

Quantum State Assignment Flows

Jonathan Schwarz ^{1,*}, Jonas Cassel ¹, Bastian Boll ¹ , Martin Gärtner ² , Peter Albers ³ and Christoph Schnörr ¹

¹ Image and Pattern Analysis Group, Institute for Mathematics, Heidelberg University, 69117 Heidelberg, Germany; cassel@math.uni-heidelberg.de (J.C.); bastian.boll@iwr.uni-heidelberg.de (B.B.); schnoerr@math.uni-heidelberg.de (C.S.)

² Physikalisches Institut, Kirchhoff Institute for Physics, Heidelberg University, 69117 Heidelberg, Germany; martin.gaertner@kip.uni-heidelberg.de

³ Research Station Geometry & Dynamics, Institute for Mathematics, Heidelberg University, 69117 Heidelberg, Germany; palbers@mathi.uni-heidelberg.de

* Correspondence: jonathan.schwarz@iwr.uni-heidelberg.de

Abstract: This paper introduces assignment flows for density matrices as state spaces for representation and analysis of data associated with vertices of an underlying weighted graph. Determining an assignment flow by geometric integration of the defining dynamical system causes an interaction of the non-commuting states across the graph, and the assignment of a pure (rank-one) state to each vertex after convergence. Adopting the Riemannian–Bogoliubov–Kubo–Mori metric from information geometry leads to closed-form local expressions that can be computed efficiently and implemented in a fine-grained parallel manner. Restriction to the submanifold of commuting density matrices recovers the assignment flows for categorical probability distributions, which merely assign labels from a finite set to each data point. As shown for these flows in our prior work, the novel class of quantum state assignment flows can also be characterized as Riemannian gradient flows with respect to a non-local, non-convex potential after proper reparameterization and under mild conditions on the underlying weight function. This weight function generates the parameters of the layers of a neural network corresponding to and generated by each step of the geometric integration scheme. Numerical results indicate and illustrate the potential of the novel approach for data representation and analysis, including the representation of correlations of data across the graph by entanglement and tensorization.



Citation: Schwarz, J.; Cassel, J.; Boll, B.; Gärtner, M.; Albers, P.; Schnörr, C. Quantum State Assignment Flows. *Entropy* **2023**, *25*, 1253. <https://doi.org/10.3390/e25091253>

Academic Editor: Antonio M. Scarfone

Received: 3 July 2023

Revised: 9 August 2023

Accepted: 14 August 2023

Published: 23 August 2023



Copyright: © 2023 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

Keywords: assignment flows; Riemannian gradient flows; density matrix; information geometry

MSC: 53B12; 62H35; 68T07

1. Introduction

1.1. Overview and Motivation

A basic task of data analysis is categorization of observed data. We consider the following scenario: On a given undirected, weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$, data $D_i \in \mathcal{X}$ are observed as points in a metric space $(\mathcal{X}, d_{\mathcal{X}})$ at each vertex $i \in \mathcal{V}$. Categorization means to determine an assignment

$$D_i \rightarrow j \in \{1, \dots, c\} =: [c] \quad (1)$$

of a *class label* j among a *finite set* of labels to each data point D_i . Depending on the application, labels carry a specific meaning, e.g., type of tissue in medical imaging data, object type in computer vision or land use in remote sensing data. The decision at any vertex typically depends on decisions at other vertices. Thus, the overall task of labeling data on a graph constitutes a particular form of *structured prediction* in the field of machine learning [1].

Assignment flows denote a particular class of approaches for data labeling on graphs [2,3]. The basic idea is to represent each possible label assignment at vertex $i \in \mathcal{V}$ by an *assignment*

vector $S_i \in \Delta_c$ in the standard probability simplex, the vertices of which encode the unique label assignment for every label by the corresponding unit vector e_j , $j \in [c]$. Data labeling is accomplished by computing the flow $S(t)$ of the dynamical system:

$$\dot{S} = R_S[\Omega S], \quad S(0) = S_0, \quad (2)$$

with the row-stochastic matrix $S(t)$ and row vectors $S_i(t)$ as the state, which, under mild conditions, converges to unique label assignment vectors (unit vectors) at every vertex $i \in \mathcal{V}$ [4]. The vector field on the right-hand side in Equation (2) is parameterized by parameters collected in a matrix Ω . These parameters strongly affect the contextual label assignments. They can be learned from data in order to take into account typical relations of data in the current field of application [5]. For a demonstration of the application of this approach to a challenging medical imaging problem, we refer to [6].

From a geometric viewpoint, system (2) can be characterized as a collection of individual flows $S_i(t)$ at each vertex that are *coupled* by the Ω parameters. Each individual flow is determined by a *replicator equation*, which constitutes a basic class of dynamical systems known from evolutionary game theory [7,8]. By restricting each vector $S_i(t)$ to the relative interior $\mathring{\Delta}_c$ of the probability simplex (i.e., the set of strictly positive discrete probability vectors) and by turning this convex set into a statistical manifold equipped with Fisher–Rao geometry [9], the assignment flow (2) becomes a Riemannian ascent flow on the corresponding product manifold. The underlying information geometry is not only important for making the flow converge to unique label assignments but also for the design of efficient algorithms that actually determine the assignments [10]. For extensions of the basic assignment flow approach to unsupervised scenarios of machine learning and for an in-depth discussion of connections to other closely related work on structured prediction on graphs, we refer to [11–13].

In this paper, we study a novel and substantial generalization of assignment flows from a different point of view: assignment of labels to metric data, where the labels are elements of a *continuous* set. This requires replacement of the simplex Δ_c as state space, which can only represent assignments of labels from a *finite* set. The substitutes for assignment vectors S_i , $i \in \mathcal{V}$ are Hermitian positive definite *density matrices* ρ_i , $i \in \mathcal{V}$ with unit trace:

$$\mathcal{D}_c = \{\rho \in \mathbb{C}^{c \times c} : \rho = \rho^*, \text{tr} \rho = 1\}. \quad (3)$$

Accordingly, the finite set of unit vectors e_j , $j \in [c]$ (vertices of Δ_c are replaced by *rank-one* density matrices ρ^∞ , also known as *pure states* in quantum mechanics [14]. The resulting *quantum state assignment flow* (QSAF)

$$\dot{\rho} = \mathfrak{R}_\rho[\Omega[\rho]], \quad \rho(0) = \rho_0, \quad (4)$$

consists of a system of nonlinear first-order differential equations, whose solution $\rho(t)$ evolves on a corresponding product of state spaces \mathcal{D}_c as given by Equation (3), with a linear averaging operator $\Omega[\cdot]$ and a generalized replicator operator $\mathfrak{R}_\rho[\cdot]$ that is linear with respect to the $\Omega[\rho]$ argument and nonlinear with respect to ρ (cf. Equation (64)). The similarity of Equations (4) and (2) can be attributed to the common underlying design strategy. System (4) couples the individual evolutions $\rho_i(t)$ at each vertex $i \in \mathcal{V}$ through Ω parameters, and the underlying information geometry causes convergence of each $\rho_i(t)$ towards a pure state. Using a different state space \mathcal{D}_c (rather than $\mathring{\Delta}_c$ in Equation (2)) requires the adoption of a different Riemannian metric, which results in a corresponding definition of the operator \mathfrak{R}_ρ .

Our approach is natural in that restricting Equation (4) to *diagonal* density matrices results in Equation (2) after identifying each vector $\text{diag}(\rho_i)$ of diagonal entries of the density matrix ρ_i with an assignment vector $S_i \in \mathring{\Delta}_c$. Conversely, Equation (4) considerably generalizes Equation (2) and enhances modeling expressivity due to the *non-commutative* interaction of

the state spaces $\rho_i, i \in \mathcal{V}$ across the underlying graph \mathcal{G} when the quantum state assignment flow is computed by applying geometric numerical integration to Equation (4).

We regard our approach merely as an *approach to data representation and analysis* rather than a contribution to quantum mechanics. For example, dynamics Equation (4) clearly differs from the Hamiltonian evolution of quantum systems, yet we adopt the term “quantum state”, since not only density matrices as state spaces but also the related information geometry are largely motivated by quantum mechanics and quantum information theory [9,15].

1.2. Contribution and Organization

Section 2 summarizes the information geometry of both the statistical manifold of categorical distributions and the manifold of strictly positive definite density matrices. Section 3 summarizes the assignment flow approach (2) as a reference for the subsequent generalization to Equation (4). This generalization is the main contribution of this paper and is presented in Section 4. Each row of Table 1 specifies the section where an increasingly general version of the original assignment flow (left column) is generalized to the corresponding quantum state assignment flow (right column, same row).

Table 1. Components of the the *Assignment Flow* approach and the corresponding components of the novel *Quantum State Assignment Flow* approach.

Assignment Flow (AF)	Quantum State AF (QSAF)
Single-vertex AF (Section 3.1)	Single-vertex QSAF (Section 4.2)
AF approach (Section 3.2)	QSAF approach (Section 4.3)
Riemannian gradient AF (Section 3.3)	Riemannian gradient QSAF (Section 4.4)
Recovery of the AF from the QSAF by restriction (Section 4.5)	

Alternative metrics on the positive definite matrix manifold that have been used in the literature are reviewed in Section 2.3 in order to position our approach from this point of view. In Section 5, we describe some academic experiments that we conducted to illustrate the properties of the novel approach. Working out a particular scenario of data analysis is beyond the scope of this paper. We conclude and indicate directions of further work in Section 6. For ease of readin, proofs are listed in Appendix A.

This paper considerably elaborates the short preliminary conference version [16].

1.3. Basic Notation

For the reader’s convenience, below, we specify the basic notational conventions used in this paper.

$[c]$	$\{1, 2, \dots, c\}, c \in \mathbb{N}$
$\mathbf{1}_c$	$(1, 1, \dots, 1)^\top \in \mathbb{R}^c$
\mathbb{R}_+^c	$\{x \in \mathbb{R}^c: x_i \geq 0, i \in [c]\}$
\mathbb{R}_{++}^c	$\{x \in \mathbb{R}^c: x_i > 0, i \in [c]\}$
e_1, e_2, \dots	Canonical basis vectors of \mathbb{R}^c
$\langle u, v \rangle$	Euclidean inner vector product
$\ u\ $	Euclidean norm $\sqrt{\langle u, u \rangle}$
I_c	Unit matrix of $\mathbb{R}^{c \times c}$
$p \cdot q$	Component-wise vector multiplication $(p \cdot q)_i = p_i q_i, i \in [c], p, q \in \mathbb{R}^c$
$\frac{q}{p}$	Component-wise division $(\frac{q}{p})_i = \frac{q_i}{p_i}, i \in [c], q \in \mathbb{R}^c, p \in \mathbb{R}_{++}^c$
\mathcal{H}_c	Space of Hermitian $c \times c$ matrices (cf. (22))
$\text{tr}(A)$	Trace $\sum_i A_{ii}$ of matrix A
$\langle A, B \rangle$	Matrix inner product $\text{tr}(AB), A, B \in \mathcal{H}_c$
$[A, B]$	Commutator $AB - BA$
$\text{Diag}(v)$	The diagonal matrix with vector v as entries

$\text{diag}(V)$	The vector of the diagonal entries of a square matrix V
\exp_m	The matrix exponential
\log_m	The matrix logarithm \exp_m^{-1}
Δ_c	The set of discrete probability vectors of dimension c (cf. (6))
\mathcal{S}_c	The relative interior of Δ_c , i.e., the set of strictly positive probability vectors (cf. (7))
\mathcal{W}_c	The product manifold $\mathcal{S}_c \times \dots \times \mathcal{S}_c$ (cf. (53))
\mathcal{P}_c	The set of symmetric positive definite $c \times c$ matrices (cf. (17))
\mathcal{D}_c	The subset of matrices in \mathcal{P}_c whose trace is equal to 1 (cf. (18))
\mathcal{Q}_c	The product manifold $\mathcal{D}_c \times \dots \times \mathcal{D}_c$ (cf. (96))
$\mathbb{1}_{\mathcal{S}_c}$	Barycenter $\frac{1}{c}\mathbb{1}_c$ of the manifold \mathcal{S}_c
$\mathbb{1}_{\mathcal{W}_c}$	Barycenter $(\mathbb{1}_{\mathcal{S}_c}, \mathbb{1}_{\mathcal{S}_c}, \dots, \mathbb{1}_{\mathcal{S}_c})^\top$ of the manifold \mathcal{W}
$\mathbb{1}_{\mathcal{D}_c}$	Matrix $\text{Diag}(\mathbb{1}_{\mathcal{S}_c}) \in \mathcal{D}_c \subset \mathbb{C}^{c \times c}$
g_p, g_W, g_ρ	The Riemannian metrics on $\mathcal{S}_c, \mathcal{W}_c, \mathcal{D}_c$ (cf. (8), (54), (25))
$T_{c,0}, \mathcal{T}_{c,0}, \mathcal{H}_{c,0}$	The tangent spaces to $\mathcal{S}_c, \mathcal{W}_c, \mathcal{D}_c$ (cf. (10), (54), (21))
$\pi_{c,0}, \Pi_{c,0}$	Orthogonal projections onto $T_0, \mathcal{H}_{c,0}$ (cf. (11), (24))
$R_p, R_W, \mathfrak{R}_\rho$	Replicator operators associated with the assignment flows on $\mathcal{S}_c, \mathcal{W}_c, \mathcal{D}_c, \mathcal{Q}_c$ (cf. (12), (58), (64), (106))
∂	Euclidean gradient operator: $\partial f(p) = (\partial_{p_1} f(p), \partial_{p_2} f(p), \dots)^\top$
grad	Riemannian gradient operator with respect to the Fisher–Rao metric
$R_W[\cdot], \Omega[\cdot], \text{etc.}$	Square brackets indicate a linear operator that acts in a non-standard way, e.g., row-wise to a matrix argument.

2. Information Geometry

Information geometry [17,18] is concerned with the representation of parametric probability distributions from a geometric viewpoint, e.g., the exponential family of distributions [19]. Specifically, an open convex set \mathcal{M} of parameters of a probability distribution becomes a Riemannian manifold (\mathcal{M}, g) when equipped with a Riemannian metric g . The Fisher–Rao metric is the canonical choice due to its invariance properties with respect to reparameterization [20]. A closely related scenario concerns the representation of the interior of compact convex bodies as Riemannian manifolds (\mathcal{M}, g) due to the correspondence between compactly supported Borel probability measures and an affine equivalence class of convex bodies [21].

A key ingredient of information geometry is the so-called α -family of affine connections introduced by Amari [17], which comprises the so-called e -connection ∇ and m -connection ∇^* as special cases. These connections are torsion-free and dual to each other in the sense that they jointly satisfy the equation that uniquely characterizes the Levi–Civita connection as a metric connection [17] (Definition 3.1, Theorem 3.1). Regarding numerical computations, working with the exponential map induced by the e connection is particularly convenient, since its domain is the entire tangent space. We refer to [9,22,23] for further reading and to [24] and [9] (Chapter 7) for the specific case of quantum state spaces.

In this paper, we are concerned with two classes of convex sets:

- The relative interior of probability simplices, each of which represents the categorical (discrete) distributions of the corresponding dimension; and
- The set of positive definite symmetric matrices with trace one.

Sections 2.1 and 2.2 introduce the information geometry for the former and the latter class of sets, respectively.

2.1. Categorical Distributions

We set

$$[c] := \{1, 2, \dots, c\}, \quad c \in \mathbb{N}. \tag{5}$$

and denote the probability simplex of distributions on $[c]$ as

$$\Delta_c := \left\{ p \in \mathbb{R}_+^c : \langle \mathbf{1}_c, p \rangle = \sum_{i \in [c]} p_i = 1 \right\}, \quad \mathbf{1}_c := (1, 1, \dots, 1)^\top \in \mathbb{R}^c. \tag{6}$$

Its relative interior equipped with the Fisher–Rao metric becomes the Riemannian manifold (\mathcal{S}_c, g) ,

$$\mathcal{S}_c := \text{rint } \Delta_c = \{ p \in \Delta_c : p_i > 0, i \in [c] \}, \tag{7}$$

$$g_p(u, v) := \sum_{i \in [c]} \frac{u_i v_i}{p_i} = \langle u, \text{Diag}(p)^{-1} v \rangle, \quad \forall u, v \in T_{c,0}, \quad p \in \mathcal{S}_c, \tag{8}$$

with the trivial tangent bundle given by

$$T\mathcal{S}_c \cong \mathcal{S}_c \times T_{c,0} \tag{9}$$

and the tangent space

$$T_{c,0} := T_{\mathbf{1}_{\mathcal{S}_c}} \mathcal{S}_c = \{ v \in \mathbb{R}^c : \langle \mathbf{1}_c, v \rangle = 0 \}. \tag{10}$$

The orthogonal projection onto $T_{c,0}$ is denoted by

$$\pi_{c,0}: \mathbb{R}^c \rightarrow T_{c,0}, \quad \pi_{c,0} v := v - \frac{1}{c} \langle \mathbf{1}_c, v \rangle \mathbf{1}_c = \left(I_c - \frac{1}{c} \mathbf{1}_c \mathbf{1}_{\mathcal{S}_c}^\top \right) v. \tag{11}$$

The mapping defined next plays a major role in all dynamical systems under consideration in this paper.

Definition 1 (replicator operator). *The replicator operator is the linear mapping of the tangent space*

$$R: \mathcal{S}_c \times T_{c,0} \rightarrow T_{c,0}, \quad R_p v := (\text{Diag}(p) - p p^\top) v, \quad p \in \mathcal{S}_c, \quad v \in T_{c,0} \tag{12}$$

parameterized by $p \in \mathcal{S}_c$.

The name ‘replicator’ is due to the role of this mapping in evolutionary game theory; see Remark 2 on page 9.

Proposition 1 (properties of R_p). *Mapping (12) satisfies*

$$R_p \mathbf{1}_c = 0, \tag{13}$$

$$\pi_{c,0} R_p = R_p \pi_{c,0} = R_p, \quad \forall p \in \mathcal{S}_c. \tag{14}$$

Furthermore, let $f: \mathcal{S}_c \rightarrow \mathbb{R}$ be a smooth function and $\tilde{f}: U \rightarrow \mathbb{R}$ be a smooth extension of f to an open neighborhood U of $\mathcal{S}_c \subset \mathbb{R}^c$ with $\tilde{f}|_{\mathcal{S}_c} = f$. Then the Riemannian gradient of f with respect to the Fisher–Rao metric (8) is given by

$$\text{grad } f(p) = R_p \tilde{\partial} f(p). \tag{15}$$

Proof. Appendix A.1 \square

Remark 1. Equations (15) and (A24) show that the replicator operator R_p is the inverse metric tensor with respect to the Fisher–Rao metric (8), as expressed in the ambient coordinates.

The exponential map induced by the e connection is defined on the entire space $T_{c,0}$ and reads [22]

$$\text{Exp}: \mathcal{S}_c \times T_{c,0} \rightarrow \mathcal{S}_c, \quad \text{Exp}_p(v) := \frac{p \cdot e^{\frac{v}{p}}}{\langle p, e^{\frac{v}{p}} \rangle}, \quad p \in \mathcal{S}_c, \quad v \in T_{c,0}. \quad (16)$$

2.2. Density Matrices

We denote the open convex cone of positive definite matrices by

$$\mathcal{P}_c := \{\rho \in \mathbb{C}^{c \times c}: \rho = \rho^*, \rho \succ 0\} \quad (17)$$

and the manifold of strictly positive definite density matrices by

$$\mathcal{D}_c := \{\rho \in \mathcal{P}_c: \text{tr} \rho = 1\}. \quad (18)$$

where \mathcal{D}_c is the intersection of \mathcal{P}_c and the hyperplane defined by the trace-one constraint. Its closure $\overline{\mathcal{D}}_c$ is convex and compact. We can identify the space \mathcal{D}_c as the space of invertible density operators in the sense of quantum mechanics on the finite-dimensional Hilbert space \mathbb{C}^c without loss of generality. Any matrix ensemble of the form

$$\{M_i\}_{i \in [n]} \subset \overline{\mathcal{P}}_c: \sum_{i \in [n]} M_i = I_c \quad (19)$$

induces the probability distribution on $[n]$ via the Born rule

$$p \in \Delta_n: \quad p_i = \langle M_i, \rho \rangle = \text{tr}(M_i \rho), \quad i \in [n]. \quad (20)$$

Equation (19) is called the *positive operator valued measure (POVM)*. We refer to [14] for the physical background and to [25] and references therein for the mathematical background.

The analog of Equation (10) is the tangent space in which at any point, $\rho \in \mathcal{D}_c$ is equal to the space of traceless symmetric matrices.

$$\mathcal{H}_{c,0} := \mathcal{H}_c \cap \{X \in \mathbb{C}^{c \times c}: \text{tr} X = 0\}, \quad (21)$$

where

$$\mathcal{H}_c := \{X \in \mathbb{C}^{c \times c}: X^* = X\}. \quad (22)$$

Therefore, the manifold \mathcal{D}_c has a trivial tangent bundle given by

$$T\mathcal{D}_c = \mathcal{D}_c \times \mathcal{H}_{c,0}, \quad (23)$$

with the tangent space $\mathcal{H}_{c,0} = T_{1_{\mathcal{D}_c}} \mathcal{D}_c$ defined in Equation (21). The corresponding orthogonal projection onto the tangent space $\mathcal{H}_{c,0}$ reads

$$\Pi_{c,0}: \mathcal{H}_c \rightarrow \mathcal{H}_{c,0}, \quad \Pi_{c,0}[X] := X - \frac{\text{tr} X}{c} I_c. \quad (24)$$

Equipping the manifold \mathcal{D}_c as defined in Equation (18) with the *Bogoliubov–Kubo–Mori (BKM) metric* [26] results in a Riemannian manifold (\mathcal{D}_c, g) . Using $T_\rho \mathcal{D}_c = \mathcal{H}_{c,0}$, this metric can be expressed by

$$g_\rho(X, Y) := \int_0^\infty \text{tr} (X(\rho + \lambda I)^{-1} Y(\rho + \lambda I)^{-1}) d\lambda, \quad X, Y \in \mathcal{H}_{c,0}, \quad \rho \in \mathcal{D}_c. \quad (25)$$

This metric uniquely ensures the existence of a symmetric e-connection ∇ on \mathcal{D}_c that is mutually dual to its m-connection ∇^* in the sense of information geometry, leading to a *dually flat* structure (g, ∇, ∇^*) [27] and [9] (Theorem 7.1).

The following map and its inverse, defined in terms of the matrix exponential \exp_m and its inverse $\log_m = \exp_m^{-1}$, are convenient.

$$\mathbb{T}: \mathcal{D}_c \times \mathcal{H}_c \rightarrow \mathcal{H}_c, \tag{26}$$

$$\mathbb{T}_\rho[X] := \frac{d}{dt} \log_m(\rho + tX)|_{t=0} = \int_0^\infty (\rho + \lambda I)^{-1} X (\rho + \lambda I)^{-1} d\lambda, \tag{27}$$

$$\mathbb{T}_\rho^{-1}[X] = \frac{d}{dt} \exp_m(H + tX)|_{t=0} = \int_0^1 \rho^{1-\lambda} X \rho^\lambda d\lambda, \quad \rho = \exp_m(H). \tag{28}$$

The inner product (25) may now be written in the form of

$$g_\rho(X, Y) = \langle \mathbb{T}_\rho[X], Y \rangle, \tag{29}$$

since the trace is invariant with respect to cyclic permutations of a matrix product as an argument. Likewise,

$$\langle \rho, X \rangle = \text{tr}(\rho X) = \text{tr} \mathbb{T}_\rho^{-1}[X]. \tag{30}$$

We also consider two subspaces on the tangent space $T_\rho \mathcal{D}_c$,

$$T_\rho^u \mathcal{D}_c := \{X \in \mathcal{H}_{c,0}: \exists \Omega = -\Omega^* \text{ such that } X = [\Omega, \rho]\}, \tag{31}$$

$$T_\rho^c \mathcal{D}_c := \{X \in \mathcal{H}_{c,0}: [\rho, X] = 0\}, \tag{32}$$

which yield the decomposition [9]

$$T_\rho \mathcal{D}_c = T_\rho^c \mathcal{D}_c \oplus T_\rho^u \mathcal{D}_c. \tag{33}$$

In Section 4.5, we use this decomposition to recover the assignment flow for categorical distributions from the quantum state assignment flow by restriction to a submanifold of commuting matrices.

2.3. Alternative Metrics and Geometries

The positive definite matrix manifold \mathcal{P}_c (We confine ourselves in this subsection to the case of real density matrices, as our main references for comparison only deal with real matrix manifolds) has become a tool for data modelling and analysis during the last two decades. Accordingly, a range of Riemannian metrics exist with varying properties. A major subclass is formed by the $O(n)$ -invariant metrics, including the log-Euclidean, affine-invariant, Bures–Wasserstein and Bogoliubov–Kubo–Mori (BKM) metrics. We refer to [28] for a comprehensive recent survey.

This section provides a brief comparison of the *BKM metric* (25) adopted in this paper with two metrics often employed in the literature: the *affine-invariant metric* and the *log-Euclidean metric*, which may be regarded as ‘antipodal points’ in the space of metrics from the geometric and the computational viewpoint, respectively.

2.3.1. Affine-Invariant Metrics

The affine-invariant metric has been derived in various ways, e.g., based on the canonical matrix inner product on the tangent space [29] (Section 6) or as Fisher–Rao metric on the statistical manifold of centered multivariate Gaussian densities [30]. The metric is given by

$$g_\rho(X, Y) = \text{tr}(\rho^{-\frac{1}{2}} X \rho^{-\frac{1}{2}} \rho^{-\frac{1}{2}} Y \rho^{-\frac{1}{2}}) = \text{tr}(\rho^{-1} X \rho^{-1} Y), \quad \rho \in \mathcal{P}_c, \quad X, Y \in T_\rho \mathcal{P}_c. \tag{34}$$

The exponential map with respect to the Levi–Civita connection reads

$$\exp_\rho^{(\text{aff})}(X) = \rho^{\frac{1}{2}} \exp_m(\rho^{-\frac{1}{2}} X \rho^{-\frac{1}{2}}) \rho^{\frac{1}{2}}, \quad \rho \in \mathcal{P}_c, \quad X \in T_\rho \mathcal{P}_c. \tag{35}$$

This Riemannian structure turns \mathcal{P}_c into a manifold with negative sectional curvature [31] (Chapter II.10), which is convenient from the geometric viewpoint due to uniquely defined Riemannian means and geodesic convexity [32] (Section 6.9). On the other hand, evaluating Equations (34) and (35) is computationally expensive, in particular when computing the quantum state assignment flow, which essentially involves geometric averaging.

2.3.2. Log-Euclidean Metric

The log-Euclidean metric introduced in [33] is the pullback of the canonical matrix inner product with respect to the matrix logarithm and is given by

$$g_\rho(X, Y) = \text{tr}(d \log_m(\rho)[X], d \log_m(\rho)[Y]) \stackrel{(27)}{=} \langle \mathbb{T}_\rho[X], \mathbb{T}_\rho[Y] \rangle, \quad \rho \in \mathcal{P}_c \quad X, Y \in T_\rho \mathcal{P}_c. \tag{36}$$

The exponential map reads

$$\exp_\rho^{(\log)}(X) = \exp_m(\log_m(\rho) + \mathbb{T}_\rho[X]), \quad \rho \in \mathcal{P}_c \quad X, Y \in T_\rho \mathcal{P}_c \tag{37}$$

and is much more convenient from the computational viewpoint. Endowed with this metric, the space \mathcal{P}_c is isometric to a Euclidean space. The log-Euclidean metric is not curved and merely invariant under orthogonal transforms and dilations [28].

2.3.3. Comparison to Bogoliubov-Kubo-Mori Metric

The BKM metric (Equations (25) and (30)) given by

$$g_\rho(X, Y) = \langle \mathbb{T}_\rho[X], Y \rangle, \quad \rho \in \mathcal{P}_c \quad X, Y \in T_\rho \mathcal{P}_c, \tag{38}$$

looks similar to the log-Euclidean metric (36). Regarding them both as members of the class of *mean kernel metrics* [28] (Definition 4.1) enables an intuitive comparison. For real-valued matrices, mean kernel metrics have the form of

$$g_\rho(X, X) = g_D(X', X') = \sum_{i,j \in [c]} \frac{(X'_{ij})^2}{\phi(D_{ii}, D_{jj})}, \quad \rho = V D V^\top, \quad V \in O(n), \quad X = V X' V^\top, \tag{39}$$

with a diagonal matrix $D = \text{Diag}(D_{11}, \dots, D_{cc})$ and a bivariate function $\phi(x, y) = a m(x, y)^\theta$, $a > 0$ in terms of a symmetric homogeneous mean $m: \mathbb{R}_+ \times \mathbb{R}_+ \rightarrow \mathbb{R}_+$. Regarding the log-Euclidean metric, $\phi(x, y) = \left(\frac{x-y}{\log x - \log y}\right)^2$, whereas for the BKM metric, $\phi(x, y) = \frac{x-y}{\log x - \log y}$.

Taking the restriction to density matrices $\mathcal{D}_c \subset \mathcal{P}_c$ into account, one has the relation

$$\exp_\rho^{(\log)}(Y) = \text{Exp}_\rho^{(e)}(X), \quad \rho \in \mathcal{D}_c, \quad X \in \mathcal{H}_{c,0}, \tag{40}$$

$$Y = X - \log \left(\text{tr} \exp_m(\log_m(\rho) + \mathbb{T}_\rho[X]) \right) \rho, \tag{41}$$

as explained in Remark 4. Here, the left-hand side of Equation (40) is the exponential map (37) induced by the log-Euclidean metric, and $\text{Exp}_\rho^{(e)}$ is the exponential map with respect to the affine e connection of information geometry, as detailed below by Proposition 4. This close relationship of the e-exponential map $\text{Exp}_\rho^{(e)}$ with the exponential map of the log-Euclidean metric highlights the computational efficiency of using the BKM metric, which we adopt for our approach. This is also motivated by the lack of an explicit formula for the exponential map with respect to the Levi-Civita connection [34]. To date, the sign of the curvature remains unknown.

We note that to the best of our knowledge, the introduction of the affine connections of information geometry as surrogates of the Riemannian connection for any statistical manifold predates the introduction of the log-Euclidean metric for the specific space \mathcal{P}_c .

3. Assignment Flows

The assignment flow approach was informally introduced in Section 1. In this section, we summarize the mathematical ingredients of this approach as a reference for the subsequent generalization to quantum states (density matrices) in Section 4. Sections 3.1 and 3.2 introduce the assignment flow on a single vertex and on an arbitrary graph, respectively. A reparameterization turns the latter into a Riemannian gradient flow (Section 3.3). Throughout this section, we refer to definitions and notions introduced in Section 2.1.

3.1. Single-Vertex Assignment Flow

Let $D = (D_1, \dots, D_c)^\top \in \mathbb{R}^c$ and consider the task of picking the smallest components of D . Formulating this operation as an optimization problem amounts to evaluating the negative support function, in the sense of convex analysis [35] (p. 28), of the probability simplex Δ_c at $-D$,

$$\min_{j \in [c]} \{D_1, \dots, D_c\} = - \max_{p \in \Delta_c} \langle -D, p \rangle. \tag{42}$$

In practice, the vector D represents real-valued noisy measurements at some vertex $i \in \mathcal{V}$ of an underlying graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and is therefore in a “general position”, that is, the minimal component is unique. If $j^* \in [c]$ indexes the minimal component D_{j^*} , then the corresponding unit vector $p^* = e_{j^*}$ maximizes the right-hand side of (42). *Assignment vectors* assign a label (index) to observed data vectors.

If D varies, the operation (42) is non-smooth. In view of a desired interaction of label assignments across the graph (cf. Section 3.2), we therefore replace this operation by a *smooth* dynamical system whose solution converges to the desired assignment vector. To this end, the vector D is represented on \mathcal{S}_c as a *likelihood vector*

$$L_p(D) := \exp_p(-\pi_{c,0}D) \stackrel{(14)}{=} \exp_p(-D), \quad p \in \mathcal{S}_c, \tag{43}$$

where

$$\exp: \mathcal{S}_c \times T_{c,0} \rightarrow \mathcal{S}_c, \quad \exp_p(v) := \text{Exp}_p \circ R_p(v) = \frac{p \cdot e^v}{\langle p, e^v \rangle}, \quad p \in \mathcal{S}_c. \tag{44}$$

The *single-vertex assignment flow* equation reads

$$\dot{p} = R_p L_p(D) = p \cdot (L_p(D) - \langle p, L_p(D) \rangle \mathbb{1}_c), \quad p(0) = \mathbb{1}_{\mathcal{S}_c}. \tag{45}$$

Its solution $p(t)$ converges to the vector that solves the label assignment problem (42) (see Corollary 1 below).

Remark 2 (replicator equation). *Differential equations of the form (45), with some \mathbb{R}^c -valued function $F(p)$ in place of $L_p(D)$, are known as replicator equation in evolutionary game theory [7].*

Lemma 1. *Let $p \in \mathcal{S}_c$. Then, the differentials of the mapping (44) with respect to p and v are given by*

$$d_v \exp_p(v)[u] = R_{\exp_p(v)} u, \tag{46}$$

$$d_p \exp_p(v)[u] = R_{\exp_p(v)} \frac{u}{p}, \quad p \in \mathcal{S}_c, \quad u, v \in T_{c,0}. \tag{47}$$

Proof. Appendix A.2. \square

Theorem 1 (single-vertex assignment flow). *The single-vertex assignment flow Equation (45) is equivalent to the system*

$$\dot{p} = R_p q, \quad p(0) = \mathbb{1}_{\mathcal{S}_c}, \tag{48}$$

$$\dot{q} = R_q q, \quad q(0) = L_{\mathbb{1}_{\mathcal{S}_c}}(D), \tag{49}$$

with solution given by

$$p(t) = \exp_{\mathbb{1}_{\mathcal{S}_c}} \left(\int_0^t q(\tau) d\tau \right). \tag{50}$$

Proof. Appendix A.2. \square

Corollary 1 (single-vertex label assignment). Let $\mathcal{J}^* := \arg \min_{j \in [c]} \{D_j : j \in [c]\} \subseteq [c]$. Then, the solution $p(t)$ to (45) satisfies

$$\lim_{t \rightarrow \infty} p(t) = \frac{1}{|\mathcal{J}^*|} \sum_{j \in \mathcal{J}^*} e_j \in \arg \max_{p \in \Delta_c} \langle -D, p \rangle. \tag{51}$$

In particular, if D has a unique minimal component D_{j^*} , then $p(t) \rightarrow e_{j^*}$ as $t \rightarrow \infty$.

Proof. Appendix A.2. \square

3.2. Assignment Flows

The assignment flow approach consists of the weighted interaction—as defined below—of single-vertex assignment flows associated with vertices $i \in \mathcal{V}$ of a weighted graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$ with a non-negative weight function

$$\omega : \mathcal{E} \rightarrow \mathbb{R}_+, \quad ik \mapsto \omega_{ik}. \tag{52}$$

The assignment vectors are denoted by $W_i, i \in \mathcal{V}$ and form the row vectors of a row-stochastic matrix.

$$W \in \mathcal{W}_c := \underbrace{\mathcal{S}_c \times \cdots \times \mathcal{S}_c}_{|\mathcal{V}| \text{ factors}}. \tag{53}$$

The product space \mathcal{W}_c is called the *assignment manifold* (\mathcal{W}_c, g) , where the metric g is defined by applying (8) row-wise,

$$g_W(U, V) := \sum_{i \in \mathcal{V}} g_{W_i}(U_i, V_i), \quad U, V \in \mathcal{T}_{c,0} := T_{c,0} \times \cdots \times T_{c,0}. \tag{54}$$

The *assignment flow equation* generalizing (45) reads

$$\dot{W} = R_W[S(W)], \tag{55}$$

where the *similarity vectors*

$$S_i(W) := \text{Exp}_{W_i} \left(\sum_{k \in \mathcal{N}_i} \omega_{ik} \text{Exp}_{W_i}^{-1} (L_{W_k}(D_k)) \right), \quad i \in \mathcal{V} \tag{56}$$

form the row vectors of the matrix $S(W) \in \mathcal{W}_c$. The neighborhoods

$$\mathcal{N}_i := \{i\} \cup \{k \in \mathcal{V} : ik \in \mathcal{E}\} \tag{57}$$

are defined by the adjacency relation of the underlying graph \mathcal{G} , and $R_W[\cdot]$ of Equation (55) applies Equation (12) row-wise,

$$R_W[S(W)]_i = R_{W_i} S_i(W), \quad i \in \mathcal{V}. \tag{58}$$

Note that the similarity vectors $S_i(W)$ given by (56) result from geometric weighted averaging of the velocity vectors $\text{Exp}_{W_i}^{-1} (L_{W_k}(D_k))$. The velocities represent given data $D_i, i \in \mathcal{V}$ via the likelihood vectors $L_{W_i}(D_i)$ given by (43). Each choice of the weights ω_{ik} in (56) associated with every edge $ik \in \mathcal{E}$ defines an assignment flow $W(t)$ solving (55). Thus,

these weight parameters determine how individual label assignments by (43) and (45) are regularized.

Well-posedness, stability and quantitative estimates of basins of attraction to integral label assignment vectors were established in [4]. Reliable and efficient algorithms for numerical computation of the assignment flow were devised in [10].

3.3. Reparameterized Assignment Flows

In [36] (Proposition 3.6), the following parameterization of the general assignment flow Equation (55) was introduced, which generalizes the parameterization (48) and (49) of the single-vertex assignment flow (45).

$$\dot{W} = R_W[\bar{S}], \quad W(0) = \mathbb{1}_{\mathcal{W}_c}, \tag{59}$$

$$\dot{\bar{S}} = R_{\bar{S}}[\Omega \bar{S}], \quad \bar{S}(0) = S(\mathbb{1}_{\mathcal{W}_c}), \tag{60}$$

with the non-negative weight matrix corresponding to the weight function (52),

$$\Omega = (\Omega_1, \dots, \Omega_{|\mathcal{V}|})^\top \in \mathbb{R}^{|\mathcal{V}| \times |\mathcal{V}|}, \quad \Omega_{ik} := \begin{cases} \omega_{ik}, & \text{if } k \in \mathcal{N}_i, \\ 0, & \text{otherwise.} \end{cases} \tag{61}$$

In terms of (60), this formulation reveals the “essential” part of the assignment flow equation, since (59) depends on (60) but not vice versa. Furthermore, the data and weights show up only in the initial point and in the vector field on the right-hand side of (60), respectively.

Henceforth, we solely focus on (60) rewritten for convenience as

$$\dot{S} = R_S[\Omega S], \quad S(0) = S_0, \tag{62}$$

where S_0 comprises the similarity vectors (56) evaluated at the barycenter $W = \mathbb{1}_{\mathcal{W}_c}$.

4. Quantum State Assignment Flows

In this section, we generalize the assignment flow Equations (55) and (62) to the product manifold \mathcal{Q}_c of density matrices as state space. The resulting equations have a similar mathematical form. Their derivation requires:

- Determination of the form of the Riemannian gradient of functions $f: \mathcal{D}_c \rightarrow \mathbb{R}$ with respect to the BKM metric (25), the corresponding replicator operator and exponential mappings Exp and exp , together with their differentials (Section 4.1);
- Definition of the single-vertex quantum state assignment flow (Section 4.2);
- Determination of the general quantum state assignment flow equation for an arbitrary graph (Section 4.3) and its alternative parameterization (Section 4.4), which generalizes Formulation (62) of the assignment flow accordingly.

A natural question is: What does “label” mean for a generalized assignment flow evolving on the product manifold \mathcal{Q}_c of density matrices? For the single vertex quantum state assignment flow, i.e., without interaction of these flows on a graph, it turns out that the pure state corresponding to the minimal eigenvalue of the initial density matrix is assigned to the given data point (Proposition 5). Coupling non-commuting density matrices over the graph through the novel quantum state assignment flow, therefore, generates interesting complex dynamics, as illustrated in Section 5. In Section 4.5, we show that the restriction of the novel quantum state assignment flow to commuting density matrices recovers the original assignment flow for discrete labels.

Throughout this section, we refer to definitions and notions introduced in Section 2.2.

4.1. Riemannian Gradient, Replicator Operator and Further Mappings

Proposition 2 (Riemannian gradient). *Let $f: \mathcal{D}_c \rightarrow \mathbb{R}$ be a smooth function defined on the manifold (18) and $\tilde{f}: U \rightarrow \mathbb{R}$ be a smooth extension of f to an open neighborhood U of*

$\mathcal{D}_c \subset \mathbb{C}^{c \times c}$ with $\tilde{f}|_{\mathcal{D}_c} = f$. Then, its Riemannian gradient with respect to the BKM metric (25) is given by

$$\text{grad}_\rho f = \mathbb{T}_\rho^{-1}[\partial \tilde{f}] - \langle \rho, \partial \tilde{f} \rangle \rho, \tag{63}$$

where \mathbb{T}_ρ^{-1} is given by (28), and $\partial \tilde{f}$ is the ordinary gradient with respect to the Euclidean structure of the ambient space $\mathbb{C}^{c \times c}$.

Proof. Appendix A.3. \square

Comparing the result (63) with (15) motivates the following:

$$\mathfrak{R}_\rho: \mathcal{H}_c \rightarrow \mathcal{H}_{c,0}, \quad \mathfrak{R}_\rho[X] := \mathbb{T}_\rho^{-1}[X] - \langle \rho, X \rangle \rho, \quad \rho \in \mathcal{D}_c \quad (\text{replicator map}) \tag{64}$$

The following lemma shows that the properties (63) extend to (64).

Lemma 2 (properties of \mathfrak{R}_ρ). Let $\Pi_{c,0}$ denote the orthogonal projection (24). Then, the replicator map (64) satisfies

$$\Pi_{c,0} \circ \mathfrak{R}_\rho = \mathfrak{R}_\rho \circ \Pi_{c,0} = \mathfrak{R}_\rho, \quad \forall \rho \in \mathcal{D}_c. \tag{65}$$

Proof. Appendix A.3. \square

Lemma 2 shows that the replicator map (64) implicitly comprises the orthogonal projection onto the tangent space. This allows for the averaging in (109) without the necessity of explicit projection, which simplifies the notation and explains the larger domain \mathcal{H}_c of \mathfrak{R}_ρ in (64).

Next, using the tangent space $\mathcal{H}_{c,0}$, we define a parameterization of the manifold \mathcal{D}_c in terms of the mapping.

$$\Gamma: \mathcal{H}_{c,0} \rightarrow \mathcal{D}_c, \quad \Gamma(X) := \frac{\exp_m(X)}{\text{tr} \exp_m(X)} = \exp_m(X - \psi(X)I), \quad (\Gamma\text{-map}) \tag{66}$$

where

$$\psi(X) := \log(\text{tr} \exp_m(X)). \tag{67}$$

The following lemma and proposition show that the domain of Γ extends to $\mathbb{R}^{c \times c}$.

Lemma 3 (extension of Γ). The extension to $\mathbb{C}^{c \times c}$ of the mapping Γ defined by (66) is well-defined and given by

$$\Gamma: \mathbb{C}^{c \times c} \rightarrow \mathcal{D}_c, \quad \Gamma(Z) = \Gamma(\Pi_{c,0}[Z]). \tag{68}$$

Proof. Appendix A.3. \square

Proposition 3 (inverse of Γ). The map Γ defined by (66) is bijective with inverse

$$\Gamma^{-1}: \mathcal{D}_c \rightarrow \mathcal{H}_{c,0}, \quad \Gamma^{-1}(\rho) = \Pi_{c,0}[\log_m \rho]. \tag{69}$$

Proof. Appendix A.3. \square

The following lemma provides the differentials of the mappings Γ and Γ^{-1} .

Lemma 4 (differentials $d\Gamma$ and $d\Gamma^{-1}$). Let $H, X \in \mathcal{H}_{c,0}$ with $\Gamma(H) = \rho$ and $Y \in T\mathcal{H}_{c,0} \cong \mathcal{H}_{c,0}$. Then,

$$d\Gamma(H)[Y] = \mathbb{T}_\rho^{-1}[Y - \langle \rho, Y \rangle I], \quad \rho = \Gamma(H), \tag{70}$$

$$d\Gamma^{-1}(\rho)[X] = \Pi_{c,0} \circ \mathbb{T}_\rho[X]. \tag{71}$$

Proof. Appendix A.3. \square

We finally compute a closed-form expression of the e-geodesic, i.e., the geodesic resp. exponential map induced by the e connection on the manifold (\mathcal{D}_c, g) .

Proposition 4 (e-geodesics). *The e-geodesic emanating at $\rho \in \mathcal{D}_c$ in the direction of $X \in \mathcal{H}_{c,0}$ and the corresponding exponential map are given by*

$$\gamma_{\rho, X}^{(e)}(t) := \text{Exp}_{\rho}^{(e)}(tX), \quad t \geq 0 \tag{e-geodesic} \tag{72}$$

$$\text{Exp}_{\rho}^{(e)}(X) := \Gamma(\Gamma^{-1}(\rho) + d\Gamma^{-1}(\rho)[X]) \tag{exponential map} \tag{73}$$

$$= \Gamma(\Gamma^{-1}(\rho) + \Pi_{c,0} \circ \mathbb{T}_{\rho}[X]). \tag{74}$$

Proof. Appendix A.3. \square

Corollary 2 (inverse exponential map). *The inverse of the exponential mapping (72) is given by*

$$(\text{Exp}_{\rho}^{(e)})^{-1}: \mathcal{D}_c \rightarrow \mathcal{H}_{c,0}, \quad (\text{Exp}_{\rho}^{(e)})^{-1}(\mu) = d\Gamma(\Gamma^{-1}(\rho))[\Gamma^{-1}(\mu) - \Gamma^{-1}(\rho)]. \tag{75}$$

Proof. Appendix A.3. \square

Analogous to (44), we define the mapping as exp_{ρ} , where both the subscript and the argument disambiguate the meaning of “exp”.

Lemma 5 (exp-map). *The mapping defined using (73) and (64) by*

$$\text{exp}_{\rho}: \mathcal{H}_{0,c} \rightarrow \mathcal{D}_c, \quad \text{exp}_{\rho}(X) := \text{Exp}_{\rho}^{(e)} \circ \mathfrak{R}_{\rho}[X], \quad \rho \in \mathcal{D}_c \tag{exp-map} \tag{76}$$

has the explicit form

$$\text{exp}_{\rho}(X) = \Gamma(\Gamma^{-1}(\rho) + X). \tag{77}$$

Proof. Appendix A.3. \square

The following lemma provides the explicit form of the differential of the mapping (77), which resembles the corresponding Formula (46) of the assignment flow.

Lemma 6 (differential d exp $_{\rho}$). *The differential of the mapping (76) and (77) reads with $\rho \in \mathcal{D}_c$, $X \in \mathcal{H}_{c,0}$ and $Y \in T\mathcal{H}_{c,0} \cong \mathcal{H}_{c,0}$*

$$d \text{exp}_{\rho}(X)[Y] = \mathfrak{R}_{\text{exp}_{\rho}(X)}[Y]. \tag{78}$$

Proof. Appendix A.3. \square

Remark 3 (comparing exp-maps-I). *Since (78) resembles (46), one may wonder about the connection between (77) and (44). In view of (66), we define*

$$\gamma: T_{c,0} \rightarrow \mathcal{S}_c, \quad \gamma(v) := \frac{e^v}{\langle \mathbf{1}, e^v \rangle} = \text{exp}_{\mathbf{1}_{\mathcal{S}_c}}(v) \tag{79}$$

and compute with the expression for its inverse (cf. [36])

$$\gamma^{-1}(p) = \pi_{c,0} \log \frac{p}{\mathbf{1}_{\mathcal{S}_c}} = \pi_{c,0}(\log p - \log \mathbf{1}_{\mathcal{S}_c}) = \pi_{c,0} \log p \tag{80}$$

$$\stackrel{(11)}{=} \log p - \langle \mathbf{1}_{\mathcal{S}_c}, \log p \rangle \mathbf{1}_c \tag{81}$$

which resembles (69). Moreover, in view of (77), the analogous expression using γ instead of Γ reads

$$\gamma(\gamma^{-1}(p) + v) = \frac{e^{\pi_{c,0} \log p + v}}{\langle \mathbb{1}, e^{\pi_{c,0} \log p + v} \rangle} = \frac{\langle \mathbb{1}_{\mathcal{S}_c}, \log p \rangle p \cdot e^v}{\langle \langle \mathbb{1}_{\mathcal{S}_c}, \log p \rangle p, e^v \rangle} = \frac{p \cdot e^v}{\langle p, e^v \rangle} \tag{82}$$

$$= \exp_p(v). \tag{83}$$

Remark 4 (comparing exp-maps-II). Using the above definitions and relations, we check Equation (40): $\exp_\rho^{(\log)}(Y) = \text{Exp}_\rho^{(e)}(X)$, where the relation (41) between Y and X can now be written in the following form:

$$Y \stackrel{(67)}{=} X - \psi(\log_m(\rho) + \mathbb{T}_\rho[X])\rho. \tag{84}$$

Direct computation yields

$$\exp_\rho^{(\log)}(Y) \stackrel{(37)}{=} \exp_m(\log_m(\rho) + \mathbb{T}_\rho[Y]) \tag{85}$$

$$\stackrel{(41)}{=} \exp_m\left(\log_m(\rho) + \mathbb{T}_\rho[X] - \psi(\log_m(\rho) + \mathbb{T}_\rho[X]) \overbrace{\mathbb{T}_\rho \circ \mathbb{T}_\rho^{-1}[I_c]}^{=I_c}\right) \tag{86}$$

$=\rho$

$$\stackrel{(66)}{=} \Gamma(\Pi_{c,0}[\log_m(\rho)] + \Pi_{c,0} \circ \mathbb{T}_\rho[X]) = \Gamma(\Gamma^{-1}(\rho) + \Pi_{c,0} \circ \mathbb{T}_\rho[X]) \tag{87}$$

$$= \text{Exp}_\rho^{(e)}(X). \tag{88}$$

4.2. Single-Vertex Density Matrix Assignment Flow

We generalize the single-vertex assignment flow Equation (45) to the manifold (\mathcal{D}_c, g_ρ) given by (18) with the BKM metric (25).

In view of (43), the likelihood matrix is defined as

$$L_\rho: \mathcal{H}_c \rightarrow \mathcal{D}_c, \quad L_\rho(D) := \exp_\rho(-\Pi_{c,0}[D]), \quad \rho \in \mathcal{D}_c, \tag{89}$$

and the corresponding single-vertex quantum state assignment flow (SQSAF) equation reads

$$\dot{\rho} = \mathfrak{R}_\rho[L_\rho(D)] \tag{SQSAF} \tag{90}$$

$$\stackrel{(64)}{=} \mathbb{T}_\rho^{-1}[L_\rho(D)] - \langle \rho, L_\rho(D) \rangle \rho, \quad \rho(0) = \mathbb{1}_{\mathcal{D}_c} = \text{Diag}(\mathbb{1}_{\mathcal{S}_c}). \tag{91}$$

Proposition 5 below specifies its properties after a preparatory Lemma.

Lemma 7. Assume

$$D = Q\Lambda_D Q^\top \in \mathcal{H}_c \quad \text{and} \quad \rho = Q\Lambda_\rho Q^\top \in \mathcal{D}_c \tag{92}$$

can be simultaneously diagonalized with $Q \in \text{O}(c)$, $\Lambda_D = \text{Diag}(\lambda_D)$, $\Lambda_\rho = \text{Diag}(\lambda_\rho)$ and $\lambda_\rho \in \mathcal{S}_c$, since $\text{tr } \rho = 1$. Then,

$$L_\rho(D) = Q \text{Diag}(\exp_{\lambda_\rho}(-\lambda_D)) Q^\top. \tag{93}$$

Proof. Appendix A.3. \square

Proposition 5 (SQSAF limit). Let $D = Q\Lambda_D Q^\top$ be the spectral decomposition of D with eigenvalues $\lambda_1 \geq \dots \geq \lambda_c$ and orthonormal eigenvectors of $Q = (q_1, \dots, q_c)$. Assume that the minimal eigenvalue λ_c is unique. Then, the solution $\rho(t)$ to (90) satisfies

$$\lim_{t \rightarrow \infty} \rho(t) = \Pi_{q_c} := q_c q_c^\top. \tag{94}$$

Proof. Appendix A.3. \square

4.3. Quantum State Assignment Flow

This section describes our main result, i.e., the definition of a novel flow of coupled density matrices in terms of a parameterized interaction of single-vertex flows of the form (90) on a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$.

We assume the weight function $\omega: \mathcal{E} \rightarrow \mathbb{R}_+$ to be non-negative with $\omega_{ij} = 0$ if $ij \notin \mathcal{E}$ and

$$\sum_{k \in \mathcal{N}_i} \omega_{ik} = 1, \tag{95}$$

where we adopt the notation (57) for neighborhoods $\mathcal{N}_i, i \in \mathcal{V}$. Analogous to (53), we define the product manifold as

$$\rho \in \mathcal{Q}_c := \underbrace{\mathcal{D}_c \times \dots \times \mathcal{D}_c}_{|\mathcal{V}| \text{ factors}} \tag{96}$$

where \mathcal{D}_c is given by (18). The corresponding factors of ρ are denoted by

$$\rho = (\rho_i)_{i \in [c]}, \quad \rho_i \in \mathcal{D}_c, \quad i \in \mathcal{V}. \tag{97}$$

where \mathcal{Q}_c becomes a Riemannian manifold when equipped with the the following metric:

$$g_\rho(X, Y) := \sum_{i \in \mathcal{V}} g_{\rho_i}(X_i, Y_i), \quad X, Y \in T\mathcal{Q}_c := \mathcal{H}_{c,0} \times \dots \times \mathcal{H}_{c,0}, \tag{98}$$

with g_{ρ_i} given by (25) for each $i \in \mathcal{V}$. We set

$$\mathbb{1}_{\mathcal{Q}_c} := (\mathbb{1}_{\mathcal{D}_c})_{i \in \mathcal{V}} \in \mathcal{Q}_c, \tag{99}$$

with $\mathbb{1}_{\mathcal{D}_c}$ given by (91). Our next step is to define a *similarity mapping* analogous to (56),

$$S: \mathcal{V} \times \mathcal{Q}_c, \quad S_i(\rho) := \text{Exp}_{\rho_i}^{(e)} \left(\sum_{k \in \mathcal{N}_i} \omega_{ik} (\text{Exp}_{\rho_i}^{(e)})^{-1} (L_{\rho_k}(D_k)) \right), \tag{100}$$

based on mappings (73) and (89). Thanks to the use of the exponential map of the e connection, the matrix $S_i(\rho)$ can be rewritten and computed in a simpler, more explicit form.

Lemma 8 (similarity map). Equation (100) is equivalent to

$$S_i(\rho) = \Gamma \left(\sum_{k \in \mathcal{N}_i} \omega_{ik} (\log_{\text{gm}} \rho_k - D_k) \right). \tag{101}$$

Proof. Appendix A.3. \square

Expression (100), which defines the similarity map, looks like a single iterative step for computing the Riemannian center of mass of the likelihood matrices $\{L_{\rho_k}(D_k): k \in \mathcal{N}_i\}$ if(!) the exponential map of the Riemannian (Levi Civita) connection were used. Instead, when using the exponential map $\text{Exp}^{(e)}$, $S_i(\rho)$ may be interpreted as carrying out a single iterative step for the corresponding *geometric mean* on the manifold \mathcal{D}_c .

Therefore, a natural idea is to define the similarity map to be this geometric mean rather than just by a single iterative step. Surprisingly, analogous to the similarity map (56) for categorical distributions (cf. [3]), both definitions are *identical*, as shown next.

Proposition 6 (geometric mean property). *Assume that $\bar{\rho} \in \mathcal{D}_c$ solves the equation*

$$0 = \sum_{k \in \mathcal{N}_i} \omega_{ik} (\text{Exp}_{\bar{\rho}}^{(e)})^{-1}(L_{\rho_k}(D_k)) \tag{102}$$

which corresponds to the optimality condition for Riemannian centers of mass [32] (Lemma 6.9.4), except using a different exponential map. Then,

$$\bar{\rho} = S_i(\rho) \tag{103}$$

with the right-hand side given by (100).

Proof. Appendix A.3. \square

We are now in the position to define the *quantum state assignment flow* along the lines of the original assignment flow (55),

$$\dot{\rho} = \mathfrak{R}_\rho[S(\rho)], \quad \rho(0) = \mathbb{1}_{\mathcal{Q}_c}, \quad \text{(QSAF)} \tag{104}$$

where both the replicator map \mathfrak{R}_ρ and the similarity map $S(\cdot)$ apply factor-wise,

$$S(\rho)_i = S_i(\rho), \tag{105}$$

$$\mathfrak{R}_\rho[S(\rho)]_i = \mathfrak{R}_{\rho_i}[S_i(\rho)], \quad i \in \mathcal{V} \tag{106}$$

with the mappings S_i given by (101) and \mathfrak{R}_{ρ_i} given by (64).

4.4. Reparameterization and Riemannian Gradient Flow

The reparameterization of the assignment flow (59) and (60) for categorical distributions described in Section 3.3 has proven to be useful for characterizing and analyzing assignment flows. Under suitable conditions on the parameter matrix Ω , the flow performs a Riemannian descent flow with respect to a non-convex potential [36] (Proposition 3.9) and has convenient stability and convergence properties [4].

In this section, we derive a similar reparameterization of the quantum state assignment flow (104).

Proposition 7 (reparameterization). *Define the linear mapping as*

$$\Omega: \mathcal{Q}_c \rightarrow \mathcal{Q}_c, \quad \Omega[\rho]_i := \sum_{k \in \mathcal{N}_i} \omega_{ik} \rho_k. \tag{107}$$

Then, the density matrix assignment flow equation (104) is equivalent to the following system:

$$\dot{\rho} = \mathfrak{R}_\rho[\mu], \quad \rho(0) = \mathbb{1}_{\mathcal{Q}_c}, \tag{108}$$

$$\dot{\mu} = \mathfrak{R}_\mu[\Omega[\mu]], \quad \mu(0) = S(\mathbb{1}_{\mathcal{Q}_c}). \tag{109}$$

Proof. Appendix A.3. \square

For the following, we adopt the *symmetry assumption*.

$$\omega_{ij} = \omega_{ji}, \quad \forall i, j \in \mathcal{V} \tag{110}$$

$$j \in \mathcal{N}_i \Leftrightarrow i \in \mathcal{N}_j, \quad i, j \in \mathcal{V}. \tag{111}$$

As a consequence, the mapping (107) is self-adjoint.

$$\langle \mu, \Omega[\rho] \rangle = \sum_{i \in \mathcal{V}} \langle \mu_i, \Omega[\rho]_i \rangle = \sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ik} \langle \mu_i, \rho_k \rangle = \sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ki} \langle \mu_i, \rho_k \rangle \tag{112}$$

$$= \sum_{k \in \mathcal{V}} \sum_{i \in \mathcal{N}_k} \omega_{ki} \langle \mu_i, \rho_k \rangle = \sum_{k \in \mathcal{N}_i} \langle \Omega[\mu]_k, \rho_k \rangle = \langle \Omega[\mu], \rho \rangle. \tag{113}$$

Proposition 8 (Riemannian gradient QSAF flow). *Suppose that the mapping $\Omega[\cdot]$ given by (107) is self-adjoint with respect to the canonical matrix inner product. Then, the solution $\mu(t)$ to (109) also solves*

$$\dot{\mu} = -\text{grad}_\mu J(\mu) \quad \text{with} \quad (\text{grad}_\mu J(\mu))_i = \text{grad}_{\mu_i} J(\mu) \tag{114}$$

with respect to the potential

$$J(\mu) := -\frac{1}{2} \langle \mu, \Omega[\mu] \rangle. \tag{115}$$

Proof. Appendix A.3. \square

We conclude this section by rewriting the potential in a more explicit, informative form.

Proposition 9 (nonconvex potential). *We define*

$$L_G: \mathcal{Q}_c \rightarrow \mathcal{Q}_c, \quad L_G := \text{id} - \Omega \tag{116}$$

with Ω given by (107). Then, the potential (115) can be rewritten as

$$J(\mu) = \frac{1}{2} (\langle \mu, L_G[\mu] \rangle - \|\mu\|^2) \tag{117}$$

$$= \frac{1}{4} \sum_{i \in \mathcal{V}} \sum_{j \in \mathcal{N}_i} \omega_{ij} \|\mu_i - \mu_j\|^2 - \frac{1}{2} \|\mu\|^2. \tag{118}$$

Proof. Appendix A.3. \square

4.5. Recovering the Assignment Flow for Categorical Distributions

In the following, we show how the assignment flow (62) for categorical distributions arises as special case of the quantum state assignment flow under suitable conditions, as detailed below.

Definition 2 (commutative submanifold). *Let*

$$\Pi = \{\pi_i: i \in [l]\}, \quad l \leq c \tag{119}$$

denote a set of operators that orthogonally project onto disjoint subspaces of \mathbb{C}^c ,

$$\pi_i^2 = \pi_i, \quad \forall i \in [l], \tag{120}$$

$$\pi_i \pi_j = 0, \quad \forall i, j \in [l], i \neq j, \tag{121}$$

and which are complete in the sense that

$$\sum_{i \in [l]} \pi_i = I_c. \tag{122}$$

Given a family Π of operators, we define

$$\mathcal{D}_\Pi := \left\{ \sum_{i \in [l]} \frac{p_i}{\text{tr } \pi_i} \pi_i : p \in \mathcal{S}_l \right\} \subset \mathcal{D}_c \tag{123}$$

the submanifold of commuting Hermitian matrices, which can be diagonalized simultaneously.

A typical example for a family (119) is

$$\Pi_{\mathcal{U}} = \{ \pi_i = u_i u_i^* : i \in [c] \}, \tag{124}$$

where $\mathcal{U} = \{u_1, \dots, u_c\}$ is an orthonormal basis of \mathbb{C}^c . The following lemma elaborates the bijection $\mathcal{D}_\Pi \leftrightarrow \mathcal{S}_l$.

Lemma 9 (properties of \mathcal{D}_Π). Let $\mathcal{D}_\Pi \subset \mathcal{D}_c$ be given by (123) and denote the corresponding inclusion map by $\iota: \mathcal{D}_\Pi \hookrightarrow \mathcal{D}_c$. Then,

- (a) The submanifold $(\mathcal{D}_\Pi, \iota^* g_{\text{BKM}})$ with the induced BKM metric is isometric to $(\mathcal{S}_l, g_{\text{FR}})$;
- (b) If $\mu \in \mathcal{D}_\Pi$, then the tangent subspace $T_\mu \mathcal{D}_\Pi$ is contained in the subspace $T_\mu^c \mathcal{D}_c \subseteq T_\mu \mathcal{D}_c$ defined by (32);
- (c) Let $\mathcal{U} = \{u_1, \dots, u_c\}$ denote an orthonormal basis of \mathbb{C}^c such that for every $\pi_i \in \Pi$, $i \in [l]$, there are $u_{i_1}, \dots, u_{i_k} \in \mathcal{U}$ that form a basis of $\text{range}(\pi_i)$. Then, there is an inclusion of commutative subsets $\mathcal{D}_\Pi \hookrightarrow \mathcal{D}_{\Pi_{\mathcal{U}}}$ that corresponds to an inclusion $\mathcal{S}_l \hookrightarrow \mathcal{S}_c$.

Proof. Appendix A.3. \square

Now, we establish that a restriction of the QSAF Equation (109) to the commutative product submanifold can be expressed in terms of the AF Equation (62). Analogous to the definition (96) of the product manifold \mathcal{Q}_c , we set

$$\mathcal{D}_{\Pi,c} = \underbrace{\mathcal{D}_\Pi \times \dots \times \mathcal{D}_\Pi}_{|\mathcal{V}| \text{ factors}}. \tag{125}$$

If Π is given by an orthonormal basis as in (124), we define the unitary matrices as

$$U = (u_1, \dots, u_c) \in \text{Un}(c), \tag{126}$$

$$U_c = \underbrace{\text{Diag}(U, \dots, U)}_{|\mathcal{V}| \text{ block-diagonal entries}}. \tag{127}$$

Proposition 10 (invariance of $\mathcal{D}_{\Pi,c}$). Let Π and \mathcal{D}_Π be given according to Definition 2. Then, the following holds.

- (i) If $\mu \in \mathcal{D}_{\Pi,c} \subset \mathcal{Q}_c$, then $\mathfrak{R}[\Omega[\mu]] \in T_\mu \mathcal{D}_{\Pi,c} \subseteq T_\mu \mathcal{Q}_c$.
- (ii) If $\Pi_{\mathcal{U}}$ has the form (124), then

$$\mathfrak{R}[\Omega[\mu]] = U_c \text{Diag} [R_S[\Omega S]] U_c^*, \tag{128}$$

where $S \in \mathcal{W}_c$ is determined by $\mu_i = U \text{Diag}(S_i) U^*$, $i \in \mathcal{V}$.

In particular, the submanifold $\mathcal{D}_{\Pi,c}$ is preserved by the quantum state assignment flow.

Proof. Appendix A.3. \square

It remains to be verified that under suitable conditions on the data matrices D_i , $i \in \mathcal{V}$ that define the initial point of (109) by similarity mapping (Lemma 8), the quantum state assignment flow becomes the ordinary assignment flow.

Corollary 3 (recovery of the AF by restriction). *In the situation of Proposition 10, assume that all data matrices $D_i, i \in \mathcal{V}$ become diagonal in the same basis \mathcal{U} , i.e.,*

$$D_i = U \text{Diag}(\lambda_i) U^*, \quad \lambda_i \in \mathbb{R}^c, \quad i \in \mathcal{V}. \tag{129}$$

Then, the solution of the QSAF

$$\dot{\mu} = \mathfrak{R}_\mu[\Omega[\mu]], \quad \mu(0) = S(\mathbb{1}_{\mathcal{Q}_c}) \tag{130}$$

is given by

$$\mu_i(t) = U \text{Diag}(S_i(t)) U^*, \quad i \in \mathcal{V}, \tag{131}$$

where $S(t)$ satisfies the ordinary AF equation

$$\dot{S} = R_S[\Omega S], \quad S(0) = S(\mathbb{1}_{\mathcal{W}_c}), \tag{132}$$

and the initial point is determined by the similarity map (56) evaluated at the barycenter $W = \mathbb{1}_{\mathcal{W}_c}$ with the vectors $\lambda_i, i \in \mathcal{V}$ as data points.

Proof. Appendix A.3. \square

5. Experiments and Discussion

In this section, we report academic experiments in order to illustrate the novelty of our approach. In comparison to the original formulation, our approach enables a continuous assignment without the need to specify explicitly prototypical labels beforehand. The experiments highlight the following properties of the novel approach, which extend the expressivity of the original assignment flow approach:

- Geometric *adaptive* feature vector averaging even when *uniform* weights are used (Section 5.2);
- Structure-preserving feature *patch* smoothing *without* accessing data at individual *pixels* (Section 5.3);
- Seamless incorporation of feature *encoding* using finite *frames* (Section 5.3).

In Section 6, we indicate the potential to represent spatial feature *context* via entanglement. However, working out the potential for various applications more thoroughly is beyond the scope of this paper.

5.1. Geometric Integration

In this section, we focus on the geometric integration of the reparameterized flow described by Equation (109). For a reasonable choice of a single-step-sized parameter, the scheme is accurate, stable and amenable to highly parallel implementations.

The e-geodesic from Proposition 4 constitutes a retraction [37] (Definition 4.1.1 and Proposition 5.4.1) onto the state manifold \mathcal{Q}_c .

Consequently, the iterative step for updating $\mu_t \in \mathcal{Q}_c, t \in \mathbb{N}_0$ and step size $\epsilon > 0$ is given by

$$(\mu_{t+1})_i = \left(\text{Exp}_{\mu_t}^{(\epsilon)} \left(\epsilon \mathfrak{R}_{\mu_t}[\Omega[\mu_t]] \right) \right)_i = \left(\text{Exp}_{(\mu_t)_i}^{(\epsilon)} \circ \mathfrak{R}_{(\mu_t)_i} \left[\epsilon(\Omega[\mu_t])_i \right] \right) \tag{133}$$

$$\stackrel{(76)}{=} \exp_{(\mu_t)_i} \left(\epsilon(\Omega[\mu_t])_i \right), \quad \forall i \in \mathcal{V}. \tag{134}$$

for all $i \in \mathcal{V}$. Using (77) and assuming

$$(\mu_t)_i = \Gamma((A_t)_i), \quad \forall i \in \mathcal{V}, \tag{135}$$

with $A_t \in \mathcal{T}_c$, we obtain

$$\Gamma((A_{t+1})_i) := \exp_{(\mu_t)_i}(\epsilon(\Omega[\mu_t])_i) \tag{136}$$

$$= \Gamma(\Gamma^{-1}((\mu_t)_i) + \epsilon(\Omega[\mu_t])_i) \tag{137}$$

$$= \Gamma(\Gamma^{-1} \circ \Gamma((A_t)_i) + \epsilon(\Omega[\mu_t])_i) \tag{138}$$

$$= \Gamma((A_t)_i + \epsilon(\Omega[\mu_t])_i), \quad i \in \mathcal{V}, \tag{139}$$

and in view of (68) and (135), we conclude that

$$A_{t+1} = A_t + \epsilon \Pi_{c,0} \Omega[\Gamma(A_t)]. \tag{140}$$

Remark 5. We note that the numerical evaluation of the replicator operator (64) is not required. This makes the geometric integration scheme summarized by Algorithm 1 quite efficient.

Algorithm 1: Geometric Integration Scheme

Initialization

Determine an initial $A_0 \in \mathcal{T}_{c,0}$ and compute μ_0 by $(\mu_0)_i = \Gamma((A_0)_i) \in \mathcal{Q}_c, \forall i \in \mathcal{V}$

while not converged do

$$\left[\begin{array}{l} (A_{t+1})_i = (A_t)_i + \epsilon \Pi_{c,0}(\Omega[\mu_t])_i \quad \forall i \in \mathcal{V} \\ (\mu_{t+1})_i = \Gamma((A_{t+1})_i), \quad \forall i \in \mathcal{V}. \end{array} \right.$$

We list a few further implementation details below.

- A reasonable convergence criterion that measures how close the states are to a rank-one matrix is $|\text{tr}(\mu_t)_i - \text{tr}(\mu_t^2)_i| \leq \epsilon, \forall i \in \mathcal{V}$;
- A reasonable range for the step size parameter is $\epsilon \leq 0.1$;
- In order to remove spurious non-Hermitian numerical rounding errors, we replace each matrix $(\Omega[\mu_t])_i$ with $\frac{1}{2}((\Omega[\mu_t])_i + (\Omega[\mu_t])_i^*)$;
- The constraint $\text{tr} \rho = 1$ of (18) can be replaced by $\text{tr} \rho = \tau$ with any constant $\tau > 1$. This ensures that for larger matrix dimensions c , the entries of ρ vary in a reasonable numerical range and that the stability of the iterative updates.

Up to moderate matrix dimensions, such as $c \leq 100$, the matrix exponential in (66) can be computed using any of the basic established algorithms [38] (Chapter 10) or available solvers. In addition, depending on the size of the neighborhood \mathcal{N}_i induced by the weighted adjacency relation of the underlying graph in (95), Algorithm 1 can be implemented in a fine-grained parallel fashion.

5.2. Labeling 3D Data on Bloch Spheres

For the purpose of visual illustration, we consider the smoothing of 3D color vectors $d = (d_1, d_2, d_3)^\top$ interpreted as Bloch vectors, which parameterize density matrices [14] (Section 5.2).

$$\rho = \rho(d) = \frac{1}{2} \left(I + d_1 \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix} + d_2 \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix} + d_3 \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix} \right) \in \mathbb{C}^{2 \times 2}, \quad \|d\| \leq 1. \tag{141}$$

Pure states ρ correspond to unit vectors $d, \|d\| = 1$, whereas vectors $d, \|d\| < 1$ parameterize mixed states ρ . Given data $d_i = (d_{i,1}, d_{i,2}, d_{i,3})^\top, i \in \mathcal{V}$ with $\|d_i\| \leq 1$, as illustrated by Figure 1 and explained in the caption, we initialized the QSAF at $\rho_i = \rho(d_i), i \in \mathcal{V}$ and integrated the flow. Each integration step involves geometric state averaging across the graph, causing mixed states $\rho_i(t) = \rho(d_i(t)), i \in \mathcal{V}$, which eventually converge towards pure states. Integration was stopped at time $t = T$, when $\min\{\|d_i(T)\| : i \in \mathcal{V}\} \geq 0.999$. The resulting vectors $d_i(T)$ are visualized as explained in the caption of Figure 1. We point out that the two experiments discussed next are supposed to illustrate the behavior of the

QSAF and the impact of the underlying geometry rather than constitute a contribution to the literature on the processing of color images.

Figure 1c shows a noisy version of the image in (b) used to initialize the quantum state assignment flow (QSAF). Panel (d) shows the labeled image, i.e., the assignment of a pure state (depicted as a Bloch vector) to each pixel of the input data (c). Although uniform weights were used and any prior information was absent, the result (d) demonstrates that the QSAF removes the noise and preserves the signal transition fairly well both for large-scale local image structure (away from the image center) and for small-scale local image structure (close to the image center). This behavior is quite unusual in comparison to traditional image denoising methods, which inevitably require *adaptation* of regularization to the scale of local image structure. In addition, we note that noise removal is ‘perfect’ for the three extreme points (red, green and blue in panel (a)) but suboptimal only for the remaining non-extreme points.

Panels (f–h) show the same results when the data are encoded in a better way, as depicted by (e) using unit vectors not only on the positive orthant but on the whole unit sphere. These data are illustrated by RGB vectors that result from translating the unit sphere (e) to the center $\frac{1}{2}(1, 1, 1)^\top$ of the RGB color cube $[0, 1]^3$ and scaling it by $\frac{1}{2}$. This improved data encoding is clearly visible in panel (g), which displays the *same* noise level as shown in panel (c). Accordingly, noise removal while preserving signal structure at *all* local scales is more effectively achieved by the QSAF in (h) in comparison to (d).

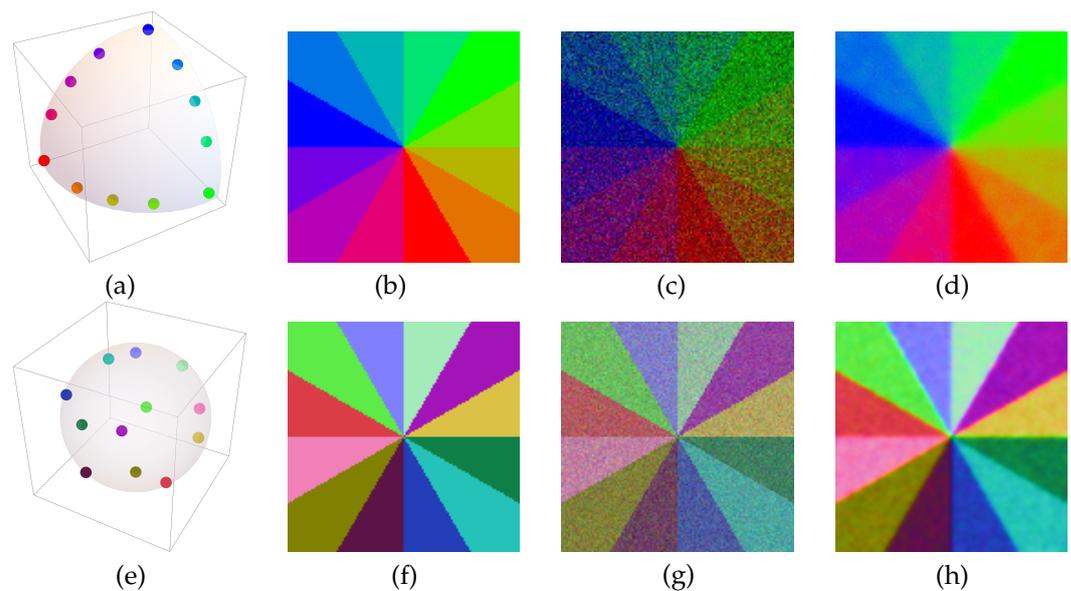


Figure 1. (a) A range of RGB unit color vectors in the positive orthant. (b) An image with data according to (a). (c) A noisy version of (b): each pixel $i \in \mathcal{V}$ displays a Bloch vector $d_i = (d_{i,1}, d_{i,2}, d_{i,3})^\top$ defined by Equation (141) as an initial density matrix $\rho_i(0)$, $i \in \mathcal{V}$ of the QSAF. (d) The labels (pure states) generated by integrating the quantum state assignment flow using uniform weights. (e) The vectors depicted by (a) are replaced by the unit vectors corresponding to the vertices of the icosahedron centered at 0. (f–h) Analogous to (b–d), based on (e) instead of (a) and using the same noise level in (g). The colors in (f–h) visualize the Bloch vectors by RGB vectors that result from translating the sphere of (e) to the center $\frac{1}{2}(1, 1, 1)^\top$ of the RGB cube and scaling it by $\frac{1}{2}$. We refer to the text for a discussion.

5.3. Basic Image Patch Smoothing

Figure 2 shows an application of the QSAF to a *random* spatial arrangement (grid graph) of normalized patches, where each vertex represents a patch, not a pixel. Applying vectorization taking the tensor product with itself, each patch is represented as a pure state in terms of a rank-one matrix D_i at the corresponding vertex $i \in \mathcal{V}$, which constitutes the input data in the similarity mapping (100). Integrating the flow causes the non-commutative

interaction of the associated state spaces $\rho_i, i \in \mathcal{V}$ through geometric averaging with uniform weights (95) until convergence towards pure states. The resulting patches are then simply given by the corresponding eigenvector, possibly after reversing the arbitrary sign of each eigenvector component, depending on the distance to the input patch.

The result shown in Figure 2 reveals an interesting behavior: structure-preserving patch smoothing without accessing explicitly individual pixels. In particular, the flow induces a *partition* of the patches without any prior assumption on the data.

Figure 3 shows a variant of the scenario depicted in Figure 2 in order to demonstrate the ability to separate local image structure by geometric smoothing at the patch level in another way.

Figure 4 generalizes the setup in two ways. First, patches were encoded using the harmonic frame given by the two-dimensional discrete Fourier matrix. Secondly, non-uniform weights $\omega_{ik} = e^{-\tau\|P_i - P_j\|_F^2}, \tau > 0$ were used depending on the distance of adjacent patches P_i, P_j .

Specifically, let P_i denote the patch at vertex $i \in \mathcal{V}$ after removing the global mean and normalization using the Frobenius norm. Then, applying the FFT to each patch and vectorization formally with the discrete two-dimensional Fourier matrix $F_2 = F \otimes F$ (Kronecker product) followed by stacking the rows $\hat{p}_i = F_2 \text{vec}(P_i)$, the input data were defined as $D_i = F_2 \text{Diag}(-|\hat{p}_i|^2)F_2^*$, where the squared magnitude $|\cdot|^2$ was computed component-wise. The flow yields were integrated again against pure states that were interpreted and decoded accordingly. The eigenvector was used as a multiplicative filter of the magnitude of the Fourier-transformed patch (keeping its phase), followed by rescaling of the norm and addition of the mean by approximating the original patch in terms of these two parameters.

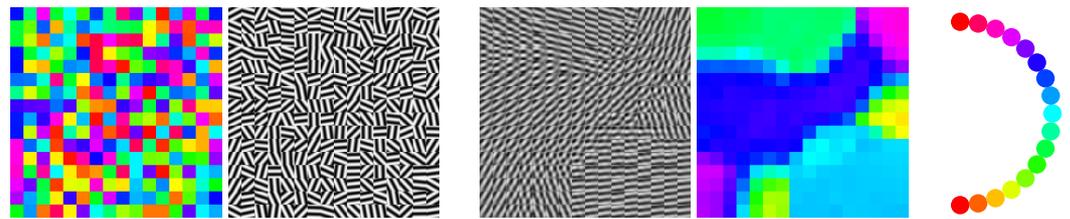


Figure 2. **Left pair:** A random collection of patches with oriented image structure. The colored image of each patch shows its orientation using the color code depicted by the rightmost panel. Each patch is represented by a rank-one matrix D in (89) obtained by vectorizing the patch and taking the tensor product. **Center pair:** The final state of the QSAF obtained by geometric integration with uniform weighting $\omega_{ik} = \frac{1}{|\mathcal{N}_i|}, \forall k \in \mathcal{N}_i, \forall i \in \mathcal{V}$ of the nearest neighbor states. It represents an image partition but preserves image structure due to geometric smoothing of patches encoded by non-commutative state spaces.

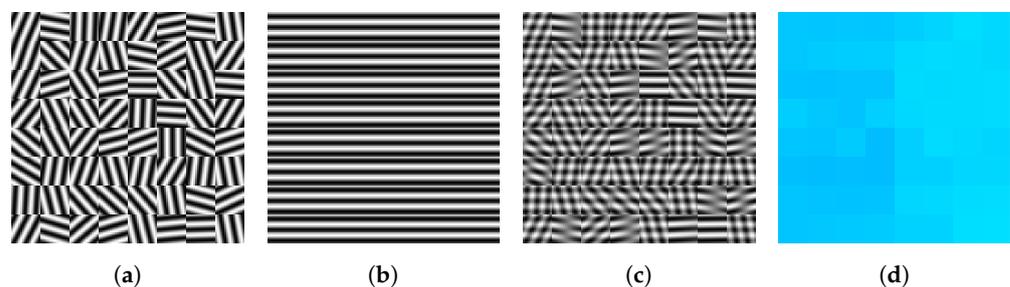


Figure 3. (a) A random collection of patches with oriented image structure. (b) A collection of patches with the same oriented image structure. (c) Pixel-wise mean of the patches (a,b) at each location. (d) The QSAF recovers a close approximation of (b) (color code: see Figure 2) by iteratively smoothing the states $\rho_k, k \in \mathcal{N}_i$ corresponding to (c) through geometric integration.

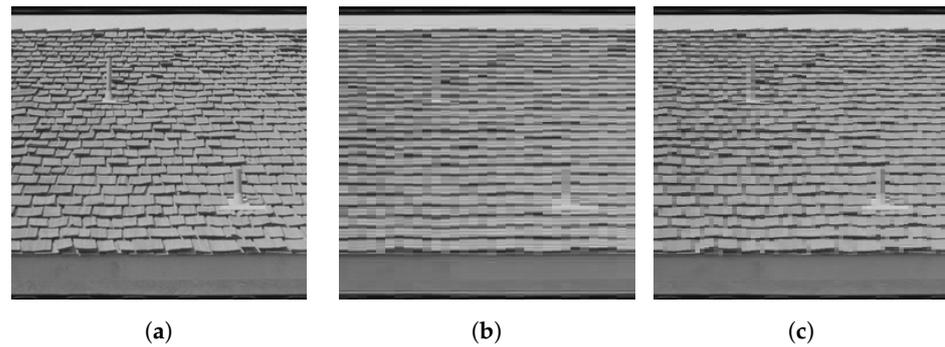


Figure 4. (a) A real image partitioned into patches of size 8×8 and 4×4 pixels, respectively. Each patch is represented as a pure state with respect to a Fourier frame (see text). Instead of the nearest neighbor adjacency on a regular grid, each patch is adjacent to its eight closest patches in the entire collection. Integrating the QSAF and decoding the resulting states (see text) yield (b) (8×8 patches) and (c) (4×4 patches), respectively. Result (b) illustrates the effect of smoothing at the patch level in the Fourier domain, whereas the smaller spatial scale used to compute (c) represents the input data fairly accurately, despite achieving significant data reduction.

The results shown in panels (b) and (c) of Figure 4 illustrate the effect of ‘geometric diffusion’ at the patch level through integration of the flow and how the input data are approximated depending on the chosen spatial scale (patch size), subject to significant data reduction.

6. Conclusions

We generalized the assignment flow approach for categorical distributions [2] to density matrices on weighted graphs. While the former flows assign each data point a label selected from a *finite* set, the latter assign each data point a generalized “label” from the *uncountable* submanifold of pure states.

Various further directions of research are indicated by numerical experiments. This includes the unusual behavior of feature vector smoothing, which parameterizes complex-valued, non-commutative state spaces (Figure 1), the structure-preserving interaction of spatially indexed feature patches without accessing individual pixels (Figures 2 and 3), the use of frames for signal representation and as observables whose expected values are governed by a quantum state assignment flow (Figure 4) and the representation of spatial correlations by entanglement and tensorization (Figure 5). Extending the representation of the original assignment flow in the broader framework of geometric mechanics to the novel quantum assignment flow approach as recently developed by [39] is another promising research project spurred by established concepts of mathematics and physics.

Based on these viewpoints, this paper adds a novel concrete approach based on information theory to the emerging literature on network design based on concepts from quantum mechanics, e.g., [40] and references therein. Our main motivation is the definition of a novel class of “neural ODEs” [41] in terms of the dynamical systems that generate a quantum state assignment flow. The layered architecture of a corresponding “neural network” is implicitly given by geometric integration. The inherent smoothness of the parameterization allows weight parameters to be learned from data. This will be explored in our future work, along the various lines of research indicated above.

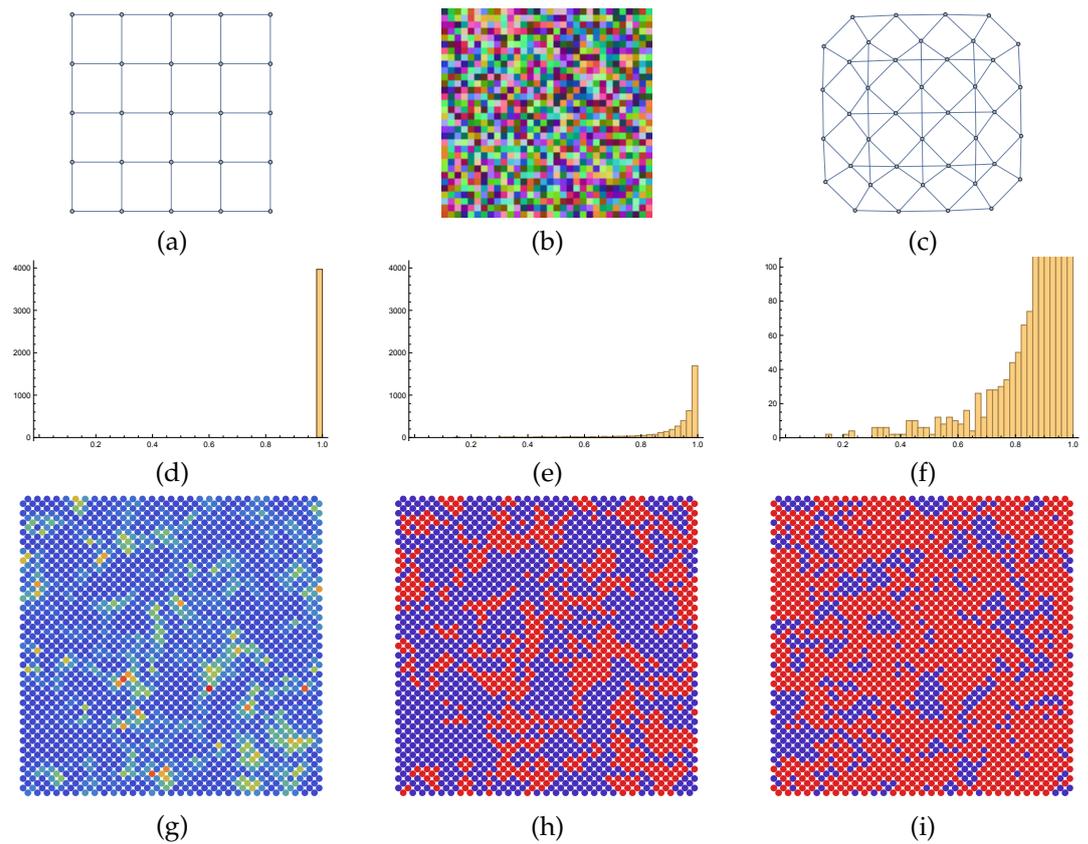


Figure 5. (a) A 5×5 grid graph. (b) Random Bloch vectors $d_i \in S^2 \subset \mathbb{R}^3$ (visualized using pseudocolor) defining states ρ_i according to Equation (141) for each vertex of a 32×32 grid graph. (c) Line graph corresponding to (a). Each vertex corresponds to an edge ij of the graph (a) and an initially separable state $\rho_{ij} = \rho_i \otimes \rho_j$. This defines a simple shallow tensor-network. The histograms display the norms of the Bloch vectors of the states $\text{tr}_j(\rho_{ij})$ and $\text{tr}_i(\rho_{ij})$ obtained by partially tracing out one factor for each state ρ_{ij} indexed by a vertex ij of the line graph of the grid graph in (b). (d) Histogram showing that in the initial state, all states are separable, while (e,f) both display a histogram of the norms of all Bloch vectors after convergence of the quantum state assignment flow with uniform weights towards pure states. (g) Using the center coordinates of each edge of the grid graph (b), the entanglement represented by ρ_{ij} is visualized by a disk and “heat map” colors (blue: low entanglement, red: large entanglement). For visual clarity, (h,i) again display the *same* information after thresholding, using two colors only: entangled states are marked with red when the norm of the Bloch vectors drops below the thresholds of 0.95 and 0.99, respectively, and otherwise with blue.

Author Contributions: Investigation, J.S., J.C., B.B., M.G., P.A. and C.S.; Supervision, C.S. All authors have read and agreed to the published version of the manuscript.

Funding: This work was funded by the Deutsche Forschungsgemeinschaft (DFG), grant SCHN 457/17-1, within the priority programme SPP 2298: Theoretical Foundations of Deep Learning. This work was funded by the Deutsche Forschungsgemeinschaft (DFG) under Germany’s Excellence Strategy EXC-2181/1-390900948 (the Heidelberg STRUCTURES Excellence Cluster).

Conflicts of Interest: The authors declare no conflict of interest.

Appendix A. Proofs

Appendix A.1. Proofs of Section 2

Proof of Proposition 1. We verify (13) and (14) by direct computation. For any $p \in \mathcal{S}_c$,

$$R_p \mathbf{1}_c = (\text{Diag}(p) - pp^\top) \mathbf{1}_c = p - \langle p, \mathbf{1}_c \rangle p = 0, \tag{A1}$$

$$R_p \pi_{c,0} = R_p (I - \mathbf{1}_c \mathbf{1}_{\mathcal{S}_c}^\top) = R_p, \tag{A2}$$

$$\pi_{c,0} R_p = (I - \mathbf{1}_c \mathbf{1}_{\mathcal{S}_c}^\top) R_p = R_p - \frac{1}{c} \mathbf{1}_c (R_p \mathbf{1}_c)^\top = R_p. \tag{A3}$$

Next, we characterize the geometric role of R_p and show (15). Let $p \in \mathcal{S}_c$ be parameterized by the local coordinates

$$\bar{p} = \varphi(p) := (p_1, p_2, \dots, p_{c-1})^\top \in \mathbb{R}_{++}^{c-1} \tag{A4}$$

$$p = \varphi^{-1}(\bar{p}) = (\bar{p}_1, \dots, \bar{p}_{c-1}, 1 - \langle \mathbf{1}_{c-1}, \bar{p} \rangle)^\top \in \mathcal{S}_c. \tag{A5}$$

Choosing the canonical basis e_1, \dots, e_c on $\mathcal{S}_c \subset \mathbb{R}^c$, we obtain a basis of the tangent space $T_{c,0}$

$$e_j - e_c = d\varphi^{-1}(e_j), \quad j \in [c-1]. \tag{A6}$$

Using these vectors, a column of the matrix is

$$B := (e_1 - e_c, \dots, e_{c-1} - e_c) = \begin{pmatrix} I_{c-1} \\ -\mathbf{1}_{c-1}^\top \end{pmatrix} \in \mathbb{R}^{c \times (c-1)}, \tag{A7}$$

one has for any $v \in T_{c,0}$

$$v = B\bar{v} = \begin{pmatrix} \bar{v} \\ v_c \end{pmatrix} = \begin{pmatrix} \bar{v} \\ -\langle \mathbf{1}_{c-1}, \bar{v} \rangle \end{pmatrix}, \quad \bar{v} = (v_1, \dots, v_{c-1})^\top \tag{A8}$$

$$\bar{v} = B^\dagger v, \quad B^\dagger = (I_{c-1} \quad 0) \pi_{c,0}, \tag{A9}$$

where $B^\dagger := (B^\top B)^{-1} B^\top$ denotes the Moore–Penrose generalized inverse of B . Substituting this parameterization and evaluating the metric (8) gives

$$g_p(u, v) = \langle \bar{u}, B^\top \text{Diag}(p)^{-1} B\bar{v} \rangle = \left\langle \bar{u}, (I_{c-1} \quad -\mathbf{1}_{c-1}) \text{Diag}(p)^{-1} \begin{pmatrix} I_{c-1} \\ -\mathbf{1}_{c-1}^\top \end{pmatrix} \bar{v} \right\rangle \tag{A10}$$

$$= \left\langle \bar{u}, \left(\text{Diag}(\bar{p})^{-1} + \frac{1}{1 - \langle \mathbf{1}_{c-1}, \bar{p} \rangle} \mathbf{1}_{c-1} \mathbf{1}_{c-1}^\top \right) \bar{v} \right\rangle \tag{A11}$$

$$=: \langle \bar{u}, G(\bar{p})\bar{v} \rangle. \tag{A12}$$

Applying the Sherman–Morrison–Woodbury matrix inversion formula [42] (p. 9)

$$(A + xy^\top)^{-1} = A^{-1} - \frac{A^{-1}xy^\top A^{-1}}{1 + \langle y, A^{-1}x \rangle} \tag{A13}$$

yields

$$G(\bar{p})^{-1} = \text{Diag}(\bar{p}) - \frac{1}{1 - \langle \mathbf{1}_{c-1}, \bar{p} \rangle} \frac{\text{Diag}(\bar{p}) \mathbf{1}_{c-1} \mathbf{1}_{c-1}^\top \text{Diag}(\bar{p})}{1 + \frac{1}{1 - \langle \mathbf{1}_{c-1}, \bar{p} \rangle} \langle \mathbf{1}_{c-1}, \bar{p} \rangle} \tag{A14}$$

$$= \text{Diag}(\bar{p}) - \text{Diag}(\bar{p}) \mathbf{1}_{c-1} \mathbf{1}_{c-1}^\top \text{Diag}(\bar{p}) = \text{Diag}(\bar{p}) - \bar{p} \bar{p}^\top \tag{A15}$$

$$= R_{\bar{p}}. \tag{A16}$$

Let $v \in T_{c,0}$. Then, using the equations

$$p_c \stackrel{(A5)}{=} 1 - \langle \mathbf{1}_{c-1}, \bar{p} \rangle, \tag{A17}$$

$$R_{\bar{p}} \mathbf{1}_{c-1} = \bar{p} - \langle \mathbf{1}_{c-1}, \bar{p} \rangle \bar{p} = p_c \bar{p}, \tag{A18}$$

we have

$$R_p v = \begin{pmatrix} R_{\bar{p}} & -p_c \bar{p} \\ -p_c \bar{p}^\top & p_c - p_c^2 \end{pmatrix} \begin{pmatrix} \bar{v} \\ v_c \end{pmatrix} = \begin{pmatrix} R_{\bar{p}} \bar{v} - v_c R_{\bar{p}} \mathbf{1}_{c-1} \\ -\langle R_{\bar{p}} \mathbf{1}_{c-1}, \bar{v} \rangle + v_c p_c \langle \mathbf{1}_{c-1}, \bar{p} \rangle \end{pmatrix} \tag{A19}$$

$$= \begin{pmatrix} R_{\bar{p}} \bar{v} \\ -\langle \mathbf{1}_{c-1}, R_{\bar{p}} \bar{v} \rangle \end{pmatrix} - v_c \begin{pmatrix} R_{\bar{p}} \mathbf{1}_{c-1} \\ -\langle \mathbf{1}_{c-1}, R_{\bar{p}} \mathbf{1}_{c-1} \rangle \end{pmatrix} \tag{A20}$$

$$\stackrel{(A7)}{=} BR_{\bar{p}}(\bar{v} - v_c \mathbf{1}_{c-1}). \tag{A21}$$

Now, consider any smooth function $f: \mathcal{S}_c \rightarrow \mathbb{R}$. Then,

$$\partial_{\bar{p}_i} (f \circ \varphi^{-1}(\bar{p})) = \sum_{j \in [c]} \partial_j f(p) \partial_{\bar{p}_i} \varphi^{-1}(\bar{p}) \stackrel{(A5)}{=} \partial_i f(p) - \partial_c f(p), \tag{A22}$$

$$\partial_{\bar{p}} (f \circ \varphi^{-1}(\bar{p})) = \overline{\partial f(p)} - \partial_c f(p) \mathbf{1}_{c-1}. \tag{A23}$$

Comparing the last equation and (A21) shows that

$$R_p \partial f(p) = BR_{\bar{p}} \partial_{\bar{p}} (f \circ \varphi^{-1}(\bar{p})) \stackrel{(A14)}{=} BG(p)^{-1} \partial_{\bar{p}} (f \circ \varphi^{-1}(\bar{p})), \tag{A24}$$

which proves (15). \square

Appendix A.2. Proofs of Section 3

Proof of Lemma 1. Let $v(t) \in T_{c,0}$ be a smooth curve with $\dot{v}(t) = u$. Then,

$$\frac{d}{dt} \exp_p(v(t)) = \frac{d}{dt} \frac{p \cdot e^{v(t)}}{\langle p, e^{v(t)} \rangle} = \frac{p \cdot u \cdot e^{v(t)}}{\langle p, e^{v(t)} \rangle} - \langle p, u \cdot e^{v(t)} \rangle \frac{p \cdot e^{v(t)}}{\langle p, e^{v(t)} \rangle^2} \tag{A25}$$

$$= \exp_p(v(t)) \cdot u - \langle u, \exp_p(v(t)) \rangle \exp_p(v(t)) = R_{\exp_p(v(t))} u. \tag{A26}$$

Similarly, for a smooth curve $p(t) \in \mathcal{S}_c$ with $\dot{p}(t) = u$, one has

$$\frac{d}{dt} \exp_{p(t)}(v) = \frac{d}{dt} \frac{p(t) \cdot e^v}{\langle p(t), e^v \rangle} = \frac{\dot{p}(t) \cdot e^v}{\langle p(t), e^v \rangle} - \langle \dot{p}(t), e^v \rangle \frac{p(t) \cdot e^v}{\langle p(t), e^v \rangle^2} \tag{A27}$$

$$= \exp_{p(t)}(v) \cdot \frac{u}{p(t)} - \left\langle \frac{u}{p(t)}, \exp_{p(t)}(v) \right\rangle \exp_{p(t)}(v) = R_{\exp_{p(t)}(v)} \frac{u}{p(t)}. \quad \square \tag{A28}$$

Proof of Theorem 1. Let

$$q(t) = L_{p(t)}(D) \tag{A29}$$

where $p(t)$ solves (45). Using (43), (47) and (A29), we obtain

$$\dot{q} = d_p L_{p(t)}(D)[\dot{p}(t)] = R_{q(t)} \left(\frac{\dot{p}(t)}{p(t)} \right) \stackrel{(45)}{=} R_{q(t)}(q(t) - \langle p(t), q(t) \rangle \mathbf{1}_c) \stackrel{(13)}{=} R_{q(t)} q(t), \tag{A30}$$

which shows (49). Then, $p(t) = \exp_{\mathbf{1}_{\mathcal{S}_c}}(r(t))$, and differentiating (50) yields $r(t) = \int_0^t q(\tau) d\tau$

$$\dot{p}(t) \stackrel{(46)}{=} R_{\exp_{\mathbf{1}_{\mathcal{S}_c}}(r(t))} \dot{r}(t) \stackrel{(50)}{=} R_{p(t)} q(t) \stackrel{(A29)}{=} R_{p(t)} L_{p(t)}(D), \tag{A31}$$

which proves the equivalence of (45), (48) and (49). \square

Proof of Corollary 1. The solution $p(t)$ to (45) is given by (48) and (49). Proposition 1 and Equation (15) show that (49) is the Riemannian ascent flow of the function $\mathcal{S}_c \ni q \mapsto \frac{1}{2} \|q\|^2$. The stationary points satisfy

$$R_q q = (q - \|q\|^2) \cdot q = 0 \tag{A32}$$

and form the set

$$Q^* := \left\{ q^* = \frac{1}{|\mathcal{J}^*|} \sum_{j \in \mathcal{J}^*} e_j : \mathcal{J}^* \subseteq [c] \right\}. \tag{A33}$$

The case of $\mathcal{J}^* = [c]$, i.e., $q^* = \mathbb{1}_{\mathcal{S}_c}$, can be ruled out if $\frac{D}{\langle \mathbb{1}_c, D \rangle} \neq \mathcal{S}_c$, which is always the case in practice, where D corresponds to real data (measurement and observation). The global maxima correspond to the vertices of $\Delta_c = \overline{\mathcal{S}_c}$, i.e., $|\mathcal{J}^*| = 1$. The remaining stationary points are local maxima and degenerate, since vectors D with non-unique minimal components form a negligible null set. In any case, $\lim_{t \rightarrow \infty} p(t) \stackrel{(50)}{=} \lim_{t \rightarrow \infty} q(t) = q^*$, depending on the index set \mathcal{J}^* determined by D . \square

Appendix A.3. Proofs of Section 4

Proof of Proposition 2. The Riemannian gradient is defined by [43] (p. 337)

$$0 = df[X] - g_\rho(\text{grad}_\rho f, X) \stackrel{(29)}{=} \langle \partial f, X \rangle - \langle \mathbb{T}_\rho[\text{grad}_\rho f], X \rangle \tag{A34}$$

$$= \langle \partial f - \mathbb{T}_\rho[\text{grad}_\rho f], X \rangle, \quad \forall X \in \mathcal{H}_{c,0}. \tag{A35}$$

Choosing the parameterization $X = Y - \text{tr}(Y)I \in \mathcal{H}_{c,0}$ with $Y \in \mathcal{H}_c$, we further obtain

$$0 = \langle \partial f - \mathbb{T}_\rho[\text{grad}_\rho f], Y \rangle - \text{tr}(Y) \text{tr}(\partial f - \mathbb{T}_\rho[\text{grad}_\rho f]) \tag{A36}$$

$$= \langle \partial f - \mathbb{T}_\rho[\text{grad}_\rho f] - \text{tr}(\partial f - \mathbb{T}_\rho[\text{grad}_\rho f])I, Y \rangle, \quad \forall Y \in \mathcal{H}_c. \tag{A37}$$

The left factor must vanish. Applying linear mapping \mathbb{T}_ρ^{-1} and solving for $\text{grad}_\rho f$ yields

$$\text{grad}_\rho f = \mathbb{T}_\rho^{-1}[\partial f] - \text{tr}(\partial f - \mathbb{T}_\rho[\text{grad}_\rho f])\mathbb{T}_\rho^{-1}[I]. \tag{A38}$$

Since $\text{grad}_\rho f \in \mathcal{H}_{c,0}$, taking the trace on both sides and using $\text{tr} \mathbb{T}_\rho^{-1}[I] = \text{tr} \rho = 1$ yields

$$0 = \text{tr} \mathbb{T}_\rho^{-1}[\partial f] - \text{tr} \partial f + \text{tr} \mathbb{T}_\rho[\text{grad}_\rho f]. \tag{A39}$$

Substituting the last two summands into the previous equation yields

$$\text{grad}_\rho f = \mathbb{T}_\rho^{-1}[\partial f] - (\text{tr} \mathbb{T}_\rho^{-1}[\partial f])\rho \tag{A40}$$

$$= \mathbb{T}_\rho^{-1}[\partial f] - \langle \rho, \partial f \rangle \rho, \tag{A41}$$

where the last equation follows from (30). \square

Proof of Lemma 2. The equation $\Pi_{c,0} \circ \mathfrak{R}_\rho = \mathfrak{R}_\rho$ follows from $\mathfrak{R}_\rho[X] \in \mathcal{H}_{c,0}$; hence,

$$\text{tr} \mathfrak{R}_\rho[X] \stackrel{(64)}{=} \text{tr} \mathbb{T}_\rho^{-1}[X] - \langle \rho, X \rangle \text{tr} \rho \stackrel{(30)}{=} \langle \rho, X \rangle - \langle \rho, X \rangle = 0. \tag{A42}$$

Thus,

$$\Pi_{c,0} \circ \mathfrak{R}_\rho[X] = \mathfrak{R}_\rho[X] = \mathfrak{R}_\rho[X] - \frac{\text{tr} X}{c} (\rho - \underbrace{\langle \rho, I \rangle}_{=1} \rho) \tag{A43}$$

$$= \mathfrak{R}_\rho[X] - \frac{\text{tr} X}{c} \mathfrak{R}_\rho[I] = \mathfrak{R}_\rho \left[X - \frac{\text{tr} X}{c} I \right] \stackrel{(24)}{=} \mathfrak{R}_\rho \circ \Pi_{c,0}[X]. \quad \square \tag{A44}$$

Proof of Lemma 3. Using (24), we compute

$$\exp_m(\Pi_{c,0}[Z]) = \exp_m\left(Z - \frac{\text{tr} Z}{c} I\right) = e^{\frac{\text{tr} Z}{c}} \exp_m(Z), \tag{A45}$$

where the last equation holds, since Z and I commute. Substitution into (66) cancels the scalar factor $e^{\frac{\text{tr} Z}{c}}$ and shows (68). \square

Proof of Proposition 3. We show $\Gamma \circ \Gamma^{-1} = \text{id}_{\mathcal{D}_c}$ and $\Gamma^{-1} \circ \Gamma = \text{id}_{\mathcal{H}_{c,0}}$. As for the first relation, we compute

$$\Gamma \circ \Gamma^{-1}(\rho) = \exp_m\left(\Gamma^{-1}(\rho) - \psi(\Gamma^{-1}(\rho))I\right) \tag{A46}$$

$$= \exp_m\left(\log_m \rho - \frac{\text{tr}(\log_m \rho)}{c} I - \log\left(\text{tr} \exp_m\left(\log_m \rho - \frac{\text{tr}(\log_m \rho)}{c} I\right)\right)I\right) \tag{A47}$$

and since $\log_m \rho$ and I commute,

$$= \exp_m\left(\log_m \rho - \frac{\text{tr}(\log_m \rho)}{c} I - \log \text{tr}\left(e^{-\frac{1}{c} \text{tr}(\log_m \rho)} \rho\right)I\right) \tag{A48}$$

$$\stackrel{\text{tr} \rho = 1}{=} \exp_m(\log_m \rho) \tag{A49}$$

$$= \rho. \tag{A50}$$

As for the second relation, we compute

$$\Gamma^{-1} \circ \Gamma(X) = \Pi_{c,0}[\log_m \circ \Gamma(X)] = \Pi_{c,0}[\log_m \circ \exp_m(X - \psi(X)I)] \tag{A51}$$

$$= \Pi_{c,0}[X] - \psi(X)\Pi_{c,0}[I] = \Pi_{c,0}[X] \tag{A52}$$

$$= X, \tag{A53}$$

since $X \in \mathcal{H}_{c,0}$ by assumption. \square

Proof of Lemma 4. In view of Definition (66) of Γ , we compute using the chain rule

$$d\Gamma(H)[Y] = \frac{d}{dt} \exp_m(H + tY - \psi(H + tY)I) \Big|_{t=0} \tag{A54}$$

$$= d \exp_m(H - \psi(H)I)[Y - d\psi(H)[Y]I] \tag{A55}$$

$$\stackrel{(28)}{=} \mathbb{T}_\rho^{-1}[Y - d\psi(H)[Y]I]. \tag{A56}$$

Furthermore,

$$d\psi(H)[Y] \stackrel{(67)}{=} \frac{1}{\text{tr} \exp_m(H)} \text{tr}(d \exp_m(H)[Y]) \tag{A57}$$

$$\stackrel{(28)}{=} \frac{1}{\text{tr} \exp_m(H)} \text{tr}(\mathbb{T}_{\exp_m(H)}^{-1}[Y]), \quad \exp_m(H) \stackrel{(66)}{=} (\text{tr} \exp_m(H))\Gamma(H) \tag{A58}$$

$$\stackrel{(30)}{=} \frac{1}{\text{tr} \exp_m(H)} \langle \exp_m(H), Y \rangle \tag{A59}$$

$$\stackrel{(66)}{=} \langle \Gamma(H), Y \rangle = \langle \rho, Y \rangle, \tag{A60}$$

where the last equation follows from the assumption that $\rho = \Gamma(H)$. Substitution into (A56) yields (70). Regarding (71), using expression (69) for Γ^{-1} , we compute

$$d\Gamma^{-1}(\rho)[X] = \Pi_{c,0} \circ d \log_m(\rho)[X] \tag{A61}$$

$$\stackrel{(27)}{=} \Pi_{c,0} \circ \mathbb{T}_\rho[X], \tag{A62}$$

which verifies (71). \square

Proof of Proposition 4. The e-geodesic connecting the two points $Q, R \in \mathcal{D}_c$ is given by [24] (Section V)

$$\Gamma(K + tA), \quad t \in [0, 1], \quad K = \log_m Q, \quad A = \log_m R - \log_m Q. \quad (A63)$$

Setting $\Gamma^{-1}(\rho) = \Pi_{c,0}[K]$ and $\mathbb{T}_\rho[X] = A$ yields (74), since the orthogonal projections $\Pi_{c,0}$ onto $\mathcal{H}_{c,0}$ are also implicitly carried out in (A63) due to Lemma 3. Expression (73) is equal to (74) due to (71). It remains to be verified that the geodesic emanates at ρ in the direction of X . We compute

$$\gamma_{\rho,X}^{(e)}(0) = \Gamma(\Gamma^{-1}(\rho)) = \rho \quad (A64)$$

$$\frac{d}{dt} \gamma_{\rho,X}^{(e)}(0) = \frac{d}{dt} \Gamma(\Gamma^{-1}(\rho) + t d\Gamma^{-1}(\rho)[X])|_{t=0} \quad (A65)$$

$$= d\Gamma(\Gamma^{-1}(\rho)) [d\Gamma^{-1}(\rho)[X]] = \text{id}[X] = X. \quad (A66)$$

\square

Proof of Corollary 2. Setting

$$\mu = \text{Exp}_\rho^{(e)}(X) \stackrel{(73)}{=} \Gamma(\Gamma^{-1}(\rho) + d\Gamma^{-1}(\rho)[X]) \quad (A67)$$

we solve for X ,

$$\Gamma^{-1}(\mu) = \Gamma^{-1}(\rho) + d\Gamma^{-1}(\rho)[X] \quad (A68)$$

$$d\Gamma^{-1}(\rho)[X] = \Gamma^{-1}(\mu) - \Gamma^{-1}(\rho) \quad (A69)$$

$$X = d\Gamma(\Gamma^{-1}(\rho)) [\Gamma^{-1}(\mu) - \Gamma^{-1}(\rho)], \quad (A70)$$

which shows (75) and where $d\Gamma(\Gamma^{-1}(\rho))^{-1} = d\Gamma^{-1}(\rho)$ was used to obtain the last equation. \square

Proof of Lemma 5. We compute

$$\text{Exp}_\rho^{(e)} \circ \mathfrak{R}_\rho[X] \stackrel{(73)}{=} \Gamma(\Gamma^{-1}(\rho) + \Pi_{c,0} \circ \mathbb{T}_\rho[\mathfrak{R}_\rho[X]]) \quad (A71)$$

$$\stackrel{(64)}{=} \Gamma(\Gamma^{-1}(\rho) + \Pi_{c,0} \circ \mathbb{T}_\rho[\mathbb{T}_\rho^{-1}[X] - \langle \rho, X \rangle \rho]) \quad (A72)$$

$$\stackrel{\rho = \mathbb{T}_\rho^{-1}[I]}{=} \Gamma(\Gamma^{-1}(\rho) + \Pi_{c,0}[X - \langle \rho, X \rangle I]) \quad (A73)$$

$$= \Gamma(\Gamma^{-1}(\rho) + X) \quad (A74)$$

and omit the projection map $\Pi_{c,0}$ in the last equation due to Lemma 2 or Lemma 3. \square

Proof of Lemma 6. We compute

$$d \exp_\rho(X)[Y] \stackrel{(77)}{=} d\Gamma(\Gamma^{-1}(\rho) + X)[Y] \stackrel{(70)}{=} \mathbb{T}_{\exp_\rho(X)}^{-1}[Y - \langle \exp_\rho(X), Y \rangle I] \quad (A75)$$

$$\stackrel{\mathbb{T}_\rho^{-1}[I] = \rho}{=} \mathbb{T}_{\exp_\rho(X)}^{-1}[Y] - \langle \exp_\rho(X), Y \rangle \exp_\rho(X) \quad (A76)$$

$$\stackrel{(64)}{=} \mathfrak{R}_{\exp_\rho(X)}[Y]. \quad \square \quad (A77)$$

Proof of Lemma 7. We compute

$$L_\rho(D) \stackrel{(65)}{=} \exp_\rho(-\Pi_{c,0}[D]) \stackrel{(76)}{=} \exp_\rho(-D) \tag{A78}$$

$$\stackrel{(77)}{=} \Gamma(\Gamma^{-1}(\rho) - D) \stackrel{(69)}{=} \Gamma(\Pi_{c,0}[\log_m \rho] - D) \tag{A79}$$

$$\stackrel{(92)}{=} \Gamma\left(Q(\log_m \Lambda_\rho)Q^\top - \frac{1}{c}(\text{tr } \Lambda_\rho)I_c - Q\Lambda_D Q^\top\right) \tag{A80}$$

$$\stackrel{(66)}{=} Q \frac{\text{Diag}(e^{\log \lambda_\rho - \langle \mathbb{1}_c, \lambda_\rho \rangle \mathbb{1}_{S_c} - \lambda_D})}{\langle \mathbb{1}_c, e^{\log \lambda_\rho - \langle \mathbb{1}_c, \lambda_\rho \rangle \mathbb{1}_{S_c} - \lambda_D} \rangle} Q^\top = Q \text{Diag} \left(\frac{e^{\log \lambda_\rho - \lambda_D}}{\langle \mathbb{1}_c, e^{\log \lambda_\rho - \lambda_D} \rangle} \right) Q^\top \tag{A81}$$

$$\stackrel{(44)}{=} Q \text{Diag} (\exp_{\lambda_\rho}(-\lambda_D)) Q^\top. \tag{A82}$$

□

Proof of Proposition 5. Writing $\rho_0 = \rho(0) = \mathbb{1}_{D_c}$ and $\text{Diag}(\lambda_D) = \Lambda_D$, we have

$$L_{\rho_0}(D) = Q \text{Diag} (\exp_{\mathbb{1}_{S_c}}(-\lambda_D)) Q^\top$$

according to Lemma 7 and

$$\dot{\rho}(0) \stackrel{(91)}{=} \mathbb{T}_{\rho_0}^{-1}[L_{\rho_0}(D)] - \langle \rho_0, L_{\rho_0}(D) \rangle \rho_0 \tag{A83}$$

$$\stackrel{(28)}{=} \int_0^1 \left(\frac{1}{c}I_c\right)^{1-s} Q \text{Diag} (\exp_{\mathbb{1}_{S_c}}(-\lambda_D)) Q^\top \left(\frac{1}{c}I_c\right)^s ds \tag{A84}$$

$$- \frac{1}{c^2} \text{tr} \left(\underbrace{Q \text{Diag} (\exp_{\mathbb{1}_{S_c}}(-\lambda_D)) Q^\top}_{=1} \right) I_c \tag{A85}$$

$$= \frac{1}{c} Q \text{Diag} (\exp_{\mathbb{1}_{S_c}}(-\lambda_D)) Q^\top - \frac{1}{c^2} I_c \tag{A86}$$

$$= \frac{1}{c} Q \text{Diag} (\exp_{\mathbb{1}_{S_c}}(-\lambda_D) - \mathbb{1}_{S_c}) Q^\top. \tag{A87}$$

Comparing this equation to the single-vertex flow (45) at time $t = 0$,

$$\dot{\rho}(0) = \frac{1}{c} (\exp_{\mathbb{1}_{S_c}}(-D) - \frac{1}{c} \underbrace{\langle \mathbb{1}_c, \exp_{\mathbb{1}_{S_c}}(-D) \rangle \mathbb{1}_c}_{=1}) = \frac{1}{c} (\exp_{\mathbb{1}_{S_c}}(-D) - \mathbb{1}_{S_c}) \tag{A88}$$

shows that

$$\dot{\rho}(0) = Q \text{Diag} (\dot{\lambda}_\rho(0)) Q^\top, \tag{A89}$$

that is

$$\rho(t) = Q \lambda_\rho(t) Q^\top, \tag{A90}$$

where $\lambda(t)$ solves the single-vertex assignment flow Equation (45) of the form

$$\dot{\lambda}_\rho = \mathfrak{A}_{\lambda_\rho} L_{\lambda_\rho}(\lambda_D). \tag{A91}$$

Corollary 1 completes the proof. □

Proof of Lemma 8. Let

$$H_i = \Gamma^{-1}(\rho_i) \stackrel{(69)}{=} \Pi_{c,0} \log_m \rho_i, \quad i \in \mathcal{V}. \tag{A92}$$

Then,

$$(\text{Exp}_{\rho_i}^{(e)})^{-1}(L_{\rho_k}(D_k)) \stackrel{(89)}{=} (\text{Exp}_{\rho_i}^{(e)})^{-1} \circ \Gamma(\Gamma^{-1}(\rho_k) - D_k) \tag{A93}$$

$$\stackrel{(77)}{=} d\Gamma(\Gamma^{-1}(\rho_i))[\Gamma^{-1}(\rho_k) - D_k - \Gamma^{-1}(\rho_i)] \tag{A94}$$

$$= d\Gamma(H_i)[H_k - D_k - H_i]. \tag{A95}$$

Substituting this expression into (100) yields

$$S_i(\rho) \stackrel{(73)}{=} \Gamma\left(H_i + \underbrace{d\Gamma^{-1}(\rho_i) \circ d\Gamma(H_i)}_{=I} \left[\sum_{k \in \mathcal{N}_i} \omega_{ik}(H_k - D_k) - H_i \right]\right) \tag{A96}$$

$$= \Gamma\left(\sum_{k \in \mathcal{N}_i} \omega_{ik}(H_k - D_k)\right). \tag{A97}$$

Substituting (A92) and omitting the projection map $\Pi_{c,0}$ due to Lemma 3 yields (101). \square

Proof of Proposition 6. Substituting as in the proof of Lemma 8, we obtain

$$0 = d\Gamma(\Gamma^{-1}(\bar{\rho})) \left[\sum_{k \in \mathcal{N}_i} \omega_{ik}(\Pi_{c,0} \log_{\text{gm}} \rho_k - D_k) - \Gamma^{-1}(\bar{\rho}) \right]. \tag{A98}$$

Since $d\Gamma$ is one-to-one, the expression inside the brackets must vanish. Solving for $\bar{\rho}$ and omitting the projection map $\Pi_{c,0}$ due to Lemma 3 yields (101). \square

Proof of Proposition 7. Let $\rho(t)$ solve (104) and denote the argument of the replicator operator \mathfrak{R}_ρ on the right-hand side by

$$\mu(t) := S(\rho(t)), \tag{A99}$$

which yields (108) and (104), respectively, whereas (109) remains to be verified. Differentiation yields

$$\dot{\mu}_i = dS_i(\rho)[\dot{\rho}] \tag{A100}$$

$$\stackrel{(101)}{=} d\Gamma\left(\sum_{k \in \mathcal{N}_i} \omega_{ik}(\log_{\text{gm}} \rho_k - D_k)\right) \left[\sum_{k \in \mathcal{N}_i} \omega_{ik} \mathbb{T}_{\rho_k}[\dot{\rho}_k] \right] \tag{A101}$$

$$\stackrel{(101)}{=} d\Gamma(\Gamma^{-1}(S_i(\rho))) \left[\sum_{k \in \mathcal{N}_i} \omega_{ik} \mathbb{T}_{\rho_k}[\dot{\rho}_k] \right] \tag{A102}$$

$$\stackrel{(70)}{=} \mathbb{T}_{S_i(\rho)}^{-1} \left[\sum_{k \in \mathcal{N}_i} \omega_{ik} \mathbb{T}_{\rho_k}[\dot{\rho}_k] - \left\langle S_i(\rho), \sum_{k \in \mathcal{N}_i} \omega_{ik} \mathbb{T}_{\rho_k}[\dot{\rho}_k] \right\rangle I \right] \tag{A103}$$

$$\stackrel{\mathbb{T}_\rho^{-1}[I]=\rho}{=} \mathbb{T}_{S_i(\rho)}^{-1} \left[\sum_{k \in \mathcal{N}_i} \omega_{ik} \mathbb{T}_{\rho_k}[\dot{\rho}_k] \right] - \left\langle S_i(\rho), \sum_{k \in \mathcal{N}_i} \omega_{ik} \mathbb{T}_{\rho_k}[\dot{\rho}_k] \right\rangle S_i(\rho) \tag{A104}$$

$$\stackrel{(64)}{=} \mathfrak{R}_{S_i(\rho)} \left[\sum_{k \in \mathcal{N}_i} \omega_{ik} \mathbb{T}_{\rho_k}[\dot{\rho}_k] \right] \stackrel{(A99)}{=} \sum_{k \in \mathcal{N}_i} \omega_{ik} \mathfrak{R}_{\mu_i}[\mathbb{T}_{\rho_k}[\dot{\rho}_k]] \tag{A105}$$

$$\stackrel{(108)}{=} \sum_{k \in \mathcal{N}_i} \omega_{ik} \mathfrak{R}_{\mu_i}[\mathbb{T}_{\rho_k}[\mathfrak{R}_{\rho_k}[\mu_k]]] \tag{A106}$$

$$\stackrel{(64)}{=} \sum_{k \in \mathcal{N}_i} \omega_{ik} \mathfrak{R}_{\mu_i}[\mathbb{T}_{\rho_k}[\mathbb{T}_{\rho_k}^{-1}[\mu_k] - \langle \rho_k, \mu_k \rangle T_{\rho_k}^{-1}[I]]] \tag{A107}$$

$$= \sum_{k \in \mathcal{N}_i} \omega_{ik} \mathfrak{R}_{\mu_i}[\mu_k - \langle \rho_k, \mu_k \rangle I] \stackrel{(65)}{=} \sum_{k \in \mathcal{N}_i} \omega_{ik} \mathfrak{R}_{\mu_i}[\mu_k] \tag{A108}$$

$$\stackrel{(107)}{=} \mathfrak{R}_{\mu_i}[\Omega[\mu]_i]. \tag{A109}$$

The initial condition for ρ is given by (104). The initial condition for μ follows from (A99). \square

Proof of Proposition 8. We compute using (110) and (111)

$$J(\mu) = -\frac{1}{2} \sum_{j \in \mathcal{V}} \langle \mu_j, \sum_{k \in \mathcal{N}_j} \omega_{jk} \mu_k \rangle \tag{A110}$$

$$\partial_{\mu_i} J(\mu) = -\frac{1}{2} \left(\sum_{k \in \mathcal{N}_i} \omega_{ik} \mu_k + \sum_{j \in \mathcal{N}_i} \omega_{ji} \mu_j \right) \stackrel{(110)}{=} -\Omega[\mu]_i. \tag{A111}$$

Consequently, according to Proposition 2 and (64),

$$-\text{grad}_{\mu_i} J(\mu) = -\mathfrak{R}_{\mu_i}[\partial_{\mu_i} J(\mu)] = \mathfrak{R}_{\mu_i}[\Omega[\mu]_i] \tag{A112}$$

which shows that the right-hand sides of (114) and (109) are equal. \square

Proof of Proposition 9. Starting with (115), we compute

$$J(\mu) = -\frac{1}{2} \langle \mu, \Omega[\mu] \rangle = \frac{1}{2} \langle \mu, (\text{id} - \Omega - \text{id})[\mu] \rangle = \frac{1}{2} (\langle \mu, L_G[\mu] \rangle - \|\mu\|^2). \tag{A113}$$

The quadratic form involving L_G remains to be rewritten.

$$\langle \mu, L_G[\mu] \rangle = \sum_{i \in \mathcal{V}} \langle \mu_i, \mu_i - \Omega[\mu]_i \rangle = \sum_{i \in \mathcal{V}} \left(\|\mu_i\|^2 - \left\langle \mu_i, \sum_{k \in \mathcal{N}_i} \omega_{ik} \mu_k \right\rangle \right) \tag{A114}$$

$$= \sum_{i \in \mathcal{V}} \left(\underbrace{\sum_{k \in \mathcal{N}_i} \omega_{ik}}_{=1} \|\mu_i\|^2 - \sum_{k \in \mathcal{N}_i} \omega_{ik} \langle \mu_i, \mu_k \rangle \right) = \sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ik} (\langle \mu_i, \mu_i - \mu_k \rangle) \tag{A115}$$

$$= \frac{1}{2} \left(\sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ik} (\langle \mu_i, \mu_i - \mu_k \rangle) + \sum_{k \in \mathcal{V}} \sum_{i \in \mathcal{N}_k} \omega_{ki} (\langle \mu_k, \mu_k - \mu_i \rangle) \right), \tag{A116}$$

where the last sum results from the first one by interchanging indices i and k . Using the symmetry relations (110) and (111), we rewrite the second sum as

$$\langle \mu, L_G[\mu] \rangle = \frac{1}{2} \left(\sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ik} (\langle \mu_i, \mu_i - \mu_k \rangle) + \sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ik} (\langle \mu_k, \mu_k - \mu_i \rangle) \right) \tag{A117}$$

$$= \frac{1}{2} \sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ik} (\langle \mu_i, \mu_i - \mu_k \rangle + \langle \mu_k, \mu_k - \mu_i \rangle) \tag{A118}$$

$$= \frac{1}{2} \sum_{i \in \mathcal{V}} \sum_{k \in \mathcal{N}_i} \omega_{ik} \|\mu_i - \mu_k\|^2 \tag{A119}$$

which completes the proof. \square

Proof of Lemma 9. We start with claim (b).

(b) Let $\mu \in \mathcal{D}_{\Pi}$ and $X \in T_{\mu} \mathcal{D}_{\Pi}$. Suppose that vector X is represented by a curve $\eta : (-\varepsilon, \varepsilon) \rightarrow \mathcal{D}_{\Pi}$ such that $\eta(0) = \mu$ and $\eta'(0) = X$. In view of the definition (123) of \mathcal{D}_{Π} , we have

$$\eta(t) = \sum_{i \in [l]} \frac{p_i(t)}{\text{tr} \pi_i} \pi_i \implies X = \sum_{i \in [l]} \frac{p'_i(0)}{\text{tr} \pi_i} \pi_i. \tag{A120}$$

Consequently, if $\mathcal{U} = \{u_1, \dots, u_c\}$ is a basis of \mathbb{C}^c that diagonalizes μ , then the tangent vector X is also diagonal in this basis \mathcal{U} , and X commutes with μ , i.e., $[\mu, X] = 0$ and $X \in T_\mu^c \mathcal{D}_c$. This proves (b).

(a) The bijection $\mathcal{D}_\Pi \rightarrow \mathcal{S}_l$ is explicitly given by

$$\Phi_\Pi: \mathcal{D}_\Pi \rightarrow \mathcal{S}_l, \quad \sum_{i \in [l]} \frac{p_i}{\text{tr} \pi_i} \pi_i \mapsto (p_1, \dots, p_l). \tag{A121}$$

This is bijective according to the definition of \mathcal{D}_Π . It remains to be shown that it is an isometry. Consider another tangent vector $Y \in T_\mu \mathcal{D}_\Pi$. We know that μ, X, Y can all be diagonalized in a common eigenbasis. This basis is denoted again by \mathcal{U} . Then, we can write

$$\mu = \sum_{i \in [c]} \tilde{p}_i u_i u_i^*, \quad X = \sum_{i \in [c]} \tilde{x}_i u_i u_i^*, \quad Y = \sum_{i \in [c]} \tilde{y}_i u_i u_i^* \tag{A122}$$

and compute

$$t^* g_{\text{BKM}, \mu}(X, Y) = \int_0^\infty \text{tr} \left(X(\mu + \lambda I)^{-1} Y(\mu + \lambda I)^{-1} \right) d\lambda \tag{A123}$$

$$= \sum_{i \in [c]} \int_0^\infty \text{tr} \left(\frac{\tilde{x}_i \tilde{y}_i}{(\tilde{p}_i + \lambda)^2} u_i u_i^* \right) d\lambda \tag{A124}$$

$$= \sum_{i \in [c]} \frac{\tilde{x}_i \tilde{y}_i}{\tilde{p}_i}. \tag{A125}$$

Note that the vector $\tilde{p} = (\tilde{p}_1, \dots, \tilde{p}_c)$ comes from $\mu \in \mathcal{D}_\Pi$. Therefore, the value $p_j / \text{tr} \pi_j$ must occur $\text{tr} \pi_j$ times in \tilde{p} for every $j \in [l]$. This observation also holds for the vectors $\tilde{x} = (\tilde{x}_1, \dots, \tilde{x}_c)$ and $\tilde{y} = (\tilde{y}_1, \dots, \tilde{y}_c)$. Thus, the sum above can be reduced to

$$\sum_{i \in [c]} \frac{\tilde{x}_i \tilde{y}_i}{\tilde{p}_i} = \sum_{j \in [l]} \frac{x_j y_j}{p_j}, \tag{A126}$$

where $(p_1, \dots, p_l) = \Phi(\mu)$, $(x_1, \dots, x_l) = d\Phi[X]$ and $(y_1, \dots, y_l) = d\Phi[Y]$. Taking into account that (x_1, \dots, x_l) and (y_1, \dots, y_l) are the images of X, Y under the differential $d\Phi$, we conclude

$$t^* g_{\text{BKM}, \mu}(X, Y) = \sum_{i \in [l]} \frac{x_i y_i}{p_i} \stackrel{(8)}{=} g_{\text{FR}, \Phi(\mu)}(d\Phi(X), d\Phi(Y)). \tag{A127}$$

This proves part (a).

(c) Part (c) is about the commutativity of the diagram.

$$\begin{array}{ccc} \mathcal{D}_\Pi & \xrightarrow{\alpha_\Pi} & \mathcal{D}_{\Pi_\mathcal{U}} \\ \downarrow \Phi_\Pi & & \downarrow \Phi_{\Pi_\mathcal{U}} \\ \mathcal{S}_l & \xrightarrow{\beta_\Pi} & \mathcal{S}_c. \end{array} \tag{A128}$$

The horizontal arrows can be described as follows. Recall that $\Pi = \{\pi_1, \dots, \pi_l\}$. $k_i = \text{tr} \pi_i$ represents the dimensions of the images of the projectors π_i . For a fixed $p = (p_1, \dots, p_l) \in \mathcal{S}_l$, set

$$P = (P_1, \dots, P_c) := \underbrace{(p_1/k_1, \dots, p_1/k_1)}_{k_1 \text{ times}}, \dots, \underbrace{(p_l/k_l, \dots, p_l/k_l)}_{k_l \text{ times}} \in \mathcal{S}_c. \tag{A129}$$

Then, α_Π is given by

$$\alpha_\Pi \left(\sum_{i \in [l]} \frac{p_i}{k_i} \pi_i \right) = \sum_{j \in [c]} P_j u_j u_j^* \in \mathcal{D}_{\Pi, \mathcal{U}} \quad \text{and} \quad \beta_\Pi(p_1, \dots, p_l) = (P_1, \dots, P_c) \quad (\text{A130})$$

Diagram (A128) commutes according to the definition of the Φ maps.

□

Proof of Lemma 10.

- (i) Due to the commutativity of the components μ_i of $\mu \in \mathcal{Q}$, we can simplify the expression for the vector field of the QSAF as follows.

$$\mathfrak{R}_\mu[\Omega[\mu]]_i \stackrel{(106)}{=} \mathfrak{R}_{\mu_i}[\Omega[\mu]_i] \quad (\text{A131})$$

$$\stackrel{(64)}{\stackrel{(28)}}{=} \sum_{k \in \mathcal{N}_i} \omega_{ik} \left(\int_0^1 \mu_i^{1-\lambda} \mu_k \mu_i^\lambda d\lambda - \text{tr}(\mu_i \mu_k) \mu_i \right) \quad (\text{A132})$$

$$= \sum_{k \in \mathcal{N}_i} \omega_{ik} (\mu_i \mu_k - \text{tr}(\mu_i \mu_k) \mu_i). \quad (\text{A133})$$

Define $\mu \in \mathcal{D}_{\Pi, c}$ such that all the components μ_i can be written as

$$\mu_i = \sum_{r \in [l]} \frac{p_r^i}{\text{tr} \pi_r} \pi_r, \quad p^i = (p_1^i, \dots, p_l^i) \in \mathcal{S}_l, \quad i \in \mathcal{V}. \quad (\text{A134})$$

Then, we can further simplify

$$\mu_i \mu_k = \sum_{r \in [l]} \frac{p_r^i p_r^k}{(\text{tr} \pi_r)^2} \pi_r \quad \text{and} \quad \text{tr}(\mu_i \mu_k) = \sum_{r \in [l]} \frac{p_r^i p_r^k}{\text{tr} \pi_r} \quad (\text{A135})$$

and, consequently,

$$\sum_{k \in \mathcal{N}_i} \omega_{ik} (\mu_i \mu_k - \text{tr}(\mu_i \mu_k) \mu_i) = \sum_{r \in [l]} \sum_{k \in \mathcal{N}_i} \omega_{ik} \left(\frac{p_r^k}{\text{tr} \pi_r} - \left(\sum_{s \in [l]} \frac{p_s^i p_s^k}{\text{tr} \pi_s} \right) \right) \frac{p_r^i}{\text{tr} \pi_r} \pi_r \quad (\text{A136})$$

$$= \sum_{r \in [l]} \frac{x_r}{\text{tr} \pi_r} \pi_r, \quad (\text{A137})$$

where

$$x_r := \sum_{k \in \mathcal{N}_i} \omega_{ik} \left(\frac{p_r^k}{\text{tr} \pi_r} - \left(\sum_{s \in [l]} \frac{p_s^i p_s^k}{\text{tr} \pi_s} \right) \right) p_r^i. \quad (\text{A138})$$

Thus,

$$\mathfrak{R}_\mu[\Omega[\mu]]_i = \sum_{r \in [l]} x_r / (\text{tr} \pi_r) \pi_r.$$

This has to be compared with the general form of a tangent vector ($X \in T_{\mu_i} \mathcal{D}_\Pi$) given by (A120). The only condition the vector $p'(0)$ in (A120) has to satisfy is that its components sum to 0. This also holds for $x = (x_1, \dots, x_l)$. We conclude that $\mathfrak{R}_\mu[\Omega[\mu]]_i$ lies in $T_{\mu_i} \mathcal{D}_\Pi$ for all $i \in \mathcal{V}$ or, equivalently, $\mathfrak{R}_\mu[\Omega[\mu]] \in T_\mu \mathcal{D}_{\Pi, c}$.

(ii) We write $\mu_i = U \text{Diag}(S_i)U^*$ for all $i \in \mathcal{V}$ with $S_i \in \mathcal{S}_c$ and express $\mathfrak{R}_\mu[\Omega[\mu]]$ in terms of $S \in \mathcal{W}$ as

$$\mathfrak{R}_\mu[\Omega[\mu]]_i = \sum_{k \in \mathcal{N}_i} \omega_{ik} (\mu_i \mu_k - \text{tr}(\mu_i \mu_k) \mu_i) \tag{A139}$$

$$= U \text{Diag} \left(\sum_{k \in \mathcal{N}_i} \omega_{ik} (S_i \cdot S_k - \langle S_i, S_k \rangle S_i) \right) U^* \tag{A140}$$

$$= U \text{Diag} (R_S[\Omega S])_i U^*. \tag{A141}$$

□

Proof of Corollary 3. We write $D_i = U \text{Diag}(\lambda_i)U^*$ for $\lambda_i \in \mathbb{R}^n$ diagonalized in the U basis. Then, the initial condition for the QSAF S flow (109) is given by

$$\mu(0)_i = S(\mathbb{1}_Q)_i \stackrel{(101)}{=} \Gamma \left(\sum_{k \in \mathcal{N}_i} \omega_{ik} (-D_k) \right). \tag{A142}$$

Then, set $\tilde{D}_i := \sum_{k \in \mathcal{N}_i} \omega_{ik} D_k = U \text{Diag}(\tilde{\lambda}_i)U^*$, where

$$\tilde{\lambda}_i = \sum_{k \in \mathcal{N}_i} \omega_{ik} \lambda_k \in \mathbb{R}^c. \tag{A143}$$

Recall further that Γ is computed in terms of the matrix exponential as specified by (66). Thus,

$$\mu(0)_i = \Gamma(-\tilde{D}_i) = \frac{\exp_m(-\tilde{D}_i)}{\text{tr} \exp_m(-\tilde{D}_i)} = \frac{U \exp_m(-\text{Diag}(\tilde{\lambda}_i))U^*}{\text{tr}(U \exp_m(-\text{Diag}(\tilde{\lambda}_i))U^*)} = U \frac{\text{Diag}(\exp(-\tilde{\lambda}_i))}{\text{tr} \exp_m(-\text{Diag}(\tilde{\lambda}_i))} U^*. \tag{A144}$$

This shows that all the $\mu(0)_i$'s are diagonalized by the same basis U and $\mu(0) \in \mathcal{D}_{\Pi_{U,c}}$, and we can apply Proposition 10 (ii). Therefore, the vector field of the quantum state assignment S flow is also diagonalized in the U basis, and we solve simply for the diagonal components. The quantum S-flow equation can be written as

$$\dot{\mu}_i = U \text{Diag}(R_{S_i}[\Omega S])U^*, \quad \mu(0)_i = U \text{Diag}(S(\mathbb{1}_W))U^* \tag{A145}$$

with the classical similarity map S defined in terms of the data vectors λ_i and μ_i related to $S_i \in \mathcal{S}_c$ by $\mu_i = U \text{diag}(S_i)U^*$. The solution to this system is

$$\mu_i(t) = U \text{Diag}(S_i(t))U^*, \tag{A146}$$

where $S \in \mathcal{W}$ solves the classical S-flow equation $\dot{S} = R_S[\Omega S]$ and $S(0) = S(\mathbb{1}_W)$. □

References

1. Bakır, G.; Hofmann, T.; Schölkopf, B.; Smolar, A.J.; Taskar, B.; Vishwanathan, S.V.N. (Eds.) *Predicting Structured Data*; MIT Press: Cambridge, MA, USA, 2007.
2. Åström, F.; Petra, S.; Schmitzer, B.; Schnörr, C. Image Labeling by Assignment. *J. Math. Imaging Vis.* **2017**, *58*, 211–238. [\[CrossRef\]](#)
3. Schnörr, C. Assignment Flows. In *Variational Methods for Nonlinear Geometric Data and Applications*; Grohs, P., Holler, M., Weinmann, A., Eds.; Springer: Berlin, Germany, 2020; pp. 235–260.
4. Zern, A.; Zeilmann, A.; Schnörr, C. Assignment Flows for Data Labeling on Graphs: Convergence and Stability. *Inf. Geom.* **2022**, *5*, 355–404. [\[CrossRef\]](#)
5. Hühnerbein, R.; Savarino, F.; Petra, S.; Schnörr, C. Learning Adaptive Regularization for Image Labeling Using Geometric Assignment. *J. Math. Imaging Vis.* **2021**, *63*, 186–215. [\[CrossRef\]](#)
6. Sitenko, D.; Boll, B.; Schnörr, C. Assignment Flow For Order-Constrained OCT Segmentation. *Int. J. Comput. Vis.* **2021**, *129*, 3088–3118. [\[CrossRef\]](#)
7. Hofbauer, J.; Siegmund, K. Evolutionary Game Dynamics. *Bull. Am. Math. Soc.* **2003**, *40*, 479–519. [\[CrossRef\]](#)
8. Sandholm, W.H. *Population Games and Evolutionary Dynamics*; MIT Press: Cambridge, MA, USA, 2010.
9. Amari, S.-I.; Nagaoka, H. *Methods of Information Geometry*; American Mathematical Society: Providence, RI, USA; Oxford University Press: Oxford, UK, 2000.
10. Zeilmann, A.; Savarino, F.; Petra, S.; Schnörr, C. Geometric Numerical Integration of the Assignment Flow. *Inverse Probl.* **2020**, *36*, 034004. [\[CrossRef\]](#)

11. Sitenko, D.; Boll, B.; Schnörr, C. A Nonlocal Graph-PDE and Higher-Order Geometric Integration for Image Labeling. *SIAM J. Imaging Sci.* **2023**, *16*, 501–567. [[CrossRef](#)]
12. Zern, A.; Zisler, M.; Petra, S.; Schnörr, C. Unsupervised Assignment Flow: Label Learning on Feature Manifolds by Spatially Regularized Geometric Assignment. *J. Math. Imaging Vis.* **2020**, *62*, 982–1006. [[CrossRef](#)]
13. Zisler, M.; Zern, A.; Petra, S.; Schnörr, C. Self-Assignment Flows for Unsupervised Data Labeling on Graphs. *SIAM J. Imaging Sci.* **2020**, *13*, 113–1156. [[CrossRef](#)]
14. Bengtsson, I.; Życzkowski, K. *Geometry of Quantum States: An Introduction to Quantum Entanglement*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2017.
15. Petz, D. *Quantum Information Theory and Quantum Statistics*; Springer: Berlin/Heidelberg, Germany, 2008.
16. Schwarz, J.; Boll, B.; Sitenko, D.; Gonzalez-Alvarado, D.; Gärttner, M.; Albers, P.; Schnörr, C. Quantum State Assignment Flows. In *Scale Space and Variational Methods in Computer Vision*; Calatroni, L., Donatelli, M., Morigi, S., Prato, M., Santacesaria, M., Eds.; Springer: Berlin, Germany, pp. 743–756.
17. Amari, S.-I. *Differential-Geometrical Methods in Statistics*, 1990 ed.; Lecture Notes in Statistics; Springer: Berlin/Heidelberg, Germany, 1985; Volume 28.
18. Lauritzen, S.L. Chapter 4: Statistical Manifolds. In *Differential Geometry in Statistical Inference*; Gupta, S.S., Amari, S.I., Barndorff-Nielsen, O.E., Kass, R.E., Lauritzen, S.L., Rao, C.R., Eds.; Institute of Mathematical Statistics: Hayward, CA, USA, 1987; pp. 163–216.
19. Brown, L.D. *Fundamentals of Statistical Exponential Families*; Institute of Mathematical Statistics: Hayward, CA, USA, 1986.
20. Āencov, N.N. *Statistical Decision Rules and Optimal Inference*; American Mathematical Society: Providence, RI, USA, 1981.
21. Brazitikos, S.; Giannopoulos, A.; Valettas, P.; Vritsiou, B.-H. *Geometry of Isotropic Convex Bodies*; American Mathematical Society: Providence, RI, USA, 2014.
22. Ay, N.; Jost, J.; Lê, H.V.; Schwachhöfer, L. *Information Geometry*; Springer: Berlin/Heidelberg, Germany, 2017.
23. Calin, O.; Udriste, C. *Geometric Modeling in Probability and Statistics*; Springer: Berlin/Heidelberg, Germany, 2014.
24. Petz, D. Geometry of Canonical Correlation on the State Space of a Quantum System. *J. Math. Phys.* **1994**, *35*, 780–795. [[CrossRef](#)]
25. Bodmann, B.G.; Haas, J.I. A Short History of Frames and Quantum Designs. In *Topological Phases of Matter and Quantum Computation*; Contemporary Mathematics; American Mathematical Society: Providence, RI, USA, 2020; Volume 747, pp. 215–226.
26. Petz, D.; Toth, G. The Bogoliubov Inner Product in Quantum Statistics. *Lett. Math. Phys.* **1993**, *27*, 205–216. [[CrossRef](#)]
27. Grasselli, M.R.; Streater, R.F. On the Uniqueness of the Chentsov Metric in Quantum Information Geometry. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **2001**, *4*, 173–182. [[CrossRef](#)]
28. Thanwerdas, Y.; Pennec, X. $O(n)$ -invariant Riemannian Metrics on SPD Matrices. *Linear Algebra Appl.* **2023**, *61*, 163–201. [[CrossRef](#)]
29. Bhatia, R. *Positive Definite Matrices*; Princeton University Press: Princeton, NJ, USA, 2006.
30. Skovgaard, L.T. A Riemannian Geometry of the Multivariate Normal Model. *Scand. J. Stat.* **1984**, *11*, 211–223.
31. Bridson, M.R.; Häflinger, A. *Metric Spaces of Non-Positive Curvature*; Springer: Berlin/Heidelberg, Germany, 1999.
32. Jost, J. *Riemannian Geometry and Geometric Analysis*, 7th ed.; Springer: Berlin/Heidelberg, Germany, 2017.
33. Arsigny, V.; Fillard, P.; Pennec, X.; Ayache, N. Geometric Means in a Novel Vector Space Structure on Symmetric Positive-Definite Matrices. *SIAM J. Matrix Anal. Appl.* **2007**, *29*, 328–347. [[CrossRef](#)]
34. Michor, P.W.; Petz, D.; Andai, A. The Curvature of the Bogoliubov-Kubo-Mori Scalar Product on Matrices. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **2000**, *3*, 1–14.
35. Rockafellar, R. *Convex Analysis*; Princeton University Press: Princeton, NJ, USA, 1970.
36. Savarino, F.; Schnörr, C. Continuous-Domain Assignment Flows. *Eur. J. Appl. Math.* **2021**, *32*, 570–597. [[CrossRef](#)]
37. Absil, P.A.; Mahony, R.; Sepulchre, R. *Optimization Algorithms on Matrix Manifolds*; Princeton University Press: Princeton, NJ, USA, 2008.
38. Higham, N.J. *Functions of Matrices: Theory and Computation*; Society for Industrial and Applied Mathematics (SIAM): Philadelphia, PA, USA, 2008.
39. Savarino, F.; Albers, P.; Schnörr, C. On the Geometric Mechanics of Assignment Flows for Metric Data Labeling. *arXiv* **2021**, arXiv:2111.02543. (to appear in: Information Geometry).
40. Levine, Y.; Yakira, D.; Cohen, N.; Shashua, A. Deep Learning and Quantum Entanglement: Fundamental Connections with Implications to Network Design. In Proceedings of the Sixth International Conference on Learning Representations ICLR, Vancouver, BC, Canada, 30 April–3 May 2018.
41. Chen, R.T.Q.; Rubanova, Y.; Bettencourt, J.; Duvenaud, D. Neural Ordinary Differential Equations. In *Advances in Neural Information Processing Systems 31, Proceedings of the 32nd Conference on Neural Information Processing Systems (NeurIPS 2018), Montreal, QC, Canada, 3–8 December 2018*.
42. Horn, R.A.; Johnson, C.R. *Matrix Analysis*, 2nd ed.; Cambridge University Press: Cambridge, UK, 2013.
43. Kobayashi, S.; Nomizu, K. *Foundations of Differential Geometry*; Interscience Publishers: New York, NY, USA; John Wiley & Sons: Hoboken, NJ, USA, 1969; Volume II.

Disclaimer/Publisher’s Note: The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.