SYMMETRIC SPACES OF GROUPS OVER INVOLUTIVE ALGEBRAS FOR GRAPH EMBEDDINGS

Anonymous authors
Paper under double-blind review

ABSTRACT

Symmetric spaces such as the Siegel upper half-space and the Siegel disk domain have proven promising for graph embedding. Recently, they have also been proposed as representation spaces for discriminative neural networks. In this paper, we are interested in generalizations of Siegel spaces for representation learning. In particular, we focus on symmetric spaces associated with symplectic groups over involutive algebras, which have recently been introduced as generalizations of the real symplectic group. Although these symmetric spaces offer the capacity to capture rich geometric structures, core components required for practical implementation of neural networks on these space are missing, making the task of learning such parametric models difficult. To this end, we further develop the theory of symplectic groups over involutive algebras by deriving closed-form formulas for some missing components, i.e., automorphisms and their inverses. We consider realizations of the studied spaces, and demonstrate their effectiveness for graph tasks.

1 INTRODUCTION

Graph embedding methods have been extensively investigated due to the ubiquity of graphs in representing structured data. These methods aim to learn a mapping that embeds nodes, or entire graphs, as points in an embedding space so that geometric relationships in the embedding space reflect the structure of the original graph Hamilton et al. (2017). Traditional models such as graph convolutional neural networks (GCNs) rely on Euclidean spaces as embedding spaces. However, many real-world graphs exhibit strongly non-Euclidean latent structures Chami et al. (2019). Therefore, applying GCNs to such graphs often results in high distortion. This motivates the introduction of Riemannian manifolds as representation spaces for GCNs, leading to the extensive development of graph embedding methods based on hyperbolic spaces. Such spaces of constant non-zero curvature offer a clear advantage over Euclidean spaces, since the former have the capability to represent data with hierarchical structure in a very compact way: objects requiring an exponential number of dimensions in Euclidean spaces can be represented in a polynomial number of dimensions in hyperbolic spaces Sarkar (2011); Weber et al. (2020). Consequently, hyperbolic embeddings achieve lower distortion compared to their Euclidean counterparts Gu et al. (2019); Chami et al. (2019). Motivated by the success of hyperbolic spaces for graph embeddings, methods based on matrix manifolds such as SPD and Grassmann manifolds have also been proposed. These Riemannian manifolds offer a convenient trade-off between structural richness and computational tractability (Cruceru et al., 2021; López et al., 2021a).

In order to provide a unified framework that is applicable to a broad class of Riemannian manifolds previously considered for graph embedding, López et al. (2021b) proposed the use of symmetric spaces for such tasks. Their work is based on nice theoretical properties of Riemannian symmetric spaces. The authors also demonstrated the applicability of their framework through the use of Siegel spaces for a variety of graph tasks.

In this paper, we further extend the work in López et al. (2021b) to symmetric spaces of groups over involutive algebras Alessandrini et al. (2021); Huang et al. (2025). In particular, we investigate symmetric spaces associated with symplectic groups that generalize the complex symplectic group Helgason (1979). We focus on two models of these symmetric spaces, i.e., the upper half-space and precompact models which are analogues of the Siegel upper half-space and the Siegel disk

Reviewers: please read the Reviewer Guidelines (iclr.cc/Conferences/2027/ReviewerGuidelines) and the AI Policy for Reviewers (iclr.cc/Conferences/2027/AIPolicyForReviewers). If you used AI to expand, edit, or polish your review, please provide the input text to the LLM. Better still, consider skipping the LLM and submitting your original text: we, and the authors, are much more interested in your unedited thoughts than in what an LLM has to say. AI-assisted or not, you are putting your name and reputation behind your review: LLM-generated falsehoods, hallucinations or misrepresentations are subject to disciplinary action, which may include desk-rejecting all papers you have authored.

domains. Since the Siegel upper half-space is the real locus of the upper half-space model associated with the complex symplectic group Alessandrini et al. (2021), our studied spaces provide additional degrees of freedom compared to Siegel spaces and may therefore benefit graph embeddings.

In summary, our contributions are the following:

- We provide mappings between upper half-space and bounded models associated with symplectic groups which generalize the complex symplectic group Helgason (1979). We identify a group action on bounded domains associated with these groups, and derive closed-form expressions for automorphisms and their inverses on these domains.
- We consider realizations of the studied symmetric spaces, and provide explicit expressions of automorphisms and their inverses on these realizations. We then propose a distance function, and a graph convolutional network on the considered bounded domains.
- To showcase our approach, we evaluate the representation power of the considered symmetric spaces on the tasks of graph reconstruction, node classification, and graph classification.

2 Mathematical Background

This section recaps important concepts of the framework proposed in Alessandrini et al. (2021) which will be useful for our exposition. We refer the reader to Alessandrini et al. (2021); Huang et al. (2025) for basic facts and more detailed discussions.

2.1 Involutive Algebras

Let $\mathbb{K}$ be a field and let A be a unital associative possibly noncommutative finite-dimensional $\mathbb{K}$-algebra. If $\mathbb{K}$ is a topological field, A has a well defined topology.

Definition 2.1. *An anti-involution on A is a $\mathbb{K}$-linear map $\sigma : A \to A$ such that $\sigma(ab) = \sigma(b)\sigma(a)$ and $\sigma^2 = \mathrm{Id}$. An involutive $\mathbb{K}$-algebra is a pair (A, σ), where A is a $\mathbb{K}$-algebra and σ is an anti-involution on A. The algebra $\mathrm{Mat}_n(A)$ of $n \times n$ matrices over A can be endowed with the anti-involution σ^T, which applies σ to each entry and then the transpose. Define*

$$A^\sigma = \{a \in A : \sigma(a) = a\}.$$

Let $A^\times$ be the group of all invertible elements of A. The set of σ-positive elements is defined as

$$A^\sigma_+ = \left\{ \sum_{l=1}^{h} a_l^2 : a_l \in (A^\sigma)^\times, h \in \mathbb{N} \right\}.$$

The set of σ-non-negative elements is defined as

$$A^\sigma_{\geq 0} = \left\{ \sum_{l=1}^{h} a_l^2 : a_l \in A^\sigma, h \in \mathbb{N} \right\}.$$

Of particular interest is the family of involutive algebras called Hermitian algebras.

Definition 2.2. *Let $\mathbb{K}$ be a real closed field. A unital associative $\mathbb{K}$-algebra with an anti-involution (A, σ) is called Hermitian if for all $x, y \in A^\sigma$, $x^2 + y^2 = 0$ implies $x = y = 0$.*

2.2 Symplectic Groups over Involutive Algebras

Before we define symplectic groups over unital rings, let us define automorphism groups associated to σ-sesquilinear forms.

Definition 2.3. *Let A be a unital associative finite dimensional algebra over a field $\mathbb{K}$ with an anti-involution σ. We view A^2 as the set of columns and endow it with the structure of a right A-module. A σ-sesquilinear form ω on A^2 is a map*

$$\omega : A^2 \times A^2 \to A$$

such that

$$\omega(x+y,z) = \omega(x,z) + \omega(y,z)$$
$$\omega(x,y+z) = \omega(x,y) + \omega(x,z)$$
$$\omega(x_1 r_1, x_2 r_2) = \sigma(r_1)\omega(x_1,x_2)r_2.$$

Define

$$\mathrm{Aut}(\omega) = \{f \in \mathrm{Aut}(A^2) : \forall x,y \in A^2 : \omega(f(x),f(y)) = \omega(x,y)\}.$$

We are now ready to define symplectic groups over unital rings, a key concept used in our work.

Definition 2.4. *Let (A,σ) be a unital ring with anti-involution. Let $\Omega = \begin{pmatrix} 0 & 1 \\ -1 & 0 \end{pmatrix}$, and let $\omega(x,y) = \sigma(x)^T \Omega y$ be a σ-sesquilinear form on A^2. The form ω is called the standard symplectic form on A^2. The symplectic group $\mathrm{Sp}_2(A,\sigma)$ over (A,σ) is defined to be the automorphism group $\mathrm{Aut}(\omega)$. By Definition 2.3,*

$$\mathrm{Sp}_2(A,\sigma) = \left\{ \begin{pmatrix} a & b \\ c & d \end{pmatrix} : \sigma(a)c, \sigma(b)d \in A^\sigma, \sigma(a)d - \sigma(c)b = 1 \right\}.$$

The constructions of symplectic groups over real algebras and their associated symmetric spaces can also be extended to more general cases. To this end, we need to define complex and quaternionic extensions of real algebras.

Definition 2.5. *Let $\mathbb{K}$ be a real closed field. We denote the algebraic closure of $\mathbb{K}$ by $\mathbb{K}_{\mathbb{C}} = \mathbb{K}[\mathbf{i}]$, where $\mathbf{i}$ is a square root of -1 called the imaginary unit of $\mathbb{K}_{\mathbb{C}}$. The field $\mathbb{K}_{\mathbb{C}}$ is called the complexification of $\mathbb{K}$. If $\mathbb{K} = \mathbb{R}$, then we denote $\mathbb{R}_{\mathbb{C}}$ as usual by $\mathbb{C}$. Let $\mathbb{K}_{\mathbb{H}}$ be a generalized quaternion skew-field defined by*

$$\mathbb{K}_{\mathbb{H}} = \{a_0 + a_1\mathbf{i} + a_2\mathbf{j} + a_3\mathbf{k} : \mathbf{i}^2 = \mathbf{j}^2 = -1, \mathbf{ij} = -\mathbf{ji} = \mathbf{k}\}.$$

If $\mathbb{K} = \mathbb{R}$, then we denote $\mathbb{R}_{\mathbb{H}}$ as usual by $\mathbb{H}$.

Definition 2.6. *If A is an algebra over a real closed field $\mathbb{K}$, then we call $A_{\mathbb{C}} = A \otimes_{\mathbb{K}} \mathbb{K}_{\mathbb{C}}$ and $A_{\mathbb{H}} = A \otimes_{\mathbb{K}} \mathbb{K}_{\mathbb{H}}$ the complexification and quaternionification of A, respectively.*

2.3 Symmetric Spaces of $\mathrm{Sp}_2(A,\sigma)$ over Complexified Algebras

Let (A,σ) be a real Hermitian algebra, and let $A_{\mathbb{C}} = A \otimes_{\mathbb{R}} \mathbb{C}$ be the complexification of A. Let $\sigma_{\mathbb{C}} : A_{\mathbb{C}} \to A_{\mathbb{C}}$ and $\bar{\sigma}_{\mathbb{C}} : A_{\mathbb{C}} \to A_{\mathbb{C}}$ be the $\mathbb{C}$-linear and $\mathbb{C}$-antilinear extensions of σ given by

$$\sigma_{\mathbb{C}}(z_1 + z_2\mathbf{i}) = \sigma(z_1) + \sigma(z_2)\mathbf{i},$$
$$\bar{\sigma}_{\mathbb{C}}(z_1 + z_2\mathbf{i}) = \sigma(z_1) - \sigma(z_2)\mathbf{i},$$

where $z_1, z_2 \in A$.

Among spaces studied in this paper, we consider two models of symmetric spaces associated with $\mathrm{Sp}_2(A_{\mathbb{C}},\sigma_{\mathbb{C}})$, i.e., the upper half-space and the precompact models. These are generalizations of the Poincaré upper half-plane model $\mathfrak{U} = \{z \in \mathbb{C} : \mathrm{Im}(z) > 0\}$ and the disk model $\mathfrak{B} = \{z \in \mathbb{C} : \bar{z}z < 1\}$, where $\bar{z}$ is the complex conjugate of z. Let $A_{\mathbb{H}} = A \otimes_{\mathbb{R}} \mathbb{H}$ be the quaternionification of A. Let $\sigma_0 : A_{\mathbb{H}} \to A_{\mathbb{H}}$ and $\sigma_1 : A_{\mathbb{H}} \to A_{\mathbb{H}}$ be the anti-involutions on $A_{\mathbb{H}}$ given by

$$\sigma_0(z_1 + z_2\mathbf{j}) = \sigma_{\mathbb{C}}(z_1) + \bar{\sigma}_{\mathbb{C}}(z_2)\mathbf{j},$$
$$\sigma_1(z_1 + z_2\mathbf{j}) = \bar{\sigma}_{\mathbb{C}}(z_1) - \sigma_{\mathbb{C}}(z_2)\mathbf{j},$$

where $z_1, z_2 \in A_{\mathbb{C}}$.

If $z = z_1 + z_2\mathbf{j}$, then the quaternionic conjugation of z on $A_{\mathbb{H}}$ is given by $\bar{z} = \bar{z}_1 - z_2\mathbf{j}$. The upper half-space model is defined as

$$\mathfrak{U} = \{z = z_1 + z_2\mathbf{j} \in A_{\mathbb{H}}^{\sigma_0} : z_2 \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+\}.$$

The precompact model is defined as

$$\mathfrak{B} = \{z \in A_{\mathbb{H}}^{\sigma_0} : 1 - \sigma_1(z)z \in (A_{\mathbb{H}}^{\sigma_1})_+\}.$$

3 PROPOSED APPROACH

In this section, we derive general formulas for some mappings in symmetric spaces associated with $\mathrm{Sp}_2(A, \sigma)$ and $\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})$. For the sake of clarity, we use bold uppercase letters to denote elements of the studied spaces, as these spaces consist of matrices.

3.1 SYMMETRIC SPACES ASSOCIATED WITH $\mathrm{Sp}_2(A, \sigma)$

We first study the connection between the upper half-space and precompact models associated with $\mathrm{Sp}_2(A, \sigma)$. In the case of the hyperbolic space, there exist bijective mappings between the upper half-plane and Poincaré ball models Helgason (1979). In our case, such mappings also exist and can be given below.

Proposition 3.1. *The following mappings convert points between $\mathfrak{U}$ and $\mathfrak{B}$*

$$\varphi(\mathbf{Z}) = (\mathbf{Z} - \mathbf{i}\mathbf{I})(\mathbf{I} - \mathbf{i}\mathbf{Z})^{-1}.$$

$$\varphi^{(-1)}(\mathbf{Z}_0) = (\mathbf{Z}_0 + \mathbf{i}\mathbf{I})(\mathbf{i}\mathbf{Z}_0 + \mathbf{I})^{-1},$$

where $\mathbf{I}$ is an identity matrix of the same size as $\mathbf{Z}$ and $\mathbf{Z}_0$.

The study of automorphisms has been one of the main focuses in complex analysis Kobayashi (1967). Recently, they have proven effective in the construction of neural networks on Siegel spaces Nguyen & Grozavu (2026). The following proposition provides the closed-form expression for an automorphism of the precompact model.

Proposition 3.2. *Let $\mathbf{X}, \mathbf{Y} \in \mathfrak{B}$. An automorphism $\phi_{\mathbf{X}}(\mathbf{Y})$ can be given by*

$$\phi_{\mathbf{X}}(\mathbf{Y}) = \left(\sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} - \mathbf{X}(\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$$
$$\left(-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} + (\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)^{-1}.$$

We have $\mathbf{X} \in \mathfrak{B}$, so that $\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X} \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+$ and thus, the mapping $\phi_{\mathbf{X}}(\mathbf{Y})$ is well-defined. The inverse map $\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y})$ of $\phi_{\mathbf{X}}(\mathbf{Y})$ can also be given in closed form.

Proposition 3.3. *Let $\mathbf{X}, \mathbf{Y} \in \mathfrak{B}$. Then the map $\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y})$ is given by*

$$\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y}) = \left(\sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} + \sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{X}\right)$$
$$\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{Y} + (\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)^{-1}.$$

3.2 SYMMETRIC SPACES ASSOCIATED WITH $\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})$

We now focus on symmetric spaces associated with $\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})$. Following the approach in the preceding section, we obtain the following results.

Proposition 3.4. *The following mappings convert points between $\mathfrak{U}$ and $\mathfrak{B}$:*

$$\psi(\mathbf{Z}) = (\mathbf{Z} - \mathbf{j}\mathbf{I})(\mathbf{I} - \mathbf{j}\mathbf{Z})^{-1}. \tag{1}$$

$$\psi^{(-1)}(\mathbf{Z}_0) = (\mathbf{Z}_0 + \mathbf{j}\mathbf{I})(\mathbf{j}\mathbf{Z}_0 + \mathbf{I})^{-1}, \tag{2}$$

It has been shown Alessandrini et al. (2021) that the group $\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})$ acts transitively on $\mathfrak{U}$ via Möbius transformations

$$\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}}) \times \mathfrak{U} \to \mathfrak{U},$$
$$\mathbf{Z} \mapsto \mathbf{M}[\mathbf{Z}] = (\mathbf{A}\mathbf{Z} + \mathbf{B})(\mathbf{C}\mathbf{Z} + \mathbf{D})^{-1}, \mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}.$$

A transitive group action on $\mathfrak{B}$ can be given below.

Proposition 3.5. *Let* $\mathbf{L} = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{I} & \mathbf{jI} \\ \mathbf{jI} & \mathbf{I} \end{pmatrix}$. *The group* $\mathbf{L}^{-1}\,\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})\mathbf{L}$ *acts transitively on* $\mathfrak{B}$ *via Möbius transformations.*

Proposition 3.6. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two points of* $\mathfrak{B}$. *Then, an automorphism* $\phi_{\mathbf{X}}(\mathbf{Y})$ *of* $\mathfrak{B}$ *is given by*

$$\phi_{\mathbf{X}}(\mathbf{Y}) = \left(\sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} - \mathbf{X}(\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right) \left(-\sigma_1(\mathbf{X})\sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} + (\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)^{-1}.$$

We have $\mathbf{X} \in \mathfrak{B}$, so that $\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X} \in (A_{\mathbb{H}}^{\sigma_1})_+$ and thus, the mapping $\phi_{\mathbf{X}}(\mathbf{Y})$ is well-defined. The inverse map $\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y})$ of $\phi_{\mathbf{X}}(\mathbf{Y})$ can also be given in closed form.

Proposition 3.7. *Let* $\mathbf{X}$ *and* $\mathbf{Y}$ *be two points of* $\mathfrak{B}$. *Then, the function* $\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y})$ *is given by*

$$\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y}) = \sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)(\mathbf{Y} + \mathbf{X})(\mathbf{I} + \sigma_1(\mathbf{X})\mathbf{Y})^{-1}(\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{\frac{1}{2}}.$$

Proposition 3.8. *Let* $\mathbf{Z} \in \mathbb{QB}_n$, *and let* $\mathbf{Z}_1$ *and* $\mathbf{Z}_2$ *be complex* $n \times n$ *matrices such that* $\psi^{(-1)}(\mathbf{Z}) = \mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j} \in \mathbb{QU}_n$. *Let* ρ *be the mapping defined by*

$$\rho(\mathbf{Z}) = \begin{pmatrix} \mathbf{Z}_2 + \mathbf{Z}_1\overline{\mathbf{Z}}_2^{-1}\bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_1) & \mathbf{Z}_1\overline{\mathbf{Z}}_2^{-1} \\ \overline{\mathbf{Z}}_2^{-1}\bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_1) & \overline{\mathbf{Z}}_2^{-1} \end{pmatrix}.$$

Then $\mathbf{Z} \in \mathfrak{U}$ *if and only if* $\rho(\mathbf{Z}) \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+$.

4 Realizations of the Studied Symmetric Spaces

In this section, we consider some realizations of the symmetric spaces introduced in Section 3 and provide explicit expressions for the corresponding automorphisms and their inverses.

4.1 Symmetric Spaces Associated with the Real Symplectic Group

Let $A = \mathrm{Mat}_n(\mathbb{R})$, and let σ be the anti-involution given by $\sigma(\mathbf{Z}) = \mathbf{P}^{-1}\mathbf{Z}^T\mathbf{P}$ for any invertible symmetric matrix $\mathbf{P}$. The upper half-space model $\mathbb{RU}_n$ is defined by

$$\mathbb{RU}_n = \{\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{i} : \mathbf{Z}_1 = \mathbf{P}^{-1}\mathbf{Z}_1^T\mathbf{P} \in A, \mathbf{Z}_2 = \mathbf{P}^{-1}\mathbf{Z}_2^T\mathbf{P} \in A_+^{\sigma}\}.$$

The precompact model $\mathbb{RB}_n$ is defined by

$$\mathbb{RB}_n = \{\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{i} : \mathbf{Z}_1 = \mathbf{P}^{-1}\mathbf{Z}_1^T\mathbf{P} \in A, \mathbf{Z}_2 = \mathbf{P}^{-1}\mathbf{Z}_2^T\mathbf{P} \in A, \mathbf{I} - \mathbf{P}^{-1}\mathbf{Z}^H\mathbf{Z}^T\mathbf{P} \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+\},$$

where $\mathbf{Z}^H$ is the conjugate transpose of $\mathbf{Z}$.

Note that when $\mathbf{P} = \mathbf{I}$, the group $\mathrm{Sp}_2(A, \sigma)$ is isomorphic to the real symplectic group $\mathrm{Sp}_{2m}(\mathbb{R})$ Helgason (1979), and the spaces $\mathbb{RU}_n$ and $\mathbb{RB}_n$ correspond to the Siegel upper half-space and the Siegel disk domain, respectively. One then recovers the closed-form expressions for the automorphism and its inverse in the Siegel disk domain Jeuris & Vandebril (2016); Nguyen & Grozavu (2026) from Propositions 3.2 and 3.3.

4.2 Symmetric Spaces Associated with the Complex Symplectic Group

Let $A_{\mathbb{C}} = \mathrm{Mat}_n(\mathbb{R}) \otimes_{\mathbb{R}} \mathbb{C}$, and let σ be the anti-involution given by the matrix transpose. Then, the group $\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})$ is isomorphic to the complex symplectic group $\mathrm{Sp}_{2m}(\mathbb{C})$ Helgason (1979). The upper half-space model $\mathbb{QU}_n$ is defined by

$$\mathbb{QU}_n = \{\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j} \in \mathrm{Mat}_n(\mathbb{H}) : \mathbf{Z}_1 \in \mathrm{Sym}_n(\mathbb{C}), \mathbf{Z}_2 \in \mathrm{Herm}_n^+\},$$

where $\mathrm{Mat}_n(\mathbb{H})$ denotes the space of $n \times n$ quaternion matrices, $\mathrm{Sym}_n(\mathbb{C})$ denotes the space of $n \times n$ complex symmetric matrices, and Herm_n^+ denotes the space of $n \times n$ Hermitian positive

definite (HPD) matrices. The precompact model (hereafter referred to as a bounded domain) $\mathbb{QB}_n$ is defined by

$$\mathbb{QB}_n = \{\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j} \in \mathrm{Mat}_n(\mathbb{H}) : \mathbf{Z}_1 \in \mathrm{Sym}_n(\mathbb{C}), \mathbf{Z}_2 \in \mathrm{Herm}_n, \mathbf{I} - (\overline{\mathbf{Z}}_1 - \overline{\mathbf{Z}}_2\mathbf{j})(\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j}) \in \mathrm{Herm}_n^+(\mathbb{H})\},$$

where Herm_n is the space of $n \times n$ Hermitian matrices, and $\mathrm{Herm}_n^+(\mathbb{H})$ is the space of $n \times n$ Hermitian quaternion positive definite matrices. The explicit expressions for the automorphism and its inverse derived in Propositions 3.6 and 3.7 can be given below.

Corollary 4.1. *An automorphism $\phi_{\mathbf{X}}(\mathbf{Y})$ on $\mathbb{QB}_n$ is given by*

$$\begin{aligned}\phi_{\mathbf{X}}(\mathbf{Y}) = ((\sigma_{\mathbb{C}}(\mathbf{Z}_1) + \bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_2)\mathbf{j})(\mathbf{Y}_1 + \mathbf{Y}_2\mathbf{j}) - (\mathbf{X}_1 + \mathbf{X}_2\mathbf{j})(\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j})) \\ \left(-(\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2\mathbf{j})(\sigma_{\mathbb{C}}(\mathbf{Z}_1) + \bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_2)\mathbf{j})(\mathbf{Y}_1 + \mathbf{Y}_2\mathbf{j}) + (\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j})\right)^{-1},\end{aligned} \tag{3}$$

where $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2\mathbf{j} \in \mathbb{QB}_n$, $\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2\mathbf{j} \in \mathbb{QB}_n$, $\mathbf{Z} = (\mathbf{I} - (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2\mathbf{j})(\mathbf{X}_1 + \mathbf{X}_2\mathbf{j}))^{-\frac{1}{2}}$, $\mathbf{Z} = \mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j}$, $\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Z}_1, \mathbf{Z}_2 \in A_{\mathbb{C}}$.

Corollary 4.2. *The inverse map $\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y})$ of the automorphism $\phi_{\mathbf{X}}(\mathbf{Y})$ is given by*

$$\begin{aligned}\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y}) = (\sigma_{\mathbb{C}}(\mathbf{Z}_1) + \bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_2)\mathbf{j})((\mathbf{Y}_1 + \mathbf{X}_1) + (\mathbf{Y}_2 + \mathbf{X}_2)\mathbf{j}) \\ (\mathbf{I} + (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2\mathbf{j})(\mathbf{Y}_1 + \mathbf{Y}_2\mathbf{j}))^{-1}(\mathbf{I} - (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2\mathbf{j})(\mathbf{X}_1 + \mathbf{X}_2\mathbf{j}))^{\frac{1}{2}},\end{aligned}$$

where $\mathbf{X} = \mathbf{X}_1 + \mathbf{X}_2\mathbf{j} \in \mathbb{QB}_n$, $\mathbf{Y} = \mathbf{Y}_1 + \mathbf{Y}_2\mathbf{j} \in \mathbb{QB}_n$, $\mathbf{Z} = (\mathbf{I} - (\overline{\mathbf{X}}_1 - \overline{\mathbf{X}}_2\mathbf{j})(\mathbf{X}_1 + \mathbf{X}_2\mathbf{j}))^{-\frac{1}{2}}$, $\mathbf{Z} = \mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j}$, $\mathbf{X}_1, \mathbf{X}_2, \mathbf{Y}_1, \mathbf{Y}_2, \mathbf{Z}_1, \mathbf{Z}_2 \in A_{\mathbb{C}}$.

5 Applications

In Section 5.1, we propose a distance function on $\mathbb{QB}_n$ that will be useful for distance-based learning in this domain. We also propose a method to improve GCNs on SPD manifolds in Section 5.2.

5.1 Distance Function

The derivation of the automorphism given in Corollary 4.1 allows us to construct an effective distance function on $\mathbb{QB}_n$. Specifically, the Kobayashi pseudodistance Kobayashi (1967) on $\mathbb{QB}_n$ can be computed as

$$d^K_{\mathbb{QB}_n}(\mathbf{X}, \mathbf{Y}) = \frac{1}{2}\log\left(\frac{1 + \|\phi_{\mathbf{X}}(\mathbf{Y})\|_2}{1 - \|\phi_{\mathbf{X}}(\mathbf{Y})\|_2}\right), \tag{4}$$

where $\mathbf{X}, \mathbf{Y} \in \mathbb{QB}_n$, $\phi_{\mathbf{X}}(\mathbf{Y})$ is the automorphism defined in Eq. (3), and $\log(\cdot)$ denotes the logarithmic function.

5.2 Improving GCNs on SPD Manifolds

In this section, we propose a method to improve GCNs on SPD manifolds (Zhao et al., 2023). Let $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ be a graph with vertex set $\mathcal{V}$ and edge set $\mathcal{E}$, $\mathbf{z}_j^l, j \in \mathcal{V}$ be the embedding of node j at layer l ($l = 0$ indicates input node features), $\mathcal{N}(j) = \{k : (j,k) \in \mathcal{E}\}$ be the set of neighbors of $j \in \mathcal{V}$, $\mathbf{W}^l$ and $\mathbf{b}^l$ be the weight and bias at layer l, v_{jk} be the relative importance of node k to node j, and δ be a non-linear activation function. A basic GCN message-passing update can be written as

$$\begin{aligned}\mathbf{p}_j^l &= \mathbf{W}^l\mathbf{z}_j^{l-1} && \text{(feature transform)} \\ \mathbf{q}_j^l &= \sum_{k \in \mathcal{N}(j)} v_{jk}\mathbf{p}_k^l && \text{(aggregation)} \\ \mathbf{z}_j^l &= \delta(\mathbf{q}_j^l + \mathbf{b}^l) && \text{(bias and nonlinearity)}\end{aligned}$$

For our network, the original features of node j are first mapped to a symmetric matrix $\mathbf{X} \in \mathrm{Mat}_n(\mathbb{R})$ which is then converted to $\mathbb{QU}_n$ as

$$\mathbf{Z} = \mathbf{X} + \mathbf{0}_n\mathbf{i} + \exp(\mathbf{X})\mathbf{j} + \mathbf{0}_n\mathbf{k},$$

Table 1: Results (mean distortion $\pm$ standard deviation) computed over five runs for graph reconstruction. S and B denote the models built on the Siegel upper half-space and Siegel disk domain, respectively. R and F_1 correspond to two types of vector-valued distance (i.e., Riemannian distance and Finsler distance inducing the l^1-metric, see López et al. (2021b) for details). The subscripts indicate the embedding dimensions.

Method	Grid	Tree	Tree $\times$ Grid	Grid $\times$ Tree	Tree $\diamond$ Grid	Grid $\diamond$ Tree
SPD_6	14.33±0.04	5.45±0.06	9.07±0.02	8.97±0.02	2.23±0.04	5.31±0.03
$\mathrm{S}_4^{F_1}$	3.34±0.08	2.02±0.13	2.64±0.07	2.67±0.04	1.08±0.09	2.33±0.06
$\mathrm{B}_4^{F_1}$	2.92±0.12	1.59±0.06	2.61±0.06	2.63±0.03	0.86±0.04	1.47±0.08
QB_4 (Ours)	**1.16±0.14**	**0.57±0.04**	**1.96±0.02**	**1.53±0.05**	**0.62±0.07**	**1.10±0.09**

where $\mathbf{0}_n$ is the $n \times n$ zero matrix, and $\exp(\cdot)$ denotes the matrix exponential. Next, the matrix $\mathbf{Z} \in \mathbb{QU}_n$ is transformed as

$$\tilde{\mathbf{Z}} = \psi^{(-1)}(\phi_{\mathbf{P}}(\psi(\mathbf{Z}))), \tag{5}$$

where $\phi_{\mathbf{P}}(\cdot)$ is the automorphism defined in Eq. (3), and $\mathbf{P} \in \mathbb{QB}_n$ is a parameter of our network. Let $\tilde{\mathbf{Z}} = \mathbf{Z}_1 + \mathbf{Z}_2\mathbf{j} \in \mathbb{QU}_n$, where $\mathbf{Z}_1 \in \mathrm{Sym}_n(\mathbb{C})$ and $\mathbf{Z}_2 \in \mathrm{Herm}_n^+$. Based on Proposition 3.8, we transform $\tilde{\mathbf{Z}}$ as

$$\mathbf{Z}' = \begin{pmatrix} \mathbf{Z}_2 + \mathbf{Z}_1\bar{\mathbf{Z}}_2^{-1}\mathbf{Z}_1^H & \mathbf{Z}_1\bar{\mathbf{Z}}_2^{-1} \\ \bar{\mathbf{Z}}_2^{-1}\mathbf{Z}_1^H & \bar{\mathbf{Z}}_2^{-1} \end{pmatrix}.$$

Then the initial embedding $\mathbf{Y}_j^0$ of node j is obtained by applying a BiMap layer (Huang & Gool, 2017) to the real embedding $\mathbf{Z}''$ of $\mathbf{Z}'$ given by

$$\mathbf{Z}'' = \begin{pmatrix} \mathrm{Re}(\mathbf{Z}') & -\mathrm{Im}(\mathbf{Z}') \\ \mathrm{Im}(\mathbf{Z}') & \mathrm{Re}(\mathbf{Z}') \end{pmatrix},$$

where $\mathrm{Re}(\mathbf{Z}')$ and $\mathrm{Im}(\mathbf{Z}')$ denote the real and imaginary parts of $\mathbf{Z}'$, respectively. Denote by $\mathbf{M}^l$ an isometry map (Zhao et al., 2023) of SPD manifolds at layer l. Let $\mathrm{Exp}_{\mathbf{I}}(\cdot)$ and $\mathrm{Log}_{\mathbf{I}}(\cdot)$ be the SPD exponential and logarithmic maps (Pennec, 2006) at the identity matrix $\mathbf{I}$, and let $v_{jk} = |\mathcal{N}(j)|^{-\frac{1}{2}}|\mathcal{N}(k)|^{-\frac{1}{2}}$. Let $\oplus$ be the binary operation defined by $\mathbf{P} \oplus \mathbf{Q} = \mathbf{P}^{\frac{1}{2}}\mathbf{Q}\mathbf{P}^{\frac{1}{2}}$ for any two points $\mathbf{P}$ and $\mathbf{Q}$ on an SPD manifold, and let $\delta(\cdot)$ be the non-linear activation function (Huang et al., 2018; Zhao et al., 2023) given by

$$\delta(\mathbf{P}) = \mathbf{U}\,\mathrm{diag}(\max(\mathbf{e}, \epsilon))\mathbf{U}^{-1},$$

where $\mathbf{P} = \mathbf{U}\,\mathrm{diag}(\mathbf{e})\mathbf{U}^{-1}$ is the eigenvalue decomposition of $\mathbf{P}$, and ϵ is a small positive constant. Let $\mathbf{B}^l$ be the SPD parameter representing the bias at layer l. Then the message-passing update of our network can be summarized as

$$\begin{aligned} \mathbf{P}_j^l &= \mathbf{M}^l \odot \mathbf{Y}_j^{l-1} && \text{(feature transform)} \\ \mathbf{Q}_j^l &= \mathrm{Exp}_{\mathbf{I}}\Big(\sum_{k \in \mathcal{N}(j)} v_{jk}\,\mathrm{Log}_{\mathbf{I}}(\mathbf{P}_k^l)\Big) && \text{(aggregation)} \\ \mathbf{Y}_j^l &= \delta(\mathbf{B}^l \oplus \mathbf{Q}_j^l) && \text{(bias and nonlinearity)} \end{aligned}$$

6 Experiments

In this section, we discuss the results of our experimental evaluation on the tasks of graph reconstruction, node classification, and graph classification. We refer the reader to Appendix B for an application of the symmetric spaces discussed in Section 4.1.

6.1 Graph Reconstruction

Siegel spaces have proven effective for graph tasks López et al. (2021b). In this section, we first investigate the representation capability of $\mathbb{QB}_n$ for the graph reconstruction task. Following the

Table 2: Node classification results (mean accuracy ± standard deviation) across ten runs.

Model	Airport	Pubmed	Cora	Disease
SPD-GCN (Zhao et al., 2023)	82.6±1.5	78.7±0.5	82.3±0.5	96.8±3.1
QB-GCN (Ours)	**92.9±0.6**	**79.6±0.4**	**82.8±0.5**	**98.1±1.9**

Table 3: Graph classification results (mean accuracy ± standard deviation) across ten runs.

Model	DHFR	COX2	PROTEINS
SPD-GCN (Zhao et al., 2023)	71.1±1.8	83.3±1.9	73.6±1.2
QB-GCN (Ours)	**72.6±1.4**	**85.9±2.1**	**75.4±1.6**

approach in López et al. (2021b), we generate datasets consisting of graphs with pure geometric features, i.e., grids, trees, and their Cartesian products as well as those with mixed geometric features, i.e., rooted products of grids and trees. We embed graph nodes in the transductive setting Chami et al. (2019) using low-dimensional embeddings. We optimize the loss proposed in Gu et al. (2019) which captures the average distortion. We use the distance function defined in Eq. (4) to learn the embeddings on $\mathbb{QB}_n$. The model built on $\mathbb{QB}_n$ is compared against those built on SPD and Siegel spaces which are strong baselines considered in López et al. (2021b). For the models based on Siegel spaces, we report their best mean accuracies obtained with different vector-valued distances López et al. (2021b). Results in Tab. 1 show that the model built on our considered domains consistently outperforms the baselines. This indicates that the use of the considered symmetric spaces can lead to enhanced graph representation for downstream tasks.

6.2 Node Classification

Having demonstrated the efficacy of the studied spaces in the graph reconstruction task, we now turn to a graph downstream task. We evaluate the performance of the method proposed in Section 5.2 on the node classification task using four datasets: Airport Zhang & Chen (2018), Pubmed Namata et al. (2012a), Cora Sen et al. (2008), and Disease Chami et al. (2019). We follow the inductive setting Chami et al. (2019). Original node features are used for fair comparison (Zhao et al., 2023). Our model and SPD-GCN learn the node embeddings on $\mathbb{QB}_8$ and Sym^+_{14}, respectively. To report the performance of each model, we compute the mean and standard deviation of its accuracies over ten runs with different seeds. Results in Tab. 2 show that in terms of mean accuracy, our model improves upon SPD-GCN on Airport and Disease datasets, and is on par with SPD-GCN on Pubmed and Cora datasets. In particular, the performance gap between our model and SDP-GCN is more than 10% on Airport dataset.

6.3 Graph Classification

We also evaluate the performance of the method proposed in Section 5.2 on the graph classification task using three datasets: DHFR, COX2 Sutherland et al. (2003), and PROTEINS Borgwardt et al. (2005). We adopt a random 80/10/10% split for the train, validation, and test sets Ji et al. (2025). We compute the arithmetic mean of the node embeddings in each graph to obtain its center which is then used to infer the graph class Morris et al. (2020). Node embeddings are learned on $\mathbb{QB}_8$ and Sym^+_{14}. The mean and standard deviation of each model are computed as in the node classification experiments. Results in Tab. 3 show that our model outperforms SPD-GCN in terms of mean accuracy across all the datasets.

6.4 Ablation Study

Impact of additional degrees of freedom of $\mathbb{QB}_n$ w.r.t. $\mathbb{SD}_n$ To further evaluate the quality of embeddings on $\mathbb{QB}_n$, we conduct graph reconstruction experiments on a tree-like graph with 40 nodes. We compute the distortions for all pairs of connected nodes from the resulting node embeddings. The same experiments are performed by embedding graph nodes in $\mathbb{SD}_n$. Fig. 1 illustrates the

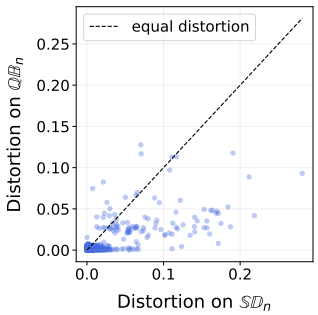


Figure 1: Comparison of distortions obtained by learning node embeddings on $\mathbb{SD}_n$ and $\mathbb{QB}_n$.

Table 4: Impact of automorphisms on node classification performance.

Model	Airport	Pubmed	Cora	Citeseer	Disease
QB-GCN-NoAut	82.6±1.5	78.7±0.5	82.3±0.5	69.9±0.8	95.9±2.1
QB-GCN (Ours)	**93.2±0.7**		**83.8±0.4**		

reconstruction distortions on $\mathbb{QB}_n$ with respect to those on $\mathbb{SD}_n$, where the distortions are averaged over five runs, and the x-coordinate and y-coordinate of each point represent the distortions for a pair of connected nodes on $\mathbb{SD}_n$ and $\mathbb{QB}_n$, respectively. Note that the points are only shown for pairs of connected nodes i and j with $i < j$. As can be observed, most of the points lie below the line indicating equal distortion. Overall, for 66% of pairs of connected nodes, the distortion computed from the node embeddings on $\mathbb{QB}_n$ is smaller than that computed from the node embeddings on $\mathbb{SD}_n$. This clearly demonstrates the advantage of the studied symmetric spaces over Siegel spaces for the graph reconstruction task. The better performance of $\mathbb{QB}_n$ may be attributed to the additional degrees of freedom in $\mathbb{QB}_n$ compared with its real locus $\mathbb{SD}_n$.

Impact of automorphisms We now investigate the impact of the automorphisms defined in Eq. (3) for the node classification task. To this end, we remove the transformation given in Eq. (5) from QB-GCN. The resulting model QB-GCN-NoAut is compared against QB-GCN on the five datasets considered in Section 6.2. Results are reported in Tab. 4.

7 Limitation of Our Method

The proposed GCN model follows the framework of SPD-GCN which is heavily based on eigenvalue decomposition operations. Moreover, these expensive operations are also required to compute the automorphism defined in Eq. (3). Therefore, our GCN model becomes computationally expensive when dealing with high-dimensional data.

8 Conclusion

In this paper, we study two families of symmetric spaces associated with symplectic groups over involutive algebras that generalize the real and complex symplectic groups. We further develop the theory of these symmetric spaces by deriving closed-form expressions for a class of automorphisms on each of these families of symmetric spaces. We then consider realizations of these spaces and their applications for graph embeddings. To this end, we introduce a distance function on the symmetric spaces associated with the complex symplectic group, and propose a GCN model that learns embeddings on these spaces. Our experimental results on graph reconstruction, node classification, and graph classification demonstrate the potential of the proposed approach.

AI USE STATEMENT

We used AI tools for our research on related works and bibliographic entries for inclusion in our .bib file.

REFERENCES

Daniele Alessandrini, Arkady Berenstein, Vladimir Retakh, Eugen Rogozinnikov, and Anna Wienhard. Symplectic Groups over Noncommutative Algebras. *Selecta Mathematica*, 28, 2021.

R. M. Anderson and R. M. May. Infectious Diseases of Humans: Dynamics and Control. *Epidemiology and Infection*, 108(1):211–211, 1992.

Karsten M. Borgwardt, Cheng Soon Ong, Stefan Schönauer, S. V. N. Vishwanathan, Alex J. Smola, and Hans-Peter Kriegel. Protein Function Prediction via Graph Kernels. *Bioinformatics*, 21(1): 47–56, 2005.

Ines Chami, Rex Ying, Christopher R, and Jure Leskovec. Hyperbolic Graph Convolutional Neural Networks. *CoRR*, abs/1910.12933, 2019. URL https://arxiv.org/abs/1910.12933.

Calin Cruceru, Gary Bécigneul, and Octavian-Eugen Ganea. Computationally Tractable Riemannian Manifolds for Graph Embeddings. In *AAAI*, pp. 7133–7141, 2021.

Albert Gu, Frederic Sala, Beliz Gunel, and Christopher Ré. Learning Mixed-Curvature Representations in Products of Model Spaces. In *ICLR*, 2019.

William L Hamilton, Rex Ying, and Jure Leskovec. Representation Learning on Graphs: Methods and Applications. *IEEE Data Engineering Bulletin*, 40(3):52–74, 2017.

S. Helgason. *Differential Geometry, Lie Groups, and Symmetric Spaces*. ISSN. Elsevier Science, 1979.

Pengfei Huang, Georgios Kydonakis, Eugen Rogozinnikov, and Anna Wienhard. Symmetric Spaces for Groups over Involutive Algebras and Applications to Higgs Bundles. *CoRR*, abs/2509.02002, 2025. URL https://arxiv.org/abs/2509.02002.

Zhiwu Huang and Luc Van Gool. A Riemannian Network for SPD Matrix Learning. In *AAAI*, pp. 2036–2042, 2017.

Zhiwu Huang, Jiqing Wu, and Luc Van Gool. Building Deep Networks on Grassmann Manifolds. In *AAAI*, pp. 3279–3286, 2018.

B. Jeuris and R. Vandebril. The Kähler Mean of Block-Toeplitz Matrices with Toeplitz Structured Blocks. *SIAM J. Matrix Anal. Appl.*, 37(3):1151–1175, 2016.

Mattie Ji, Amauri H. Souza, and Vikas K. Garg. On Topological Descriptors for Graph Products. In *NeurIPS*, 2025.

Shôshichi Kobayashi. Invariant Distances on Complex Manifolds and Holomorphic Mappings. *Journal of The Mathematical Society of Japan*, 19:460–480, 1967.

Federico López, Beatrice Pozzetti, Steve Trettel, Michael Strube, and Anna Wienhard. Vector-valued Distance and Gyrocalculus on the Space of Symmetric Positive Definite Matrices. In *NeurIPS*, pp. 18350–18366, 2021a.

Federico López, Beatrice Pozzetti, Steve Trettel, Michael Strube, and Anna Wienhard. Symmetric Spaces for Graph Embeddings: A Finsler-Riemannian Approach. In *ICML*, pp. 7090–7101, 2021b.

Christopher Morris, Nils M. Kriege, Franka Bause, Kristian Kersting, Petra Mutzel, and Marion Neumann. TUDataset: A Collection of Benchmark Datasets for Learning with Graphs. *CoRR*, abs/2007.08663, 2020. URL https://arxiv.org/abs/2007.08663.

Galileo Namata, Ben London, Lise Getoor, and Bert Huang. Query-driven Active Surveying for Collective Classification. In *10th International Workshop on Mining and Learning with Graphs*, volume 8, pp. 1, 2012a.

Galileo Namata, Ben London, Lise Getoor, and Bert Huang. Query-driven Active Surveying for Collective Classification. In *Workshop on Mining and Learning with Graphs*, 2012b.

Xuan Son Nguyen and Nistor Grozavu. Batch normalization for neural networks on complex domains. *CoRR*, abs/2605.00467, 2026. URL https://arxiv.org/abs/2605.00467.

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

Rik Sarkar. Low Distortion Delaunay Embedding of Trees in Hyperbolic Plane. In *International Symposium on Graph Drawing and Network Visualization*, pp. 355–366, 2011.

Prithviraj Sen, Galileo Mark Namata, Mustafa Bilgic, Lise Getoor, Brian Gallagher, and Tina Eliassi-Rad. Collective Classification in Network Data. *AI Magazine*, 29(3):93–106, 2008.

Jeffrey J. Sutherland, Lee A. O'Brien, and Donald F. Weaver. Spline-Fitting with a Genetic Algorithm: A Method for Developing Classification Structure-Activity Relationships. *Journal of Chemical Information and Computer Sciences*, 43(6):1906–1915, 2003.

Melanie Weber, Manzil Zaheer, Ankit Singh Rawat, Aditya Krishna Menon, and Sanjiv Kumar. Robust Large-margin Learning in Hyperbolic Space. In *NeurIPS*, pp. 13459–13471, 2020.

Muhan Zhang and Yixin Chen. Link Prediction Based on Graph Neural Networks. *CoRR*, abs/1802.09691, 2018. URL https://arxiv.org/abs/1802.09691.

Wei Zhao, Federico Lopez, J. Maxwell Riestenberg, Michael Strube, Diaaeldin Taha, and Steve Trettel. Modeling Graphs Beyond Hyperbolic: Graph Neural Networks in Symmetric Positive Definite Matrices. *CoRR*, abs/2306.14064, 2023. URL https://arxiv.org/abs/2306.14064.

A Experiments

A.1 Graph Reconstruction

A.1.1 Datasets and Experimental Settings

We generate datasets of graphs of different geometric structures and sizes. The statistics of the datasets are summarized in Tab. 5.

A.1.2 Optimization and Hyperparameters

We create graph node embeddings on a space $\mathcal{M}$ by optimizing a distance-based loss function (Gu et al., 2019; López et al., 2021b). Given the distances $\{d_G(l_1, l_2)\}_{l_1,l_2=1}^N$ between all pairs of connected nodes l_1 and l_2, the loss function is computed as:

$$\mathcal{L}(\mathbf{X}) = \sum_{l_1,l_2=1}^{N} \left| \left(\frac{d_{\mathcal{M}}(\mathbf{X}_{l_1}, \mathbf{X}_{l_2})}{d_G(l_1, l_2)} \right)^2 - 1 \right|,$$

where $d_{\mathcal{M}}(\cdot, \cdot)$ is the distance function on $\mathcal{M}$, $\mathbf{X}_{l_1}$ and $\mathbf{X}_{l_2}$ are the embeddings of nodes l_1 and l_2, respectively. The input distances are the shortest distances between pairs of connected nodes.

Our evaluation metric is the average of distortions over all pairs of points (López et al., 2021b). The distortion of a pair of connected nodes l_1 and l_2 is given as

$$\text{distortion}(l_1, l_2) = \frac{\|d_{\mathcal{M}}(\mathbf{X}_{l_1}, \mathbf{X}_{l_2}) - d_G(l_1, l_2)\|}{d_G(l_1, l_2)}.$$

Table 5: Statistics of the datasets for graph reconstruction experiments.

Graph	Nodes	Edges	Triples	Grid Layout	Tree Branching	Tree Height
Tree	40	39	780		3	3
Tree $\times$ Grid	1080	3213	582660	3×27	3	3
Grid $\times$ Tree	1080	3213	582660	3×27	3	3
Tree $\diamond$ Grid	364	363	66066	4×4	3	5
Grid $\diamond$ Tree	63	62	1953	4×4	2	5

Table 6: Statistics of the datasets for the node classification experiments.

Dataset	#Nodes	#Edges	#Classes	#Features
Disease	1044	1043	2	1000
Airport	3188	18631	4	4
Pubmed	19717	44338	3	500
Cora	2708	5429	7	1433

We parameterize the node embeddings on $\mathbb{QU}_n$ and convert them to $\mathbb{QB}_n$. Given $\mathbf{X} + \mathbf{Yj} \in \mathbb{QU}_n$, we have $\mathbf{X} \in \mathrm{Sym}_n(\mathbb{C})$ and $\mathbf{Y} \in \mathrm{Herm}_n^+$. Thus the matrix $\mathbf{X}$ can be parameterized by two vectors $\mathbf{u}, \mathbf{v} \in \mathbb{R}^{\frac{n(n+1)}{2}}$, and the matrix $\mathbf{Y}$ can be parameterized by two vectors $\mathbf{u}_1 \in \mathbb{R}^{\frac{n(n+1)}{2}}$ and $\mathbf{v}_1 \in \mathbb{R}^{\frac{n(n-1)}{2}}$.

Our experiments are based on the official code of López et al. (2021b)[1]. Our networks are implemented in the Pytorch framework and trained for a maximum of 3000 epochs. The batch size and the learning rate are set to 2048 and $1e-3$, respectively. The learning rate is reduced by a factor of 5 if the model does not improve the performance after 50 epochs López et al. (2021b). We use Riemannian SGD optimizer for SPD and Siegel models, and SGD optimizer for our models. Early stopping based on the average distortion is used when the model does not improve after 150 epochs. Results are computed over 5 random parameter initializations for each model. These experiments are performed using a machine with Intel(R) Xeon(R) Gold 6230R CPU @ 2.10GHz.

A.2 Node Classification

A.2.1 Datasets and Experimental Settings

Airport (Chami et al., 2019) It is a flight network dataset from OpenFlights.org where nodes represent airports, edges represent the airline Routes, and node labels are the populations of the country where the airport belongs.

Pubmed (Namata et al., 2012b) It is a standard benchmark describing citation networks where nodes represent scientific papers in the area of medicine, edges are citations between them, and node labels are academic (sub)areas.

Cora (Sen et al., 2008) It is a citation network where nodes represent scientific papers in the area of machine learning, edges are citations between them, and node labels are academic (sub)areas. Each publication in the dataset is described by a 0/1-valued word vector indicating the absence/presence of the corresponding word from the dictionary.

Disease (Anderson & May, 1992) It models the spread of infectious disease on a graph using the SusceptibleInfectedRecovered (SIR) epidemic model. Each node represents an individual, while edges represent disease transmission pathways between individuals. The class of a node indicates whether it is infected or uninfected individual.

[1] https://github.com/fedelopez77/sympa

We use the train, validation, and test splits as in Zhao et al. (2023). The statistics of the datasets are summarized in Tab. 6.

A.2.2 Optimization and Hyperparameters

The BiMap layer of our network reduces the dimensions of the input matrices by a factor of two. The isometry map used for the feature transform operation of our network is given (Zhao et al., 2023) by

$$\mathbf{M} \odot \mathbf{P} = \mathbf{M}\mathbf{P}\mathbf{M}^T,$$

where $\mathbf{M}$ is an invertible real matrix, and $\mathbf{P}$ is an SPD matrix.

Our experiments are based on the official code of Zhao et al. (2023)[2]. The networks are implemented in the Pytorch framework and trained using cross-entropy loss and Adam optimizer for a maximum of 500 epochs. The batch size is set to the total number of graph nodes in a dataset (Chami et al., 2019; Zhao et al., 2023). The learning rate is set to $1e-2$. Early stopping is used when the loss on the validation set has not decreased for 200 epochs. Each network has two layers that perform message passing twice at one iteration (Zhao et al., 2023). Our experiments are performed using a machine with Intel(R) Xeon(R) Silver 4510.

B More Applications

We consider the upper half-space and the precompact model associated with the matrix $\mathbf{P} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & -\mathbf{I} \end{pmatrix} \in \mathrm{Mat}_{2n}(\mathbb{R})$. In this case, the upper half-space model $\mathbb{RU}_{2n}$ is defined as

$$\mathbb{RU}_{2n} = \{\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{i} : \mathbf{Z}_1 = \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1 \\ -\mathbf{B}_1^T & \mathbf{C}_1 \end{pmatrix}, \mathbf{Z}_2 = \begin{pmatrix} \mathbf{A}_2 & \mathbf{B}_2 \\ -\mathbf{B}_2^T & \mathbf{C}_2 \end{pmatrix} \in A_+^\sigma, \mathbf{A}_1, \mathbf{C}_1, \mathbf{A}_2, \mathbf{C}_2 \in \mathrm{Sym}_n\}.$$

The precompact model $\mathbb{RB}_{2n}$ is defined by

$$\mathbb{RB}_{2n} = \{\mathbf{Z}_1 + \mathbf{Z}_2\mathbf{i} : \mathbf{Z}_1 = \begin{pmatrix} \mathbf{A}_1 & \mathbf{B}_1 \\ -\mathbf{B}_1^T & \mathbf{C}_1 \end{pmatrix}, \mathbf{Z}_2 = \begin{pmatrix} \mathbf{A}_2 & \mathbf{B}_2 \\ -\mathbf{B}_2^T & \mathbf{C}_2 \end{pmatrix}, \mathbf{I}_{2n} - \mathbf{P}^{-1}\mathbf{Z}^H\mathbf{Z}^T\mathbf{P} \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+, \mathbf{A}_1, \mathbf{C}_1, \mathbf{A}_2, \mathbf{C}_2 \in \mathrm{Sym}_n\}.$$

Since $\mathbf{Z}_1$ and $\mathbf{Z}_2$ in the definition of $\mathbb{RU}_{2n}$ can be parameterized by fewer parameters than matrices in Sym_{2n} (by selecting specific forms of $\mathbf{B}_1$), the spaces $\mathbb{RU}_{2n}$ and $\mathbb{RB}_{2n}$ can be good alternatives for the Siegel spaces in applications such as those investigated in Nguyen & Grozavu (2026).

C Mathematical Proofs

C.1 Proof of Proposition 3.1

Proof. The proof of this proposition proceeds in the same way as that of Proposition 3.4.

□

C.2 Proof of Proposition 3.2

Proof. Let $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ where $\mathbf{A} = \sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, $\mathbf{B} = -\mathbf{X}(\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$, $\mathbf{C} = -\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, and $\mathbf{D} = (\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$. Then

$$\sigma_{\mathbb{C}}(\mathbf{A})\mathbf{C} = -(\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right),$$

$$\sigma_{\mathbb{C}}(\mathbf{C})\mathbf{A} = -(\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\sigma_{\mathbb{C}}(\bar{\sigma}_{\mathbb{C}}(\mathbf{X}))\sigma_{\mathbb{C}}\left((\mathbf{I} - \bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right),$$

[2] https://github.com/andyweizhao/SPD4GNNs

$$\sigma_{\mathbb{C}}(\mathbf{B})\mathbf{D} = -\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\sigma_{\mathbb{C}}(\mathbf{X})(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}},$$

$$\sigma_{\mathbb{C}}(\mathbf{D})\mathbf{B} = -\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{X}(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}},$$

$$\sigma_{\mathbb{C}}(\mathbf{A})\mathbf{D} = (\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-1},$$

$$\sigma_{\mathbb{C}}(\mathbf{C})\mathbf{B} = (\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\sigma_{\mathbb{C}}(\bar{\sigma}_{\mathbb{C}}(\mathbf{X}))\mathbf{X}(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}.$$

Let $\mathbf{X} = \mathbf{U} + \mathbf{V}\mathbf{i} \in A_{\mathbb{C}}^{\sigma_{\mathbb{C}}}$. Then $\sigma_{\mathbb{C}}(\mathbf{X}) = \sigma(\mathbf{U}) + \sigma(\mathbf{V})\mathbf{i}$ and $\bar{\sigma}_{\mathbb{C}}(\mathbf{X}) = \sigma(\mathbf{U}) - \sigma(\mathbf{V})\mathbf{i}$. Since $\sigma_{\mathbb{C}}(\mathbf{X}) = \mathbf{X}$, we have $\sigma(\mathbf{U}) = \mathbf{U}$ and $\sigma(\mathbf{V}) = \mathbf{V}$. Thus $\sigma_{\mathbb{C}}(\bar{\sigma}_{\mathbb{C}}(\mathbf{X})) = \sigma(\sigma(\mathbf{U})) - \sigma(\sigma(\mathbf{V}))\mathbf{i} = \mathbf{U} - \mathbf{V}\mathbf{i} = \sigma(\mathbf{U}) - \sigma(\mathbf{V})\mathbf{i} = \bar{\sigma}_{\mathbb{C}}(\mathbf{X})$. Therefore,

$$\sigma_{\mathbb{C}}(\mathbf{A})\mathbf{C} = \sigma_{\mathbb{C}}(\mathbf{C})\mathbf{A}, \quad \sigma_{\mathbb{C}}(\mathbf{B})\mathbf{D} = \sigma_{\mathbb{C}}(\mathbf{D})\mathbf{B}, \quad \sigma_{\mathbb{C}}(\mathbf{A})\mathbf{D} - \sigma_{\mathbb{C}}(\mathbf{C})\mathbf{B} = \mathbf{I}.$$

The above equalities show that $\mathbf{M}$ satisfies the property that $\sigma_{\mathbb{C}}(\mathbf{M})^T\Omega\mathbf{M} = \Omega$. The action of $\mathbf{M}$ on $\mathbf{Y} \in \mathfrak{B}$ results in the map $\phi_{\mathbf{X}}(\mathbf{Y})$ given by

$$\begin{aligned}\phi_{\mathbf{X}}(\mathbf{Y}) = &\left(\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} - \mathbf{X}(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\\ &\left(-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} + (\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)^{-1}.\end{aligned}$$

□

C.3 Proof of Proposition 3.3

Proof. We first prove the following lemma.

Lemma C.1. *Let* $\mathbf{M} = \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix}$ *such that* $\sigma(\mathbf{M})^T\Omega\mathbf{M} = \Omega$. *Then*

$$\mathbf{M}^{-1} = \begin{pmatrix}\sigma(\mathbf{D}) & -\sigma(\mathbf{B})\\ -\sigma(\mathbf{C}) & \sigma(\mathbf{A})\end{pmatrix}$$

Proof. Since $\sigma(\mathbf{M})^T\Omega\mathbf{M} = \Omega$ and $\Omega^{-1} = -\Omega$, we have $\mathbf{M}^{-1}\Omega(\sigma(\mathbf{M})^T)^{-1} = \Omega$. Thus $\mathbf{M}\Omega\sigma(\mathbf{M})^T = \Omega$. This results in the following equalities:

$$\mathbf{A}\sigma(\mathbf{B}) - \mathbf{B}\sigma(\mathbf{A}) = \mathbf{0}, \quad \mathbf{C}\sigma(\mathbf{D}) - \mathbf{D}\sigma(\mathbf{C}) = \mathbf{0}, \quad \mathbf{A}\sigma(\mathbf{D}) - \mathbf{B}\sigma(\mathbf{C}) = \mathbf{I}.$$

It is then easy to see that $\mathbf{M}\begin{pmatrix}\sigma(\mathbf{D}) & -\sigma(\mathbf{B})\\ -\sigma(\mathbf{C}) & \sigma(\mathbf{A})\end{pmatrix} = \begin{pmatrix}\mathbf{I} & \mathbf{0}\\ \mathbf{0} & \mathbf{I}\end{pmatrix}$.

□

Let $\mathbf{M} = \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix} \in \begin{pmatrix}\mathbf{I} & \mathbf{i}\mathbf{I}\\ \mathbf{i}\mathbf{I} & \mathbf{I}\end{pmatrix}^{-1}\mathrm{Sp}_2(A,\sigma)\begin{pmatrix}\mathbf{I} & \mathbf{i}\mathbf{I}\\ \mathbf{i}\mathbf{I} & \mathbf{I}\end{pmatrix}$, then $\sigma_{\mathbb{C}}(\mathbf{M})^T\Omega\mathbf{M} = \Omega$. Let $\mathbf{Y} \in \mathfrak{B}$ and $\mathbf{Z} = (\mathbf{A}\mathbf{Y}+\mathbf{B})(\mathbf{C}\mathbf{Y}+\mathbf{D})^{-1}$, then

$$\begin{aligned}\sigma_{\mathbb{C}}(\mathbf{Z}) &= \sigma_{\mathbb{C}}\left((\mathbf{C}\mathbf{Y}+\mathbf{D})^{-1}\right)\sigma_{\mathbb{C}}(\mathbf{A}\mathbf{Y}+\mathbf{B})\\ &= (\sigma_{\mathbb{C}}(\mathbf{C}\mathbf{Y}+\mathbf{D}))^{-1}\sigma_{\mathbb{C}}(\mathbf{A}\mathbf{Y}+\mathbf{B})\\ &= (\sigma_{\mathbb{C}}(\mathbf{Y})\sigma_{\mathbb{C}}(\mathbf{C}) + \sigma_{\mathbb{C}}(\mathbf{D}))^{-1}(\sigma_{\mathbb{C}}(\mathbf{Y})\sigma_{\mathbb{C}}(\mathbf{A}) + \sigma_{\mathbb{C}}(\mathbf{B})).\end{aligned}$$

Hence

$$(\sigma_{\mathbb{C}}(\mathbf{Y})\sigma_{\mathbb{C}}(\mathbf{C}) + \sigma_{\mathbb{C}}(\mathbf{D}))\sigma_{\mathbb{C}}(\mathbf{Z}) = \sigma_{\mathbb{C}}(\mathbf{Y})\sigma_{\mathbb{C}}(\mathbf{A}) + \sigma_{\mathbb{C}}(\mathbf{B}).$$

We obtain

$$\sigma_{\mathbb{C}}(\mathbf{Y})(-\sigma_{\mathbb{C}}(\mathbf{C})\sigma_{\mathbb{C}}(\mathbf{Z})+\sigma_{\mathbb{C}}(\mathbf{A}))=\sigma_{\mathbb{C}}(\mathbf{D})\sigma_{\mathbb{C}}(\mathbf{Z})-\sigma_{\mathbb{C}}(\mathbf{B}).$$

Since $\begin{pmatrix}\mathbf{I} & \mathbf{iI}\\ \mathbf{iI} & \mathbf{I}\end{pmatrix}^{-1}\mathrm{Sp}_2(A,\sigma)\begin{pmatrix}\mathbf{I} & \mathbf{iI}\\ \mathbf{iI} & \mathbf{I}\end{pmatrix}$ acts on $\mathfrak{B}$ by Möbius transformations, we have $\mathbf{Z}\in\mathfrak{B}$. Thus $\sigma_{\mathbb{C}}(\mathbf{Z})=\mathbf{Z}$. By Lemma C.2, $\mathbf{M}^{-1}=\begin{pmatrix}\sigma_{\mathbb{C}}(\mathbf{D}) & -\sigma_{\mathbb{C}}(\mathbf{B})\\ -\sigma_{\mathbb{C}}(\mathbf{C}) & \sigma_{\mathbb{C}}(\mathbf{A})\end{pmatrix}\in\begin{pmatrix}\mathbf{I} & \mathbf{iI}\\ \mathbf{iI} & \mathbf{I}\end{pmatrix}^{-1}\mathrm{Sp}_2(A,\sigma)\begin{pmatrix}\mathbf{I} & \mathbf{iI}\\ \mathbf{iI} & \mathbf{I}\end{pmatrix}$. Thus $-\sigma_{\mathbb{C}}(\mathbf{C})\mathbf{Z}+\sigma_{\mathbb{C}}(\mathbf{A})$ is invertible and we have

$$\mathbf{Y}=(\sigma_{\mathbb{C}}(\mathbf{D})\mathbf{Z}-\sigma_{\mathbb{C}}(\mathbf{B}))(-\sigma_{\mathbb{C}}(\mathbf{C})\mathbf{Z}+\sigma_{\mathbb{C}}(\mathbf{A}))^{-1}.$$

The automorphism $\phi_{\mathbf{X}}(\mathbf{Y})$ is obtained by the action of $\mathbf{M}=\begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix}$ on $\mathbf{Y}$, where $\mathbf{A}=\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, $\mathbf{B}=-\mathbf{X}(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$, $\mathbf{C}=-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, and $\mathbf{D}=(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$. Thus the inverse map $\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y})$ is obtained by the action of $\begin{pmatrix}\mathbf{A}^{(-1)} & \mathbf{B}^{(-1)}\\ \mathbf{C}^{(-1)} & \mathbf{D}^{(-1)}\end{pmatrix}$ on $\mathbf{Y}$, where

$$\begin{aligned}\mathbf{A}^{(-1)}&=\sigma_{\mathbb{C}}(\mathbf{D})=\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right),\\ \mathbf{B}^{(-1)}&=-\sigma_{\mathbb{C}}(\mathbf{B})=\sigma_{\mathbb{C}}\left((\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{X},\\ \mathbf{C}^{(-1)}&=-\sigma_{\mathbb{C}}(\mathbf{C})=(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\bar{\sigma}_{\mathbb{C}}(\mathbf{X}),\\ \mathbf{D}^{(-1)}&=\sigma_{\mathbb{C}}(\mathbf{A})=(\mathbf{I}-\bar{\sigma}_{\mathbb{C}}(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}.\end{aligned}$$

This completes the proof. □

C.4 Proof of Proposition 3.4

Proof. Let $\mathbf{Z}_0\in\mathfrak{B}$, then

$$\begin{aligned}\sigma_0(\mathbf{Z}_0+\mathbf{jI})(\mathbf{I}+\mathbf{jZ}_0)&=(\mathbf{Z}_0+\mathbf{jI})(\mathbf{I}+\mathbf{jZ}_0)\\ &=(\mathbf{I}+\mathbf{Z}_0\mathbf{j})(\mathbf{Z}_0+\mathbf{jI})\\ &=(\mathbf{I}+\sigma_0(\mathbf{Z}_0)\mathbf{j})(\mathbf{Z}_0+\mathbf{jI})\\ &=(\mathbf{I}+\sigma_0(\mathbf{jZ}_0))(\mathbf{Z}_0+\mathbf{jI})\\ &=\sigma_0(\mathbf{I}+\mathbf{jZ}_0)(\mathbf{Z}_0+\mathbf{jI}).\end{aligned}\tag{6}$$

Let $\mathbf{R}$ be a solution of $(\mathbf{I}+\mathbf{jZ}_0)\mathbf{R}=\mathbf{0}$, then $\mathbf{R}=-\mathbf{jZ}_0\mathbf{R}$ and we have

$$\begin{aligned}\sigma_1(\mathbf{R})(\mathbf{I}-\sigma_1(\mathbf{Z}_0)\mathbf{Z}_0)\mathbf{R}&=\sigma_1(\mathbf{R})\mathbf{R}-\sigma_1(\mathbf{Z}_0\mathbf{R})\mathbf{Z}_0\mathbf{R}\\ &=\sigma_1(-\mathbf{jZ}_0\mathbf{R})(-\mathbf{jZ}_0\mathbf{R})-\sigma_1(\mathbf{Z}_0\mathbf{R})\mathbf{Z}_0\mathbf{R}\\ &=\sigma_1(\mathbf{Z}_0\mathbf{R})\mathbf{j}(-\mathbf{jZ}_0\mathbf{R})-\sigma_1(\mathbf{Z}_0\mathbf{R})\mathbf{Z}_0\mathbf{R}\\ &=\mathbf{0}.\end{aligned}$$

Since $\mathbf{I}-\sigma_1(\mathbf{Z}_0)\mathbf{Z}_0\in(A_{\mathbb{H}}^{\sigma_1})_+$, we deduce that $\mathbf{I}+\mathbf{jZ}_0$ is invertible. Thus, from Eq. (6) we obtain

$$\begin{aligned}(\mathbf{Z}_0+\mathbf{jI})(\mathbf{I}+\mathbf{jZ}_0)^{-1}&=\sigma_0((\mathbf{I}+\mathbf{jZ}_0)^{-1})\sigma_0(\mathbf{Z}_0+\mathbf{jI})\\ &=\sigma_0((\mathbf{Z}_0+\mathbf{jI})(\mathbf{I}+\mathbf{jZ}_0)^{-1}).\end{aligned}$$

Let $\mathbf{Z}=(\mathbf{Z}_0+\mathbf{jI})(\mathbf{I}+\mathbf{jZ}_0)^{-1}$, then the above equation shows that $\mathbf{Z}=\sigma_0(\mathbf{Z})$ and thus $\mathbf{Z}\in A_{\mathbb{H}}^{\sigma_0}$. Let $\mathbf{Z}=\mathbf{X}+\mathbf{Yj}$, then

$$\begin{aligned}-2\mathbf{Y}&=\mathbf{Zj}-\mathbf{j}\sigma_1(\mathbf{Z})\\ &=(\mathbf{Z}_0+\mathbf{jI})(\mathbf{I}+\mathbf{jZ}_0)^{-1}\mathbf{j}-\mathbf{j}\sigma_1((\mathbf{I}+\mathbf{jZ}_0)^{-1})\sigma_1(\mathbf{Z}_0+\mathbf{jI}).\end{aligned}$$

Therefore

$$\begin{aligned}-2\sigma_1(\mathbf{I}+\mathbf{j}\mathbf{Z}_0)\mathbf{j}\mathbf{Y}\mathbf{j}(\mathbf{I}+\mathbf{j}\mathbf{Z}_0) &= \sigma_1(\mathbf{Z}_0+\mathbf{j}\mathbf{I})\mathbf{j}(\mathbf{I}+\mathbf{j}\mathbf{Z}_0) - \sigma_1(\mathbf{I}+\mathbf{j}\mathbf{Z}_0)\mathbf{j}(\mathbf{Z}_0+\mathbf{j}\mathbf{I})\\ &= (\sigma_1(\mathbf{Z}_0)-\mathbf{j})\mathbf{j}(\mathbf{I}+\mathbf{j}\mathbf{Z}_0) - (\mathbf{I}-\sigma_1(\mathbf{Z}_0)\mathbf{j})\mathbf{j}(\mathbf{Z}_0+\mathbf{j}\mathbf{I})\\ &= (\sigma_1(\mathbf{Z}_0)\mathbf{j}+\mathbf{I})(\mathbf{I}+\mathbf{j}\mathbf{Z}_0) - (\sigma_1(\mathbf{Z}_0)+\mathbf{j}\mathbf{I})(\mathbf{Z}_0+\mathbf{j}\mathbf{I})\\ &= 2(\mathbf{I}-\sigma_1(\mathbf{Z}_0)\mathbf{Z}_0),\end{aligned}$$

which shows that $\mathbf{Y} \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+$ and therefore $\mathbf{Z} \in \mathfrak{U}$.

Let $\mathbf{Z} \in \mathfrak{U}$, then

$$\begin{aligned}\sigma_0(\mathbf{I}-\mathbf{j}\mathbf{Z})(\mathbf{Z}-\mathbf{j}\mathbf{I}) &= (\mathbf{I}-\sigma_0(\mathbf{j}\mathbf{Z}))(\mathbf{Z}-\mathbf{j}\mathbf{I})\\ &= (\mathbf{I}-\sigma_0(\mathbf{Z})\mathbf{j})(\mathbf{Z}-\mathbf{j}\mathbf{I})\\ &= (\mathbf{I}-\mathbf{Z}\mathbf{j})(\mathbf{Z}-\mathbf{j}\mathbf{I})\\ &= (\mathbf{Z}-\mathbf{j}\mathbf{I})(\mathbf{I}-\mathbf{j}\mathbf{Z})\\ &= (\sigma_0(\mathbf{Z})-\mathbf{j}\mathbf{I})(\mathbf{I}-\mathbf{j}\mathbf{Z})\\ &= \sigma_0(\mathbf{Z}-\mathbf{j}\mathbf{I})(\mathbf{I}-\mathbf{j}\mathbf{Z}).\end{aligned} \tag{7}$$

Let $\mathbf{R}$ be a solution of $(\mathbf{I}-\mathbf{j}\mathbf{Z})\mathbf{R} = \mathbf{0}$, then $\mathbf{R} = \mathbf{j}\mathbf{Z}\mathbf{R}$ and we have

$$\begin{aligned}\sigma_1(\mathbf{R})(\sigma_1(\mathbf{Z})\mathbf{j}-\mathbf{j}\mathbf{Z})\mathbf{R} &= \sigma_1(\mathbf{Z}\mathbf{R})\mathbf{j}\mathbf{R} - \sigma_1(\mathbf{R})\mathbf{j}\mathbf{Z}\mathbf{R}\\ &= \sigma_1(\mathbf{Z}\mathbf{R})\mathbf{j}\mathbf{j}\mathbf{Z}\mathbf{R} + \sigma_1(\mathbf{Z}\mathbf{R})\mathbf{j}\mathbf{j}\mathbf{Z}\mathbf{R}\\ &= -2\sigma_1(\mathbf{Z}\mathbf{R})\mathbf{Z}\mathbf{R}.\end{aligned}$$

Since $\sigma_1(\mathbf{Z}\mathbf{R})\mathbf{Z}\mathbf{R} \geq 0$ and $\sigma_1(\mathbf{Z})\mathbf{j}-\mathbf{j}\mathbf{Z} = \mathbf{j}(\mathbf{Z}\mathbf{j}-\mathbf{j}\sigma_1(\mathbf{Z}))\mathbf{j} = -\mathbf{j}\,\mathrm{Im}(\mathbf{Z})\mathbf{j}$, we deduce that $\mathbf{I}-\mathbf{j}\mathbf{Z}$ is invertible. This shows the existence of $\mathbf{Z}_0 = (\mathbf{Z}-\mathbf{j}\mathbf{I})(\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1}$. From Eq. (7) we obtain $\sigma_0(\mathbf{Z}_0) = \mathbf{Z}_0$ and therefore $\mathbf{Z}_0 \in A_{\mathbb{H}}^{\sigma_0}$. Note that

$$\begin{aligned}\mathbf{I}-\sigma_1(\mathbf{Z}_0)\mathbf{Z}_0 &= \mathbf{I}-\sigma_1((\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1})\sigma_1(\mathbf{Z}-\mathbf{j}\mathbf{I})(\mathbf{Z}-\mathbf{j}\mathbf{I})(\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1}\\ &= \sigma_1((\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1})\left(\sigma_1(\mathbf{I}-\mathbf{j}\mathbf{Z})(\mathbf{I}-\mathbf{j}\mathbf{Z}) - \sigma_1(\mathbf{Z}-\mathbf{j}\mathbf{I})(\mathbf{Z}-\mathbf{j}\mathbf{I})\right)(\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1}\\ &= \sigma_1((\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1})\left((\mathbf{I}+\sigma_1(\mathbf{Z})\mathbf{j})(\mathbf{I}-\mathbf{j}\mathbf{Z}) - (\sigma_1(\mathbf{Z})+\mathbf{j}\mathbf{I})(\mathbf{Z}-\mathbf{j}\mathbf{I})\right)(\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1}\\ &= 2\sigma_1((\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1})(\sigma_1(\mathbf{Z})\mathbf{j}-\mathbf{j}\mathbf{Z})(\mathbf{I}-\mathbf{j}\mathbf{Z})^{-1}.\end{aligned}$$

Therefore

$$\begin{aligned}\mathbf{j}\sigma_1(\mathbf{I}-\mathbf{j}\mathbf{Z})(\mathbf{I}-\sigma_1(\mathbf{Z}_0)\mathbf{Z}_0)(\mathbf{I}-\mathbf{j}\mathbf{Z})\mathbf{j} &= 2\mathbf{j}(\sigma_1(\mathbf{Z})\mathbf{j}-\mathbf{j}\mathbf{Z})\mathbf{j}\\ &= 2(\mathbf{Z}\mathbf{j}-\mathbf{j}\sigma_1(\mathbf{Z}))\\ &= -2\,\mathrm{Im}(\mathbf{Z}),\end{aligned}$$

which shows that $\mathbf{I}-\sigma_1(\mathbf{Z}_0)\mathbf{Z}_0 \in (A_{\mathbb{H}}^{\sigma_1})_+$ and thus $\mathbf{Z}_0 \in \mathfrak{B}$.

□

C.5 Proof of Proposition 3.5

Proof. Let $\mathbf{L} = \frac{1}{\sqrt{2}}\begin{pmatrix}\mathbf{I} & \mathbf{j}\mathbf{I}\\ \mathbf{j}\mathbf{I} & \mathbf{I}\end{pmatrix}$, $\mathbf{M} = \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix} \in \mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})$, $\mathbf{M}_0 = \begin{pmatrix}\mathbf{A}_0 & \mathbf{B}_0\\ \mathbf{C}_0 & \mathbf{D}_0\end{pmatrix} = \mathbf{L}^{-1}\mathbf{M}\mathbf{L}$, and let $\mathbf{Z}_0 \in \mathfrak{B}$, then

$$\begin{aligned}\mathbf{M}_0\begin{pmatrix}\mathbf{Z}_0\\ \mathbf{I}\end{pmatrix} &= \mathbf{L}^{-1}\mathbf{M}\mathbf{L}\begin{pmatrix}\mathbf{Z}_0\\ \mathbf{I}\end{pmatrix}\\ &= \frac{1}{\sqrt{2}}\mathbf{L}^{-1}\mathbf{M}\begin{pmatrix}\mathbf{Z}_0+\mathbf{j}\mathbf{I}\\ \mathbf{j}\mathbf{Z}_0+\mathbf{I}\end{pmatrix}\\ &= \frac{1}{\sqrt{2}}\mathbf{L}^{-1}\mathbf{M}\begin{pmatrix}(\mathbf{Z}_0+\mathbf{j}\mathbf{I})(\mathbf{j}\mathbf{Z}_0+\mathbf{I})^{-1}\\ \mathbf{I}\end{pmatrix}(\mathbf{j}\mathbf{Z}_0+\mathbf{I}).\end{aligned}$$

Let $\mathbf{Z} = (\mathbf{Z}_0 + \mathbf{jI})(\mathbf{jZ}_0 + \mathbf{I})^{-1}$, then $\mathbf{Z} \in \mathfrak{U}$ and we have

$$\begin{aligned}\mathbf{M}_0 \begin{pmatrix}\mathbf{Z}_0 \\ \mathbf{I}\end{pmatrix} &= \frac{1}{\sqrt{2}}\mathbf{L}^{-1}\mathbf{M}\begin{pmatrix}\mathbf{Z} \\ \mathbf{I}\end{pmatrix}(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{\sqrt{2}}\mathbf{L}^{-1}\begin{pmatrix}\mathbf{AZ} + \mathbf{B} \\ \mathbf{CZ} + \mathbf{D}\end{pmatrix}(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{\sqrt{2}}\mathbf{L}^{-1}\begin{pmatrix}(\mathbf{AZ} + \mathbf{B})(\mathbf{CZ} + \mathbf{D})^{-1} \\ \mathbf{I}\end{pmatrix}(\mathbf{CZ} + \mathbf{D})(\mathbf{jZ}_0 + \mathbf{I}).\end{aligned}$$

Let $\mathbf{Z}' = (\mathbf{AZ} + \mathbf{B})(\mathbf{CZ} + \mathbf{D})^{-1}$, then $\mathbf{Z}' \in \mathfrak{U}$ and we have

$$\begin{aligned}\mathbf{M}_0 \begin{pmatrix}\mathbf{Z}_0 \\ \mathbf{I}\end{pmatrix} &= \frac{1}{\sqrt{2}}\mathbf{L}^{-1}\begin{pmatrix}\mathbf{Z}' \\ \mathbf{I}\end{pmatrix}(\mathbf{CZ} + \mathbf{D})(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{2}\begin{pmatrix}\mathbf{Z}' - \mathbf{jI} \\ \mathbf{I} - \mathbf{jZ}'\end{pmatrix}(\mathbf{CZ} + \mathbf{D})(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{2}\begin{pmatrix}(\mathbf{Z}' - \mathbf{jI})(\mathbf{I} - \mathbf{jZ}')^{-1} \\ \mathbf{I}\end{pmatrix}(\mathbf{I} - \mathbf{jZ}')(\mathbf{CZ} + \mathbf{D})(\mathbf{jZ}_0 + \mathbf{I}).\end{aligned}$$

Let $\mathbf{Z}'_0 = (\mathbf{Z}' - \mathbf{jI})(\mathbf{I} - \mathbf{jZ}')^{-1}$, then $\mathbf{Z}'_0 \in \mathfrak{B}$ and we have

$$\begin{aligned}\mathbf{M}_0 \begin{pmatrix}\mathbf{Z}_0 \\ \mathbf{I}\end{pmatrix} &= \frac{1}{2}\begin{pmatrix}\mathbf{Z}'_0 \\ \mathbf{I}\end{pmatrix}(\mathbf{I} - \mathbf{jZ}')(\mathbf{CZ} + \mathbf{D})(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{2}\begin{pmatrix}\mathbf{Z}'_0 \\ \mathbf{I}\end{pmatrix}(\mathbf{CZ} + \mathbf{D} - \mathbf{jZ}'(\mathbf{CZ} + \mathbf{D}))(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{2}\begin{pmatrix}\mathbf{Z}'_0 \\ \mathbf{I}\end{pmatrix}(\mathbf{CZ} + \mathbf{D} - \mathbf{j}(\mathbf{AZ} + \mathbf{B}))(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{2}\begin{pmatrix}\mathbf{Z}'_0 \\ \mathbf{I}\end{pmatrix}((\mathbf{C} - \mathbf{jA})\mathbf{Z} + \mathbf{D} - \mathbf{jB})(\mathbf{jZ}_0 + \mathbf{I}) \\ &= \frac{1}{2}\begin{pmatrix}\mathbf{Z}'_0 \\ \mathbf{I}\end{pmatrix}((\mathbf{C} - \mathbf{jA})(\mathbf{Z}_0 + \mathbf{jI}) + (\mathbf{D} - \mathbf{jB})(\mathbf{jZ}_0 + \mathbf{I})) \\ &= \frac{1}{2}\begin{pmatrix}\mathbf{Z}'_0 \\ \mathbf{I}\end{pmatrix}((\mathbf{C} - \mathbf{jA} + (\mathbf{D} - \mathbf{jB})\mathbf{j})\mathbf{Z}_0 + (\mathbf{C} - \mathbf{jA})\mathbf{j} + \mathbf{D} - \mathbf{jB}) \\ &= \begin{pmatrix}\mathbf{Z}'_0 \\ \mathbf{I}\end{pmatrix}(\mathbf{C}_0\mathbf{Z}_0 + \mathbf{D}_0).\end{aligned}$$

From the above chain of equations, we obtain $\mathbf{C}_0\mathbf{Z}_0 + \mathbf{D}_0 = \frac{1}{2}(\mathbf{I} - \mathbf{jZ}')(\mathbf{CZ} + \mathbf{D})(\mathbf{jZ}_0 + \mathbf{I})$ and $\mathbf{A}_0\mathbf{Z}_0 + \mathbf{B}_0 = \mathbf{Z}'_0(\mathbf{C}_0\mathbf{Z}_0 + \mathbf{D}_0)$. Thus $\mathbf{C}_0\mathbf{Z}_0 + \mathbf{D}_0$ is invertible and $\mathbf{Z}'_0 = (\mathbf{A}_0\mathbf{Z}_0 + \mathbf{B}_0)(\mathbf{C}_0\mathbf{Z}_0 + \mathbf{D}_0)^{-1}$. This shows that the Möbius transformation associated with $\mathbf{M}_0$ maps any point $\mathbf{Z}_0 \in \mathfrak{B}$ to a point $\mathbf{Z}'_0 \in \mathfrak{B}$. This action is transitive since the action of $\mathrm{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})$ on $\mathfrak{U}$ is transitive, and points can be converted between $\mathfrak{U}$ and $\mathfrak{B}$ using the mappings in Eqs. (1) and (2). This completes the proof.

□

C.6 Proof of Proposition 3.6

Proof. Let $\mathbf{M} = \begin{pmatrix}\mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D}\end{pmatrix}$ where $\mathbf{A} = \sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, $\mathbf{B} = -\mathbf{X}(\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$, $\mathbf{C} = -\sigma_1(\mathbf{X})\sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, and $\mathbf{D} = (\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$. Then

$$\sigma_0(\mathbf{A})\mathbf{C} = -(\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\sigma_1(\mathbf{X})\sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right),$$

$$\sigma_0(\mathbf{C})\mathbf{A} = -(\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\sigma_0(\sigma_1(\mathbf{X}))\sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right),$$

$$\sigma_0(\mathbf{B})\mathbf{D} = -\sigma_0\left((\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\sigma_0(\mathbf{X})(\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}},$$

$$\sigma_0(\mathbf{D})\mathbf{B} = -\sigma_0\left((\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{X}(\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}},$$

$$\sigma_0(\mathbf{A})\mathbf{D} = (\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-1},$$

$$\sigma_0(\mathbf{C})\mathbf{B} = (\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\sigma_0(\sigma_1(\mathbf{X}))\mathbf{X}(\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}.$$

Let $\mathbf{X} = \mathbf{U} + \mathbf{V}\mathbf{j} \in A^{\sigma_0}_{\mathbb{H}}$, then $\sigma_0(\mathbf{X}) = \sigma_{\mathbb{C}}(\mathbf{U}) + \bar{\sigma}_{\mathbb{C}}(\mathbf{V})\mathbf{j}$ and $\sigma_1(\mathbf{X}) = \bar{\sigma}_{\mathbb{C}}(\mathbf{U}) - \sigma_{\mathbb{C}}(\mathbf{V})\mathbf{j}$. Since $\sigma_0(\mathbf{X}) = \mathbf{X}$, we have $\sigma_{\mathbb{C}}(\mathbf{U}) = \mathbf{U}$ and $\bar{\sigma}_{\mathbb{C}}(\mathbf{V}) = \mathbf{V}$. Thus $\sigma_0(\sigma_1(\mathbf{X})) = \sigma_{\mathbb{C}}(\bar{\sigma}_{\mathbb{C}}(\mathbf{U})) - \bar{\sigma}_{\mathbb{C}}(\sigma_{\mathbb{C}}(\mathbf{V}))\mathbf{j}$.

Let $\mathbf{U} = \mathbf{U}_0 + \mathbf{U}_1\mathbf{i}$, then $\bar{\sigma}_{\mathbb{C}}(\mathbf{U}) = \sigma(\mathbf{U}_0) - \sigma(\mathbf{U}_1)\mathbf{i}$ and we have

$$\begin{aligned}\sigma_{\mathbb{C}}(\bar{\sigma}_{\mathbb{C}}(\mathbf{U})) &= \sigma(\sigma(\mathbf{U}_0)) - \sigma(\sigma(\mathbf{U}_1))\mathbf{i}\\ &= \mathbf{U}_0 - \mathbf{U}_1\mathbf{i}\\ &= \sigma(\mathbf{U}_0) - \sigma(\mathbf{U}_1)\mathbf{i}\\ &= \bar{\sigma}_{\mathbb{C}}(\mathbf{U}),\end{aligned}$$

where the third equality follows from the fact that $\sigma_{\mathbb{C}}(\mathbf{U}) = \mathbf{U}$. Let $\mathbf{V} = \mathbf{V}_0 + \mathbf{V}_1\mathbf{i}$, then $\sigma_{\mathbb{C}}(\mathbf{V}) = \sigma(\mathbf{V}_0) + \sigma(\mathbf{V}_1)\mathbf{i}$ and we have

$$\begin{aligned}\bar{\sigma}_{\mathbb{C}}(\sigma_{\mathbb{C}}(\mathbf{V})) &= \sigma(\sigma(\mathbf{V}_0)) - \sigma(\sigma(\mathbf{V}_1))\mathbf{i}\\ &= \mathbf{V}_0 - \mathbf{V}_1\mathbf{i}\\ &= \sigma(\mathbf{V}_0) + \sigma(\mathbf{V}_1)\mathbf{i}\\ &= \sigma_{\mathbb{C}}(\mathbf{V}),\end{aligned}$$

where the third equality follows from the fact that $\bar{\sigma}_{\mathbb{C}}(\mathbf{V}) = \mathbf{V}$. We thus have

$$\begin{aligned}\sigma_0(\sigma_1(\mathbf{X})) &= \bar{\sigma}_{\mathbb{C}}(\mathbf{U}) - \sigma_{\mathbb{C}}(\mathbf{V})\mathbf{j}\\ &= \sigma_1(\mathbf{X}).\end{aligned}$$

Therefore,

$$\sigma_0(\mathbf{A})\mathbf{C} = \sigma_0(\mathbf{C})\mathbf{A}, \quad \sigma_0(\mathbf{B})\mathbf{D} = \sigma_0(\mathbf{D})\mathbf{B}, \quad \sigma_0(\mathbf{A})\mathbf{D} - \sigma_0(\mathbf{C})\mathbf{B} = \mathbf{I}.$$

The above equalities show that $\mathbf{M}$ satisfies the equation $\sigma_0(\mathbf{M})^{\mathrm{T}}\Omega\mathbf{M} = \Omega$. By Proposition 3.5, this equation characterizes the group acting on $\mathfrak{B}$ via Möbius transformations. The action of $\mathbf{M}$ on a point $\mathbf{Y} \in \mathfrak{B}$ results in the map $\phi_{\mathbf{X}}(\mathbf{Y})$ given by

$$\begin{aligned}\phi_{\mathbf{X}}(\mathbf{Y}) = &\left(\sigma_0\left((\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} - \mathbf{X}(\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\\ &\left(-\sigma_1(\mathbf{X})\sigma_0\left((\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{Y} + (\mathbf{I}-\sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)^{-1}.\end{aligned}$$

This completes the proof.

□

C.7 Proof of Proposition 3.7

Proof. We first prove the following lemma.

Lemma C.2. *Let* $\mathbf{M} = \begin{pmatrix}\mathbf{A} & \mathbf{B}\\ \mathbf{C} & \mathbf{D}\end{pmatrix}$ *be such that* $\sigma(\mathbf{M})^T\Omega\mathbf{M} = \Omega$. *Then*

$$\mathbf{M}^{-1} = \begin{pmatrix}\sigma(\mathbf{D}) & -\sigma(\mathbf{B})\\ -\sigma(\mathbf{C}) & \sigma(\mathbf{A})\end{pmatrix}$$

Proof. Since $\sigma(\mathbf{M})^{\mathrm{T}}\Omega\mathbf{M} = \Omega$ and $\Omega^{-1} = -\Omega$, we have $\mathbf{M}^{-1}\Omega(\sigma(\mathbf{M})^{\mathrm{T}})^{-1} = \Omega$. Thus $\mathbf{M}\Omega\sigma(\mathbf{M})^{\mathrm{T}} = \Omega$. This results in the following equalities:

$$\mathbf{A}\sigma(\mathbf{B}) - \mathbf{B}\sigma(\mathbf{A}) = \mathbf{0}, \quad \mathbf{C}\sigma(\mathbf{D}) - \mathbf{D}\sigma(\mathbf{C}) = \mathbf{0}, \quad \mathbf{A}\sigma(\mathbf{D}) - \mathbf{B}\sigma(\mathbf{C}) = \mathbf{I}.$$

It is then easy to see that $\mathbf{M}\begin{pmatrix} \sigma(\mathbf{D}) & -\sigma(\mathbf{B}) \\ -\sigma(\mathbf{C}) & \sigma(\mathbf{A}) \end{pmatrix} = \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{pmatrix}$.

□

Let $\mathbf{L} = \frac{1}{\sqrt{2}}\begin{pmatrix} \mathbf{I} & \mathbf{jI} \\ \mathbf{jI} & \mathbf{I} \end{pmatrix}$, $\mathbf{M} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix} \in \mathbf{L}^{-1}\operatorname{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})\mathbf{L}$, $\mathbf{Y} \in \mathfrak{B}$, and let $\mathbf{Z} = (\mathbf{AY} + \mathbf{B})(\mathbf{CY} + \mathbf{D})^{-1}$, then

$$\begin{aligned} \sigma_0(\mathbf{Z}) &= \sigma_0\left((\mathbf{CY} + \mathbf{D})^{-1}\right)\sigma_0(\mathbf{AY} + \mathbf{B}) \\ &= (\sigma_0(\mathbf{CY} + \mathbf{D}))^{-1}\sigma_0(\mathbf{AY} + \mathbf{B}) \\ &= (\sigma_0(\mathbf{Y})\sigma_0(\mathbf{C}) + \sigma_0(\mathbf{D}))^{-1}(\sigma_0(\mathbf{Y})\sigma_0(\mathbf{A}) + \sigma_0(\mathbf{B})). \end{aligned}$$

Therefore

$$(\sigma_0(\mathbf{Y})\sigma_0(\mathbf{C}) + \sigma_0(\mathbf{D}))\sigma_0(\mathbf{Z}) = \sigma_0(\mathbf{Y})\sigma_0(\mathbf{A}) + \sigma_0(\mathbf{B}),$$

which results in

$$\sigma_0(\mathbf{Y})(-\sigma_0(\mathbf{C})\sigma_0(\mathbf{Z}) + \sigma_0(\mathbf{A})) = \sigma_0(\mathbf{D})\sigma_0(\mathbf{Z}) - \sigma_0(\mathbf{B}). \tag{8}$$

Since $\mathbf{L}^{-1}\operatorname{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})\mathbf{L}$ acts on $\mathfrak{B}$ by Möbius transformations, we have $\mathbf{Z} \in \mathfrak{B}$, $\sigma_0(\mathbf{Z}) = \mathbf{Z}$. Note that $\sigma_0(\mathbf{M})^{\mathrm{T}}\Omega\mathbf{M} = \Omega$. By Lemma C.2, $\mathbf{M}^{-1} = \begin{pmatrix} \sigma_0(\mathbf{D}) & -\sigma_0(\mathbf{B}) \\ -\sigma_0(\mathbf{C}) & \sigma_0(\mathbf{A}) \end{pmatrix} \in \mathbf{L}^{-1}\operatorname{Sp}_2(A_{\mathbb{C}}, \sigma_{\mathbb{C}})\mathbf{L}$. Thus $-\sigma_0(\mathbf{C})\mathbf{Z} + \sigma_0(\mathbf{A})$ is invertible and from Eq. (8) we get

$$\mathbf{Y} = (\sigma_0(\mathbf{D})\mathbf{Z} - \sigma_0(\mathbf{B}))(-\sigma_0(\mathbf{C})\mathbf{Z} + \sigma_0(\mathbf{A}))^{-1}.$$

For any $\mathbf{X}, \mathbf{Y} \in \mathfrak{B}$, the automorphism $\phi_{\mathbf{X}}(\mathbf{Y})$ is obtained by the action of $\begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$ on $\mathbf{Y}$, where $\mathbf{A} = \sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, $\mathbf{B} = -\mathbf{X}(\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$, $\mathbf{C} = -\sigma_1(\mathbf{X})\sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)$, and $\mathbf{D} = (\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}$. Thus the inverse map $\phi_{\mathbf{X}}^{(-1)}(\mathbf{Y})$ is obtained by the action of $\begin{pmatrix} \mathbf{A}^{(-1)} & \mathbf{B}^{(-1)} \\ \mathbf{C}^{(-1)} & \mathbf{D}^{(-1)} \end{pmatrix}$ on $\mathbf{Y}$, where

$$\begin{aligned} \mathbf{A}^{(-1)} &= \sigma_0(\mathbf{D}) = \sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right), \\ \mathbf{B}^{(-1)} &= -\sigma_0(\mathbf{B}) = \sigma_0\left((\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\right)\mathbf{X}, \\ \mathbf{C}^{(-1)} &= -\sigma_0(\mathbf{C}) = (\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}\sigma_1(\mathbf{X}), \\ \mathbf{D}^{(-1)} &= \sigma_0(\mathbf{A}) = (\mathbf{I} - \sigma_1(\mathbf{X})\mathbf{X})^{-\frac{1}{2}}. \end{aligned}$$

This completes the proof.

□

C.8 Proof of Proposition 3.8

Proof. First, note that

$$\rho(\mathbf{Z}) = \begin{pmatrix} \mathbf{I} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{I} \end{pmatrix}\begin{pmatrix} \mathbf{Z}_2 & \mathbf{0} \\ \mathbf{0} & \overline{\mathbf{Z}}_2^{-1} \end{pmatrix}\begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_1) & \mathbf{I} \end{pmatrix}. \tag{9}$$

If $\mathbf{Z} \in \mathfrak{U}$, then $\mathbf{Z}_2 \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+$ and there exists $\mathbf{P} \in A_{\mathbb{C}}^{\times}$ such that $\mathbf{Z}_2 = \mathbf{P}\bar{\sigma}_{\mathbb{C}}(\mathbf{P})$. Thus

$$\begin{aligned}\rho(\mathbf{Z}) &= \begin{pmatrix} \mathbf{I} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{P}\bar{\sigma}_{\mathbb{C}}(\mathbf{P}) & \mathbf{0} \\ \mathbf{0} & \sigma_{\mathbb{C}}(\mathbf{P})^{-1}\bar{\sigma}_{\mathbb{C}}(\sigma_{\mathbb{C}}(\mathbf{P})^{-1}) \end{pmatrix} \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_1) & \mathbf{I} \end{pmatrix} \\ &= \mathbf{Q}\bar{\sigma}_{\mathbb{C}}(\mathbf{Q}),\end{aligned}$$

where $\mathbf{Q} = \begin{pmatrix} \mathbf{I} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{I} \end{pmatrix} \begin{pmatrix} \mathbf{P} & \mathbf{0} \\ \mathbf{0} & \sigma_{\mathbb{C}}(\mathbf{P})^{-1} \end{pmatrix} \in A_{\mathbb{C}}^{\times}$. Therefore, $\rho(\mathbf{Z}) \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+$.

If $\rho(\mathbf{Z}) \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+$, then there exists $\mathbf{P} \in A_{\mathbb{C}}^{\times}$ such that $\rho(\mathbf{Z}) = \mathbf{P}\bar{\sigma}_{\mathbb{C}}(\mathbf{P})$. By Eq. (9),

$$\begin{aligned}\begin{pmatrix} \mathbf{Z}_2 & \mathbf{0} \\ \mathbf{0} & \bar{\mathbf{Z}}_2^{-1} \end{pmatrix} &= \begin{pmatrix} \mathbf{I} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{I} \end{pmatrix}^{-1} \mathbf{P}\bar{\sigma}_{\mathbb{C}}(\mathbf{P}) \begin{pmatrix} \mathbf{I} & \mathbf{0} \\ \bar{\sigma}_{\mathbb{C}}(\mathbf{Z}_1) & \mathbf{I} \end{pmatrix}^{-1} \\ &= \mathbf{Q}\bar{\sigma}_{\mathbb{C}}(\mathbf{Q}),\end{aligned}$$

where $\mathbf{Q} = \begin{pmatrix} \mathbf{I} & \mathbf{Z}_1 \\ \mathbf{0} & \mathbf{I} \end{pmatrix}^{-1} \mathbf{P} \in A_{\mathbb{C}}^{\times}$. Let $\mathbf{Q} = \begin{pmatrix} \mathbf{A} & \mathbf{B} \\ \mathbf{C} & \mathbf{D} \end{pmatrix}$. Then $\mathbf{Z}_2 = \mathbf{A}\bar{\sigma}_{\mathbb{C}}(\mathbf{A}) + \mathbf{B}\bar{\sigma}_{\mathbb{C}}(\mathbf{B})$. Since $\mathbf{Z}_2 \in A_{\mathbb{C}}^{\times}$, we deduce that $\mathbf{Z}_2 \in (A_{\mathbb{C}}^{\bar{\sigma}_{\mathbb{C}}})_+$, which concludes the proposition.

□

C.9 Proof of Corollary 4.1

Proof. The result can be proved using the definition of the precompact model on $\mathbb{QB}_n$ and Proposition 3.6.

□

C.10 Proof of Corollary 4.2

Proof. The result can be proved using the definition of the precompact model on $\mathbb{QB}_n$ and Proposition 3.7.

□