropoints (SPD matrices) in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$. The SPD gyrocosine of the measure of the SPD gyroangle α, $0 \le \alpha \le \pi$, between $\ominus_g \mathbf{P} \oplus_g \mathbf{Q}$ and $\ominus_g \mathbf{P} \oplus_g \mathbf{R}$ is given by the equation

$$\cos \alpha = \frac{\langle \ominus_g \mathbf{P} \oplus_g \mathbf{Q}, \ominus_g \mathbf{P} \oplus_g \mathbf{R} \rangle}{\|\ominus_g \mathbf{P} \oplus_g \mathbf{Q}\|.\|\ominus_g \mathbf{P} \oplus_g \mathbf{R}\|}.$$

The SPD gyroangle α is denoted by $\alpha = \angle \mathbf{QPR}$.

Notice that our definition of the SPD gyrocosine of a SPD gyroangle is not based on unit gyrovectors (Ungar, 2014) and thus is not the same as that of the gyrocosine of a gyroangle in hyperbolic spaces. Similarly to Euclidean and hyperbolic spaces, one can state the Law of SPD gyrocosines. It will be useful later on when we introduce the concept of SPD pseudo-gyrodistance from a SPD matrix to a SPD hypergyroplane.

Theorem 2.21 (**The Law of SPD Gyrocosines**)**.** *Let* $\mathbf{P}, \mathbf{Q}$, *and* $\mathbf{R}$ *be three distinct SPD gyropoints in a gyrovector space* $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$ *where* $g \in \{le, lc\}$. *Let* $\widetilde{\mathbf{P}} = \ominus_g \mathbf{Q} \oplus_g \mathbf{R}$, $\widetilde{\mathbf{Q}} = \ominus_g \mathbf{P} \oplus_g \mathbf{R}$, *and* $\widetilde{\mathbf{R}} = \ominus_g \mathbf{P} \oplus_g \mathbf{Q}$ *be the SPD gyrosides of the SPD gyrotriangle formed by the three SPD gyropoints. Let* $p = \|\widetilde{\mathbf{P}}\|$, $q = \|\widetilde{\mathbf{Q}}\|$, *and* $r = \|\widetilde{\mathbf{R}}\|$. *Let* $\alpha = \angle \mathbf{QPR}$, $\beta = \angle \mathbf{PQR}$, *and* $\gamma = \angle \mathbf{PRQ}$ *be the SPD gyroangles of the SPD gyrotriangle. Then*

$$p^2 = q^2 + r^2 - 2qr \cos \alpha.$$
$$q^2 = p^2 + r^2 - 2pr \cos \beta.$$
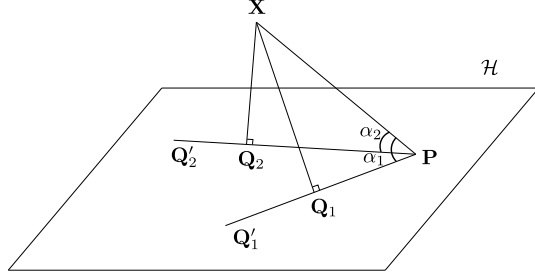$$r^2 = p^2 + q^2 - 2pq \cos \gamma.$$

Figure 2. Illustration of the distance from a point $\mathbf{X}$ to a hyperplane $\mathcal{H}$ in $\mathbb{R}^n$. Here $\mathbf{P} \in \mathcal{H}$, $\mathbf{Q}'_1$ and $\mathbf{Q}'_2$ are two distinct points such that $\mathbf{Q}'_1, \mathbf{Q}'_2 \in \mathcal{H} \setminus \{\mathbf{P}\}$, $\mathbf{Q}_1$ and $\mathbf{Q}_2$ are the projections of $\mathbf{X}$ on lines $\mathbf{PQ}'_1$ and $\mathbf{PQ}'_2$ that are supposed to belong to these lines, respectively, α_1 and α_2 are the angles that lines $\mathbf{PQ}'_1$ and $\mathbf{PQ}'_2$ make with line $\mathbf{PX}$, respectively. If $\cos(\alpha_2) \geq \cos(\alpha_1)$, then $\|\mathbf{XQ}_2\|_F \leq \|\mathbf{XQ}_1\|_F$. The distance from $\mathbf{X}$ to hyperplane $\mathcal{H}$ is obtained when $\cos(\alpha)$, α is the angle between lines $\mathbf{PX}$ and $\mathbf{PQ}$, $\mathbf{Q} \in \mathcal{H} \setminus \{\mathbf{P}\}$, gets the maximum value.

Proof   See Appendix O.

Fig. 1 illustrates the notions given in Theorem 2.21. It states that one can calculate a SPD gyroside of a SPD gyrotriangle when the SPD gyroangle opposite to the SPD gyroside and the other two SPD gyrosides are known. This result is fully analogous with those in Euclidean and hyperbolic spaces. The Law of SPD gyrosines is given in Appendix D.

We now introduce the concept of SPD pseudo-gyrodistance from a SPD matrix to a SPD hypergyroplane that is inspired from a property of the distance from a point in $\mathbb{R}^n$ to a hyperplane in $\mathbb{R}^n$. The key idea is illustrated in Fig. 2.

Definition 2.22 (The SPD Pseudo-gyrodistance from a SPD Matrix to a SPD Hypergyroplane). Let $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ be a SPD hypergyroplane, and $\mathbf{X} \in \mathrm{Sym}_n^+$. The SPD pseudo-gyrodistance from $\mathbf{X}$ to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ is defined as

$$\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \sin(\angle \mathbf{X}\mathbf{P}\bar{\mathbf{Q}})d(\mathbf{X}, \mathbf{P}),$$

where $\bar{\mathbf{Q}}$ is given by

$$\bar{\mathbf{Q}} = \operatorname*{arg\,max}_{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \setminus \{\mathbf{P}\}} \left(\frac{\langle \ominus_g \mathbf{P} \oplus_g \mathbf{Q}, \ominus_g \mathbf{P} \oplus_g \mathbf{X} \rangle}{\| \ominus_g \mathbf{P} \oplus_g \mathbf{Q}\|.\| \ominus_g \mathbf{P} \oplus_g \mathbf{X}\|} \right).$$

By convention, $\sin(\angle \mathbf{X}\mathbf{P}\mathbf{Q}) = 0$ for any $\mathbf{X}, \mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}}$.

The SPD gyrodistance from $\mathbf{X}$ to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ is defined as

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \min_{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}}} d(\mathbf{X}, \mathbf{Q}).$$

From Theorem 2.21, it turns out that the SPD pseudo-gyrodistance agrees with the SPD gyrodistance in certain cases. In particular, we have the following results.

Theorem 2.23 (The SPD Gyrodistance from a SPD Matrix to a SPD Hypergyroplane in a LE Gyrovector Space). *Let $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ be a SPD hypergyroplane in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_{le}, \otimes_{le})$, and $\mathbf{X} \in \mathrm{Sym}_n^+$. Then the SPD pseudo-gyrodistance from $\mathbf{X}$ to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ is equal to the SPD gyrodistance from $\mathbf{X}$ to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ and is given by*

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \frac{|\langle \log(\mathbf{X}) - \log(\mathbf{P}), D \log_{\mathbf{P}}(\mathbf{W}) \rangle_F|}{\|D \log_{\mathbf{P}}(\mathbf{W})\|_F}.$$

Proof   See Appendix P.

Theorem 2.24 (The SPD Gyrodistance from a SPD Matrix to a SPD Hypergyroplane in a LC Gyrovector Space). *Let $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ be a SPD hypergyroplane in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_{lc}, \otimes_{lc})$, and $\mathbf{X} \in \mathrm{Sym}_n^+$. Then the SPD pseudo-gyrodistance from $\mathbf{X}$ to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ is equal to the SPD gyrodistance from $\mathbf{X}$ to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ and is given by*

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \frac{|\langle \mathbf{A}, \mathbf{B} \rangle_F|}{\|\mathbf{B}\|_F},$$

where

$$\mathbf{A} = -\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{X}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1}\mathbb{D}(\varphi(\mathbf{X}))),$$

$$\mathbf{B} = \lfloor \widetilde{\mathbf{W}} \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1}\mathbb{D}(\widetilde{\mathbf{W}}),$$

$$\widetilde{\mathbf{W}} = \varphi(\mathbf{P})\left(\varphi(\mathbf{P})^{-1}\mathbf{W}(\varphi(\mathbf{P})^{-1})^T \right)_{\frac{1}{2}},$$

where $\lfloor \mathbf{Y} \rfloor$ is a matrix of the same size as matrix $\mathbf{Y} \in \mathrm{M}_{n,n}$ whose (i, j) element is $\mathbf{Y}_{ij}$ if $i > j$ and is zero otherwise, $\mathbb{D}(\mathbf{Y})$ is a diagonal matrix of the same size as matrix $\mathbf{Y}$ whose (i, i) element is $\mathbf{Y}_{ii}$, $\mathbf{Y}_{\frac{1}{2}}$ is the lower triangular part of $\mathbf{Y}$ with the diagonal entries halved, and $\varphi(\mathbf{Q})$ denotes the Cholesky factor of $\mathbf{Q} \in \mathrm{Sym}_n^+$, i.e., $\varphi(\mathbf{Q})$ is a lower triangular matrix with positive diagonal entries such that $\mathbf{Q} = \varphi(\mathbf{Q})\varphi(\mathbf{Q})^T$.

Proof   See Appendix Q.

We cannot establish an equivalent result in the case of AI gyrovector spaces. Nevertheless, a closed-form expression for the SPD pseudo-gyrodistance can still be obtained in this case.

Theorem 2.25 (The SPD Pseudo-gyrodistance from a SPD Matrix to a SPD Hypergyroplane in an AI Gyrovector Space). *Let $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ be a SPD hypergyroplane in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_{ai}, \otimes_{ai})$, and $\mathbf{X} \in \mathrm{Sym}_n^+$. Then the SPD pseudo-gyrodistance from $\mathbf{X}$ to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ is given by*

$$\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \frac{|\langle \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{X}\mathbf{P}^{-\frac{1}{2}}), \mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}} \rangle_F|}{\|\mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}}\|_F}.$$

Dataset	SPDNet	SPDNetBN	GyroLE	GyroLC	GyroAI
HDM05	72.83	**76.42**	72.64	63.78	73.34
#HDM05	6.58	6.68	6.53	6.53	6.53
FPHA	89.25	91.34	**94.61**	82.43	93.39
#FPHA	0.99	1.03	0.95	0.95	0.95
NTU60	77.82	79.61	81.68	72.26	**82.75**
#NTU60	1.80	2.06	1.49	1.49	1.49

Table 1. Accuracy comparison (%) of our SPD models against SPDNet and SPDNetBN with comparable model sizes (MB).

Dataset	GyroAI-HAUNet	MLR-LE	MLR-LC	MLR-AI
HDM05	78.14	77.62	71.35	**79.84**
#HDM05	0.31	0.60	0.60	0.60
FPHA	96.00	**96.44**	88.62	96.26
#FPHA	0.11	0.21	0.21	0.21
NTU60	94.72	95.87	88.24	**96.48**
#NTU60	0.02	0.05	0.05	0.05

Table 2. Accuracy comparison (%) of our SPD models against GyroAI-HAUNet.

Proof See Appendix R.

The results in Theorems 2.23, 2.24, and 2.25 lead to Corollary 2.26 that concerns with the SPD gyrodistance and pseudo-gyrodistance from a set of SPD matrices to a SPD hypergyroplane.

Corollary 2.26. *Let* $\mathbf{P}_1, \ldots, \mathbf{P}_N, \mathbf{Q}_1, \ldots, \mathbf{Q}_N,$ *and* $\mathbf{X}_1, \ldots, \mathbf{X}_N \in \mathrm{Sym}_n^+$. *Let* $\mathbf{W}_1, \ldots, \mathbf{W}_N \in \mathrm{Sym}_n$. *Denote by* $\mathrm{diag}(\mathbf{P}_1, \ldots, \mathbf{P}_N)$ *the following matrix:*

$$\mathrm{diag}(\mathbf{P}_1, \ldots, \mathbf{P}_N) = \begin{bmatrix} \mathbf{P}_1 \cdots \cdots \\ \cdots \mathbf{P}_2 \cdots \\ \cdots \cdots \mathbf{P}_N \end{bmatrix},$$

where the diagonal entries of $\mathbf{P}_i, i = 1, \ldots, N$ *belong to the diagonal entries of* $\mathrm{diag}(\mathbf{P}_1, \ldots, \mathbf{P}_N)$. *Let* $\mathbf{P} = \mathrm{diag}(\mathbf{P}_1, \ldots, \mathbf{P}_N)$, $\mathbf{Q} = \mathrm{diag}(\mathbf{Q}_1, \ldots, \mathbf{Q}_N)$, $\mathbf{X} = \mathrm{diag}(\mathbf{X}_1, \ldots, \mathbf{X}_N)$, *and* $\mathbf{W} = \mathrm{diag}(\mathbf{W}_1, \ldots, \mathbf{W}_N)$.

(1) Denote by $\mathcal{H}_{\mathbf{W}, \mathbf{P}}$ *a SPD hypergyroplane in a gyrovector space* $(\mathrm{Sym}_n^+, \oplus_{le}, \otimes_{le})$. *Then the SPD gyrodistance from* $\mathbf{X}$ *to* $\mathcal{H}_{\mathbf{W}, \mathbf{P}}$ *is given by*

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W}, \mathbf{P}}) = \frac{|\sum_{i=1}^N \langle \log(\mathbf{X}_i) - \log(\mathbf{P}_i), D\log_{\mathbf{P}_i}(\mathbf{W}_i)\rangle_F|}{\sqrt{\sum_{i=1}^N \|D\log_{\mathbf{P}_i}(\mathbf{W}_i)\|_F^2}}. \tag{11}$$

(2) Denote by $\mathcal{H}_{\mathbf{W}, \mathbf{P}}$ *a SPD hypergyroplane in a gyrovector space* $(\mathrm{Sym}_n^+, \oplus_{lc}, \otimes_{lc})$. *Then the SPD gyrodistance from* $\mathbf{X}$ *to* $\mathcal{H}_{\mathbf{W}, \mathbf{P}}$ *is given by*

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W}, \mathbf{P}}) = \frac{|\sum_{i=1}^N \langle \mathbf{A}_i, \mathbf{B}_i\rangle_F|}{\sqrt{\sum_{i=1}^N \|\mathbf{B}_i\|_F^2}}, \tag{12}$$

where

$$\mathbf{A}_i = -\lfloor \varphi(\mathbf{P}_i)\rfloor + \lfloor \varphi(\mathbf{X}_i)\rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}_i))^{-1}\mathbb{D}(\varphi(\mathbf{X}_i))),$$

$$\mathbf{B}_i = \lfloor \widetilde{\mathbf{W}}_i\rfloor + \mathbb{D}(\varphi(\mathbf{P}_i))^{-1}\mathbb{D}(\widetilde{\mathbf{W}}_i),$$

$$\widetilde{\mathbf{W}}_i = \varphi(\mathbf{P}_i)\Big(\varphi(\mathbf{P}_i)^{-1}\mathbf{W}_i(\varphi(\mathbf{P}_i)^{-1})^T\Big)_{\frac{1}{2}}.$$

(3) Denote by $\mathcal{H}_{\mathbf{W}, \mathbf{P}}$ *a SPD hypergyroplane in a gyrovector space* $(\mathrm{Sym}_n^+, \oplus_{ai}, \otimes_{ai})$. *Then the SPD pseudo-*

gyrodistance from $\mathbf{X}$ *to* $\mathcal{H}_{\mathbf{W}, \mathbf{P}}$ *is given by*

$$\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W}, \mathbf{P}}) = \frac{|\sum_{i=1}^N \langle \log(\mathbf{P}_i^{-\frac{1}{2}}\mathbf{X}_i\mathbf{P}_i^{-\frac{1}{2}}), \mathbf{P}_i^{-\frac{1}{2}}\mathbf{W}_i\mathbf{P}_i^{-\frac{1}{2}}\rangle_F|}{\sqrt{\sum_{i=1}^N \|\mathbf{P}_i^{-\frac{1}{2}}\mathbf{W}_i\mathbf{P}_i^{-\frac{1}{2}}\|_F^2}}. \tag{13}$$

Proof See Appendix S.

3. Experiments

In this section, we report results of our experiments for two applications, i.e., human action recognition and knowledge graph completion. Details on the datasets and our experimental settings are given in Appendix A.

3.1. Human Action Recognition

We use three datasets, i.e., HDM05 (Müller et al., 2007), FPHA (Garcia-Hernando et al., 2018), and NTU60 (Shahroudy et al., 2016).

3.1.1. SPD NEURAL NETWORKS

We design three networks, each of them is composed of a layer based on the Affine-Invariant translation model and of a MLR (Log-Euclidean, Log-Cholesky, and Affine-Invariant, see Section 2.4.2). These networks are compared against SPDNet (Huang & Gool, 2017)[2] and SPDNetBN (Brooks et al., 2019)[3]. Temporal pyramid representation is used as in Nguyen (2022b). Each sequence is then represented by a set of SPD matrices. We use Eqs. (11), (12), and (13) to compute the SPD gyrodistances and pseudo-gyrodistances for MLR. Results of the five networks are given in Tab. 1. On HDM05 dataset, GyroLE is on par with SPDNet while GyroAI outperforms SPDNet. On FPHA and NTU60 datasets, GyroLE and GyroAI outperform both SPDNet and SPDNetBN.

We also compare GyroAI-HAUNet in Nguyen (2022b)

---

[2] https://github.com/zhiwu-huang/SPDNet
[3] https://papers.nips.cc/paper/2019/hash/6e69ebbfad976d4637bb4b39de261bf7-Abstract.html

Dataset	LSTM	ST-TR	HypGRU	ST-GCN	Shift-GCN	MLR-LE	MLR-LC	MLR-AI
HDM05	72.82	76.12	58.50	76.58	**80.28**	77.62	71.35	79.84
#HDM05	0.54	27.73	0.61	17.73	4.20	0.60	0.60	0.60
FPHA	81.22	91.34	61.42	78.78	91.08	96.44	88.62	96.26
#FPHA	0.41	27.55	0.47	17.60	3.84	0.21	0.21	0.21
NTU60	87.27	93.78	88.03	91.75	95.01	95.87	88.24	96.48
#NTU60	0.035	27.50	0.039	17.66	3.90	0.05	0.05	0.05

Table 3. Accuracy comparison (%) of our SPD models against state-of-the-art models.

against three other networks in which we replace the classification layer of GyroAI-HAUNet with a MLR based on Log-Euclidean, Log-Cholesky, and Affine-Invariant metrics, respectively. Results of the four networks are shown in Tab. 2. The best results are obtained by our models MLR-AI or MLR-LE. However, these models have 2x more parameters than GyroAI-HAUNet.

Tab. 3 reports results of our SPD models and those of some state-of-the-art models from four categories of neural networks: recurrent neural networks (i.e., LSTM), hyperbolic neural networks (i.e., HypGRU (Ganea et al., 2018)), graph neural networks (i.e., ST-GCN (Yan et al., 2018) and Shift-GCN (Cheng et al., 2020)), and transformers (i.e., ST-TR (Plizzari et al., 2021)). MLR-LE and MLR-AI outperform the other networks on FPHA and NTU60 datasets. Also, our networks use far fewer parameters than the transformer and graph neural networks.

3.1.2. GRASSMANN NEURAL NETWORKS

Huang et al. (Huang et al., 2018) proposed a discriminative Grassmann neural network called GrNet. The network applies the FRmap layer to reduce the dimension of input matrices. Orthogonal matrices are then obtained from the outputs of the FRmap layer via QR-decomposition performed by the ReOrth layer. This creates an issue in the backward pass of the ReOrth layer, where the inverse of upper-triangular matrices must be computed. In practice, these matrices are often ill-conditioned and cannot be returned by popular deep learning frameworks like Tensorflow[4] and Pytorch. We address this issue by replacing the FRmap and ReOrth layers with a layer based on the Grassmann translation model (see Section 2.3.2). The resulting network GyroGr[5] is compared against GrNet based on its official Matlab code[6]. Results of the two networks are given in Tab. 4. GyroGr outperforms GrNet by 3.60%, 2.79%, and 2.14% on HDM05, FPHA, and NTU60 datasets, respec-

Method	HDM05	FPHA	NTU60
GrNet	52.71	81.91	65.45
GyroGr	**56.32**	**84.70**	**67.60**

Table 4. Accuracy comparison (%) of GyroGr against GrNet.

Method	HDM05	FPHA	NTU60
GyroGr-Scaling	45.69	65.74	55.26
GyroGr	**56.32**	**84.70**	**67.60**

Table 5. Accuracy comparison (%) of GyroGr against the Grassmann scaling model in Nguyen (2022b).

tively. These results clearly demonstrate the effectiveness of the Grassmann translation model in a discriminative Grassmann neural network like GrNet.

We also conduct another experiment in order to compare the Grassmann translation model and Grassmann scaling model in Nguyen (2022b) within the framework of GrNet. To this end, we design a new network from GyroGr by replacing the Grassmann translation layer with a Grassmann scaling layer. Results of the two networks are presented in Tab. 5. GyroGr significantly outperforms GyroGr-Scaling on all the datasets, showing that the Grassmann translation operation is much more effective than the matrix scaling within the framework of GrNet.

3.2. Knowledge Graph Completion

The goal of this experiment is to compare the Grassmann model based on the projector perspective in Nguyen (2022b) against the one based on the ONB perspective. We use two datasets, i.e., WN18RR (Miller, 1995) and FB15k-237 (Toutanova et al., 2015).

Following Balažević et al. (2019); Nguyen (2022b), we design a model that learns a scoring function

$$\phi_{kgc}(e_s, r, e_o) = -d((\mathbf{A}\widetilde{\otimes}\mathbf{S})\widetilde{\oplus}_{gr}\mathbf{R}, \mathbf{O})^2 + b_s + b_o,$$

where $\mathbf{S}$ and $\mathbf{O}$ are embeddings of the subject and object entities, respectively, $\mathbf{R}$ and $\mathbf{A}$ are matrices associated with relation r, $b_s, b_o \in \mathbb{R}$ are scalar biases for the subject and

---

[4]https://github.com/master/tensorflow-riemopt/tree/master/examples/grnet.

[5]Code available at https://github.com/spratmnt/har

[6]https://github.com/zhiwu-huang/GrNet.

DOF	Model	MRR	H@1	H@3	H@10
144	GyroGr-KGC$_{\text{proj+sca}}$	44.2	38.5	46.8	54.6
	GyroGr-KGC$_{\text{onb+sca}}$	**44.9**	**39.5**	**47.2**	54.6

Table 6. Comparison of our Grassmann model against the Grassmann model in Nguyen (2022b) on the validation set of WN18RR dataset. GyroGr-KGC$_{\text{onb+sca}}$ learns embeddings in $\widetilde{\text{Gr}}_{24,12}$. GyroGr-KGC$_{\text{proj+sca}}$ learns embeddings in $\text{Gr}_{24,12}$ (DOF stands for degrees of freedom).

DOF	Model	MRR	H@1	H@3	H@10
144	GyroGr-KGC$_{\text{proj+sca}}$	29.3	20.5	32.4	46.8
	GyroGr-KGC$_{\text{onb+sca}}$	**29.9**	**20.8**	**33.2**	**48.2**

Table 7. Comparison of our Grassmann model against the Grassmann model in Nguyen (2022b) on the validation set of FB15k-237 dataset. GyroGr-KGC$_{\text{onb+sca}}$ learns embeddings in $\widetilde{\text{Gr}}_{24,12}$. GyroGr-KGC$_{\text{proj+sca}}$ learns embeddings in $\text{Gr}_{24,12}$.

object entities, respectively. The operation $\widetilde{\otimes}$ is defined as

$$\mathbf{A}\widetilde{\oplus}_{gr}\mathbf{P} = \exp\left(\begin{bmatrix} 0 & \mathbf{A}*\mathbf{B} \\ -(\mathbf{A}*\mathbf{B})^T & 0 \end{bmatrix}\right)\widetilde{\mathbf{I}}_{n,p},$$

where $\mathbf{A} \in \text{M}_{p,n-p}$, and $\mathbf{P}$ is given by

$$\mathbf{P} = \exp\left(\begin{bmatrix} 0 & \mathbf{B} \\ -\mathbf{B}^T & 0 \end{bmatrix}\right)\widetilde{\mathbf{I}}_{n,p}.$$

The binary operation $\widetilde{\oplus}_{gr}$ is defined in Eq. (7). We use the distance function (Edelman et al., 1998)

$$d(\mathbf{P},\mathbf{Q}) = \|\theta\|_2, \tag{14}$$

where $\theta_i, i = 1,\ldots,p$ are the principle angles between two subspaces spanned by the columns of $\mathbf{P}$ and $\mathbf{Q}$, i.e., $\mathbf{U}\,\text{diag}(\cos(\theta_1),\ldots,\cos(\theta_p))\mathbf{V}^T$ is the SVD of $\mathbf{P}^T\mathbf{Q}$.

Results of our model[7] and the Grassmann model in Nguyen (2022b) on the validation sets of WN18RR and FB15k-237 datasets are shown in Tabs. 6 and 7, respectively. Our model GyroGr-KGC$_{\text{onb+sca}}$ gives the same or better performance than GyroGr-KGC$_{\text{proj+sca}}$ in all cases. In particular, on WN18RR dataset, GyroGr-KGC$_{\text{onb+sca}}$ outperforms GyroGr-KGC$_{\text{proj+sca}}$ by 1% in terms of H@1. On FB15k-237 dataset, GyroGr-KGC$_{\text{onb+sca}}$ outperforms GyroGr-KGC$_{\text{proj+sca}}$ by 1.4% in terms of H@10.

3.3. Complexity Analysis

Let n be the size of input matrices (SPD or projection matrices), n_c be the number of action classes, n_s be the number

---

[7]Code available at https://github.com/spratmnt/kgc

---

of SPD matrices used by GyroAI and MLR-AI for representing an action sequence, n_t and n_p be the number of transformation matrices and the number of projection matrices for the W-ProjPooling layer in GyroGr, respectively. For the sake of simplicity, we analyze the complexity of the models for one training sample and one iteration.

- GyroAI: The binary operation has time complexity $O(n^3)$ and memory complexity $O(n^2)$. The MLR has time complexity $O(n_c n_s n^3)$ and memory complexity $O(n_c n_s n^2)$.

- MLR-AI: The RNN cell has time complexity $O(n^3)$ and memory complexity $O(n^2)$. The MLR has time complexity $O(n_c n_s n^3)$ and memory complexity $O(n_c n_s n^2)$.

- GyroGr: The Grassmann translation layer and OrthMap layer of GrNet have time complexity $O(n_t n^3)$ and memory complexity $O(n_t n^2)$. The W-ProjPooling layer (pooling within one projection matrix) has time complexity $O(n_p n^2)$ and memory complexity $O(n_p n^2)$. The classification layer has time complexity $O(n_t n_c n^2)$ and memory complexity $O(n_t n_c n^2)$.

- GyroGr-KGC: The computation of the scoring function has time complexity $O(n^3)$ and memory complexity $O(n^2)$.

4. Limitation

As pointed out in Shimizu et al. (2021), the hyperbolic MLR (Ganea et al., 2018) is over-parameterized because of the reparameterization of the scalar term b_k in Eq. (9) as a vector $\mathbf{p}_k \in \mathbf{R}^n$. Since our definition of SPD hypergyroplanes follows that of Poincaré hyperplane (Ganea et al., 2018), our MLR suffers from the same problem. More precisely, in order to parameterize a SPD hypergyroplane, we use two symmetric matrices, i.e., $\mathbf{P}$ and $\mathbf{W}$ for each class (see Eq. (10)). Thus our MLR requires $n(n+1)K$ parameters, while a linear layer with input SPD matrices of the same size requires only $(n(n+1)/2+1)K$ parameters. This problem should be addressed in future work.

5. Conclusion

We have generalized the notions of inner product and gyroangles in gyrovector spaces for SPD and Grassmann manifolds. We have studied some isometric models on SPD and Grassmann manifolds, and reformulated MLR on SPD manifolds. We have compared our models against state-of-the-art models for the tasks of human action recognition and knowledge graph completion.

References

Absil, P.-A., Mahony, R., and Sepulchre, R. *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2007.

Arsigny, V., Fillard, P., Pennec, X., and Ayache, N. Fast and Simple Computations on Tensors with Log-Euclidean Metrics. Technical Report RR-5584, INRIA, 2005.

Balažević, I., Allen, C., and Hospedales, T. Multi-relational Poincaré Graph Embeddings. In *NeurIPS*, pp. 4465–4475, 2019.

Banerjee, M., Chakraborty, R., Bouza, J., and Vemuri, B. C. VolterraNet: A Higher Order Convolutional Network With Group Equivariance for Homogeneous Manifolds. *IEEE Trans. Pattern Anal. Mach. Intell.*, 44(2):823–833, 2022.

Bendokat, T., Zimmermann, R., and Absil, P. A. A Grassmann Manifold Handbook: Basic Geometry and Computational Aspects. *CoRR*, abs/2011.13699, 2020.

Bollacker, K., Evans, C., Paritosh, P., Sturge, T., and Taylor, J. Freebase: A Collaboratively Created Graph Database for Structuring Human Knowledge. In *Proceedings of the 2008 ACM SIGMOD International Conference on Management of Data*, pp. 1247–1250, 2008.

Bordes, A., Usunier, N., Garcia-Durán, A., Weston, J., and Yakhnenko, O. Translating Embeddings for Modeling Multi-Relational Data. In *NIPS*, pp. 2787–2795, 2013.

Botelho, F., Jamison, J., and Molnár, L. Surjective Isometries on Grassmann Spaces. *Journal of Functional Analysis*, 265(10):2226–2238, 2013.

Bronstein, M. M., Bruna, J., LeCun, Y., Szlam, A., and Vandergheynst, P. Geometric Deep Learning: Going beyond Euclidean Data. *IEEE Signal Processing Magazine*, 34 (4):18–42, 2017.

Brooks, D. A., Schwander, O., Barbaresco, F., Schneider, J.-Y., and Cord, M. Riemannian Batch Normalization for SPD Neural Networks. In *NeurIPS*, pp. 15463–15474, 2019.

Chakraborty, R., Yang, C.-H., Zhen, X., Banerjee, M., Archer, D., Vaillancourt, D. E., Singh, V., and Vemuri, B. C. A Statistical Recurrent Model on the Manifold of Symmetric Positive Definite Matrices. In *NeurIPS*, pp. 8897–8908, 2018.

Chakraborty, R., Bouza, J., Manton, J., and Vemuri, B. C. ManifoldNet: A Deep Neural Network for Manifold-valued Data with Applications. *TPAMI*, 44(2):799–810, 2020.

Cheng, K., Zhang, Y., He, X., Chen, W., Cheng, J., and Lu, H. Skeleton-Based Action Recognition With Shift Graph Convolutional Network. In *CVPR*, pp. 180–189, 2020.

Dong, Z., Jia, S., Zhang, C., Pei, M., and Wu, Y. Deep Manifold Learning of Symmetric Positive Definite Matrices with Application to Face Recognition. In *AAAI*, pp. 4009–4015, 2017.

Edelman, A., Arias, T. A., and Smith, S. T. The Geometry of Algorithms with Orthogonality Constraints. *SIAM Journal on Matrix Analysis and Applications*, 20(2):303–353, 1998.

Gallier, J. and Quaintance, J. *Differential Geometry and Lie Groups*. Springer International Publishing, 2020.

Ganea, O., Becigneul, G., and Hofmann, T. Hyperbolic neural networks. In *NeurIPS*, pp. 5350–5360, 2018.

Garcia-Hernando, G., Yuan, S., Baek, S., and Kim, T.-K. First-Person Hand Action Benchmark with RGB-D Videos and 3D Hand Pose Annotations. In *CVPR*, pp. 409–419, 2018.

Gehér, G. P. and Semrl, P. Isometries of Grassmann Spaces. *Journal of Functional Analysis*, 270(4):1585–1601, 2016.

Gehér, G. P. and Semrl, P. Isometries of Grassmann Spaces, II. *Advances in Mathematics*, 332:287–310, 2018.

Harandi, M., Salzmann, M., and Hartley, R. Dimensionality Reduction on SPD Manifolds: The Emergence of Geometry-Aware Methods. *TPAMI*, 40:48–62, 2018.

Helmke, U. and Moore, J. B. *Optimization and Dynamical Systems*. Springer London, 1994.

Huang, Z. and Gool, L. V. A Riemannian Network for SPD Matrix Learning. In *AAAI*, pp. 2036–2042, 2017.

Huang, Z., Wu, J., and Gool, L. V. Building Deep Networks on Grassmann Manifolds. In *AAAI*, pp. 3279–3286, 2018.

Kim, S. Ordered Gyrovector Spaces. *Symmetry*, 12(6), 2020.

Lebanon, G. and Lafferty, J. Hyperplane Margin Classifiers on the Multinomial Manifold. In *ICML*, pp. 66, 2004.

Lin, Z. Riemannian Geometry of Symmetric Positive Definite Matrices via Cholesky Decomposition. *SIAM Journal on Matrix Analysis and Applications*, 40(4):1353–1370, 2019.

López, F., Pozzetti, B., Trettel, S., Strube, M., and Wienhard, A. Vector-valued Distance and Gyrocalculus on the Space of Symmetric Positive Definite Matrices. In *NeurIPS*, pp. 18350–18366, 2021.

Miller, G. A. WordNet: A Lexical Database for English. *Communications of the ACM*, 38(11):39–41, 1995.

Molnár, L. Jordan Triple Endomorphisms and Isometries of Spaces of Positive Definite Matrices. *Linear and Multilinear Algebra*, 63(1):12–33, 2015.

Molnár, L. and Szokol, P. Transformations on Positive Definite Matrices Preserving Generalized Distance Measures. *Linear Algebra and its Applications*, 466:141–159, 2015.

Müller, M., Röder, T., Clausen, M., Eberhardt, B., Krüger, B., and Weber, A. Documentation Mocap Database HDM05. Technical Report CG-2007-2, Universität Bonn, June 2007.

Nguyen, X. S. GeomNet: A Neural Network Based on Riemannian Geometries of SPD Matrix Space and Cholesky Space for 3D Skeleton-Based Interaction Recognition. In *ICCV*, pp. 13379–13389, 2021.

Nguyen, X. S. A Gyrovector Space Approach for Symmetric Positive Semi-definite Matrix Learning. In *ECCV*, pp. 52–68, 2022a.

Nguyen, X. S. The Gyro-Structure of Some Matrix Manifolds. In *NeurIPS*, 2022b.

Nguyen, X. S., Brun, L., Lézoray, O., and Bougleux, S. A Neural Network Based on SPD Manifold Learning for Skeleton-based Hand Gesture Recognition. In *CVPR*, pp. 12036–12045, 2019a.

Nguyen, X. S., Brun, L., Lézoray, O., and Bougleux, S. Skeleton-Based Hand Gesture Recognition by Learning SPD Matrices with Neural Networks. In *FG*, pp. 1–5, 2019b.

Nguyen, X. S., Brun, L., Lézoray, O., and Bougleux, S. Learning Recurrent High-order Statistics for Skeleton-based Hand Gesture Recognition. In *ICPR*, pp. 975–982, 2020.

Pennec, X. *Statistical Computing on Manifolds for Computational Anatomy*. Habilitation à diriger des recherches, Université Nice Sophia-Antipolis, 2006.

Pennec, X., Fillard, P., and Ayache, N. A Riemannian Framework for Tensor Computing. Technical Report RR-5255, INRIA, 2004.

Plizzari, C., Cannici, M., and Matteucci, M. Skeleton-based Action Recognition via Spatial and Temporal Transformer Networks. *Computer Vision and Image Understanding*, 208:103219, 2021.

Qian, W., Shen, J., Shi, W., Wu, W., and Yuan, W. Surjective L^p-isometries of Grassmann spaces. *CoRR*, abs/2104.07027, 2021.

Shahroudy, A., Liu, J., Ng, T.-T., and Wang, G. NTU RGB+D: A Large Scale Dataset for 3D Human Activity Analysis. In *CVPR*, pp. 1010–1019, 2016.

Shimizu, R., Mukuta, Y., and Harada, T. Hyperbolic Neural Networks++. In *ICLR*, 2021.

Skopek, O., Ganea, O.-E., and Bécigneul, G. Mixed-curvature Variational Autoencoders. In *ICLR*, 2020.

Toutanova, K., Chen, D., Pantel, P., Poon, H., Choudhury, P., and Gamon, M. Representing Text for Joint Embedding of Text and Knowledge Bases. In *Conference on Empirical Methods in Natural Language Processing*, pp. 1499–1509, 2015.

Ungar, A. A. *Beyond the Einstein Addition Law and Its Gyroscopic Thomas Precession: The Theory of Gyrogroups and Gyrovector Spaces*. Fundamental Theories of Physics, vol. 117, Springer, Netherlands, 2002.

Ungar, A. A. *Analytic Hyperbolic Geometry: Mathematical Foundations and Applications*. World Scientific Publishing Co. Pte. Ltd., Hackensack, NJ, 2005.

Ungar, A. A. *Analytic Hyperbolic Geometry in N Dimensions: An Introduction*. CRC Press, 2014.

Wang, R., Wu, X.-J., and Kittler, J. SymNet: A Simple Symmetric Positive Definite Manifold Deep Learning Method for Image Set Classification. *IEEE Transactions on Neural Networks and Learning Systems*, pp. 1–15, 2021.

Weiler, M., Forré, P., Verlinde, E., and Welling, M. Coordinate Independent Convolutional Networks - Isometry and Gauge Equivariant Convolutions on Riemannian Manifolds. *CoRR*, abs/2106.06020, 2021.

Xu, Y., Lei, J., Dobriban, E., and Daniilidis, K. Unified Fourier-based Kernel and Nonlinearity Design for Equivariant Networks on Homogeneous Spaces. In *ICML*, pp. 24596–24614, 2022.

Yan, S., Xiong, Y., and Lin, D. Spatial Temporal Graph Convolutional Networks for Skeleton-Based Action Recognition. In *AAAI*, pp. 7444–7452, 2018.

A. Experimental Details

A.1. Human Action Recognition

HDM05 (Müller et al., 2007)   It has 2337 sequences of 3D skeleton data classified into 130 classes. Each frame contains the 3D coordinates of 31 body joints. We use all the action classes and follow the experimental protocol of Harandi et al. (2018) in which 2 subjects are used for training and the remaining 3 subjects are used for testing.

FPHA (Garcia-Hernando et al., 2018)   It has 1175 sequences of 3D skeleton data classified into 45 classes. Each frame contains the 3D coordinates of 21 hand joints. We follow the experimental protocol of Garcia-Hernando et al. (2018) in which 600 sequences are used for training and 575 sequences are used for testing.

NTU60 (Shahroudy et al., 2016)   It has 56880 sequences of 3D skeleton data classified into 60 classes. Each frame contains the 3D coordinates of 25 or 50 body joints. We use the mutual actions and follow the cross-subject experimental protocol of Shahroudy et al. (2016) in which data from 20 subjects are used for training, and those from the other 20 subjects are used for testing.

SPD Neural Networks   As in Huang & Gool (2017); Brooks et al. (2019); Nguyen (2022b), each sequence is represented by a covariance matrix. The sizes of the covariance matrices are 93×93, 60×60, and 150×150 for HDM05, FPHA, and NTU60 datasets, respectively. For SPDNet, the same architecture as the one in Huang & Gool (2017) is used with three Bimap layers. For SPDNetBN, the same architecture as the one in Brooks et al. (2019) is used with three Bimap layers. For all the networks, the sizes of the transformation matrices for the experiments on HDM05, FPHA, and NTU60 datasets are set to 93×93, 60×60, and 150×150, respectively.

Grassmann Neural Networks   Following Huang et al. (2018), action sequences are represented by linear subspaces of order 10 which belong to $\mathrm{Gr}_{93,10}$, $\mathrm{Gr}_{60,10}$, and $\mathrm{Gr}_{150,10}$ for the experiments on HDM05, FPHA, and NTU datasets, respectively. For GrNet, the sizes of the connection weights are set respectively to 93×93, 60×60, and 150×150 for the experiments on HDM05, FPHA, and NTU datasets.

Our networks are implemented with Pytorch framework. They are trained using cross-entropy loss and Adadelta optimizer for 2000 epochs. The learning rate is set to 10^{-3}. We use a batch size of 32 for HDM05 and FPHA datasets, and a batch size of 256 for NTU60 dataset. We run each model three times and report the best accuracy from these three runs (Ganea et al., 2018; Nguyen, 2022b).

A.2. Knowledge Graph Completion

WN18RR (Miller, 1995)   It is a subset of WordNet (Miller, 1995), a hierarchical collection of relations between words, created from WN18 (Bordes et al., 2013) by removing the inverse of many relations from validation and test sets. It contains 40,943 entities and 11 relations.

FB15k-237 (Toutanova et al., 2015)   It is a subset of Freebase (Bollacker et al., 2008), a collection of real world facts, created in the same way as WN18RR from FB15k (Bordes et al., 2013). It contains 14,541 entities and 237 relations.

The networks are implemented with Pytorch framework. They are trained using binary cross-entropy loss and SGD optimizer for 2000 epochs. The learning rate is set to 10^{-3} with weight decay of 10^{-5}. The batch size is set to 4096. The number of negative samples is set to 10. These settings are taken from López et al. (2021). We test with embeddings in $\mathrm{Gr}_{n,p}$ and $\widetilde{\mathrm{Gr}}_{n,p}$ where $(n,p) \in \{(2k,k)\}, k = 5, 6, \ldots, 14$. The models give the best results with $(n,p) = (24, 12)$. The MRR and hits at K (H@K, $K = 1, 3, 10$) are used as evaluation metrics (Balažević et al., 2019). Early stopping is used when the MRR score of the model on the validation set does not improve after 500 epochs. In all experiments, the models that obtain the best MRR scores on the validation set are used for testing.

B. Gyrogroups and Gyrovector Spaces

Gyrovector spaces form the setting for hyperbolic geometry in the same way that vector spaces form the setting for Euclidean geometry (Ungar, 2002; 2005; 2014). We recap the definitions of gyrogroups and gyrocommutative gyrogroups proposed

in (Ungar, 2002; 2005; 2014). For greater mathematical detail and in-depth discussion, we refer the interested reader to these papers.

Definition B.1 (Gyrogroups (Ungar, 2014)). A pair $(G, \oplus)$ is a groupoid in the sense that it is a nonempty set, G, with a binary operation, $\oplus$. A groupoid $(G, \oplus)$ is a gyrogroup if its binary operation satisfies the following axioms for $a, b, c \in G$:

(G1) There is at least one element $e \in G$ called a left identity such that $e \oplus a = a$.

(G2) There is an element $\ominus a \in G$ called a left inverse of a such that $\ominus a \oplus a = e$.

(G3) There is an automorphism $\mathrm{gyr}[a, b] : G \to G$ for each $a, b \in G$ such that

$$a \oplus (b \oplus c) = (a \oplus b) \oplus \mathrm{gyr}[a, b]c \quad \text{(Left Gyroassociative Law)}.$$

The automorphism $\mathrm{gyr}[a, b]$ is called the gyroautomorphism, or the gyration of G generated by a, b.

(G4) $\mathrm{gyr}[a, b] = \mathrm{gyr}[a \oplus b, b]$ (Left Reduction Property).

Definition B.2 (Gyrocommutative Gyrogroups (Ungar, 2014)). A gyrogroup $(G, \oplus)$ is gyrocommutative if it satisfies

$$a \oplus b = \mathrm{gyr}[a, b](b \oplus a) \quad \text{(Gyrocommutative Law)}.$$

The following definition of gyrovector spaces is slightly different from Definition 3.2 in (Ungar, 2014).

Definition B.3 (Gyrovector Spaces). A gyrocommutative gyrogroup $(G, \oplus)$ equipped with a scalar multiplication

$$(t, x) \to t \odot x : \mathbb{R} \times G \to G$$

is called a gyrovector space if it satisfies the following axioms for $s, t \in \mathbb{R}$ and $a, b, c \in G$:

(V1) $1 \odot a = a, 0 \odot a = t \odot e = e$, and $(-1) \odot a = \ominus a$.

(V2) $(s + t) \odot a = s \odot a \oplus t \odot a$.

(V3) $(st) \odot a = s \odot (t \odot a)$.

(V4) $\mathrm{gyr}[a, b](t \odot c) = t \odot \mathrm{gyr}[a, b]c$.

(V5) $\mathrm{gyr}[s \odot a, t \odot a] = \mathrm{Id}$, where Id is the identity map.

C. Gyrovector Spaces of SPD Matrices with a Log-Cholesky Geometry

The recent work (Nguyen, 2022a) has shown the gyro-structure of SPD manifolds with a Log-Cholesky geometry (Lin, 2019). Here we present another method based on Lemmas 2.1, 2.2, and 2.3 for deriving closed-form expressions of the basic operations and gyroautomorphism of these manifolds.

Using Eqs. (1), (2), and (3), we first derive closed-form expressions of the basic operations and gyroautomorphism for L_n^+, the space of $n \times n$ lower triangular matrices with positive diagonal entries.

Let $\mathbf{U}, \mathbf{V}, \mathbf{W} \in \mathrm{L}_n^+$ and $t \in \mathbb{R}$. Then

$$\mathbf{U} \oplus_{lt} \mathbf{V} = \lfloor \mathbf{U} \rfloor + \lfloor \mathbf{V} \rfloor + \mathbb{D}(\mathbf{U})\mathbb{D}(\mathbf{V}),$$

$$t \otimes_{lt} \mathbf{U} = t\lfloor \mathbf{U} \rfloor + \mathbb{D}(\mathbf{U})^t,$$

$$\mathrm{gyr}_{lt}[\mathbf{U}, \mathbf{V}]\mathbf{W} = Id,$$

where $\oplus_{lt}, \otimes_{lt}$, and $\mathrm{gyr}_{lt}[.,.]$ denote the binary operation, scalar multiplication, and gyroautomorphism of L_n^+, respectively.

As shown in Lin (2019), there exists a diffeomorphism between L_n^+ and Sym_n^+ given by:

$$\xi : \mathrm{Sym}_n^+ \to \mathrm{L}_n^+, \quad \mathbf{P} \to \mathbf{U}, \mathbf{U}\mathbf{U}^T = \mathbf{P}.$$

This diffeomorphism gives us a simple way to obtain closed-form expressions of the basic operations and gyroautomorphism for SPD manifolds with a Log-Cholesky geometry, that is,

$$\mathbf{P} \oplus_{lc} \mathbf{Q} = (\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{Q}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))\mathbb{D}(\varphi(\mathbf{Q}))).(\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{Q}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))\mathbb{D}(\varphi(\mathbf{Q})))^T,$$

$$t \otimes_{lc} \mathbf{P} = (t \lfloor \varphi(\mathbf{P}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^t).(t \lfloor \varphi(\mathbf{P}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^t)^T,$$

$$\mathrm{gyr}_{lc}[\mathbf{P}, \mathbf{Q}]\mathbf{R} = Id,$$

where $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \mathrm{Sym}_n^+$ and $t \in \mathbb{R}$.

D. The Law of SPD Gyrosines

Theorem D.1 (**The Law of SPD Gyrosines**). *Let $\mathbf{P}, \mathbf{Q}$, and $\mathbf{R}$ be three distinct SPD gyropoints in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$ where $g \in \{le, lc\}$. Let $\widetilde{\mathbf{P}} = \ominus_g \mathbf{Q} \oplus_g \mathbf{R}$, $\widetilde{\mathbf{Q}} = \ominus_g \mathbf{P} \oplus_g \mathbf{R}$, and $\widetilde{\mathbf{R}} = \ominus_g \mathbf{P} \oplus_g \mathbf{Q}$ be the SPD gyrosides of the SPD gyrotriangle formed by the three SPD gyropoints. Let $p = \|\widetilde{\mathbf{P}}\|$, $q = \|\widetilde{\mathbf{Q}}\|$, and $r = \|\widetilde{\mathbf{R}}\|$. Let $\alpha = \angle \mathbf{QPR}$, $\beta = \angle \mathbf{PQR}$, and $\gamma = \angle \mathbf{PRQ}$ be the SPD gyroangles of the SPD gyrotriangle. Then*

$$\frac{\sin(\alpha)}{p} = \frac{\sin(\beta)}{q} = \frac{\sin(\gamma)}{r}.$$

Proof. This is a direct consequence of the Law of SPD gyrocosines. $\qquad\square$

E. Proof of Lemma 2.1

Proof. We first recall some results from Gallier & Quaintance (2020).

Proposition E.1 ((Gallier & Quaintance, 2020))**.** *Let M and N be two Riemannian manifolds. If $\phi : M \to N$ is a local isometry, then the following concepts are preserved:*

(1) Parallel translation along a curve. If $\mathcal{T}_\delta$ denotes parallel transport along the curve δ and if $\mathcal{T}_{\phi \circ \delta}$ denotes parallel transport along the curve $\phi \circ \delta$, then

$$D\phi_{\delta(1)} \circ \mathcal{T}_\delta = \mathcal{T}_{\phi \circ \delta} \circ D\phi_{\delta(0)}. \tag{15}$$

(2) Exponential maps. We have

$$\phi \circ \mathrm{Exp}_{\mathbf{P}} = \mathrm{Exp}_{\phi(\mathbf{P})} \circ D\phi_{\mathbf{P}}. \tag{16}$$

We also need to prove the following result.

Proposition E.2. *Let $\phi : M \to N$ be an isometry. Then*

$$\mathrm{Log}_{\mathbf{P}}(\mathbf{Q}) = (D\phi_{\phi(\mathbf{P})}^{-1})(\mathrm{Log}_{\phi(\mathbf{P})}(\phi(\mathbf{Q}))). \tag{17}$$

Proof. Since ϕ is an isometry, its inverse ϕ^{-1} is an isometry. Therefore, from Eq. (16) we have

$$\phi^{-1} \circ \mathrm{Exp}_{\phi(\mathbf{P})}(\mathrm{Log}_{\phi(\mathbf{P})}(\phi(\mathbf{Q}))) = \mathrm{Exp}_{\phi^{-1}(\phi(\mathbf{P}))} \circ (D\phi_{\phi(\mathbf{P})}^{-1})(\mathrm{Log}_{\phi(\mathbf{P})}(\phi(\mathbf{Q})))$$
$$= \mathrm{Exp}_{\mathbf{P}} \circ (D\phi_{\phi(\mathbf{P})}^{-1})(\mathrm{Log}_{\phi(\mathbf{P})}(\phi(\mathbf{Q}))).$$

Hence

$$\phi^{-1} \circ \phi(\mathbf{Q}) = \mathrm{Exp}_{\mathbf{P}} \circ (D\phi_{\phi(\mathbf{P})}^{-1})(\mathrm{Log}_{\phi(\mathbf{P})}(\phi(\mathbf{Q}))),$$

which is equivalent to

$$\mathbf{Q} = \mathrm{Exp}_{\mathbf{P}} \circ (D\phi_{\phi(\mathbf{P})}^{-1})(\mathrm{Log}_{\phi(\mathbf{P})}(\phi(\mathbf{Q}))).$$

Therefore

$$\mathrm{Log}_{\mathbf{P}}(\mathbf{Q}) = (D\phi_{\phi(\mathbf{P})}^{-1})(\mathrm{Log}_{\phi(\mathbf{P})}(\phi(\mathbf{Q}))).$$

$$\square$$

According to the definition of the binary operation $\oplus_m$ in Eq. (1),

$$
\begin{aligned}
\mathbf{P} \oplus_m \mathbf{Q} &= \mathrm{Exp}_{\mathbf{P}}(\mathcal{T}_{\bar{\mathbf{I}} \to \mathbf{P}}(\mathrm{Log}_{\bar{\mathbf{I}}}(\mathbf{Q}))) \\
&\overset{(1)}{=} \mathrm{Exp}_{\mathbf{P}}(\mathcal{T}_{\bar{\mathbf{I}} \to \mathbf{P}}((D\phi_{\phi(\bar{\mathbf{I}})}^{-1})(\mathrm{Log}_{\phi(\bar{\mathbf{I}})}(\phi(\mathbf{Q}))))) \\
&\overset{(2)}{=} \mathrm{Exp}_{\mathbf{P}}\left((D\phi_{\phi(\mathbf{P})}^{-1})\big(\mathcal{T}_{\phi(\bar{\mathbf{I}}) \to \phi(\mathbf{P})}(\mathrm{Log}_{\phi(\bar{\mathbf{I}})}(\phi(\mathbf{Q})))\big)\right) \\
&\overset{(3)}{=} \phi^{-1}\left(\mathrm{Exp}_{\phi(\mathbf{P})}\big(\mathcal{T}_{\phi(\bar{\mathbf{I}}) \to \phi(\mathbf{P})}(\mathrm{Log}_{\phi(\bar{\mathbf{I}})}(\phi(\mathbf{Q})))\big)\right) \\
&\overset{(4)}{=} \phi^{-1}(\phi(\mathbf{P}) \oplus_n \phi(\mathbf{Q})).
\end{aligned}
\tag{18}
$$

The derivation of Eq. (18) follows.

(1) follows from Proposition E.2.

(2) follows from Eq. (15).

(3) follows from Eq. (16).

(4) follows from the definition of the binary operation $\oplus_n$.

$\square$

F. Proof of Lemma 2.2

Proof. According to the definition of the scalar multiplication $\otimes_m$ in Eq. (2),

$$
\begin{aligned}
t \otimes_m \mathbf{P} &= \mathrm{Exp}_{\bar{\mathbf{I}}}(t\,\mathrm{Log}_{\bar{\mathbf{I}}}(\mathbf{P})) \\
&\overset{(1)}{=} \mathrm{Exp}_{\bar{\mathbf{I}}}\big(t(D\phi_{\phi(\bar{\mathbf{I}})}^{-1})(\mathrm{Log}_{\phi(\bar{\mathbf{I}})}(\phi(\mathbf{P})))\big) \\
&\overset{(2)}{=} \mathrm{Exp}_{\bar{\mathbf{I}}}\big((D\phi_{\phi(\bar{\mathbf{I}})}^{-1})(t\,\mathrm{Log}_{\phi(\bar{\mathbf{I}})}(\phi(\mathbf{P})))\big) \\
&\overset{(3)}{=} \phi^{-1}\big(\mathrm{Exp}_{\phi(\bar{\mathbf{I}})}(t\,\mathrm{Log}_{\phi(\bar{\mathbf{I}})}(\phi(\mathbf{P})))\big) \\
&\overset{(4)}{=} \phi^{-1}(t \otimes_n \phi(\mathbf{P})).
\end{aligned}
\tag{19}
$$

The derivation of Eq. (19) follows.

(1) follows from Proposition E.2.

(2) follows from the fact that $D\phi^{-1}$ is a linear operator.

(3) follows from Eq. (16).

(4) follows from the definition of the scalar multiplication $\otimes_n$.

$\square$

G. Proof of Lemma 2.3

Proof. For any $\mathbf{Y} \in M$, we have

$$
\begin{aligned}
\phi(\bar{\mathbf{I}}) &= \phi(\ominus_m \mathbf{Y} \oplus_m \mathbf{Y}) \\
&\overset{(1)}{=} \phi(\ominus_m \mathbf{Y}) \oplus_n \phi(\mathbf{Y}),
\end{aligned}
\tag{20}
$$

where (1) follows from Eq. (4).

Note that

$$\begin{aligned}
\operatorname{gyr}_m[\mathbf{P},\mathbf{Q}]\mathbf{R} &\stackrel{(1)}{=} \big(\ominus_m(\mathbf{P}\oplus_m\mathbf{Q})\big)\oplus_m\big(\mathbf{P}\oplus_m(\mathbf{Q}\oplus_m\mathbf{R})\big)\\
&\stackrel{(2)}{=} \phi^{-1}\Big(\phi\big(\ominus_m(\mathbf{P}\oplus_m\mathbf{Q})\big)\oplus_n\phi\big(\mathbf{P}\oplus_m(\mathbf{Q}\oplus_m\mathbf{R})\big)\Big)\\
&\stackrel{(3)}{=} \phi^{-1}\Big(\ominus_n\big(\phi(\mathbf{P}\oplus_m\mathbf{Q})\big)\oplus_n\phi\big(\mathbf{P}\oplus_m(\mathbf{Q}\oplus_m\mathbf{R})\big)\Big)\\
&\stackrel{(4)}{=} \phi^{-1}\Big(\ominus_n\big(\phi(\mathbf{P})\oplus_n\phi(\mathbf{Q})\big)\oplus_n\phi\big(\mathbf{P}\oplus_m(\mathbf{Q}\oplus_m\mathbf{R})\big)\Big)\\
&\stackrel{(5)}{=} \phi^{-1}\Big(\ominus_n\big(\phi(\mathbf{P})\oplus_n\phi(\mathbf{Q})\big)\oplus_n\big(\phi(\mathbf{P})\oplus_n\phi(\mathbf{Q}\oplus_m\mathbf{R})\big)\Big)\\
&\stackrel{(6)}{=} \phi^{-1}\Big(\ominus_n\big(\phi(\mathbf{P})\oplus_n\phi(\mathbf{Q})\big)\oplus_n\big(\phi(\mathbf{P})\oplus_n(\phi(\mathbf{Q})\oplus_n\phi(\mathbf{R}))\big)\Big)\\
&\stackrel{(7)}{=} \phi^{-1}(\operatorname{gyr}_n[\phi(\mathbf{P}),\phi(\mathbf{Q})]\phi(\mathbf{R})).
\end{aligned} \tag{21}$$

The derivation of Eq. (21) follows.

(1) follows from Eq. (3).

(2) follows from Eq. (4).

(3) follows from Eq. (20).

(4), (5), and (6) follow from Eq. (4).

(7) follows from Eq. (3).

$\square$

H. Proof of Theorem 2.4

Axiom (G1)

Proof. Let $\phi(\bar{\mathbf{I}})$ be a left identity in G_n where $\bar{\mathbf{I}} \in G_m$. Then for $\mathbf{P} \in G_m$, we have

$$\phi(\mathbf{P}) = \phi(\bar{\mathbf{I}})\oplus_n\phi(\mathbf{P}) = \phi(\bar{\mathbf{I}}\oplus_m\mathbf{P}),$$

which shows that $\mathbf{P} = \bar{\mathbf{I}}\oplus_m\mathbf{P}$ and therefore $\bar{\mathbf{I}}$ is a left identity in G_m.

$\square$

Axiom (G2)

Proof. For $\mathbf{P} \in G_m$, by the assumption that $(G_n, \oplus_n, \otimes_n)$ is a gyrovector space, there exists a left inverse $\ominus_n\phi(\mathbf{P})$ of $\phi(\mathbf{P})$ such that

$$\ominus_n\phi(\mathbf{P})\oplus_n\phi(\mathbf{P}) = \phi(\bar{\mathbf{I}}).$$

Hence

$$\bar{\mathbf{I}} = \phi^{-1}(\ominus_n\phi(\mathbf{P})\oplus_n\phi(\mathbf{P})) = \phi^{-1}(\ominus_n\phi(\mathbf{P}))\oplus_m\mathbf{P},$$

which shows that $\phi^{-1}(\ominus_n\phi(\mathbf{P})) \in G_m$ is a left inverse of $\mathbf{P}$.

$\square$

Axiom (G3)

Proof. For $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in G_m$, we have

$$
\begin{aligned}
\mathbf{P} \oplus_m (\mathbf{Q} \oplus_m \mathbf{R}) &\stackrel{(1)}{=} \mathbf{P} \oplus_m \phi^{-1}(\phi(\mathbf{Q}) \oplus_n \phi(\mathbf{R})) \\
&\stackrel{(2)}{=} \phi^{-1}\big(\phi(\mathbf{P}) \oplus_n (\phi(\mathbf{Q}) \oplus_n \phi(\mathbf{R}))\big) \\
&\stackrel{(3)}{=} \phi^{-1}\big((\phi(\mathbf{P}) \oplus_n \phi(\mathbf{Q})) \oplus_n \mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\mathbf{R})\big) \\
&\stackrel{(4)}{=} \phi^{-1}\big((\phi(\mathbf{P}) \oplus_n \phi(\mathbf{Q})) \oplus_n \phi(\mathrm{gyr}_m[\mathbf{P}, \mathbf{Q}]\mathbf{R})\big) \\
&\stackrel{(5)}{=} \phi^{-1}(\phi(\mathbf{P}) \oplus_n \phi(\mathbf{Q})) \oplus_m \mathrm{gyr}_m[\mathbf{P}, \mathbf{Q}]\mathbf{R} \\
&\stackrel{(6)}{=} (\mathbf{P} \oplus_m \mathbf{Q}) \oplus_m \mathrm{gyr}_m[\mathbf{P}, \mathbf{Q}]\mathbf{R}.
\end{aligned}
\tag{22}
$$

The derivation of Eq. (22) follows.

(1) follows from the definition of the binary operation $\oplus_m$.

(2) follows from the definition of the binary operation $\oplus_m$.

(3) follows from Axiom (G3) verified by gyrovector space $(G_n, \oplus_n, \otimes_n)$.

(4) follows from the definition of the gyroautomorphism $\mathrm{gyr}_m[.,.]$.

(5) follows from the definition of the binary operation $\oplus_m$.

(6) follows from the definition of the binary operation $\oplus_m$.

$\square$

Axiom (G4)

Proof. For $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in G_m$, we have

$$
\begin{aligned}
\mathrm{gyr}_m[\mathbf{P}, \mathbf{Q}]\mathbf{R} &\stackrel{(1)}{=} \phi^{-1}\big(\mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\mathbf{R})\big) \\
&\stackrel{(2)}{=} \phi^{-1}\big(\mathrm{gyr}_n[\phi(\mathbf{P}) \oplus_n \phi(\mathbf{Q}), \phi(\mathbf{Q})]\phi(\mathbf{R})\big) \\
&\stackrel{(3)}{=} \phi^{-1}\big(\mathrm{gyr}_n[\phi(\mathbf{P} \oplus_m \mathbf{Q}), \phi(\mathbf{Q})]\phi(\mathbf{R})\big) \\
&\stackrel{(4)}{=} \phi^{-1}\big(\phi(\mathrm{gyr}_m[\mathbf{P} \oplus_m \mathbf{Q}, \mathbf{Q}]\mathbf{R})\big) \\
&\stackrel{(5)}{=} \mathrm{gyr}_m[\mathbf{P} \oplus_m \mathbf{Q}, \mathbf{Q}]\mathbf{R}.
\end{aligned}
\tag{23}
$$

The derivation of Eq. (23) follows.

(1) follows from the definition of the gyroautomorphism $\mathrm{gyr}_m[.,.]$.

(2) follows from Axiom (G4) verified by gyrovector space $(G_n, \oplus_n, \otimes_n)$.

(3) follows from the definition of the binary operation $\oplus_m$.

(4) follows from the definition of the gyroautomorphism $\mathrm{gyr}_m[.,.]$.

$\square$

Gyrocommutative Law

Proof. For $\mathbf{P}, \mathbf{Q} \in G_m$, we have

$$
\begin{aligned}
\mathbf{P} \oplus_m \mathbf{Q} &\overset{(1)}{=} \phi^{-1}\big(\phi(\mathbf{P}) \oplus_n \phi(\mathbf{Q})\big) \\
&\overset{(2)}{=} \phi^{-1}\big(\operatorname{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})](\phi(\mathbf{Q}) \oplus_n \phi(\mathbf{P}))\big) \\
&\overset{(3)}{=} \phi^{-1}\big(\operatorname{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\mathbf{Q} \oplus_m \mathbf{P})\big) \\
&\overset{(4)}{=} \operatorname{gyr}_m[\mathbf{P}, \mathbf{Q}](\mathbf{Q} \oplus_m \mathbf{P}).
\end{aligned}
\tag{24}
$$

The derivation of Eq. (24) follows.

(1) follows from the definition of the binary operation $\oplus_m$.

(2) follows from the Gyrocommutative Law verified by gyrovector space $(G_n, \oplus_n, \otimes_n)$.

(3) follows from the definition of the binary operation $\oplus_m$.

(4) follows from the definition of the gyroautomorphism $\operatorname{gyr}_m[.,.]$.

$\square$

Axiom (V1)

Proof. For $t \in \mathbb{R}$ and $\mathbf{P} \in G_m$, by the assumption that $(G_n, \oplus_n, \otimes_n)$ is a gyrovector space and from Eqs. (5) and (20), we have

$$
1 \otimes_m \mathbf{P} = \phi^{-1}(1 \otimes_n \phi(\mathbf{P})) = \phi^{-1}(\phi(\mathbf{P})) = \mathbf{P}.
$$

$$
0 \otimes_m \mathbf{P} = \phi^{-1}(0 \otimes_n \phi(\mathbf{P})) = \phi^{-1}(\phi(\bar{\mathbf{I}})) = \bar{\mathbf{I}}.
$$

$$
t \otimes_m \bar{\mathbf{I}} = \phi^{-1}(t \otimes_n \phi(\bar{\mathbf{I}})) = \phi^{-1}(\phi(\bar{\mathbf{I}})) = \bar{\mathbf{I}}.
$$

$$
(-1) \otimes_m \mathbf{P} = \phi^{-1}((-1) \otimes_n \phi(\mathbf{P})) = \phi^{-1}(\ominus_n \phi(\mathbf{P})) = \phi^{-1}(\phi(\ominus_m \mathbf{P})) = \ominus_m \mathbf{P}.
$$

$\square$

Axiom (V2)

Proof. For $s, t \in \mathbb{R}$ and $\mathbf{P} \in G_m$, we have

$$
\begin{aligned}
(s + t) \otimes_m \mathbf{P} &\overset{(1)}{=} \phi^{-1}((s + t) \otimes_n \phi(\mathbf{P})) \\
&\overset{(2)}{=} \phi^{-1}(s \otimes_n \phi(\mathbf{P}) \oplus_n t \otimes_n \phi(\mathbf{P})) \\
&\overset{(3)}{=} \phi^{-1}(\phi(\phi^{-1}(s \otimes_n \phi(\mathbf{P}))) \oplus_n \phi(\phi^{-1}(t \otimes_n \phi(\mathbf{P})))) \\
&\overset{(4)}{=} \phi^{-1}(s \otimes_n \phi(\mathbf{P})) \oplus_m \phi^{-1}(t \otimes_n \phi(\mathbf{P})) \\
&\overset{(5)}{=} s \otimes_m \mathbf{P} \oplus_m t \otimes_m \mathbf{P}.
\end{aligned}
\tag{25}
$$

The derivation of Eq. (25) follows.

(1) follows from the definition of the scalar multiplication $\otimes_m$.

(2) follows from Axiom (V2) verified by gyrovector space $(G_n, \oplus_n, \otimes_n)$.

(3) follows from the fact that ϕ is an isometry.

(4) follows from the definition of the binary operation $\oplus_m$.

(5) follows from the definition of the scalar multiplication $\otimes_m$.

$\square$

Axiom (V3)

Proof. For $s, t \in \mathbb{R}$ and $\mathbf{P} \in G_m$, we have

$$
\begin{aligned}
(st) \otimes_m \mathbf{P} &\overset{(1)}{=} \phi^{-1}((st) \otimes_n \phi(\mathbf{P})) \\
&\overset{(2)}{=} \phi^{-1}(s \otimes_n (t \otimes_n \phi(\mathbf{P}))) \\
&\overset{(3)}{=} \phi^{-1}(s \otimes_n \phi(\phi^{-1}(t \otimes_n \phi(\mathbf{P})))) \\
&\overset{(4)}{=} s \otimes_m \phi^{-1}(t \otimes_n \phi(\mathbf{P})) \\
&\overset{(5)}{=} s \otimes_m (t \otimes_m \mathbf{P}).
\end{aligned}
\tag{26}
$$

The derivation of Eq. (26) follows.

(1) follows from the definition of the scalar multiplication $\otimes_m$.

(2) follows from Axiom (V3) verified by gyrovector space $(G_n, \oplus_n, \otimes_n)$.

(3) follows from the fact that ϕ is an isometry.

(4) follows from the definition of the scalar multiplication $\otimes_m$.

(5) follows from the definition of the scalar multiplication $\otimes_m$.

$\square$

Axiom (V4)

Proof. For $t \in \mathbb{R}$ and $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in G_m$, we have

$$
\begin{aligned}
\mathrm{gyr}_m[\mathbf{P}, \mathbf{Q}](t \otimes_m \mathbf{R}) &\overset{(1)}{=} \phi^{-1}(\mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(t \otimes_m \mathbf{R})) \\
&\overset{(2)}{=} \phi^{-1}(\mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\phi^{-1}(t \otimes_n \phi(\mathbf{R})))) \\
&\overset{(3)}{=} \phi^{-1}(\mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})](t \otimes_n \phi(\mathbf{R}))) \\
&\overset{(4)}{=} \phi^{-1}(t \otimes_n \mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\mathbf{R})) \\
&\overset{(5)}{=} \phi^{-1}(t \otimes_n \phi(\phi^{-1}(\mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\mathbf{R})))) \\
&\overset{(6)}{=} t \otimes_m \phi^{-1}(\mathrm{gyr}_n[\phi(\mathbf{P}), \phi(\mathbf{Q})]\phi(\mathbf{R})) \\
&\overset{(7)}{=} t \otimes_m \mathrm{gyr}_m[\mathbf{P}, \mathbf{Q}]\mathbf{R}.
\end{aligned}
\tag{27}
$$

The derivation of Eq. (27) follows.

(1) follows from the definition of the gyroautomorphism $\mathrm{gyr}_m[.,.]$.

(2) follows from the definition of the scalar multiplication $\otimes_m$.

(3) follows from the fact that ϕ is an isometry.

(4) follows from Axiom (V4) verified by gyrovector space $(G_n, \oplus_n, \otimes_n)$.

(5) follows from the fact that ϕ is an isometry.

(6) follows from the definition of the scalar multiplication $\otimes_m$.

(7) follows from the definition of the gyroautomorphism $\mathrm{gyr}_m[.,.]$.

$\square$

Axiom (V5)

Proof. For $s, t \in \mathbb{R}$ and $\mathbf{P}, \mathbf{Q} \in G_m$, we have

$$
\begin{aligned}
\mathrm{gyr}_m[s \otimes_m \mathbf{P}, t \otimes_m \mathbf{P}]\mathbf{Q} &\overset{(1)}{=} \phi^{-1}(\mathrm{gyr}_n[\phi(s \otimes_m \mathbf{P}), \phi(t \otimes_m \mathbf{P})]\phi(\mathbf{Q})) \\
&\overset{(2)}{=} \phi^{-1}(\mathrm{gyr}_n[\phi(\phi^{-1}(s \otimes_n \phi(\mathbf{P}))), \phi(\phi^{-1}(t \otimes_n \phi(\mathbf{P})))]\phi(\mathbf{Q})) \\
&\overset{(3)}{=} \phi^{-1}(\mathrm{gyr}_n[s \otimes_n \phi(\mathbf{P}), t \otimes_n \phi(\mathbf{P})]\phi(\mathbf{Q})) \\
&\overset{(4)}{=} \phi^{-1}(\phi(\mathbf{Q})) \\
&\overset{(5)}{=} \mathbf{Q}.
\end{aligned}
\tag{28}
$$

The derivation of Eq. (28) follows.

(1) follows from the definition of the gyroautomorphism $\mathrm{gyr}_m[.,.]$.

(2) follows from the definition of the scalar multiplication $\otimes_m$.

(3) follows from the fact that ϕ is an isometry.

(4) follows from Axiom (V5) verified by gyrovector space $(G_n, \oplus_n, \otimes_n)$.

(5) follows from the fact that ϕ is an isometry.

$\square$

I. Proof of Theorem 2.12

Proof. We first prove the following lemma:

Lemma I.1. *Gyrogroups* $(\mathrm{Gr}_{n,p}, \oplus_{gr})$ *verify the Left Gyrotranslation Law ([Ungar, 2014](#)), that is,*

$$
\ominus_{gr}(\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} (\mathbf{P} \oplus_{gr} \mathbf{R}) = \mathrm{gyr}[\mathbf{P}, \mathbf{Q}](\ominus_{gr}\mathbf{Q} \oplus_{gr} \mathbf{R}),
$$

where $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \mathrm{Gr}_{n,p}$.

Proof. First, note that gyrogroups $(\mathrm{Gr}_{n,p}, \oplus_{gr})$ verify the Left Cancellation Law ([Ungar, 2014](#)), i.e.,

$$
\ominus_{gr}\mathbf{P} \oplus_{gr} (\mathbf{P} \oplus_{gr} \mathbf{Q}) = \mathbf{Q},
$$

where $\mathbf{P}, \mathbf{Q} \in \mathrm{Gr}_{n,p}$.

We have

$$
\begin{aligned}
(\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} \mathrm{gyr}[\mathbf{P}, \mathbf{Q}](\ominus_{gr}\mathbf{Q} \oplus_{gr} \mathbf{R}) &\overset{(1)}{=} \mathbf{P} \oplus_{gr} (\mathbf{Q} \oplus_{gr} (\ominus_{gr}\mathbf{Q} \oplus_{gr} \mathbf{R})) \\
&\overset{(2)}{=} \mathbf{P} \oplus_{gr} \mathbf{R},
\end{aligned}
$$

where (1) follows from the Left Gyroassociative Law, and (2) follows from the Left Cancellation Law. Hence

$$
\begin{aligned}
\ominus_{gr}(\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} (\mathbf{P} \oplus_{gr} \mathbf{R}) &= \ominus_{gr}(\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} ((\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} \mathrm{gyr}[\mathbf{P}, \mathbf{Q}](\ominus_{gr}\mathbf{Q} \oplus_{gr} \mathbf{R})) \\
&\overset{(1)}{=} \mathrm{gyr}[\mathbf{P}, \mathbf{Q}](\ominus_{gr}\mathbf{Q} \oplus_{gr} \mathbf{R}),
\end{aligned}
$$

where (1) follows from the Left Cancellation Law. $\square$

We also need to prove the following lemma:

Lemma I.2. *Gyroautomorphisms* $\mathrm{gyr}_{gr}[.,.]$ *preserve the norm.*

Proof. Denote by O_n the space of $n \times n$ orthogonal matrices, $\langle.\rangle_F$ the Frobenius inner product, $\|.\|_F$ the Frobenius norm. Let $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \mathrm{Gr}_{n,p}$. Then notice that

$$\mathrm{gyr}_{gr}[\mathbf{P}, \mathbf{Q}]\mathbf{R} = \begin{bmatrix} \mathbf{O}_1 & 0 \\ 0 & \mathbf{O}_2 \end{bmatrix} \mathbf{R} \begin{bmatrix} \mathbf{O}_1 & 0 \\ 0 & \mathbf{O}_2 \end{bmatrix}^T,$$

where $\mathbf{O}_1 \in \mathrm{O}_p, \mathbf{O}_2 \in \mathrm{O}_{n-p}$. Let $\mathbf{O} = \begin{bmatrix} \mathbf{O}_1 & 0 \\ 0 & \mathbf{O}_2 \end{bmatrix}$. Then

$$
\begin{aligned}
\| \mathrm{gyr}_{gr}[\mathbf{P}, \mathbf{Q}]\mathbf{R}\| &= \| \mathrm{Log}^{gr}_{\mathbf{I}_{n,p}}(\mathrm{gyr}_{gr}[\mathbf{P}, \mathbf{Q}]\mathbf{R})\|_F \\
&\overset{(1)}{=} \| \mathrm{Log}^{gr}_{\mathbf{OI}_{n,p}\mathbf{O}^T}(\mathbf{ORO}^T)\|_F \\
&\overset{(2)}{=} \|\mathbf{O}\, \mathrm{Log}^{gr}_{\mathbf{I}_{n,p}}(\mathbf{R})\mathbf{O}^T\|_F \\
&= \| \mathrm{Log}^{gr}_{\mathbf{I}_{n,p}}(\mathbf{R})\|_F \\
&= \|\mathbf{R}\|.
\end{aligned}
\tag{29}
$$

The derivation of Eq. (29) follows.

(1) follows from the fact that $\mathbf{I}_{n,p} = \mathbf{OI}_{n,p}\mathbf{O}^T$.

(2) follows from Nguyen (2022b) (see Lemma 3.19).

$\square$

We now have the following chain of equations:

$$\| \ominus_{gr} (\mathbf{A} \oplus_{gr} \mathbf{P}) \oplus_{gr} (\mathbf{A} \oplus_{gr} \mathbf{Q})\| \overset{(1)}{=} \| \mathrm{gyr}[\mathbf{A}, \mathbf{P}](\ominus_{gr}\mathbf{P} \oplus_{gr} \mathbf{Q})\|$$
$$\overset{(2)}{=} \| \ominus_{gr} \mathbf{P} \oplus_{gr} \mathbf{Q}\|,$$

where (1) follows from the Left Gyrotranslation Law, and (2) follows from the invariance of the norm under gyroautomorphisms (Lemma I.2).

$\square$

J. Proof of Theorem 2.13

Proof. Let $\mathbf{P}, \mathbf{Q}, \mathbf{R}, \mathbf{S} \in \mathrm{Gr}_{n,p}$. Then by the Left Gyroassociative Law and Left Cancellation Law,

$$\mathrm{gyr}[\mathbf{P}, \mathbf{Q}]\mathbf{R} = \ominus_{gr}(\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} (\mathbf{P} \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{R})),$$

$$\mathrm{gyr}[\mathbf{P}, \mathbf{Q}]\mathbf{S} = \ominus_{gr}(\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} (\mathbf{P} \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{S})).$$

Let $\mathbf{X} = \ominus_{gr}(\mathbf{P} \oplus_{gr} \mathbf{Q})$, $\mathbf{Y} = \mathbf{P} \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{R})$, and $\mathbf{Z} = \mathbf{P} \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{S})$. Then we have the following chain of

equations:

$$
\begin{aligned}
d(\mathrm{gyr}[\mathbf{P},\mathbf{Q}]\mathbf{R}, \mathrm{gyr}[\mathbf{P},\mathbf{Q}]\mathbf{S}) &= \| \ominus_{gr} \mathrm{gyr}[\mathbf{P},\mathbf{Q}]\mathbf{R} \oplus_{gr} \mathrm{gyr}[\mathbf{P},\mathbf{Q}]\mathbf{S} \| \\
&= \| \ominus_{gr} (\mathbf{X} \oplus_{gr} \mathbf{Y}) \oplus_{gr} (\mathbf{X} \oplus_{gr} \mathbf{Z}) \| \\
&\overset{(1)}{=} \| \mathrm{gyr}[\mathbf{X},\mathbf{Y}](\ominus_{gr}\mathbf{Y} \oplus_{gr} \mathbf{Z}) \| \\
&\overset{(2)}{=} \| \ominus_{gr} \mathbf{Y} \oplus_{gr} \mathbf{Z} \| \\
&= \| \ominus_{gr} (\mathbf{P} \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{R})) \oplus_{gr} (\mathbf{P} \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{S})) \| \\
&\overset{(3)}{=} \| \mathrm{gyr}[\mathbf{P}, \mathbf{Q} \oplus_{gr} \mathbf{R}](\ominus_{gr}(\mathbf{Q} \oplus_{gr} \mathbf{R}) \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{S})) \| \\
&\overset{(4)}{=} \| \ominus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{R}) \oplus_{gr} (\mathbf{Q} \oplus_{gr} \mathbf{S}) \| \\
&\overset{(5)}{=} \| \mathrm{gyr}[\mathbf{Q},\mathbf{R}](\ominus_{gr}\mathbf{R} \oplus_{gr} \mathbf{S}) \| \\
&\overset{(6)}{=} \| \ominus_{gr} \mathbf{R} \oplus_{gr} \mathbf{S} \| \\
&= d(\mathbf{R}, \mathbf{S}).
\end{aligned}
\tag{30}
$$

The derivation of Eq. (30) follows.

(1) follows from the Left Gyrotranslation Law.

(2) follows from Lemma I.2.

(3) follows from the Left Gyrotranslation Law.

(4) follows from Lemma I.2.

(5) follows from the Left Gyrotranslation Law.

(6) follows from Lemma I.2.

$\square$

K. Proof of Theorem 2.14

Proof. We first prove the following Lemma:

Lemma K.1. *Grassmann inverse maps preserve the norm.*

Proof. For $\mathbf{P} \in \mathrm{Gr}_{n,p}$, we have

$$
\begin{aligned}
\| \ominus_{gr} \mathbf{P} \| &= \| \mathrm{Log}^{gr}_{\mathbf{I}_{n,p}}(\ominus_{gr}\mathbf{P}) \|_F \\
&= \| - \mathrm{Log}^{gr}_{\mathbf{I}_{n,p}}(\mathbf{P}) \|_F \\
&= \| \mathrm{Log}^{gr}_{\mathbf{I}_{n,p}}(\mathbf{P}) \|_F \\
&= \| \mathbf{P} \|.
\end{aligned}
$$

$\square$

For $\mathbf{P}, \mathbf{Q} \in \mathrm{Gr}_{n,p}$, we have

$$
\begin{aligned}
\| \ominus_{gr} \mathbf{P} \oplus_{gr} \mathbf{Q} \| &\overset{(1)}{=} \| \ominus_{gr} (\ominus\mathbf{P} \oplus_{gr} \mathbf{Q}) \| \\
&= \| \ominus_{gr} (\ominus_{gr}\mathbf{P} \oplus_{gr} \mathbf{Q}) \oplus_{gr} (\ominus_{gr}\mathbf{P} \oplus_{gr} \mathbf{P}) \| \\
&\overset{(2)}{=} \| \mathrm{gyr}[\ominus_{gr}\mathbf{P}, \mathbf{Q}](\ominus_{gr}\mathbf{Q} \oplus_{gr} \mathbf{P}) \| \\
&\overset{(3)}{=} \| \ominus_{gr} \mathbf{Q} \oplus_{gr} \mathbf{P} \|.
\end{aligned}
\tag{31}
$$

The derivation of Eq. (31) follows.

(1) follows from Lemma K.1.

(2) follows from the Left Gyrotranslation Law.

(3) follows from Lemma I.2.

Notice that

$$\|\mathbf{P} \oplus_{gr} (\ominus_{gr}\mathbf{Q})\| \overset{(1)}{=} \| \operatorname{gyr}[\mathbf{P}, \ominus_{gr}\mathbf{Q}](\ominus_{gr}\mathbf{Q} \oplus_{gr} \mathbf{P})\|$$
$$\overset{(2)}{=} \| \ominus_{gr} \mathbf{Q} \oplus_{gr} \mathbf{P}\|, \tag{32}$$

where (1) follows from the Gyrocommutative Law, and (2) follows from Lemma I.2.

Combining Eqs. (31) and (32) results in

$$\| \ominus_{gr} \mathbf{P} \oplus_{gr} \mathbf{Q}\| = \|\mathbf{P} \oplus_{gr} (\ominus_{gr}\mathbf{Q})\|$$
$$= \| \ominus_{gr} (\ominus_{gr}\mathbf{P}) \oplus_{gr} (\ominus\mathbf{Q})\|,$$

which leads to the conclusion of the theorem.

$\square$

L. Proof of Theorem 2.16

Proof. We need to prove the following lemmas:

Lemma L.1. *Gyrovector spaces* $(\operatorname{Sym}_n^+, \oplus_g, \otimes_g)$ *verify the Left Gyrotranslation Law, that is,*

$$\ominus_g(\mathbf{P} \oplus_g \mathbf{Q}) \oplus_g (\mathbf{P} \oplus_g \mathbf{R}) = \operatorname{gyr}[\mathbf{P}, \mathbf{Q}](\ominus_g\mathbf{Q} \oplus_g \mathbf{R}),$$

where $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \operatorname{Gr}_{n,p}$.

Proof. Note that gyrovector spaces $(\operatorname{Sym}_n^+, \oplus_g, \otimes_g)$ verify the Left Cancellation Law and Left Gyroassociative Law. Then the lemma can be proved by using the same arguments as those in the proof of Lemma I.1.

$\square$

Lemma L.2. *Gyroautomorphisms* $\operatorname{gyr}_g[.,.]$ *preserve the norm.*

Proof. The lemma can be easily proved by using the expressions of gyroautomorphisms $\operatorname{gyr}_g[.,.]$.

$\square$

We now have the following chain of equations:

$$\| \ominus_g (\mathbf{A} \oplus_g \mathbf{P}) \oplus_g (\mathbf{A} \oplus_g \mathbf{Q})\| \overset{(1)}{=} \| \operatorname{gyr}[\mathbf{A}, \mathbf{P}](\ominus_g\mathbf{P} \oplus_g \mathbf{Q})\|$$
$$\overset{(2)}{=} \| \ominus_g \mathbf{P} \oplus_g \mathbf{Q}\|,$$

where (1) follows from the Left Gyrotranslation Law (Lemma L.1), and (2) follows from the invariance of the norm under gyroautomorphisms (Lemma L.2).

$\square$

M. Proof of Theorem 2.17

Proof. Note that gyrovector spaces $(\operatorname{Sym}_n^+, \oplus_g, \otimes_g)$ verify the Left Gyroassociative Law, Left Cancellation Law, and Left Gyrotranslation Law (Lemma L.1) and that gyroautomorphisms $\operatorname{gyr}_g[.,.]$ preserve the norm (Lemma L.2). Then the lemma can be proved by using the same arguments as those in the proof of Theorem 2.13. $\square$

N. Proof of Theorem 2.18

Proof. The following lemma can be proved by using the same arguments as those in the proof of Lemma K.1.

Lemma N.1. *SPD inverse maps preserve the norm.*

Note that gyrovector spaces $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$ verify the Gyrocommutative Law and Left Gyrotranslation Law (Lemma L.1). Note also that SPD inverse maps preserve the norm (Lemma N.1) and that gyroautomorphisms $\mathrm{gyr}_g[.,.]$ preserve the norm (Lemma L.2). Then the lemma can be proved by using the same arguments as those in the proof of Theorem 2.14.

$\square$

O. Proof of Theorem 2.21

LC Gyrovector Spaces

Proof. We only need to prove the fist identity. For $\mathbf{P}, \mathbf{Q}, \mathbf{R} \in \mathrm{Sym}_n^+$, by the Left Gyrotranslation Law,

$$\ominus_g(\ominus_g\mathbf{P} \oplus_g \mathbf{Q}) \oplus_g (\ominus_g\mathbf{P} \oplus_g \mathbf{R}) = \mathrm{gyr}[\ominus_g\mathbf{P}, \mathbf{Q}](\ominus_g\mathbf{Q} \oplus_g \mathbf{R}).$$

Since gyroautomorphisms preserve the norm, we have

$$\| \ominus_g (\ominus_g\mathbf{P} \oplus_g \mathbf{Q}) \oplus_g (\ominus_g\mathbf{P} \oplus_g \mathbf{R})\| = \| \mathrm{gyr}[\ominus_g\mathbf{P}, \mathbf{Q}](\ominus_g\mathbf{Q} \oplus_g \mathbf{R})\| = \| \ominus_g \mathbf{Q} \oplus_g \mathbf{R}\|,$$

which results in

$$\| \ominus_g \widetilde{\mathbf{R}} \oplus_g \widetilde{\mathbf{Q}}\| = \|\widetilde{\mathbf{P}}\|. \tag{33}$$

Hence

$$
\begin{aligned}
p^2 = \|\widetilde{\mathbf{P}}\|^2 &= \| \ominus_{lc} \widetilde{\mathbf{R}} \oplus_{lc} \widetilde{\mathbf{Q}}\|^2 \\
&= \langle \lfloor\varphi(\widetilde{\mathbf{Q}})\rfloor - \lfloor\varphi(\widetilde{\mathbf{R}})\rfloor + \log(\mathbb{D}(\varphi(\widetilde{\mathbf{R}}))^{-1}\mathbb{D}(\varphi(\widetilde{\mathbf{Q}}))), \lfloor\varphi(\widetilde{\mathbf{Q}})\rfloor - \lfloor\varphi(\widetilde{\mathbf{R}})\rfloor + \log(\mathbb{D}(\varphi(\widetilde{\mathbf{R}}))^{-1}\mathbb{D}(\varphi(\widetilde{\mathbf{Q}})))\rangle_F \\
&= \|\lfloor\varphi(\widetilde{\mathbf{Q}})\rfloor\|_F^2 + \|\lfloor\varphi(\widetilde{\mathbf{R}})\rfloor\|_F^2 - 2\langle\lfloor\varphi(\widetilde{\mathbf{Q}})\rfloor, \lfloor\varphi(\widetilde{\mathbf{R}})\rfloor\rangle_F + \\
&\quad \| \log(\mathbb{D}(\varphi(\widetilde{\mathbf{Q}})))\|_F^2 + \| \log(\mathbb{D}(\varphi(\widetilde{\mathbf{R}})))\|_F^2 - 2\langle\log(\mathbb{D}(\varphi(\widetilde{\mathbf{Q}}))), \log(\mathbb{D}(\varphi(\widetilde{\mathbf{R}})))\rangle_F.
\end{aligned}
\tag{34}
$$

Notice that

$$q^2 + r^2 = \|\widetilde{\mathbf{Q}}\|^2 + \|\widetilde{\mathbf{R}}\|^2 = \|\lfloor\varphi(\widetilde{\mathbf{Q}})\rfloor\|_F^2 + \|\lfloor\varphi(\widetilde{\mathbf{R}})\rfloor\|_F^2 + \| \log(\mathbb{D}(\varphi(\widetilde{\mathbf{Q}})))\|_F^2 + \| \log(\mathbb{D}(\varphi(\widetilde{\mathbf{R}})))\|_F^2, \tag{35}$$

and

$$2\langle\widetilde{\mathbf{Q}}, \widetilde{\mathbf{R}}\rangle = 2\langle\lfloor\varphi(\widetilde{\mathbf{Q}})\rfloor, \lfloor\varphi(\widetilde{\mathbf{R}})\rfloor\rangle_F + 2\langle\log(\mathbb{D}(\varphi(\widetilde{\mathbf{Q}}))), \log(\mathbb{D}(\varphi(\widetilde{\mathbf{R}})))\rangle_F. \tag{36}$$

Combining Eqs. (34), (35), and (36), we get

$$p^2 = q^2 + r^2 - 2\langle\widetilde{\mathbf{Q}}, \widetilde{\mathbf{R}}\rangle = q^2 + r^2 - 2qr\cos\alpha.$$

$\square$

LE Gyrovector Spaces

Proof. From Eq. (33),

$$
\begin{aligned}
p^2 = \|\widetilde{\mathbf{P}}\|^2 &= \| \ominus_{le} \widetilde{\mathbf{R}} \oplus_{le} \widetilde{\mathbf{Q}} \|^2 \\
&= \| \log(\widetilde{\mathbf{Q}}) - \log(\widetilde{\mathbf{R}}) \|_F^2 \\
&= \| \log(\widetilde{\mathbf{Q}}) \|_F^2 + \| \log(\widetilde{\mathbf{R}}) \|_F^2 - 2\langle \log(\widetilde{\mathbf{Q}}), \log(\widetilde{\mathbf{R}}) \rangle_F \\
&= \| \widetilde{\mathbf{Q}} \|^2 + \| \widetilde{\mathbf{R}} \|^2 - 2\langle \widetilde{\mathbf{Q}}, \widetilde{\mathbf{R}} \rangle \\
&= q^2 + r^2 - 2qr \cos \alpha.
\end{aligned}
$$

$\square$

P. Proof of Theorem 2.23

Proof. We need to prove the following lemmas:

Lemma P.1. *Let* $\mathbf{P}$ *and* $\mathbf{Q}$ *be two distinct points in a gyrovector space* $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$. *Then the geodesic* $\delta_{\mathbf{P} \to \mathbf{Q}}(t), 0 \leq t \leq 1$ *joining* $\mathbf{P}$ *and* $\mathbf{Q}$ *that passes through* $\mathbf{P}$ *when* $t = 0$ *and passes through* $\mathbf{Q}$ *when* $t = 1$ *is given by*

$$
\delta_{\mathbf{P} \to \mathbf{Q}}(t) = \mathbf{P} \oplus_g t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q}).
$$

Proof. The lemma can be proved using the expressions of the binary operation, inverse operation, and scalar multiplication in LE, LC, and AI gyrovector spaces given in Nguyen (2022a;b). $\square$

Lemma P.2. *Let* $\mathbf{P}$ *and* $\mathbf{Q}$ *be two distinct points in a gyrovector space* $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$. *Let* $\mathbf{S}$ *be a point on the geodesic* $\delta_{\mathbf{P} \to \mathbf{Q}}(t), 0 \leq t \leq 1$ *joining* $\mathbf{P}$ *and* $\mathbf{Q}$, *and* $\mathbf{R} \notin \delta_{\mathbf{P} \to \mathbf{Q}}(t)$. *Then*

$$
\angle \mathbf{RPS} = \angle \mathbf{RPQ}.
$$

Proof. By Lemma P.1, $\delta_{\mathbf{P} \to \mathbf{Q}}(t)$ can be given as

$$
\delta_{\mathbf{P} \to \mathbf{Q}}(t) = \mathbf{P} \oplus_g t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q}). \tag{37}
$$

By the definition of the scalar multiplication,

$$
t \otimes_g \mathbf{P} = \mathrm{Exp}_{\mathbf{I}_n}^g (t \, \mathrm{Log}_{\mathbf{I}_n}^g (\mathbf{P})),
$$

where $t \in \mathbb{R}$, which results in

$$
\mathrm{Log}_{\mathbf{I}_n}(t \otimes_g \mathbf{P}) = t \, \mathrm{Log}_{\mathbf{I}_n}^g (\mathbf{P}). \tag{38}
$$

From Eq. (38), we get

$$
\| \mathrm{Log}_{\mathbf{I}_n}(t \otimes_g \mathbf{P}) \|_F = \| t \, \mathrm{Log}_{\mathbf{I}_n}^g (\mathbf{P}) \|_F = t \| \mathrm{Log}_{\mathbf{I}_n}^g (\mathbf{P}) \|_F,
$$

which leads to

$$
\| t \otimes_g \mathbf{P} \| = t \| \mathbf{P} \|. \tag{39}
$$

By the definition of the SPD inner product and Eq. (38), we also have

$$
\begin{aligned}
\langle \mathbf{P}, t \otimes_g \mathbf{Q} \rangle &= \langle \mathrm{Log}_{\mathbf{I}_n}(\mathbf{P}), \mathrm{Log}_{\mathbf{I}_n}(t \otimes_g \mathbf{Q}) \rangle_F \\
&= \langle \mathrm{Log}_{\mathbf{I}_n}(\mathbf{P}), t \, \mathrm{Log}_{\mathbf{I}_n}(\mathbf{Q}) \rangle_F \\
&= t \langle \mathrm{Log}_{\mathbf{I}_n}(\mathbf{P}), \mathrm{Log}_{\mathbf{I}_n}(\mathbf{Q}) \rangle_F \\
&= t \langle \mathbf{P}, \mathbf{Q} \rangle.
\end{aligned} \tag{40}
$$

Therefore

$$
\begin{aligned}
\cos(\angle \mathbf{RPS}) &\overset{(1)}{=} \frac{\langle \ominus_g \mathbf{P} \oplus_g \mathbf{R}, \ominus_g \mathbf{P} \oplus_g \mathbf{S} \rangle}{\| \ominus_g \mathbf{P} \oplus_g \mathbf{R} \|.\| \ominus_g \mathbf{P} \oplus_g \mathbf{S} \|} \\
&\overset{(2)}{=} \frac{\langle \ominus_g \mathbf{P} \oplus_g \mathbf{R}, \ominus_g \mathbf{P} \oplus_g (\mathbf{P} \oplus_g t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q})) \rangle}{\| \ominus_g \mathbf{P} \oplus_g \mathbf{R} \|.\| \ominus_g \mathbf{P} \oplus_g (\mathbf{P} \oplus_g t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q})) \|} \\
&\overset{(3)}{=} \frac{\langle \ominus_g \mathbf{P} \oplus_g \mathbf{R}, t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q}) \rangle}{\| \ominus_g \mathbf{P} \oplus_g \mathbf{R} \|.\| t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q}) \|} \\
&\overset{(4)}{=} \frac{t \langle \ominus_g \mathbf{P} \oplus_g \mathbf{R}, \ominus_g \mathbf{P} \oplus_g \mathbf{Q} \rangle}{t \| \ominus_g \mathbf{P} \oplus_g \mathbf{R} \|.\| \ominus_g \mathbf{P} \oplus_g \mathbf{Q} \|} \\
&\overset{(5)}{=} \cos(\angle \mathbf{RPQ}).
\end{aligned}
\tag{41}
$$

The derivation of Eq. (41) follows.

(1) follows from the definition of SPD gyroangles.

(2) follows from Eq. (37).

(3) follows from the Left Cancellation Law.

(4) follows from Eqs. (39) and (40).

(5) follows from the definition of SPD gyroangles.

This leads to the conclusion of the lemma.

$\square$

Lemma P.3. *Let $\mathbf{P}$ and $\mathbf{Q}$ be two distinct points in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$, $g \in \{le, lc\}$. Denote by $\delta_{\mathbf{P} \to \mathbf{Q}}(t), 0 \le t \le 1$ the geodesic joining $\mathbf{P}$ and $\mathbf{Q}$, $\mathbf{P}' = \delta_{\mathbf{P} \to \mathbf{Q}}(-\infty)$, $\mathbf{Q}' = \delta_{\mathbf{P} \to \mathbf{Q}}(\infty)$, and $\mathbf{R} \in (\mathrm{Sym}_n^+, \oplus_g, \otimes_g)$ such that $\mathbf{R} \notin \delta_{\mathbf{P} \to \mathbf{Q}}(t), t \in \mathbb{R}$. Then there exists a unique $\mathbf{S} \in \delta_{\mathbf{P} \to \mathbf{Q}}(t), t \in \mathbb{R}$ such that $\angle \mathbf{RSQ}' = \frac{\pi}{2}$.*

LE Gyrovector Spaces

Proof. First, it is easy to see that any points $\mathbf{P}', \mathbf{Q}'$ such that $\mathbf{Q} \in \delta_{\mathbf{P} \to \mathbf{Q}'}(t)$ and $\mathbf{P} \in \delta_{\mathbf{P}' \to \mathbf{Q}}(t), 0 \le t \le 1$ can be written as $\delta_{\mathbf{P} \to \mathbf{Q}}(t)$ in Eq. (37) where $t \in \mathbb{R}$. We have

$$
\begin{aligned}
\cos(\angle \mathbf{RSQ}') &\overset{(1)}{=} \frac{\langle \ominus_{le} \mathbf{S} \oplus_{le} \mathbf{R}, \ominus_{le} \mathbf{S} \oplus_{le} \mathbf{Q}' \rangle}{\| \ominus_{le} \mathbf{S} \oplus_{le} \mathbf{R} \|.\| \ominus_{le} \mathbf{S} \oplus_{le} \mathbf{Q}' \|} \\
&= \frac{\langle \log(\mathbf{R}) - \log(\mathbf{S}), \log(\mathbf{Q}') - \log(\mathbf{S}) \rangle_F}{\| \log(\mathbf{R}) - \log(\mathbf{S}) \|_F.\| \log(\mathbf{Q}') - \log(\mathbf{S}) \|_F} \\
&\overset{(2)}{=} \frac{\langle \log(\mathbf{R}) - (1-t)\log(\mathbf{P}') - t\log(\mathbf{Q}'), \log(\mathbf{Q}') - (1-t)\log(\mathbf{P}') - t\log(\mathbf{Q}') \rangle_F}{\| \log(\mathbf{R}) - (1-t)\log(\mathbf{P}') - t\log(\mathbf{Q}') \|_F.\| \log(\mathbf{Q}') - (1-t)\log(\mathbf{P}') - t\log(\mathbf{Q}') \|_F} \\
&= \frac{\langle \log(\mathbf{R}) - \log(\mathbf{P}') - t(\log(\mathbf{Q}') - \log(\mathbf{P}')), \log(\mathbf{Q}') - \log(\mathbf{P}') \rangle_F}{\| \log(\mathbf{R}) - \log(\mathbf{P}') - t(\log(\mathbf{Q}') - \log(\mathbf{P}')) \|_F.\| \log(\mathbf{Q}') - \log(\mathbf{P}') \|_F},
\end{aligned}
$$

where (1) follows from the definition of SPD gyroangles, and (2) follows from the equation of geodesics in LE gyrovector spaces.

It can be seen that there exists $t \in \mathbb{R}$ such that $\langle \log(\mathbf{R}) - \log(\mathbf{P}') - t(\log(\mathbf{Q}') - \log(\mathbf{P}')), \log(\mathbf{Q}') - \log(\mathbf{P}') \rangle_F = 0$ and thus $\cos(\angle \mathbf{RSQ}') = 0$, or equivalently, $\angle \mathbf{RSQ}' = \frac{\pi}{2}$. Now, assuming that there exists two distinct points $\mathbf{S}, \mathbf{S}' \in \delta_{\mathbf{P} \to \mathbf{Q}}(t), t \in \mathbb{R}$ such that $\angle \mathbf{RSQ}' = \angle \mathbf{RS}'\mathbf{Q}' = \frac{\pi}{2}$. Let $p = \| \ominus_{le} \mathbf{S}' \oplus_{le} \mathbf{R} \|$, $q = \| \ominus_{le} \mathbf{S} \oplus_{le} \mathbf{R} \|$, and $r = \| \ominus_{le} \mathbf{S} \oplus_{le} \mathbf{S}' \|$. By the Law of SPD gyrocosines (see Theorem 2.21),

$$
p^2 = q^2 + r^2,
$$

and

$$
q^2 = p^2 + r^2,
$$

which leads to contradiction as $r > 0$. We conclude that there exists a unique $\mathbf{S}$ that verifies the property in Lemma P.3.

$\square$

LC Gyrovector Spaces

Note that the geodesic $\delta_{\mathbf{P} \to \mathbf{Q}}(t), 0 \le t \le 1$ joining $\mathbf{P}$ and $\mathbf{Q}$ can be written as

$$\delta_{\mathbf{P} \to \mathbf{Q}}(t) = \big(\varphi(\mathbf{P}) \oplus_{lc} t \otimes_{lc} (\ominus_{lc}\varphi(\mathbf{P}) \oplus_{lc} \varphi(\mathbf{Q}))\big)\big(\varphi(\mathbf{P}) \oplus_{lc} t \otimes_{lc} (\ominus_{lc}\varphi(\mathbf{P}) \oplus_{lc} \varphi(\mathbf{Q}))\big)^T.$$

Some manipulations lead to

$$\cos(\angle \mathbf{RSQ'}) = \frac{\langle \mathbf{A}, \mathbf{B} \rangle_F}{\|\mathbf{A}\|_F . \|\mathbf{B}\|_F},$$

where

$$\mathbf{A} = (\lfloor \varphi(\mathbf{R}) \rfloor - \lfloor \varphi(\mathbf{P}) \rfloor) + \log(\mathbb{D}(\varphi(\mathbf{R}))) - \log(\mathbb{D}(\varphi(\mathbf{P}))) + t(\lfloor \varphi(\mathbf{P}) \rfloor - \lfloor \varphi(\mathbf{Q'}) \rfloor) + t\big(\log(\mathbb{D}(\varphi(\mathbf{P}))) - \log(\mathbb{D}(\varphi(\mathbf{Q'})))\big),$$

$$\mathbf{B} = \lfloor \varphi(\mathbf{P}) \rfloor - \lfloor \varphi(\mathbf{Q'}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))) - \log(\mathbb{D}(\varphi(\mathbf{Q'}))).$$

It can be seen that there exists $t \in \mathbb{R}$ such that $\langle \mathbf{A}, \mathbf{B} \rangle_F = 0$ and thus $\cos(\angle \mathbf{RSQ'}) = 0$, or equivalently, $\angle \mathbf{RSQ'} = \frac{\pi}{2}$. The uniqueness of $\mathbf{S}$ can be proved by using the same arguments as for LE gyrovector spaces.

Lemma P.4. *Let $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ be a SPD hypergyroplane in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_g, \oplus_g)$, and $\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \setminus \{\mathbf{P}\}$. Then all points on the geodesic $\delta_{\mathbf{P} \to \mathbf{Q}}(t)$ belong to $\mathcal{H}_{\mathbf{W},\mathbf{P}}$.*

Proof. We have

$$
\begin{aligned}
\mathrm{Log}_{\mathbf{P}}(\delta_{\mathbf{P} \to \mathbf{Q}}(t)) &\overset{(1)}{=} \mathrm{Log}_{\mathbf{P}}(\mathbf{P} \oplus_g t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q})) \\
&\overset{(2)}{=} \mathrm{Log}_{\mathbf{P}} \big(\mathrm{Exp}_{\mathbf{P}}(\mathcal{T}_{\mathbf{I}_n \to \mathbf{P}}(\mathrm{Log}_{\mathbf{I}_n}(t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q})))) \big) \\
&= \mathcal{T}_{\mathbf{I}_n \to \mathbf{P}}(\mathrm{Log}_{\mathbf{I}_n}(t \otimes_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q}))) \\
&\overset{(3)}{=} \mathcal{T}_{\mathbf{I}_n \to \mathbf{P}}(t \, \mathrm{Log}_{\mathbf{I}_n}(\ominus_g \mathbf{P} \oplus_g \mathbf{Q})) \\
&= t \mathcal{T}_{\mathbf{I}_n \to \mathbf{P}}(\mathrm{Log}_{\mathbf{I}_n}(\ominus_g \mathbf{P} \oplus_g \mathbf{Q})) \\
&= t \, \mathrm{Log}_{\mathbf{P}} \Big(\mathrm{Exp}_{\mathbf{P}} \big(\mathcal{T}_{\mathbf{I}_n \to \mathbf{P}}(\mathrm{Log}_{\mathbf{I}_n}(\ominus_g \mathbf{P} \oplus_g \mathbf{Q})) \big) \Big) \\
&\overset{(4)}{=} t \, \mathrm{Log}_{\mathbf{P}}(\mathbf{P} \oplus_g (\ominus_g \mathbf{P} \oplus_g \mathbf{Q})) \\
&\overset{(5)}{=} t \, \mathrm{Log}_{\mathbf{P}}(\mathbf{Q}).
\end{aligned}
\tag{42}
$$

The derivation of Eq. (42) follows.

(1) follows from Eq. (37).

(2) follows from the definition of the binary operation in Eq. (1).

(3) follows from Eq. (38).

(4) follows from the definition of the binary operation in Eq. (1).

(5) follows from the Left Cancellation Law.

Therefore

$$\langle \mathrm{Log}_{\mathbf{P}}(\delta_{\mathbf{P} \to \mathbf{Q}}(t)), \mathbf{W} \rangle_{\mathbf{P}} = \langle t \, \mathrm{Log}_{\mathbf{P}}(\mathbf{Q}), \mathbf{W} \rangle_{\mathbf{P}},$$

which results in $\langle \mathrm{Log}_{\mathbf{P}}(\delta_{\mathbf{P} \to \mathbf{Q}}(t)), \mathbf{W} \rangle_{\mathbf{P}} = 0$. This shows that all points on the geodesic $\delta_{\mathbf{P} \to \mathbf{Q}}(t)$ belong to SPD hypergyroplane $\mathcal{H}_{\mathbf{W},\mathbf{P}}$. $\square$

Let $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ be a SPD hypergyroplane in a gyrovector space $(\mathrm{Sym}_n^+, \oplus_{le}, \otimes_{le})$, $\mathbf{X} \notin \mathcal{H}_{\mathbf{W},\mathbf{P}}$, and $\mathbf{Q}^* \in \mathcal{H}_{\mathbf{W},\mathbf{P}}$ such that

$$d(\mathbf{X}, \mathbf{Q}^*) = \min_{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}}} d(\mathbf{X}, \mathbf{Q}) = d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}). \tag{43}$$

We prove the first part of the theorem, i.e.,

$$\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}).$$

We consider two cases:

Case 1: $\mathbf{Q}^* \neq \mathbf{P}$.

If $\angle \mathbf{X}\mathbf{Q}^*\mathbf{P} \neq \frac{\pi}{2}$, then by Lemma P.3, there exists a unique $\mathbf{Q}^{**} \in \delta_{\mathbf{P} \rightarrow \mathbf{Q}^*}(t), t \in \mathbf{R}, \mathbf{Q}^{**} \neq \mathbf{Q}^*$ such that $\angle \mathbf{X}\mathbf{Q}^{**}\mathbf{Q}' = \frac{\pi}{2}$ where $\mathbf{Q}' = \delta_{\mathbf{P} \rightarrow \mathbf{Q}^*}(\infty)$. By Lemma P.2, $\angle \mathbf{X}\mathbf{Q}^{**}\mathbf{Q}^* = \frac{\pi}{2}$. By the Law of SPD gyrosines,

$$d(\mathbf{X}, \mathbf{Q}^{**}) = \sin(\angle \mathbf{X}\mathbf{Q}^*\mathbf{Q}^{**})d(\mathbf{X}, \mathbf{Q}^*),$$

which means that $d(\mathbf{X}, \mathbf{Q}^{**}) < d(\mathbf{X}, \mathbf{Q}^*)$. By Lemma P.4, $\mathbf{Q}^{**} \in \mathcal{H}_{\mathbf{W},\mathbf{P}}$. This leads to a contradiction because of the definition of $\mathbf{Q}^*$. Therefore, we must have $\angle \mathbf{X}\mathbf{Q}^*\mathbf{P} = \frac{\pi}{2}$. Now, by the Law of SPD gyrosines,

$$d(\mathbf{X}, \mathbf{Q}^*) = \sin(\angle \mathbf{X}\mathbf{P}\mathbf{Q}^*)d(\mathbf{X}, \mathbf{P}).$$

We thus deduce that

$$\sin(\angle \mathbf{X}\mathbf{P}\mathbf{Q}^*) = \min_{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \backslash \{\mathbf{P}\}} \sin(\angle \mathbf{X}\mathbf{P}\mathbf{Q}),$$

or equivalently,

$$\cos(\angle \mathbf{X}\mathbf{P}\mathbf{Q}^*) = \max_{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \backslash \{\mathbf{P}\}} \cos(\angle \mathbf{X}\mathbf{P}\mathbf{Q}),$$

Therefore

$$d(\mathbf{X}, \mathbf{Q}^*) = \bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}). \tag{44}$$

Combining Eqs. (43) and (44) leads to

$$\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}).$$

Case 2: $\mathbf{Q}^* = \mathbf{P}$.

For any $\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \backslash \{\mathbf{P}\}$, by the same arguments as above, we must have $\angle \mathbf{X}\mathbf{P}\mathbf{Q} = \frac{\pi}{2}$ and therefore

$$\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \sin(\angle \mathbf{X}\mathbf{P}\mathbf{Q})d(\mathbf{X}, \mathbf{P}) = d(\mathbf{X}, \mathbf{P}) = d(\mathbf{X}, \mathbf{Q}^*) = d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}),$$

which concludes the first part of the theorem.

We now prove the second part of the theorem, i.e.,

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \frac{|\langle \log(\mathbf{X}) - \log(\mathbf{P}), D\log_{\mathbf{P}}(\mathbf{W})\rangle_F|}{\|D\log_{\mathbf{P}}(\mathbf{W})\|_F}.$$

Again, we consider two cases:

Case 1: $\mathbf{Q}^* \neq \mathbf{P}$.

For $\mathbf{Q} \in \mathrm{Sym}_n^+$, note that

$$\mathbf{Q} = \exp_{\mathbf{P}}(\mathrm{Log}_{\mathbf{P}}^{le}(\mathbf{Q})) \overset{(1)}{=} \exp(\log(\mathbf{P}) + D\log_{\mathbf{P}}(\mathrm{Log}_{\mathbf{P}}^{le}(\mathbf{Q}))),$$

where (1) follows from the expression of the exponential map associated with Log-Euclidean metrics.

Hence

$$D \log_{\mathbf{P}}(\text{Log}_{\mathbf{P}}^{le}(\mathbf{Q})) = \log(\mathbf{Q}) - \log(\mathbf{P}). \tag{45}$$

We then have

$$\langle \text{Log}_{\mathbf{P}}^{le}(\mathbf{Q}), \mathbf{W} \rangle_{\mathbf{P}} \overset{(1)}{=} \langle D \log_{\mathbf{P}}(\text{Log}_{\mathbf{P}}^{le}(\mathbf{Q})), D \log_{\mathbf{P}}(\mathbf{W}) \rangle_F$$

$$\overset{(2)}{=} \langle \log(\mathbf{Q}) - \log(\mathbf{P}), D \log_{\mathbf{P}}(\mathbf{W}) \rangle_F,$$

where (1) follows from the fact that LE metrics are bi-invariant, and (2) follows from Eq. (45). Thus, for $\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}}$, we have

$$\langle \log(\mathbf{Q}) - \log(\mathbf{P}), D \log_{\mathbf{P}}(\mathbf{W}) \rangle_F = 0. \tag{46}$$

We need to find $\mathbf{Q}^* \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \setminus \{\mathbf{P}\}$ such that

$$\mathbf{Q}^* = \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \setminus \{\mathbf{P}\}}{\arg \max} \frac{\langle \ominus_{le}\mathbf{P} \oplus_{le} \mathbf{Q}, \ominus_{le}\mathbf{P} \oplus_{le} \mathbf{X} \rangle}{\| \ominus_{le} \mathbf{P} \oplus_{le} \mathbf{Q}\| . \| \ominus_{le} \mathbf{P} \oplus_{le} \mathbf{X} \|}$$

$$\overset{(1)}{=} \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \setminus \{\mathbf{P}\}}{\arg \max} \frac{\langle \text{Log}_{\mathbf{I}_n}^{le}(\ominus_{le}\mathbf{P} \oplus_{le} \mathbf{Q}), \text{Log}_{\mathbf{I}_n}^{le}(\ominus_{le}\mathbf{P} \oplus_{le} \mathbf{X}) \rangle_F}{\| \text{Log}_{\mathbf{I}_n}^{le}(\ominus_{le}\mathbf{P} \oplus_{le} \mathbf{Q})\|_F . \| \text{Log}_{\mathbf{I}_n}^{le}(\ominus_{le}\mathbf{P} \oplus_{le} \mathbf{X})\|_F}$$

$$\overset{(2)}{=} \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \setminus \{\mathbf{P}\}}{\arg \max} \frac{\langle \log(\mathbf{Q}) - \log(\mathbf{P}), \log(\mathbf{X}) - \log(\mathbf{P}) \rangle_F}{\| \log(\mathbf{Q}) - \log(\mathbf{P})\|_F . \| \log(\mathbf{X}) - \log(\mathbf{P})\|_F},$$

where (1) follows from the definition of the SPD inner product, and (2) follows from the expressions of the binary operation $\oplus_{le}$ and inverse operation $\ominus_{le}$.

Our problem returns to the one of finding the minimum angle between the vector $\log(\mathbf{X}) - \log(\mathbf{P})$ and the Euclidean hyperplane described by Eq. (46). The SPD gyrodistance $d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}})$ thus can be obtained as

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \frac{|\langle \log(\mathbf{X}) - \log(\mathbf{P}), \frac{D \log_{\mathbf{P}}(\mathbf{W})}{\|D \log_{\mathbf{P}}(\mathbf{W})\|_F} \rangle_F|}{\| \log(\mathbf{X}) - \log(\mathbf{P})\|_F} . \| \log(\mathbf{X}) - \log(\mathbf{P})\|_F$$

$$= \frac{|\langle \log(\mathbf{X}) - \log(\mathbf{P}), D \log_{\mathbf{P}}(\mathbf{W}) \rangle_F|}{\|D \log_{\mathbf{P}}(\mathbf{W})\|_F}.$$

Case 2: $\mathbf{Q}^* = \mathbf{P}$.

This case is trivial.

$\square$

Q. Proof of Theorem 2.24

Proof. Let $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ be a SPD hypergyroplane in a gyrovector space $(\text{Sym}_n^+, \oplus_{lc}, \otimes_{lc})$, $\mathbf{X} \notin \mathcal{H}_{\mathbf{W},\mathbf{P}}$, and $\mathbf{Q}^* \in \mathcal{H}_{\mathbf{W},\mathbf{P}}$ such that

$$d(\mathbf{X}, \mathbf{Q}^*) = \min_{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}}} d(\mathbf{X}, \mathbf{Q}) = d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}).$$

The first part of the theorem can be proved using the same arguments as those in Appendix P. For the second part, we will only consider the case where $\mathbf{Q}^* \neq \mathbf{P}$, as the case where $\mathbf{Q}^* = \mathbf{P}$ is trivial (see Appendix P). We have

$$\mathbf{Q}^* = \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \setminus \{\mathbf{P}\}}{\arg \max} \frac{\langle \ominus_{lc}\mathbf{P} \oplus_{lc} \mathbf{Q}, \ominus_{lc}\mathbf{P} \oplus_{lc} \mathbf{X} \rangle}{\| \ominus_{lc} \mathbf{P} \oplus_{lc} \mathbf{Q}\| . \| \ominus_{lc} \mathbf{P} \oplus_{lc} \mathbf{X}\|}.$$

Let $\widetilde{\mathbf{Q}} = \ominus_{lc}\mathbf{P} \oplus_{lc} \mathbf{Q}$, $\widetilde{\mathbf{X}} = \ominus_{lc}\mathbf{P} \oplus_{lc} \mathbf{X}$. Then

$$\widetilde{\mathbf{Q}} = \big(-\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{Q}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1}\mathbb{D}(\varphi(\mathbf{Q})) \big) \big(-\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{Q}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1}\mathbb{D}(\varphi(\mathbf{Q})) \big)^T,$$

$$\widetilde{\mathbf{X}} = \big(-\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{X}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{X})) \big) \big(-\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{X}) \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{X})) \big)^T.$$

Using the definition of the SPD inner product, we get

$$\langle \widetilde{\mathbf{Q}}, \widetilde{\mathbf{X}} \rangle = \langle -\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{Q}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{Q}))), -\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{X}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{X}))) \rangle_F.$$

Therefore

$$\mathbf{Q}^* = \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W}, \mathbf{P}} \backslash \{\mathbf{P}\}}{\arg \max} \frac{\langle \mathbf{Z}_1, \mathbf{Z}_2 \rangle_F}{\|\mathbf{Z}_1\|_F . \|\mathbf{Z}_2\|_F},$$

where $\mathbf{Z}_1 = -\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{Q}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{Q})))$ and $\mathbf{Z}_2 = -\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{X}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{X})))$.

By the definition of Log-Cholesky metrics,

$$\langle \text{Log}_{\mathbf{P}}^{lc}(\mathbf{Q}), \mathbf{W} \rangle_{\mathbf{P}} = \langle \varphi(\mathbf{P}) \big(\varphi(\mathbf{P})^{-1} \text{Log}_{\mathbf{P}}^{lc}(\mathbf{Q}) (\varphi(\mathbf{P})^{-1})^T \big)_{\frac{1}{2}}, \varphi(\mathbf{P}) \big(\varphi(\mathbf{P})^{-1} \mathbf{W} (\varphi(\mathbf{P})^{-1})^T \big)_{\frac{1}{2}} \rangle_{\varphi(\mathbf{P})}.$$

Note that

$$\varphi(\mathbf{P})^{-1} \text{Log}_{\mathbf{P}}^{lc}(\mathbf{Q}) (\varphi(\mathbf{P})^{-1})^T = \varphi(\mathbf{P})^{-1} \Big(\varphi(\mathbf{P}) \big(\widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})) \big)^T + \widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})) \varphi(\mathbf{P})^T \Big) (\varphi(\mathbf{P})^{-1})^T$$

$$= \big(\widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})) \big)^T (\varphi(\mathbf{P})^{-1})^T + \varphi(\mathbf{P})^{-1} \widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})),$$

where $\widetilde{\text{Log}}_{\mathbf{L}}(\mathbf{K}) = \lfloor \mathbf{K} \rfloor - \lfloor \mathbf{L} \rfloor + \mathbb{D}(\mathbf{L}) \log(\mathbb{D}(\mathbf{L})^{-1} \mathbb{D}(\mathbf{K}))$ denotes the exponential map on the space of lower triangular matrices with positive diagonal entries (Lin, 2019).

Hence

$$\langle \text{Log}_{\mathbf{P}}^{lc}(\mathbf{Q}), \mathbf{W} \rangle_{\mathbf{P}} = \langle \varphi(\mathbf{P}) \Big(\big(\widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})) \big)^T (\varphi(\mathbf{P})^{-1})^T + \varphi(\mathbf{P})^{-1} \widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})) \Big)_{\frac{1}{2}},$$

$$\varphi(\mathbf{P}) \big(\varphi(\mathbf{P})^{-1} \mathbf{W} (\varphi(\mathbf{P})^{-1})^T \big)_{\frac{1}{2}} \rangle_{\varphi(\mathbf{P})}$$

$$= \langle \varphi(\mathbf{P}) \varphi(\mathbf{P})^{-1} \widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})), \varphi(\mathbf{P}) \big(\varphi(\mathbf{P})^{-1} \mathbf{W} (\varphi(\mathbf{P})^{-1})^T \big)_{\frac{1}{2}} \rangle_{\varphi(\mathbf{P})}$$

$$= \langle \widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})), \varphi(\mathbf{P}) \big(\varphi(\mathbf{P})^{-1} \mathbf{W} (\varphi(\mathbf{P})^{-1})^T \big)_{\frac{1}{2}} \rangle_{\varphi(\mathbf{P})}.$$

Let $\widetilde{\mathbf{W}} = \varphi(\mathbf{P}) \big(\varphi(\mathbf{P})^{-1} \mathbf{W} (\varphi(\mathbf{P})^{-1})^T \big)_{\frac{1}{2}}$. Then

$$\langle \text{Log}_{\mathbf{P}}^{lc}(\mathbf{Q}), \mathbf{W} \rangle_{\mathbf{P}} = \langle \widetilde{\text{Log}}_{\varphi(\mathbf{P})}(\varphi(\mathbf{Q})), \widetilde{\mathbf{W}} \rangle_{\varphi(\mathbf{P})}$$

$$= \langle \lfloor \varphi(\mathbf{Q}) \rfloor - \lfloor \varphi(\mathbf{P}) \rfloor + \mathbb{D}(\varphi(\mathbf{P})) \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{Q}))), \widetilde{\mathbf{W}} \rangle_{\varphi(\mathbf{P})}$$

$$= \langle \lfloor \varphi(\mathbf{Q}) \rfloor - \lfloor \varphi(\mathbf{P}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{Q}))), \lfloor \widetilde{\mathbf{W}} \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\widetilde{\mathbf{W}}) \rangle_F.$$

Thus, for $\mathbf{Q} \in \mathcal{H}_{\mathbf{W}, \mathbf{P}}$, we have

$$\langle \lfloor \varphi(\mathbf{Q}) \rfloor - \lfloor \varphi(\mathbf{P}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{Q}))), \lfloor \widetilde{\mathbf{W}} \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\widetilde{\mathbf{W}}) \rangle_F = 0.$$

The SPD gyrodistance $d(\mathbf{X}, \mathcal{H}_{\mathbf{W}, \mathbf{P}})$ is therefore given by

$$d(\mathbf{X}, \mathcal{H}_{\mathbf{W}, \mathbf{P}}) = \frac{|\langle -\lfloor \varphi(\mathbf{P}) \rfloor + \lfloor \varphi(\mathbf{X}) \rfloor + \log(\mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\varphi(\mathbf{X}))), \lfloor \widetilde{\mathbf{W}} \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\widetilde{\mathbf{W}}) \rangle_F|}{\| \lfloor \widetilde{\mathbf{W}} \rfloor + \mathbb{D}(\varphi(\mathbf{P}))^{-1} \mathbb{D}(\widetilde{\mathbf{W}}) \|_F},$$

where $\widetilde{\mathbf{W}} = \varphi(\mathbf{P}) \big(\varphi(\mathbf{P})^{-1} \mathbf{W} (\varphi(\mathbf{P})^{-1})^T \big)_{\frac{1}{2}}$.

$\square$

R. Proof of Theorem 2.25

Proof. Note that

$$
\begin{aligned}
\langle \mathrm{Log}_{\mathbf{P}}^{ai}(\mathbf{Q}), \mathbf{W} \rangle_{\mathbf{P}} &\overset{(1)}{=} \langle \mathbf{P}^{-\frac{1}{2}} \mathrm{Log}_{\mathbf{P}}^{ai}(\mathbf{Q})\mathbf{P}^{-\frac{1}{2}}, \mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}} \rangle_F \\
&\overset{(2)}{=} \langle \mathbf{P}^{-\frac{1}{2}}\mathbf{P}^{\frac{1}{2}} \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{Q}\mathbf{P}^{-\frac{1}{2}})\mathbf{P}^{\frac{1}{2}}\mathbf{P}^{-\frac{1}{2}}, \mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}} \rangle_F \\
&= \langle \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{Q}\mathbf{P}^{-\frac{1}{2}}), \mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}} \rangle_F,
\end{aligned}
\tag{47}
$$

where (1) follows from the definition of Affine-Invariant metrics, and (2) follows from the expression of the exponential map associated with Affine-Invariant metrics.

Thus, for $\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}}$, we have

$$
\langle \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{Q}\mathbf{P}^{-\frac{1}{2}}), \mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}} \rangle_F = 0.
\tag{48}
$$

By the definition of the SPD pseudo-gyrodistance,

$$
\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \sin(\angle \mathbf{X}\mathbf{P}\bar{\mathbf{Q}})d(\mathbf{X}, \mathbf{P}),
$$

where

$$
\begin{aligned}
\bar{\mathbf{Q}} &= \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \backslash \{\mathbf{P}\}}{\arg\max} \frac{\langle \ominus_{ai}\mathbf{P} \oplus_{ai} \mathbf{Q}, \ominus_{ai}\mathbf{P} \oplus_{ai} \mathbf{X} \rangle}{\| \ominus_{ai} \mathbf{P} \oplus_{ai} \mathbf{Q} \|.\| \ominus_{ai} \mathbf{P} \oplus_{ai} \mathbf{X} \|} \\
&\overset{(1)}{=} \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \backslash \{\mathbf{P}\}}{\arg\max} \frac{\langle \mathrm{Log}_{\mathbf{I}_n}^{ai}(\ominus_{ai}\mathbf{P} \oplus_{ai} \mathbf{Q}), \mathrm{Log}_{\mathbf{I}_n}^{ai}(\ominus_{ai}\mathbf{P} \oplus_{ai} \mathbf{X}) \rangle_F}{\| \mathrm{Log}_{\mathbf{I}_n}^{ai}(\ominus_{ai}\mathbf{P} \oplus_{ai} \mathbf{Q}) \|_F.\| \mathrm{Log}_{\mathbf{I}_n}^{ai}(\ominus_{ai}\mathbf{P} \oplus_{ai} \mathbf{X}) \|_F} \\
&= \underset{\mathbf{Q} \in \mathcal{H}_{\mathbf{W},\mathbf{P}} \backslash \{\mathbf{P}\}}{\arg\max} \frac{\langle \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{Q}\mathbf{P}^{-\frac{1}{2}}), \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{X}\mathbf{P}^{-\frac{1}{2}}) \rangle_F}{\| \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{Q}\mathbf{P}^{-\frac{1}{2}}) \|_F.\| \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{X}\mathbf{P}^{-\frac{1}{2}})\mathbf{X} \|_F}
\end{aligned}
$$

Our problem becomes the one of finding the minimum angle between the vector $\log(\mathbf{P}^{-\frac{1}{2}}\mathbf{X}\mathbf{P}^{-\frac{1}{2}})$ and the Euclidean hyperplane described by Eq. (48). Therefore, the SPD pseudo-gyrodistance $\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}})$ is given by

$$
\bar{d}(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) = \frac{|\langle \log(\mathbf{P}^{-\frac{1}{2}}\mathbf{X}\mathbf{P}^{-\frac{1}{2}}), \mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}} \rangle_F|}{\|\mathbf{P}^{-\frac{1}{2}}\mathbf{W}\mathbf{P}^{-\frac{1}{2}}\|_F}.
$$

$\square$

S. Proof of Corollary 2.26

Proof. We first recall a result from Pennec (2006).

Proposition S.1. *Let* $\mathbf{P} \in \mathrm{Sym}_n^+$ *and* $\mathbf{W} \in \mathcal{T}_{\mathbf{P}}\,\mathrm{Sym}_n^+$. *Then*

$$
D\log_{\mathbf{P}}(\mathbf{W}) = \mathbf{O}(D\log_{\mathbf{\Sigma}}(\mathbf{O}^T\mathbf{W}\mathbf{O}))\mathbf{O}^T,
$$

where $\mathbf{P} = \mathbf{O}\mathbf{\Sigma}\mathbf{O}^T$ *is the eigenvalue decomposition of* $\mathbf{P}$.

First, it is easy to see that

$$
\log(\mathbf{X}) = \mathrm{diag}(\log(\mathbf{X}_1), \dots, \log(\mathbf{X}_N)) \,, \log(\mathbf{P}) = \mathrm{diag}(\log(\mathbf{P}_1), \dots, \log(\mathbf{P}_N)).
$$

Let $\mathbf{P}_i = \mathbf{O}_i\mathbf{\Sigma}_i\mathbf{O}_i^T$ be the eigenvalue decomposition of $\mathbf{P}_i, i = 1, \dots, N$. Then $\mathbf{P} = \mathbf{O}\mathbf{\Sigma}\mathbf{O}^T$ where $\mathbf{O} =$

$\mathrm{diag}(\mathbf{O}_1, \ldots, \mathbf{O}_N)$, and $\boldsymbol{\Sigma} = \mathrm{diag}(\boldsymbol{\Sigma}_1, \ldots, \boldsymbol{\Sigma}_N)$. Note that

$$
\begin{aligned}
D \log_{\mathbf{P}}(\mathbf{W}) &\overset{(1)}{=} \mathbf{O}(D \log_{\boldsymbol{\Sigma}}(\mathbf{O}^T \mathbf{W} \mathbf{O})) \mathbf{O}^T \\
&= \mathbf{O}\Big(D \log_{\boldsymbol{\Sigma}} \big(\mathrm{diag}(\mathbf{O}_1^T \mathbf{W}_1 \mathbf{O}_1, \ldots, \mathbf{O}_N^T \mathbf{W}_N \mathbf{O}_N))\big)\Big) \mathbf{O}^T \\
&\overset{(2)}{=} \mathbf{O}\Big(\mathrm{diag} \big(D \log_{\boldsymbol{\Sigma}_1}(\mathbf{O}_1^T \mathbf{W}_1 \mathbf{O}_1), \ldots, D \log_{\boldsymbol{\Sigma}_N}(\mathbf{O}_N^T \mathbf{W}_N \mathbf{O}_N))\big)\Big) \mathbf{O}^T \qquad (49) \\
&= \mathrm{diag} \Big(\mathbf{O}_1 \big(D \log_{\boldsymbol{\Sigma}_1}(\mathbf{O}_1^T \mathbf{W}_1 \mathbf{O}_1)\big)\mathbf{O}_1^T, \ldots, \mathbf{O}_N \big(D \log_{\boldsymbol{\Sigma}_N}(\mathbf{O}_N^T \mathbf{W}_N \mathbf{O}_N)\big)\mathbf{O}_N^T\Big) \\
&\overset{(3)}{=} \mathrm{diag} \big(D \log_{\mathbf{P}_1}(\mathbf{W}_1), \ldots, D \log_{\mathbf{P}_N}(\mathbf{W}_N)\big).
\end{aligned}
$$

The derivation of Eq. (49) follows.

(1) follows from Proposition S.1.

(2) follows from the property of $D \log_{\boldsymbol{\Sigma}}(.)$.

(3) follows from Proposition S.1.

According to Theorem 2.23, the SPD gyrodistance from $\mathbf{X}$ to a hyperplanes $\mathcal{H}_{\mathbf{W},\mathbf{P}}$ in a gyrovector space $(\mathrm{Sym}^+_{n \times N}, \oplus_{le}, \otimes_{le})$ is given by

$$
\begin{aligned}
d(\mathbf{X}, \mathcal{H}_{\mathbf{W},\mathbf{P}}) &= \frac{|\langle \log(\mathbf{X}) - \log(\mathbf{P}), D \log_{\mathbf{P}}(\mathbf{W}) \rangle_F|}{\|D \log_{\mathbf{P}}(\mathbf{W})\|_F} \\
&= \frac{|\sum_{i=1}^N \langle \log(\mathbf{X}_i) - \log(\mathbf{P}_i), D \log_{\mathbf{P}_i}(\mathbf{W}_i) \rangle_F|}{\sqrt{\sum_{i=1}^N \|D \log_{\mathbf{P}_i}(\mathbf{W}_i)\|_F^2}}.
\end{aligned}
$$

Note that each term in the expressions of the SPD gyrodistance and pseudo-gyrodistance given in Theorems 2.24 and 2.25 is formed from the operation $\mathrm{diag}(.)$ defined in Corollary 2.26. Then the results for LC and AI gyrovector spaces can be easily proved.

$\square$