Quantum State Assignment Flows

Jonathan Schwarz[3][0000-0003-1825-3826], Bastian Boll[3][0000-0002-3490-3350], Daniel Gonzalez-Alvarado[3][0000-0002-4636-3697], Dmitrij Sitenko[3][0000-0002-0022-3891], Martin Gärttner[1][0000-0003-1914-7099], Peter Albers[2][0000-0002-7324-0609], and Christoph Schnörr[3][0000-0002-8999-2338]

[1] Kichhoff Institute for Physics, Heidelberg University, Germany
[2] Mathematical Institute, Heidelberg University, Germany
[3] Institute for Applied Mathematics, Heidelberg University, Germany
jonathan.schwarz@iwr.uni-heidelberg.de

Abstract. This paper extends the assignment flow approach from categorial distributions to complex-valued Hermitian density matrices, used as state spaces for representing and analyzing data associated with vertices of an underlying graph. Determining the flow of the resulting dynamical system by geometric integration causes an interaction of these non-commuting states across the graph, and the assignment of a pure (rank-one) state to each vertex after convergence. Experiments with toy systems indicate the potential of the novel approach for data representation and analysis.

Keywords: Assignment Flows · Density Matrix · Information Geometry

1 Introduction

The *assignment flow approach* provides data models in terms of state spaces that interact geometrically across an underlying graph. In principle, any open convex set can serve as a state space which becomes a Riemannian manifold when endowed with a Riemannian metric. The canonical cases are parameter spaces of distributions of the exponential family and the Fisher-Rao metric, which is the subject of *information geometry* [2].

The assignment flow approach has been introduced by [3] for the basic family of *categorial distributions*, in order to assign a unique element of a finite set of labels set to each data point observed in a metric space. We refer to [17] for more details and a review of related work.

Contribution. In this paper, we apply the assignment flow approach to a novel class of state spaces, the class of *complex valued, Hermitian positive semidefinite matrices*, known as density matrices in quantum mechanics, where they represent a physical system [5]. Even though this extension appears to be straightforward from an abstract mathematical viewpoint, details matter with regards to both the components of the approach and the scope of applications.

Specifically, a *key difference* is the *non-commutative* interaction of density matrices, opposed to the multiplicative interaction of discrete probability vectors.

Furthermore, regarding the objects to be assigned to data, the finite set of labels is replaced by the uncountable set of rank-one density matrices, i.e. the set of orthogonal projectors onto one-dimensional subspaces. We show that the original assignment flow can be recovered by restriction to the submanifold of *diagonal* density matrices.

Approach Component	**Discrete Labeling Assignment Flow**	**Quantum State Assignment Flow**
state space	product manifold of categorial distributions	product manifold of density matrices
state evolution	$\dot{S} = R_S[\Omega S]$	$\dot{\rho} = R_\rho[\Omega[\rho]]$
assigned limit states	unit vector at each vertex of $\mathcal{G}$	pure (rank one) state at each vertex of $\mathcal{G}$

Scope. Regarding applications, we have in mind data modeling and analysis as well as applications in quantum mechanics. Regarding the former aspect, we confine ourselves to a few proof of concept examples that demonstrate the different character and indicate the enhanced flexibility for data modeling of the novel assignment flow approach. Regarding the latter aspect, we point out that the *representation* of image data for quantum computing is an active field of research [7]. This paper contributes an approach for data *modeling and analysis* based on concepts of information geometry and quantum mechanics.

Organization. Section 2 collects concepts of information geometry. The novel approach is introduced in Section 3. Its properties are illustrated and discussed in Section 4. We conclude and point out further work in Section 5.

Due to the page limit, we have to omit many details and almost all proofs, and refer to [18].

2 Information Geometry

Information geometry [1,12] is concerned with the representation of parametric probability distributions like, e.g., the exponential familiy of distributions [6], from a geometric viewpoint. 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 reparametrization [19].

A key ingredient of information geometry is the so-called α-family of affine connections introduced by Amari, 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 which uniquely characterizes the Levi-Civita connection as metric connection [1, Def. 3.1, Thm. 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 [2,8,4] for further reading and to [13], [2, Ch. 7] for the specific case of the state spaces of quantum mechanics.

In this paper, we are concerned with two classes of convex sets, the relative interior of *probability simplices*, each of which represents the categorial (discrete) distributions of the corresponding dimension, and *density matrices*, i.e. the set of positive-definite Hermitian matrices with trace equal to one. Sections 2.1 and 2.2 introduce the information geometry for the former and the latter class of sets, respectively.

2.1 Categorial Distributions

We set $[c] := \{1, 2, \ldots, c\}$ and $\mathbb{1}_c := (1, \ldots, 1)^\top \in \mathbb{R}^c$ for $c \in \mathbb{N}$. The probability simplex of distributions on $[c]$ is denoted by $\Delta_c := \{p \in \mathbb{R}^c_+ \colon \langle \mathbb{1}_c, p\rangle = \sum_{i\in[c]} p_i = 1\}$. Its relative interior equipped with the Fisher-Rao metric g becomes the Riemannian manifold $(\mathcal{S}_c, g)$, where

$$\mathcal{S}_c := \operatorname{rint} \Delta_c = \{p \in \Delta_c \colon p_i > 0,\ i \in [c]\}, \tag{2.1a}$$

$$g_p(u, v) = \langle u, \operatorname{Diag}(p)^{-1} v\rangle, \qquad \forall u, v \in T_{c,0}, \quad p \in \mathcal{S}_c \tag{2.1b}$$

with the tangent space (with the barycenter of $\mathcal{S}_c$ denoted by $\mathbb{1}_{\mathcal{S}_c} = \frac{1}{c}\mathbb{1}_c$)

$$T_{c,0} := T_{\mathbb{1}_{\mathcal{S}_c}} \mathcal{S}_c = \{v \in \mathbb{R}^c \colon \langle \mathbb{1}_c, v\rangle = 0\} \tag{2.2}$$

and the trivial tangent bundle $T\mathcal{S}_c \cong \mathcal{S}_c \times T_{c,0}$. The orthogonal projection onto $T_{c,0}$ reads

$$\pi_0 \colon \mathbb{R}^c \to T_{c,0}, \qquad \pi_0 v = (I_c - \mathbb{1}_c \mathbb{1}_{\mathcal{S}_c}^\top) v, \tag{2.3}$$

A key role plays the *replicator mapping*

$$R \colon \mathcal{S}_c \times T_{c,0} \to T_{c,0}, \qquad R_p v := (\operatorname{Diag}(p) - pp^\top) v, \tag{2.4}$$

which is parametrized by $p \in \mathcal{S}_c$ and has the properties

$$R_p \mathbb{1}_c = 0 \qquad \text{and} \qquad \pi_0 R_p = R_p \pi_0 = R_p, \qquad \forall p \in \mathcal{S}_c. \tag{2.5}$$

In particular, R_p is the inverse metric tensor expressed in the ambient coordinates p and its restriction to the tangent space $T_{c,0}$ is a linear isomorphism [16, Lemma 3.1]. Accordingly, given a smooth function $f \colon \mathcal{S}_c \to \mathbb{R}$, its Riemannian gradient with respect to the Fisher-Rao metric (2.1b) is given by

$$\operatorname{grad} f(p) = R_p \partial f(p). \tag{2.6}$$

We list two further mappings required below. The exponential map induced by the e-connection is defined on the entire space $T_{c,0}$ and reads [4]

$$\operatorname{Exp} \colon \mathcal{S}_c \times T_{c,0} \to \mathcal{S}_c, \qquad \operatorname{Exp}_p(v) := \langle p, e^{\frac{v}{p}}\rangle^{-1} \big(p \cdot e^{\frac{v}{p}}\big), \tag{2.7}$$

where $\cdot$ denotes *componentwise* multiplication of vectors (Hadamard product). The so-called *lifting map* introduced in [3] reads

$$\exp \colon \mathcal{S}_c \times T_{c,0} \to \mathcal{S}_c, \qquad \exp_p(v) := \operatorname{Exp}_p \circ R_p(v) = \langle p, e^v\rangle^{-1} \big(p \cdot e^v\big). \tag{2.8}$$

The subscript of $\exp_p$ disambiguates its meaning in view of the ordinary exponential *function* e^v written without subscripts, and from the symbol $\exp_{\mathrm{m}}$ which always means the *matrix exponential* function.

2.2 Density Matrices

We denote by $\rho^* = \overline{\rho}^\top$ the conjugate transpose of a matrix $\rho \in \mathbb{C}^{c\times c}$. The inner products on $\mathbb{C}^c$ and $\mathbb{C}^{c\times c}$, respectively, are denoted by $\langle a,b\rangle = a^*b$ and $\langle A,B\rangle = \operatorname{tr}(A^*B)$. We denote the open convex cone of positive definite Hermitian matrices by $\mathcal{P}_c := \{\rho \in \mathbb{C}^{c\times c} \colon \rho = \rho^*,\ \rho \succ 0\}$ and its intersection with the hyperplane defined by constraint $\operatorname{tr}\rho = 1$, the space of *density matrices*, by

$$\mathcal{D}_c := \{\rho \in \mathcal{P}_c \colon \operatorname{tr}\rho = 1\}. \tag{2.9}$$

We refer to [5] for the physical background and to [14] for mathematical aspects related to quantum information theory. Denoting the vector space of Hermitian matrices by $\mathcal{H}_c := \{X \in \mathbb{C}^{c\times c} \colon X^* = X\}$, we have analogous to (2.2) the tangent space (with $\mathbb{1}_{\mathcal{D}_c} := \operatorname{Diag}(\mathbb{1}_{\mathcal{S}_c})$)

$$\mathcal{H}_{c,0} := T_{\mathbb{1}_{\mathcal{D}_c}}\mathcal{D}_c = \mathcal{H}_c \cap \{X \in \mathbb{C}^{c\times c} \colon \operatorname{tr} X = 0\} \tag{2.10}$$

and the trivial tangent bundle $T\mathcal{D}_c \cong \mathcal{D}_c \times \mathcal{H}_{c,0}$. The corresponding orthogonal projection reads[4]

$$\pi_0 \colon \mathcal{H}_c \to \mathcal{H}_{c,0}, \qquad \pi_0 X := X - (\operatorname{tr} X) I_{\mathcal{D}_c}. \tag{2.11}$$

The metric g is the *Bogoliubov-Kubo-Mori (BKM) metric* [15]

$$g_\rho(X,Y) := \int_0^\infty \operatorname{tr}\big(X(\rho+\lambda I)^{-1}Y(\rho+\lambda I)^{-1}\big)d\lambda,\ X,Y \in \mathcal{H}_{c,0},\ \rho \in \mathcal{D}_c. \tag{2.12}$$

This metric uniquely ensures that the e-connection ∇ induced on $\mathcal{D}_c$ is symmetric and the connections ∇, ∇^* are mutually dual to each other in the sense of information geometry [9], [2, Thm. 7.1].

The following map and its inverse, defined in terms of the matrix exponential $\exp_{\mathrm{m}}$ and its inverse $\log_{\mathrm{m}} = \exp_{\mathrm{m}}^{-1}$ will be convenient: $\mathbb{T} \colon \mathcal{D}_c \times \mathcal{H}_c \to \mathcal{H}_c$, with

$$\mathbb{T}_\rho[X] := \frac{d}{dt}\log_{\mathrm{m}}(\rho + tX)\big|_{t=0} = \int_0^\infty (\rho+\lambda I)^{-1}X(\rho+\lambda I)^{-1}d\lambda, \tag{2.13a}$$

$$\mathbb{T}_\rho^{-1}[X] = \frac{d}{dt}\exp_{\mathrm{m}}(H + tX)\big|_{t=0} = \int_0^1 \rho^{1-\lambda}X\rho^\lambda d\lambda, \qquad \rho = \exp_m(H). \tag{2.13b}$$

The inner product (2.12) may now be written in the form $g_\rho(X,Y) = \langle X, \mathbb{T}_\rho[Y]\rangle$ since the trace is invariant with respect to cyclic permutations of a matrix product as argument. Likewise, the relation $\langle \rho, X\rangle = \operatorname{tr}(\rho X) = \operatorname{tr}\mathbb{T}_\rho^{-1}[X]$ holds.

3 Quantum State Assignment Flows

This section summarizes our results regarding the extension of the assignment flow on the manifold of categorial distributions to the manifold of density matrices. The extension significantly generalizes the state space and the corresponding

[4] We keep the symbol π_0 for notational simplicity. The argument disambiguates the projections (2.3) and (2.11), respectively.

assignment flow. The positivity condition imposed on discrete probability vectors is replaced by the positive definiteness condition imposed on Hermitian matrices, and mass conservation of categorial distributions is replaced by trace normalization. The assignment flow has to be generalized accordingly, which is accomplished within the framework of information geometry. Rather than assigning a single label from a *finite* set of labels to each data point, the resulting quantum-state assignment flow (QSAF) assigns a pure (rank-one) state from an *uncountable* set to each data point. And analogous to the encoding of labels by unit vectors on the boundary of the product of probability simplices (assignment manifold), the QSAF converges towards the boundary of the corresponding density matrix product manifold.

This section is organized as follows. A few basic relations are collected in Section 3.1. Section 3.2 introduces the flow for a single state space which is generalized in Section 3.3 to the case of multiple states whose evolutions interact across a graph. This approach generalizes the assignment flow approach introduced by [3]. Section 3.4 generalizes accordingly the reparametrization introduced by [16] that enables to characterize the approach as a Riemannian gradient flow with respect to a nonconvex potential. Finally, Section 3.5 elucidates that the original assignment flow can be recovered as special case by restricting the quantum state assignment flow to diagonal density matrices.

Due to the page limit, we omit the proofs, except for the short one of Prop. 6, and refer to the journal version of this paper [18].

3.1 Basic Relations

A parametrization of the manifold $\mathcal{D}_c$ is given by

$$\Gamma\colon \mathcal{H}_{c,0} \to \mathcal{D}_c, \qquad \Gamma(X) := \frac{\exp_{\mathrm{m}}(X)}{\operatorname{tr}\exp_{\mathrm{m}}(X)}. \tag{3.1}$$

Lemma 1. *The mapping* (3.1) *is bijective with inverse*

$$\Gamma^{-1}\colon \mathcal{D}_c \to \mathcal{H}_{c,0}, \qquad \Gamma^{-1}(\rho) = \pi_0 \log_{\mathrm{m}} \rho, \tag{3.2}$$

Furthermore, for $H, X \in \mathcal{H}_{c,0}$ with $\Gamma(H) = \rho$ and $Y \in T\mathcal{H}_{c,0} \cong \mathcal{H}_{c,0}$, the respective differential mappings are given by

$$d\Gamma(H)[Y] = \mathbb{T}_\rho^{-1}\big[Y - \langle \rho, Y\rangle I\big], \qquad \rho = \Gamma(H) \tag{3.3a}$$

$$d\Gamma^{-1}(\rho)[X] = \pi_0 \circ \mathbb{T}_\rho[X]. \tag{3.3b}$$

A key concept of information geometry is the affine e-connection and the corresponding exponential map. It rarely occurs that pairs of points on a Riemannian manifold can be connected in closed form by a corresponding geodesic.

Proposition 1. *The e-geodesic emanating at $\rho \in \mathcal{D}_c$ in the direction $X \in \mathcal{H}_{c,0}$ and the corresponding exponential map are given by*

$$\gamma^{(e)}_{\rho,X}(t) := \mathrm{Exp}^{(e)}_{\rho}(tX), \quad t \geq 0 \tag{3.4a}$$

$$\mathrm{Exp}^{(e)}_{\rho}(X) := \Gamma\big(\Gamma^{-1}(\rho) + d\Gamma^{-1}(\rho)[X]\big) \tag{3.4b}$$

$$= \Gamma\big(\Gamma^{-1}(\rho) + \pi_0 \circ \mathbb{T}_{\rho}[X]\big). \tag{3.4c}$$

We next consider the evaluation of Riemannian gradients.

Proposition 2. *The Riemannian gradient of a smooth function $f\colon \mathcal{D}_c \to \mathbb{R}$ with respect to the BKM-metric (2.12) is given by*

$$\mathrm{grad}_{\rho} f = \mathbb{T}^{-1}_{\rho}[\partial f] - \langle \rho, \partial f\rangle \rho, \tag{3.5}$$

where $\mathbb{T}^{-1}_{\rho}$ is given by (2.13b) and ∂f is the ordinary gradient with respect to the Euclidean structure of the ambient space $\mathbb{R}^{c\times c}$.

The general defining formula $g_{\rho}(\mathrm{grad}\, f, X) = df_{\rho}X$, $\forall X \in \mathcal{H}_{c,0}$ for the Riemannian gradient [11, pp. 337] generally yields an expression of the form $\mathrm{grad}\, f_{\rho} = G(\rho)^{-1}\partial f(\rho)$ given in local coordinates. The role of the inverse metric tensor G^{-1} is played by the replicator map (2.4) which, in the present context and in view of the result (3.5), takes the more general form

$$R_{\rho}\colon \mathcal{H}_c \to \mathcal{H}_{c,0}, \quad R_{\rho}[X] := \mathbb{T}^{-1}_{\rho}[X] - \langle \rho, X\rangle \rho \qquad (\textbf{replicator map}) \tag{3.6}$$

where $\rho \in \mathcal{D}_c$ and $X \in \mathcal{H}_{c,0}$.

For a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$, we finally introduce the product spaces

$$\mathcal{H} := \mathcal{H}_c \times \cdots \times \mathcal{H}_c, \qquad \mathcal{H}_0 := \mathcal{H}_{c,0} \times \cdots \times \mathcal{H}_{c,0}, \tag{3.7}$$

each with $|\mathcal{V}|$ factors.

3.2 Single-Vertex Quantum State Assignment Flow

Let $D \in \mathcal{H}_c$ denote a given Hermitian matrix. Then we define the corresponding *likelihood matrix* by

$$L_{\rho}\colon \mathcal{H}_c \to \mathcal{D}_c, \qquad L_{\rho}(D) := \exp_{\rho}(-\pi_0 D), \qquad \rho \in \mathcal{D}_c. \tag{3.8}$$

The *single-vertex quantum state assignment flow* equation reads

$$\dot{\rho} = R_{\rho}[L_{\rho}(D)], \qquad \rho(0) = \mathbb{1}_{\mathcal{D}_c}, \tag{3.9}$$

where R_{ρ} is given by (3.6). The evolution of $\rho(t)$ solving (3.9) behaves as follows.

Proposition 3. *Let $D = Q\Lambda_D Q^{\top}$ be the spectral decomposition of D with eigenvalues $\lambda_1 \geq \cdots \geq \lambda_c$ and orthonormal eigenvectors $Q = (q_1, \ldots, q_c)$. Assume the minimal eigenvalue λ_c is unique. Then the solution $\rho(t)$ to (3.9) satisfies*

$$\lim_{t\to\infty} \rho(t) = \Pi_{q_c} := q_c q_c^{\top}. \tag{3.10}$$

We refer to [18] for a proof. It relies on a decomposition of $\mathcal{H}_{c,0}$ [2, Section 7.1] that allows for a reduction of the single-vertex quantum state assignment flow to the standard assignment flow [3]. The convergence of the latter is discussed in detail in [21].

A natural question is how multiple states which evolve in this way, can interact so as to influence their limit points, but not their property of being pure (rank one) states. Such a dynamical system is provided next.

3.3 Quantum State Assignment Flow

Let $\mathcal{G} = (\mathcal{V}, \mathcal{E}, \omega)$ be a given graph with nonnegative weight function $\omega\colon \mathcal{E} \to \mathbb{R}_+$, satifying $\sum_{k\in\mathcal{N}_i} \omega_{ik} = 1$ with respect to the neighborhood system $\mathcal{N}_i := \{i\} \cup \{k \in \mathcal{V}\colon k \sim i\},\ i \in \mathcal{V}$, induced by the adjacency relation $\mathcal{E}$. The motivation as well as one possible choice will be further clarified in the experimental section. Based on (2.9), we define the product manifold $(\mathcal{Q}_c, g)$ where

$$\rho = (\dots, \rho_i, \dots) \in \mathcal{Q}_c := \underbrace{\mathcal{D}_c \times \dots \times \mathcal{D}_c}_{|\mathcal{V}|\text{ factors}} \tag{3.11}$$

and the Riemannian metric in view of (2.12) and (3.7) is given by

$$g_\rho(X, Y) := \sum_{i\in\mathcal{V}} g_{\rho_i}(X_i, Y_i), \qquad X, Y \in T_\rho\mathcal{Q}_c := \mathcal{H}_0, \quad \forall \rho. \tag{3.12}$$

Let $\mathbb{1}_{\mathcal{Q}_c}$ denote the barycenter of $\mathcal{Q}_c$ given by $(\mathbb{1}_{\mathcal{Q}_c})_i = \mathbb{1}_{\mathcal{D}_c}$ for all $i \in \mathcal{V}$. Then the geometric interaction of likelihood matrices of the form (3.8) is defined by the *similarity mapping*

$$S\colon \mathcal{V}\times\mathcal{Q}_c \to \mathcal{D}_c, \qquad S_i(\rho) := \operatorname{Exp}^{(e)}_{\rho_i}\Big(\sum_{k\in\mathcal{N}_i} \omega_{ik}\big(\operatorname{Exp}^{(e)}_{\rho_i}\big)^{-1}\big(L_{\rho_k}(D_k)\big)\Big). \tag{3.13}$$

A characterization of the similarity map and a formula for evaluating it conveniently follow. They illustrate the benefit of using information geometry and the e-connection, rather than the Riemannian connection, from the computational viewpoint.

Proposition 4. *An equivalent expression of the similarity mapping (3.13) is given by*

$$S_i(\rho) = \Gamma\Big(\sum_{k\in\mathcal{N}_i} \omega_{ik}(\log_{\mathrm{m}} \rho_k - D_k)\Big), \qquad i \in \mathcal{V}. \tag{3.14}$$

Furthermore, if $\overline{\rho} \in \mathcal{D}_c$ solves the equation $0 = \sum_{k\in\mathcal{N}_i} \omega_{ik}\big(\operatorname{Exp}^{(e)}_{\overline{\rho}}\big)^{-1}\big(L_{\rho_k}(D_k)\big)$, which corresponds to the optimality condition for Riemannian centers of mass [10, Lemma 6.9.4], except for using a different exponential map, then

$$\overline{\rho} = S_i(\rho), \tag{3.15}$$

with $S_i(\rho)$ given by (3.13) and (3.14), respectively.

We are now in the position to define the

$$\dot{\rho} = R_\rho[S(\rho)], \quad \rho(0) = \mathbb{1}_{\mathcal{Q}_c} \qquad (\textbf{quantum state assignment flow}) \tag{3.16a}$$

where both the replicator map R_ρ and the similarity map $S(\rho)$ apply factorwise,

$$S(\rho)_i = S_i(\rho), \qquad R_\rho[S(\rho)]_i = R_{\rho_i}[S_i(\rho)], \qquad i \in \mathcal{V}. \tag{3.16b}$$

3.4 Riemannian Gradient Flow Parametrization

The proposition below provides a reparametrization of the quantum state flow equation (3.16a) that constitutes a Riemannian gradient flow with respect to a nonconvex potential.

Based on the weight function $\omega\colon \mathcal{E} \to \mathbb{R}_+$ of a given graph $\mathcal{G} = (\mathcal{V}, \mathcal{E}, w)$, we define the linear mapping

$$\Omega\colon \mathcal{Q}_c \to \mathcal{Q}_c, \qquad \Omega[\rho]_i := \sum_{k \in \mathcal{N}_i} \omega_{ik}\rho_k \in \mathcal{D}_c, \qquad i \in \mathcal{V}. \tag{3.17}$$

In addition, we adopt the *symmetry assumption*

$$\omega_{ij} = \omega_{ji}, \qquad j \in \mathcal{N}_i \;\Leftrightarrow\; i \in \mathcal{N}_j, \qquad \forall i, j \in \mathcal{V} \tag{3.18}$$

which makes the mapping (3.17) self-adjoint: $\langle \mu, \Omega[\rho] \rangle := \sum_{i \in \mathcal{V}} \langle \mu_i, \Omega[\rho]_i \rangle = \langle \Omega[\mu], \rho \rangle$ for all $\mu, \rho \in \mathcal{Q}_c$. Then the following holds.

Proposition 5. *The flow equation* (3.16a) *is equivalent to the system*

$$\dot{\rho} = R_\rho[\mu], \qquad \rho(0) = \mathbb{1}_{\mathcal{Q}_c}, \tag{3.19a}$$

$$\dot{\mu} = R_\mu\big[\Omega[\mu]\big], \qquad \mu(0) = S(\mathbb{1}_{\mathcal{Q}_c}). \tag{3.19b}$$

Furthermore, (3.19b) *is the Riemannian gradient flow*

$$\dot{\mu} = -\operatorname{grad}_\mu J(\mu) \tag{3.20a}$$

with respect to the potential $J(\mu)$ *given by*

$$J(\mu) := -\frac{1}{2}\langle \mu, \Omega[\mu] \rangle = \frac{1}{2}\big(\langle \mu, L_{\mathcal{G}}[\mu] \rangle - \|\mu\|^2\big), \tag{3.20b}$$

and with the 'Laplacian' operator $L_{\mathcal{G}}\colon \mathcal{Q}_c \to \mathcal{Q}_c,\ L_{\mathcal{G}} := \mathrm{id} - \Omega$.

The crucial point of Proposition (5) is that the evolution of $\mu(t)$ described by (3.19b) represents the 'essential' part of the quantum state flow equation (3.16a), since $\rho(t)$ solving (3.19a) is a function of $\mu(t)$ but *not* vice versa. Hence the geometric potential flow (3.20) provides a suitable basis for analyzing 'deep' quantum state assignment flows that result from the geometric integration of $\mu(t)$ (where each time step defines a 'layer') and a task-dependent choice of a time-variant mapping $\Omega(t)$, typically to be learnt from data.

3.5 Recovering the Assignment Flow for Categorial Distributions

In this section, we show that the quantum state assignment flow on a product manifold of density matrices contains as special case the assignment flow for categorial distributions, when the former flow is restricted to diagonal density matrices. This is quite natural because $\rho \succ 0$ implies positive diagonal elements and $\operatorname{tr} \rho = 1$ implies $\operatorname{diag}(\rho) \in \mathcal{S}_c$.

We confine ourselves to the formulation of quantum state assignment flow provided by Proposition 5 whose flow corresponds one-to-one to the flow generated by Equation (3.16a).

Proposition 6. *Let*

$$\mathcal{Q}_c^d := \mathcal{D}_c^d \times \cdots \times \mathcal{D}_c^d \subset \mathcal{Q}_c, \qquad \mathcal{D}_c^d := \{\operatorname{Diag}(p) : p \in \mathcal{S}_c\} \tag{3.21}$$

denote the product submanifold of diagonal density matrices of the manifold $\mathcal{Q}_c$ given by (3.11). Then the quantum state flow equation in the form (3.19b) reduces to the dynamical system[5]

$$\dot{S} = R_S[\Omega S], \qquad S(0) = S(\mathbb{1}_{\mathcal{W}_c}) \tag{3.22}$$

called 'S-flow' in [16, Prop. 3.6], where

$$S \in \mathbb{R}_+^{|\mathcal{V}|\times c}, \quad S_i = \operatorname{diag}(\mu_i), \quad i \in \mathcal{V} \tag{3.23a}$$

$$\Omega \in \mathbb{R}_+^{n\times n}, \quad \Omega_{ij} = \omega(ij), \quad ij \in \mathcal{E} \tag{3.23b}$$

$$R_S[\Omega S]_i = R_{S_i}(\Omega S)_i, \tag{3.23c}$$

with R_{S_i} given by (2.4), with $\mathbb{1}_{\mathcal{W}_c}$ denoting the barycenter of the assignment manifold and with the inital point $S(0)$ defined as in [16, Prop. 3.6].

Proof. The proof basically reduces to identifying (i) the restriction of product states of the form (3.11) to diagonal density matrices as factors and (ii) matrices $S \in \mathbb{R}_{|V|\times c}$ with row vectors

$$S_i = \operatorname{diag}(\mu_i), \qquad i \in \mathcal{V}. \tag{3.24}$$

Then the mapping (3.17) takes the form ΩS with Ω given by (3.23b), whereas the right-hand side in (3.19b) takes for *diagonal* – and hence *commuting* – density matrices $\mu_i,\ i \in \mathcal{V}$, the form

$$R_\mu[\Omega\mu]_i \overset{(3.16b)}{=} R_{\mu_i}\big[\Omega[\mu]_i\big] = R_{\mu_i}\Big[\sum_{k\in\mathcal{N}_i} \omega_{ik}\rho_k\Big] \tag{3.25a}$$

$$\overset{\substack{(3.6)\\(2.13b)}}{=} \sum_{k\in\mathcal{N}_i} \omega_{ik}\Big(\int_0^1 \mu_i^{1-\lambda}\mu_k\mu_i^{\lambda} d\lambda - \operatorname{tr}(\mu_i\mu_k)\mu_i\Big) \tag{3.25b}$$

[5] The use of the symbol S in the present context should *not* be confused with the similarity mapping (3.13). We just adhere to the notation used in prior work in order to reference clearly.

$$\overset{(3.24)}{=} \sum_{k\in\mathcal{N}_i} \omega_{ik}\big(S_i \cdot S_k - \langle S_i, S_k\rangle S_i\big) = R_{S_i}\Big(\sum_{k\in\mathcal{N}_i} \omega_{ik} S_k\Big) \tag{3.25c}$$

$$= R_{S_i}(\Omega S)_i \overset{(3.23c)}{=} R_S[\Omega S]_i. \qquad \square \tag{3.25d}$$

Proposition 6 basically says that the quantum assignment flow introduced in this paper is the *natural* generalization of the assignment flow approach introduced by [3,17] to *non-commutative* state spaces.

4 Experiments and Discussion

This section presents few academical results which illustrate properties of the novel ***QSAF*** approach (3.16). *We do not intend to consider and discuss any serious and fully worked out application in this paper* (see Section 5). Rather, the focus is on properties of the novel approach that cannot be achieved with the original assignment flow of categorial distributions. All computations were done using a geometric numerical Euler scheme adapted to the QSAF, after generalizing the approach presented in [20] accordingly.

Basic patch smoothing. Figure 4.1 shows an application of the QSAF to a *random* spatial arrangement (grid graph) of patches, where each vertex represents a patch, not a pixel. We refer to the caption for a description. The result demontrates

- that *geometric* smoothing of image data at the *patch level* can preserve spatial image structure;
- that after convergence the final state constitutes a piecewise constant labeling with pure (rank-one) states.

Thus the QSAF, directly applied to the raw data, performs image partitioning which is *not* piecewise *constant* at the pixel level.

Noise separation at the patch level. Figure 4.2 shows an application of the QSAF to a spatial collection of patches, each of which is pixelwise the mean of a randomly oriented patch and a patch with a fixed orientation. The result demonstrates that the QSAF effectively separates and removes 'random patch noise' at the *patch level without* any prior information or accessing the pixel level.

Patch smoothing using harmonic frames. Any matrix ensemble of the form $\{M_j\}_{j\in[c]} \subset \overline{\mathcal{P}}_c\colon \sum_{j\in[c]} M_j = I_c$ induces the categorial probability distribution $p \in \Delta_c$ on $[c]$ by taking inner products: $p_j = \langle M_j, \rho\rangle = \operatorname{tr}(M_j\rho),\ j \in [c]$, for any $\rho \in \mathcal{D}_c$. A simple instance are the projection operators $M_j = F_2^j (F_2^j)^*$ formed by the columns F_2^j of the unitary discrete Fourier matrix $F_2 = F \otimes F \in \mathbb{C}^{c\times c}$ which performs the 2D discrete Fourier transform when applied to a vectorized patch with c pixels. Subtracting the mean of a vectorized patch followed by normalization using the $\|\cdot\|_2$ norm, the patch at vertex $i \in \mathcal{V}$ was encoded by a matrix D_i in (3.8) of the form $D_i = F_2 \operatorname{Diag}(-|\widehat{p}_i|^2) F_2^*$ with $|\widehat{p}_i|_j = |(F_2 \operatorname{vec}(P_i))_j|$, $\forall j$, for patch P_i.

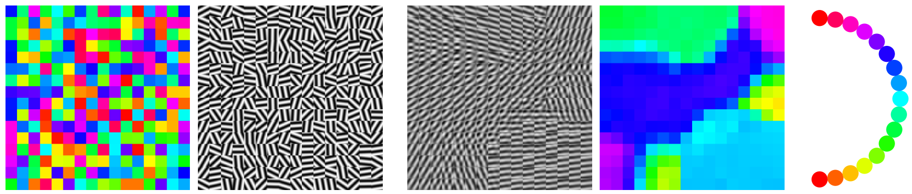
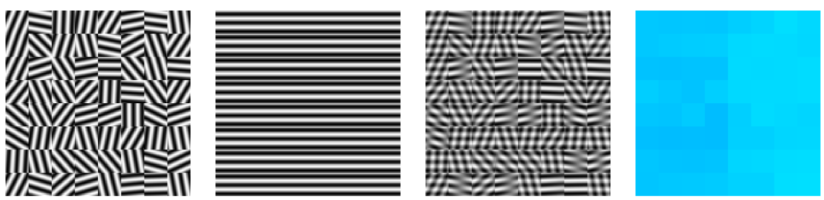


Fig. 4.1: **Left pair:** A random collection of patches with oriented image structure. The colored image displays for each patch its orientation using the color code depicted by the rightmost panel. Each patch is represented by a rank-one matrix D in (3.8), obtained by vectorizing the patch and taking 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 neighbors states. It represents an image partition but preserves image structure, due to geometric smoothing of patches encoded by non-commutative state spaces.



(a) (b) (c) (d)

Fig. 4.2: **(a)** A random collection of patches with oriented image structure. **(b)** A collection of patches with the same oriented image structure. **(c)** Pixelwise mean of the patches (a) (b) at each location. **(d)** The QSAF recovers a close approximation of (b) (color code: see Fig. 4.1) by iteratively smoothing the states $\rho_k,\ k \in \mathcal{N}_i$ corresponding to (c) through geometric integration.

Fig. 4.3: **Left:** A real roof texture. **Right:** 8×8 patches were encoded using the discrete Fourier frame (see text). Integrating the QSAF yields the same effect as shown by Figure 4.2, here with respect to the Fourier frame, however.

Integrating the QSAF yields a denoising effect at the patch level similar to Figure 4.2, here in the Fourier domain, however. After convergence, each state has the form $\rho_i = q_i q_i^\top$ for some unit vector q_i which was used to filter the Fourier-transformed patch vector using the Hadamard product, followed by decoding the patch. Accordingly, “assignment” here means scale and orientation in a spatial context, as encoded by the harmonic Fourier frame; see Figure 4.3.

Translation-invariant patch smoothing in a harmonic frame on a non-grid graph. The scenario of Figure 4.3 was extended: rather than using nearest-neighborhoods, the 8 most similar patches in the *entire* collection of image patches were selected for each patch, to define corresponding non-grid edges and irregular neighborhoods. Similarity was defined in terms of the distance between the orbits of patches, generated by 2D cyclic translation. The resulting unitary ‘registration operator’ was attached to each edge. Such elementary preprocessing changes the data encoding in terms of the matrix D in (3.8), but not the QSAF flow, which yields a very sparse but sufficiently detailed representation of image structure, although the Fourier frame only represents two orientations at various scales; see Figure 4.4.

(a) (b) (c)

Fig. 4.4: **(a)** A real scene. **(b)** A section of (a). **(c)** QSAF-filtered patches using a single harmonic frame, irregular non-local neighborhoods and translation invariant patch encoding (see text). Due to partitioning the image into patches, using a single harmonic frame and a shift-invariant patch distance, image structure is encoded very sparsely but sufficiently detailed.

5 Conclusion

We introduced a novel dynamical system for data representation and analysis, by extending the assignment flow approach to density matrices. Few numerical examples illustrated context-sensitive patch smoothing by geometric averaging of the non-commuting state spaces.

The model expressivity of the approach which performs the assignment of rank-one density matrices as ‘labels’, is larger than our preliminary experiments indicate. For instance, *latent* states may be used to parametrize Parseval frames which in turn transform the primary states. Moreover, by learning the parameters Ω in (3.13) and (3.17), respectively, from data, our approach may be seen as a novel ‘neural ODE’ from the viewpoint of machine learning.

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

References

1. Amari, S.I.: Differential-Geometrical Methods in Statistics, Lect. Not. Statistics, vol. 28. Springer, corrected 2nd printing 1990 edn. (1985)
2. Amari, S.I., Nagaoka, H.: Methods of Information Geometry. Amer. Math. Soc. and Oxford Univ. Press (2000)
3. Åström, F., Petra, S., Schmitzer, B., Schnörr, C.: Image Labeling by Assignment. Journal of Mathematical Imaging and Vision **58**(2), 211–238 (2017)
4. Ay, N., Jost, J., Lê, H.V., Schwachhöfer, L.: Information Geometry. Springer (2017)
5. Bengtsson, I., Zyczkowski, K.: Geometry of Quantum States: An Introduction to Quantum Entanglement. Cambridge University Press, 2nd edn. (2017)
6. Brown, L.D.: Fundamentals of Statistical Exponential Families. Institute of Mathematical Statistics, Hayward, CA (1986)
7. Cai, Y., Lu, X., Jiang, N.: A Survey on Quantum Image Processing. Chinese J. Electronics **27**(4), 667–888 (2018)
8. Calin, O., Udriste, C.: Geometric Modeling in Probability and Statistics. Springer (2014)
9. Grasselli, M., Streater, R.: On the Uniqueness of the Chentsov Metric in Quantum Information Geometry. Infinite Dimensional Analysis, Quantum Probability and Related Topics **4**(2), 173–182 (2001)
10. Jost, J.: Riemannian Geometry and Geometric Analysis. Springer-Verlag Berlin Heidelberg, 7th edn. (2017)
11. Kobayashi, S., Nomizu, K.: Foundations of Differential Geometry, vol. II. Interscience Publ., J. Wiley & Sons (1969)
12. Lauritzen, S.L.: Chapter 4: Statistical Manifolds. In: Gupta, S.S., Amari, S.I., Barndorff-Nielsen, O.E., Kass, R.E., Lauritzen, S.L., Rao, C.R. (eds.) Differential Geometry in Statistical Inference, pp. 163–216. IMS, Hayward, CA (1987)
13. Petz, D.: Geometry of Canonical Correlation on the State Space of a Quantum System. J. Math. Physics **35**(2), 780–795 (1994)
14. Petz, D.: Quantum Information Theory and Quantum Statistics. Springer (2008)
15. Petz, D., Toth, G.: The Bogoliubov Inner Product in Quantum Statistics. Letters in Mathematical Physics **27**(3), 205–216 (1993)
16. Savarino, F., Schnörr, C.: Continuous-Domain Assignment Flows. European Journal of Applied Mathematics **32**(3), 570–597 (2021)
17. Schnörr, C.: Assignment Flows. In: Grohs, P., Holler, M., Weinmann, A. (eds.) Variational Methods for Nonlinear Geometric Data and Applications, pp. 235–260. Springer (2020)
18. Schwarz, J., Boll, B., Gonzalez-Alvarado, D., Sitenko, D., Gärttner, M., Albers, P., Schnörr, C.: The present paper. Journal version, in preparation (2023)
19. Čencov, N.N.: Statistical Decision Rules and Optimal Inference. AMS (1981)
20. Zeilmann, A., Savarino, F., Petra, S., Schnörr, C.: Geometric Numerical Integration of the Assignment Flow. Inverse Problems **36**(3), 034004 (33pp) (2020)
21. Zern, A., Zeilmann, A., Schnörr, C.: Assignment Flows for Data Labeling on Graphs: Convergence and Stability. Information Geometry pp. 1–50 (2021)