On the Correspondence between Replicator Dynamics and Assignment Flows

Bastian Boll[1,2], Jonathan Schwarz[1,2], and Christoph Schnörr[1,2]

[1] Heidelberg Collaboratory for Image Processing, Heidelberg University, Germany
[2] Image and Pattern Analysis Group, Heidelberg University, Germany
bastian.boll@iwr.uni-heidelberg.de

Abstract. Assignment flows are smooth dynamical systems for data labeling on graphs. Although they exhibit structural similarities with the well-studied class of replicator dynamics, it is nontrivial to apply existing tools to their analysis. We propose an embedding of the underlying assignment manifold into the interior of a single probability simplex. Under this embedding, a large class of assignment flows are pushed to much higher-dimensional replicator dynamics. We demonstrate the applicability of this result by transferring a spectral decomposition of replicator dynamics to assignment flows.

Keywords: image labeling · assignment flows · replicator dynamics · spectral decomposition.

1 Introduction

Given a graph G with nodes $\mathcal{I}$ and a weighted adjacency matrix Ω, data labeling is the task of assigning a label from a discrete set $\mathcal{J}$ to each node in $\mathcal{I}$ such that both consistency with given data on $\mathcal{I}$ and spatial regularity with respect to Ω are simultaneously maximized. This constitutes a basic problem in image processing and formalizes e.g. image segmentation. In [1], a class of smooth geometric labeling systems is introduced which evolves high-entropy assignment states towards hard node–label decisions. *Assignment flows* are applicable to data in any metric space and regularize via geometric averaging according to Ω. If nodes are decoupled (i.e. $\Omega = \mathbb{I}$), assignment flows reduce to simple node-wise replicator equations which have been extensively studied as models of evolution in mathematical biology [6,10,9]. However, established theory is difficult to apply to the study of assignment flows, because it is unclear how to incorporate the coupling of nodes via geometric averaging which is at the core of their expressiveness. In the present work, we show how a large class of assignment flows can be seen as marginalization of replicator dynamics. We subsequently leverage this perspective to transfer a spectral decomposition result from the replicator setting to assignment flows.

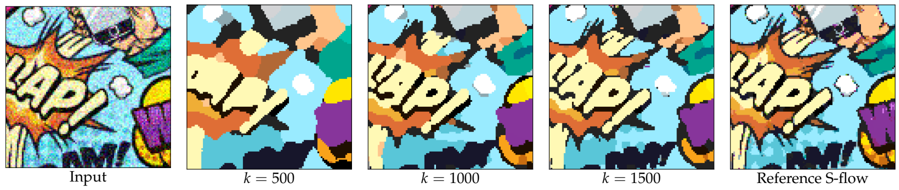

Fig. 1. Spectral approximations to the (S)-assignment flow limit (right) by approximating (4.23) with $k \in \{500, 1000, 1500\}$ most dominant eigenvectors of Ω. Averaging weights are computed analogous to nonlocal means denoising [3].

1.1 Related Work

A nonlinear spectral framework has recently been developed for the treatment of regularizations induced by one-homogeneous functions [7,5] such as total variation (TV). Applicability of this framework was demonstrated for diverse data processing tasks such as image fusion [2]. However, because the underlying spectral theory is nonlinear, specialized methods are required to compute eigenfunctions [4]. The spectral decomposition we propose here can be seen as a tool to study *nonlinear* assignment flows which have been used in similar areas of application. However, unlike the above framework, the spectral analysis we propose is *linear*, making it amenable to standard methods.

2 Assignment Manifold and Assignment Flows

We briefly summarize assignment flows as introduced in [1] and refer to the recent survey [12] for more background, more details and a review of recent related work.

Let $(\mathcal{F}, d_{\mathcal{F}})$ be a metric space and $\mathcal{F}_n = \{f_i \in \mathcal{F} \colon i \in \mathcal{I}\}$, $|\mathcal{I}| = n$ be given data. Assume that a predefined set of $|\mathcal{J}| = c$ prototypes $\mathcal{F}_* = \{f^*_j \in \mathcal{F} \colon j \in \mathcal{J}\}$ is given. *Data labeling* denotes the task of finding assignments $j \to i$, $f^*_j \to f_i$ in a spatially regularized fashion. At each node $i \in \mathcal{I}$, a distribution $W_i = (W_{i1}, \dots, W_{ic})^\top$ in the relative interior $\mathcal{S} = \operatorname{rint} \Delta_c$ of the probability simplex encodes assignment of $\mathcal{J}$ to i. We view $\mathcal{S}$ as a Riemannian manifold $(\mathcal{S}, g)$ endowed with the Fisher-Rao metric g from information geometry. The *assignment manifold* $(\mathcal{W}, g)$, $\mathcal{W} = \mathcal{S} \times \cdots \times \mathcal{S}$ ($n = |\mathcal{I}|$ factors) is the product manifold whose points encode label assignments at all nodes. By comparing given data to prototypes, the distance vector field $D_{\mathcal{F};i} = \big(d_{\mathcal{F}}(f_i, f^*_1), \dots, d_{\mathcal{F}}(f_i, f^*_c)\big)^\top$, $i \in \mathcal{I}$ is a data representation which abstracts the specific feature space $\mathcal{F}$. It is lifted to the assignment manifold by the *likelihood map* and the *likelihood vectors*, respectively,

$$L_i \colon \mathcal{S} \to \mathcal{S}, \quad L_i(W_i) = \frac{W_i \odot e^{-\frac{1}{\rho} D_{\mathcal{F};i}}}{\langle W_i, e^{-\frac{1}{\rho} D_{\mathcal{F};i}} \rangle}, \quad i \in \mathcal{I}. \tag{2.1}$$

This map is based on the affine e-connection of information geometry and $\rho > 0$ is used to normalize the application-specific scale of distances $D_{\mathcal{F};i}$. Likelihood vectors are spatially regularized by the *similarity map* and *similarity vectors*, respectively,

$$S_i\colon \mathcal{W} \to \mathcal{S}, \quad S_i(W) = \operatorname{Exp}_{W_i}\Big(\sum_{k\in\mathcal{N}_i} w_{ik} \operatorname{Exp}_{W_i}^{-1}\big(L_k(W_k)\big)\Big), \quad i \in \mathcal{I}, \tag{2.2}$$

where $\operatorname{Exp}_p(v) = \frac{pe^{v/p}}{\langle p, e^{v/p}\rangle}$ is the exponential map corresponding to the e-connection. Positive weights ω_{ik}, $k \in \mathcal{N}_i$, that sum up to 1 in every neighborhood $\mathcal{N}_i$, determine the regularization properties. We collect these weights into an adjacency matrix $\Omega \in \mathbb{R}^{n\times n}$. The ($W$-)*assignment flow* is induced on the assignment manifold $\mathcal{W}$ by the locally coupled system of nonlinear ODEs

$$\dot{W}_i = R_{W_i} S_i(W), \qquad W_i(0) = \mathbb{1}_{\mathcal{S}}, \quad i \in \mathcal{I}, \tag{2.3}$$

where the map $R_p = \operatorname{Diag}(p) - pp^\top$, $p \in \mathcal{S}$ turns the right-hand side into a tangent vector field and $\mathbb{1}_{\mathcal{W}} \in \mathcal{W}$ denotes the barycenter of the assignment manifold $\mathcal{W}$. The solution $W(t) \in \mathcal{W}$ is numerically computed by geometric integration [13] and determines a labeling $W(T)$ for sufficiently large T after a trivial rounding operation. Convergence and stability of the assignment flow have been studied by [14]. For symmetric weights $\omega_{ij} = \omega_{ji}$, it has been shown [11] that (2.3) can be parameterized entirely by similarity vectors

$$\dot{W}(t) = R_{W(t)}[S(t)], \qquad W(0) = \mathbb{1}_{\mathcal{W}} \tag{2.4}$$

$$\dot{S}(t) = R_{S(t)}[\Omega S(t)], \qquad S(0) = \exp_{\mathbb{1}_{\mathcal{W}}}(-\Omega D)\ . \tag{2.5}$$

We call the dynamics (2.5) *S-assignment flow* or *S-flow*. Here, the *lifting map* $\exp_W\colon T_0\mathcal{W} \to \mathcal{W}$ is defined by

$$\exp_{W_i}(V_i)_i = \operatorname{Exp}_{W_i}(R_{W_i} V_i) = \frac{W_i \odot \exp(V_i)}{\langle W_i, \exp(V_i)\rangle}, \qquad i \in \mathcal{I} \tag{2.6}$$

on the flat tangent space $T_0\mathcal{W} = \{V \in \mathbb{R}^{n\times c}\colon V\mathbb{1}_c = 0\}$ and R applies to each row of W resp. S separately. Note that $\exp_p(v + \alpha\mathbb{1}) = \exp_p(v)$ for any $\alpha \in \mathbb{R}$ as one easily checks.

Notation In the following, we consider a graph G with n nodes and encode assignment of c classes to these nodes as assignment matrices $S \in \mathcal{W} \subseteq \mathbb{R}^{n\times c}$. For tensors $U \in \mathbb{R}^{n_1\times\cdots\times n_r}$, $U^v = \operatorname{vec} U \in \mathbb{R}^{n_1\cdots n_r}$ denotes vectorization. The action of linear operators T on U is written as $T(U)$ if U is regarded as a tensor and by juxtaposition if U is vectorized, i.e. $T(U)^v = TU^v$. We define the symbol Δ to refer to the relative interior of a single probability simplex with $N = c^n$ corners. Vectors $P \in \Delta$ are identified with tensors in $\mathbb{R}^{c\times\cdots\times c}$ such that their entries can be referred to by multi-indices $\alpha \in [c]^n$. The symbol $\mathbb{1}_{n\times c}$ denotes the matrix of size $n \times c$ filled with 1 and $\odot$ denotes componentwise multiplication.

3 S-Flow Embedding

Assignment matrices $S \in \mathcal{W}$ associate each node $i \in [n]$ with a distribution $S_i \in \mathcal{S}$. Up to a negligible set of pathological cases, assignment flows converge to integral assignments, i.e. $S_i(t) \to e_{k(i)} \in \mathbb{R}^c$ for $t \to \infty$ as shown in [14]. By enumerating all c^n possible assignments of c classes to n nodes, we may equivalently view the integral limit point as approached within Δ. In keeping with this perspective, the aim of this section is to find an embedding of $\mathcal{W}$ into Δ such that S-flows in $\mathcal{W}$ translate to replicator dynamics in Δ. To this end, define the maps

$$T\colon \mathcal{W} \to \Delta, \qquad T(S)_\alpha := \prod_{i\in[n]} S_{i,\alpha_i} \tag{3.1}$$

$$Q\colon \mathcal{W} \to \Delta, \qquad Q(S)_\gamma := \sum_{l\in[n]} S_{l,\gamma_l} \ . \tag{3.2}$$

T maps $S \in \mathcal{W}$ to a rank-1 tensor in Δ. The inverse process is marginalization for each node.

Lemma 1. *The map T defined by (3.1) is a diffeomorphism between $\mathcal{W}$ and a subset of Δ with inverse*

$$T^{-1}(P)_{i,j} = \sum_{\alpha_i=j} P_\alpha := \sum_{\alpha\in[c]^n} \delta_{\alpha_i=j} P_\alpha, \qquad (i,j) \in [n]\times[c] \ . \tag{3.3}$$

Proof. We check that the inverse of T has the form (3.3).

$$T^{-1}(T(S))_{i,j} = \sum_{\alpha_i=j}\prod_{r\in[n]} S_{r,\alpha_r} \tag{3.4a}$$

$$= \sum_{k_1\in[c]} \dots \sum_{k_{i-1}\in[c]} \sum_{k_{i+1}\in[c]} \dots \sum_{k_n\in[c]} S_{1,k_1}\dots S_{i-1,k_{i-1}} S_{i,j} S_{i+1,k_{i+1}} \dots S_{n,k_n} \tag{3.4b}$$

$$= S_{i,j} \prod_{l\in[n]\setminus\{i\}} \langle S_l, \mathbb{1}\rangle = S_{i,j} \ . \tag{3.4c}$$

Clearly, both T and T^{-1} are smooth. □

If individual nodes are decoupled, i.e. $\Omega = \mathbb{I}_n$, assignment flows reduce to simple replicator dynamics for each node. We will show that the case of coupled nodes via more general choices of Ω may likewise be seen as a single, much larger replicator equation. We start with a preparatory lemma.

Lemma 2 (Adjoint of Q). *T^{-1} and Q are adjoint linear operators.*

Proof. Let $P \in \mathbb{R}^N$ and $V \in \mathbb{R}^{n\times c}$, then

$$\langle P, Q(V)\rangle = \sum_{\gamma} P_\gamma Q(V)_\gamma = \sum_{\gamma} P_\gamma \sum_{l\in[n]} V_{l,\gamma_l} = \sum_{l\in[n]}\sum_{j\in[c]}\sum_{\gamma_l=j} P_\gamma V_{l,\gamma_l} \tag{3.5a}$$

$$= \sum_{l\in[n]}\sum_{j\in[c]} V_{l,j}\sum_{\gamma_l=j} P_\gamma \overset{(3.3)}{=} \sum_{l\in[n]}\sum_{j\in[c]} V_{l,j} T^{-1}(P)_{l,j} \tag{3.5b}$$

$$= \langle T^{-1}(P), V\rangle \ . \tag{3.5c}$$

□

Theorem 1 (Replicator dynamics induce S-Flow). *For any S-flow*

$$\dot{S}(t) = R_{S(t)}[\Omega S(t)] =: X(S(t)), \qquad S(0) = S_0 \tag{3.6}$$

on $\mathcal{W}$ exists a matrix $\overline{\Omega} \in \mathbb{R}^{N\times N}$ such that (3.6) is induced by marginalization of the replicator dynamics

$$\dot{P}(t) = (T_\sharp X)(P) = R_{P(t)}[\overline{\Omega} P(t)], \qquad P(0) = T(S_0) \tag{3.7}$$

on Δ. $\overline{\Omega}$ is symmetric exactly if Ω is symmetric.

Proof. We push forward the S-flow vector field $X(S) := R_S[\Omega S]$ via T. The components of the differential dT read

$$\frac{\partial T_\alpha}{\partial S_{l,m}} = \frac{\partial}{\partial S_{l,m}}\prod_{i\in[n]} S_{i,\alpha_i} = \prod_{i\in[n]\setminus\{l\}} S_{i,\alpha_i}\frac{\partial}{\partial S_{l,m}} S_{l,\alpha_l} = \delta_{\alpha_l=m}\prod_{i\in[n]\setminus\{l\}} S_{i,\alpha_i} \ . \tag{3.8}$$

By setting $P = T(S)$, we may rewrite this as

$$\frac{\partial T_\alpha}{\partial S_{l,m}} = \delta_{\alpha_l=m}\prod_{i\in[n]\setminus\{l\}} S_{i,\alpha_i} \ . \tag{3.9}$$

In the following, every occurrence of S is meant as $S(P) = T^{-1}(P)$ (Lemma 1). The S-flow field X given by (3.6) has components

$$X(T^{-1}(P))_{l,m} = S_{l,m}\Big(\sum_{j\in[n]} \omega_{lj} S_{j,m} - \langle S_l, (\Omega S)_l\rangle\Big) \tag{3.10}$$

and the pushforward via T consequently reads

$$(T_\sharp X)(P)_\gamma = \sum_{l\in[n]}\sum_{m\in[c]} \frac{\partial T_\gamma}{\partial S_{l,m}} X(T^{-1}(P))_{l,m} \tag{3.11a}$$

$$= \sum_{l\in[n]}\Big(\prod_{i\in[n]\setminus\{l\}} S_{i,\gamma_i}\Big) S_{l,\gamma_l}\Big[\sum_{j\in[n]}\omega_{lj} S_{j,\gamma_l} - \langle S_l, (\Omega S)_l\rangle\Big] \tag{3.11b}$$

$$= \Big(\prod_{i\in[n]} S_{i,\gamma_i}\Big)\sum_{l\in[n]}\Big[(\Omega S)_{l,\gamma_l} - \langle S_l, (\Omega S)_l\rangle\Big] \tag{3.11c}$$

$$= P_\gamma\Big(\sum_{l\in[n]} (\Omega S)_{l,\gamma_l} - \langle S, (\Omega S)\rangle\Big) \ . \tag{3.11d}$$

We define the linear operator $\overline{\Omega} = Q(\Omega \otimes \mathbb{I}_c)Q^\top$ on Δ with Q defined by (3.2). Clearly, $\overline{\Omega}$ is symmetric exactly if Ω is symmetric. Lemma 2 now implies

$$\langle P, \overline{\Omega}P\rangle = \langle TS^v, Q(\Omega \otimes \mathbb{I}_c)Q^\top TS^v\rangle = \langle S^v, (\Omega \otimes \mathbb{I}_c)S^v\rangle = \langle S, \Omega S\rangle \tag{3.12}$$

because $S \in \mathcal{W}$, as well as

$$\sum_{l\in[n]} (\Omega S)_{l,\gamma_l} = Q(\Omega S)_\gamma = \big(Q(\Omega \otimes \mathbb{I})Q^\top TS^v\big)_\gamma = (\overline{\Omega}P)_\gamma\ . \tag{3.13}$$

Returning to (3.11), this gives

$$(T_\sharp X)(P)_\gamma = P_\gamma\big((\overline{\Omega}P)_\gamma - \langle P, \overline{\Omega}P\rangle\big) \tag{3.14}$$

which may be written more compactly as $(T_\sharp X)(P) = R_P[\overline{\Omega}P]$. Because the inverse map T^{-1} performs marginalization for each node, this shows the assertion. □

4 *S*-Flow Spectral Decomposition

Theorem 1 allows to view assignment flows as marginal dynamics of replicator systems. We aim to leverage this insight to transfer results from the study of replicator equations to the assignment flow setting. To this end, we first briefly describe a spectral decomposition result for replicator dynamics in Section 4.1 and subsequently apply it to the S-flow (2.5) in Section 4.2.

4.1 Selection Systems

Given a simplex $\Delta \subset \mathbb{R}^N$ and a symmetric matrix $\overline{\Omega} \in \mathbb{R}^{N\times N}$

$$\dot{P}(t) = R_{P(t)}[\overline{\Omega}P(t)], \qquad P(0) = P_0 \tag{4.1}$$

is called replicator equation [6] for linear "fitness" $\overline{\Omega}$ and initial value P_0. Dynamics of this type have been studied in mathematical biology by [8]. Let $\overline{\Omega} = \sum_{k=1}^{\operatorname{rank}\overline{\Omega}} \lambda_k h_k h_k^\top$ denote the spectral decomposition of $\overline{\Omega}$. In mathematical models of evolution, P describes the relative frequencies of N traits in a given population. Let $l(t) \in \mathbb{R}^N$ model the absolute number of individuals exhibiting each trait such that $P(t) = \frac{l(t)}{\langle l(t), \mathbb{1}_N\rangle}$, $l(0) = l_0$. The selection system equivalent to (4.1) reads

$$l(t) = l_0 \odot K(t), \qquad K(t) = \exp\Big(\sum_k s_k(t)h_k\Big) \tag{4.2}$$

with coefficients $s_k(t)$ following the so-called *escort system* dynamics [8]

$$\dot{s}_i(t) = \lambda_i \frac{\Big\langle P_0, h_i \odot \exp\big(\sum_k s_k(t)h_k\big)\Big\rangle}{\langle P_0, \exp\big(\sum_k s_k(t)h_k\big)\rangle}\ . \tag{4.3}$$

Proposition 1. *The replicator dynamics (4.1) is equivalent to*

$$P(t) = \exp_{P_0}\Big(\sum_k s_k(t)h_k\Big) \tag{4.4}$$

$$\dot{s}_k(t) = \lambda_k\langle h_k, P(t)\rangle, \qquad s_k(0)=0, \qquad k\in[\operatorname{rank}\overline{\Omega}] \tag{4.5}$$

Proof. The quantity l can be normalized to yield a corresponding assignment $P\in\Delta$. Let $l_0=l(0)$ be such that $\langle l_0,\mathbb{1}\rangle=1$. Then $P_0=l_0$ and we find

$$P(t) = \frac{P_0\odot\exp\big(\sum_k s_k(t)h_k\big)}{\Big\langle \mathbb{1},\, P_0\odot\exp\big(\sum_k s_k(t)h_k\big)\Big\rangle} \overset{(2.6)}{=} \exp_{P_0}\Big(\sum_k s_k(t)h_k\Big)\ . \tag{4.6}$$

Escort system dynamics can be transformed to

$$\begin{aligned}\dot{s}_i(t) &= \lambda_i\frac{\langle h_i,\, P_0\odot\exp\big(\sum_k s_k(t)h_k\big)\rangle}{\langle P_0, \exp\big(\sum_k s_k(t)h_k\big)\rangle} && (4.7a)\\ &= \lambda_i\Big\langle h_i,\, \exp_{P_0}\big(\sum_k s_k(t)h_k\big)\Big\rangle && (4.7b)\\ &= \lambda_i\langle h_i,\, P(t)\rangle && (4.7c)\end{aligned}$$

and the initial conditions (4.5) are consistent with (4.1). ☐

4.2 Spectral Decomposition

We aim to transfer the spectral decomposition of Proposition 1 to S-flows. To this end, the following lemmata describe further behavior of the maps T and Q introduced in (3.1) and (3.2).

Lemma 3. *For any matrix $V\in\mathbb{R}^{n\times c}$ it holds that $\langle Q(V),\mathbb{1}_N\rangle = c^{n-1}\langle V,\mathbb{1}_{n\times c}\rangle$.*

Proof. We directly compute

$$\begin{aligned}\sum_\gamma\sum_{l\in[n]} V_{l,\gamma_l} &= \sum_{l\in[n]}\sum_{m\in[c]}\sum_{\gamma_l=m} V_{l,\gamma_l} = \sum_{l\in[n]}\sum_{m\in[c]} c^{n-1}V_{l,m} && (4.8a)\\ &= c^{n-1}\langle V,\mathbb{1}_{n\times c}\rangle\ . && (4.8b)\end{aligned}$$

☐

Lemma 4. *For any $V\in T_0\mathcal{W}$ it holds that $T^{-1}(Q(V)) = c^{n-1}V$.*

Proof. For arbitrary indices $(i,j)\in[n]\times[c]$, we find

$$\begin{aligned}T^{-1}(Q(V))_{i,j} &\overset{(3.2),(3.3)}{=} \sum_{\gamma_i=j}\sum_{l\in[n]} V_{l,\gamma_l} = \sum_{\gamma_i=j}\Big(V_{i,\gamma_i} + \sum_{l\in[n]\setminus\{i\}} V_{l,\gamma_l}\Big) && (4.9a)\\ &= c^{n-1}V_{i,j} + \sum_{\gamma_i=j}\sum_{l\in[n]\setminus\{i\}} V_{l,\gamma_l}\ . && (4.9b)\end{aligned}$$

Let $\tilde{V}$ contain only the rows with indices $[n] \setminus \{i\}$ of V and let $\tilde{\gamma}$ denote multi-indices of its rows. Then the last sum in (4.9b) is proportional to

$$\sum_{l\in[n]\setminus\{i\}} \sum_{\tilde{\gamma}} \tilde{V}_{l,\tilde{\gamma}_l} = \sum_{l\in[n]\setminus\{i\}} \sum_{m\in[c]} \sum_{\tilde{\gamma}_l=m} \tilde{V}_{l,\tilde{\gamma}_l} \propto \sum_{l\in[n]\setminus\{i\}} \sum_{m\in[c]} \tilde{V}_{l,m} = 0 \tag{4.10}$$

□

Lemma 5. *If $V \in T_0\mathcal{W}$ and $\lambda \in \mathbb{R}$ satisfy $\Omega V = \lambda V$ then $\overline{V} = QV^v$ is an eigenvector of $\overline{\Omega} = Q(\Omega \otimes \mathbb{I}_c)Q^\top$ for eigenvalue $c^{n-1}\lambda$. Additionally, $\mathbb{1}_N$ is an eigenvalue of $\overline{\Omega}$ and $QU^v \propto \mathbb{1}_N$ if $U \propto \mathbb{1}_{n\times c}$.*

Proof. Let $V \in T_0\mathcal{W}$ and $\lambda \in \mathbb{R}$ satisfy $\Omega V = \lambda V$. Then by Lemma 4 it holds that

$$\overline{\Omega} QV^v = Q(\Omega \otimes \mathbb{I}_n)Q^\top QV^v = c^{n-1}Q(\Omega \otimes \mathbb{I}_n)V^v = c^{n-1}\lambda QV^v \ . \tag{4.11}$$

Now let $U \propto \mathbb{1}_{n\times c}$. Then $QU^v \propto \mathbb{1}_N$ by Lemma 3 and we find

$$\overline{\Omega}\mathbb{1}_N = Q(\Omega \otimes \mathbb{I}_n)Q^\top \mathbb{1}_N \propto Q(\Omega \otimes \mathbb{I}_n)\mathbb{1}_{nc} = Q\mathbb{1}_{nc} \propto \mathbb{1}_N \tag{4.12}$$

by using Lemma 2. □

Lemma 6. *It holds $\ker Q = \{\mathrm{Diag}(d)\mathbb{1}_{n\times c} \colon d \in \mathbb{R}^n,\ \langle d, \mathbb{1}_n\rangle = 0\}$ as well as $\operatorname{rank} Q = nc - (n-1)$.*

Proof. Let $V \in \ker Q$ and let $\gamma, \tilde{\gamma}$ be two fixed multi-indices which differ exactly at position i but are otherwise arbitrary. We have $(QV)_\gamma = (QV)_{\tilde{\gamma}} = 0$ by assumption. Thus

$$\begin{aligned}(QV)_{\tilde{\gamma}} &= V_{i,\tilde{\gamma}_i} + \sum_{l\in[n]\setminus\{i\}} V_{l,\tilde{\gamma}_l} = V_{i,\tilde{\gamma}_i} + \sum_{l\in[n]\setminus\{i\}} V_{l,\gamma_l} &(4.13a)\\ &= (QV)_\gamma = V_{i,\gamma_i} + \sum_{l\in[n]\setminus\{i\}} V_{l,\gamma_l} &(4.13b)\end{aligned}$$

which implies $V_{i,\tilde{\gamma}_i} = V_{i,\gamma_i}$, i.e. $V = \mathrm{Diag}(d)\mathbb{1}_{n\times c}$ for some $d \in \mathbb{R}^n$ since i was arbitrary. Let V have this form. Then

$$(QV)_\gamma = \sum_{l\in[n]} V_{l,\gamma_l} = \sum_{l\in[n]} d_l = \langle d, \mathbb{1}_n\rangle \ . \tag{4.14}$$

so V is in the kernel of Q exactly if $\langle d, \mathbb{1}_n\rangle = 0$. There are $(n-1)$ linearly independent vectors $d \in \mathbb{R}^n$ with this property, therefore Q has the specified rank. □

Proposition 2 (Eigenvectors of $\overline{\Omega}$). *Let $\Omega \in \mathbb{R}^{n\times n}$ be a symmetric matrix of full rank such that $\Omega\mathbb{1}_n = \mathbb{1}_n$. Then $\overline{\Omega}$ has rank $nc - (n-1)$ and there is an orthogonal matrix $V \in \mathbb{R}^{nc\times n(c-1)}$ such that the columns of V are eigenvectors of $\Omega \otimes \mathbb{I}_c$ and the columns of $\overline{V} = \sqrt{\frac{c}{N}}QV$ are pairwise orthonormal eigenvectors of $\overline{\Omega}$. The columns of $\overline{V}$ together with the vector $\frac{1}{\sqrt{N}}\mathbb{1}_N$ form an orthonormal basis of $\operatorname{img}\overline{\Omega}$.*

Proof. Let $G \in \mathbb{R}^{n\times n}$ be an orthogonal matrix of eigenvectors of Ω and let $\Lambda = \mathrm{Diag}(\lambda_1, \dots, \lambda_n)$ have the respective eigenvalues as diagonal entries. Then $\Omega = G\Lambda G^\top$. Now let $H \in \mathbb{R}^{c\times c}$ be an orthogonal matrix with first column $h_1 = \frac{1}{\sqrt{c}}\mathbb{1}_c$ such that the remaining columns span $\{d \in \mathbb{R}^c \colon \langle d, \mathbb{1}_c\rangle = 0\}$. We find

$$\begin{aligned}(G\otimes H)(\Lambda\otimes \mathbb{I}_c)(G\otimes H)^\top &= (G\otimes H)(\Lambda\otimes \mathbb{I}_c)(G^\top\otimes H^\top) && (4.15a)\\ &= (G\otimes H)((\Lambda G^\top)\otimes H^\top) && (4.15b)\\ &= (G\Lambda G^\top)\otimes (HH^\top) && (4.15c)\\ &= \Omega\otimes \mathbb{I}_c && (4.15d)\end{aligned}$$

as well as $(G\otimes H)(G\otimes H)^\top = (GG^\top)\otimes(HH^\top) = \mathbb{I}_{cn}$ so $(G\otimes H)$ is an orthogonal matrix whose columns are eigenvectors of $\Omega\otimes\mathbb{I}_c$. Viewing the eigenvectors V_i with indices $i \in I_1 = \{1 + c(k-1) \colon k \in [n]\}$ as matrices $V_i \in \mathbb{R}^{n\times c}$, we have $V_i^v = g_k \otimes h_1$ which gives $V_i = g_k h_1^\top = c^{-\frac{1}{2}}\,\mathrm{Diag}(g_k)\mathbb{1}_{n\times c}$ and thus $Q(V_i) = \frac{1}{\sqrt{c}}\langle g_k, \mathbb{1}_n\rangle \mathbb{1}_N \propto \mathbb{1}_N$. Therefore, the rank of $\overline{\Omega}$ is at most $nc - (n-1)$. For the remaining indices $j = l + c(k-1)$, $k\in[n]$, $l \in [c]\setminus\{1\}$ it holds that

$$V_j\mathbb{1}_c = g_k h_l^\top \mathbb{1}_c = g_k\langle h_l, \mathbb{1}_c\rangle = 0 \tag{4.16}$$

thus $V_j \in T_0\mathcal{W}$. Denote the set of these indices by $I_2 = [nc]\setminus I_1$. By Lemma 4 it holds that

$$\langle \overline{V}_{j_1}, \overline{V}_{j_2}\rangle = \frac{c}{N}\langle QV_{j_1}^v, QV_{j_2}^v\rangle = \frac{c}{N}\langle V_{j_1}^v, Q^\top Q V_{j_2}^v\rangle = \langle V_{j_1}, V_{j_2}\rangle = 0 \tag{4.17}$$

for all $j_1, j_2 \in I_2$ and $\|\overline{V}_j\|_2^2 = \frac{c}{N}\langle QV_j^v, QV_j^v\rangle = \|V_j^v\|_2^2 = 1$ for $j \in I_2$. We additionally find

$$\langle \overline{V}_j, \frac{1}{\sqrt{N}}\mathbb{1}_N\rangle = \frac{\sqrt{c}}{N}\langle QV_j^v, \mathbb{1}_N\rangle = \frac{1}{\sqrt{c}}\langle V_j, \mathbb{1}_{n\times c}\rangle = 0 \qquad j \in I_2 \tag{4.18}$$

by using Lemma 3. Because all columns of $\overline{V}$ are eigenvectors of $\overline{\Omega}$ by Lemma 5, this shows the assertion. □

Lemma 7 (Lifting Map Lemma). *Let $S \in \mathcal{W}$ and $V \in \mathbb{R}^{n\times c}$. Then*

$$T(\exp_S(V)) = \exp_{T(S)}(Q(V))\ . \tag{4.19}$$

Proof. We have $T(\exp(V)) = \exp(Q(V))$ because for any multi-index γ

$$\exp(Q(V))_\gamma = \exp\Big(\sum_{l\in[n]} V_{l,\gamma_l}\Big) = \prod_{l\in[n]}\exp\Big(V_{l,\gamma_l}\Big) = T(\exp(V))_\gamma\ . \tag{4.20}$$

Let $D \in \mathbb{R}^{n\times n}$ be a diagonal matrix with nonzero diagonal entries. Then $T(DR) \propto T(R)$ for any $R \in \mathbb{R}^{n\times c}$ because

$$T(DR)_\gamma = \prod_{l\in[n]}(DR)_{l,\gamma_l} = \Big(\prod_{l\in[n]} D_{ll}\Big)\Big(\prod_{l\in[n]} R_{l,\gamma_l}\Big) \propto T(R)_\gamma\ . \tag{4.21}$$

Due to $T(S \odot V) = T(S) \odot T(V)$ it follows

$$T(\exp_S(V)) \propto T(S \odot \exp(V)) = T(S) \odot \exp(Q(V)) \propto \exp_{T(S)}(Q(V)) \ . \tag{4.22}$$

Because both first and last term in (4.22) are clearly elements of Δ, this implies the assertion. □

Theorem 2 (Spectral Decomposition). *Let $\Omega \in \mathbb{R}^{n\times n}$ be a symmetric matrix of full rank such that $\Omega \mathbb{1}_n = \mathbb{1}_n$ and let $V \in \mathbb{R}^{nc\times nc}$ be given by Proposition 2. Then the S-flow (3.6) is equivalent to*

$$S(t) = \exp_{S_0}\Big(\sum_{k\in I_2} s_k(t) V_k\Big) \tag{4.23}$$

$$\dot{s}_k(t) = \lambda_k \langle V_k, S(t)\rangle, \qquad s_k(0) = 0, \qquad k \in I_2 \tag{4.24}$$

where $I_2 := [nc] \setminus \{1 + c(k-1) \colon k \in [n]\}$ and $V_k \in \mathbb{R}^{n\times c}$ denotes the unique matrix such that V_k^v is the k-th column of V.

Proof. By Propositions 1 and 2 it holds that

$$S(t) = T^{-1}(P(t)) = T^{-1}\Big(\exp_{P_0}\Big(\sum_{k\in I_1\cup I_2} s_k(t)\overline{V}_k\Big)\Big) \tag{4.25a}$$

$$= T^{-1}\Big(\exp_{P_0}\Big(\sum_{k\in I_2} s_k(t)\overline{V}_k\Big)\Big) = \exp_{S_0}\Big(\sum_{k\in I_2} s_k(t) V_k\Big) \tag{4.25b}$$

where we used Lemma 7 as well as $\overline{V}_k \propto \mathbb{1}_N$ for $k \in I_1$ in (4.25b). Additionally, for $k \in I_2$ it holds

$$\dot{s}_k(t) = c^{n-1}\lambda_k \langle P(t), \overline{V}_k\rangle = \lambda_k\langle P(t), QV_k\rangle \overset{\text{Lemma 2}}{=} \lambda_k \langle S(t), V_k\rangle \ . \tag{4.26}$$

□

5 Experiments

Because $\mathcal{W}$ is not a flat space, standard methods are not canonically suitable for numerical integration of assignment flows. A remedy is to construct a diffeomorphism $\varphi \colon \mathcal{W} \to V$ and integrate the φ-related vector field in some flat space V. Theorem 2 implicitly achieves this by parameterizing S-flows through the coefficients $s_k(t)$, $k \in I_2$ which live in the unbounded flat space $\mathbb{R}$. Standard methods of numerical integration such as Runge-Kutta or linear multistep methods are therefore directly applicable to the dynamics (4.24). In our empirical examination, we consider a graph of 100×100 grid pixel nodes with adjacency Ω computed analogous to nonlocal means denoising [3], i.e. the averaging weight Ω_{i_1,i_2} is large if i_1 is close to i_2 in the image plane and if the patch of pixels around i_1 is similar to the patch of pixels around i_2. We added independent noise drawn from a normal distribution to each channel of the original cartoon image in Fig. 1 and aim to recover the original prototype colors

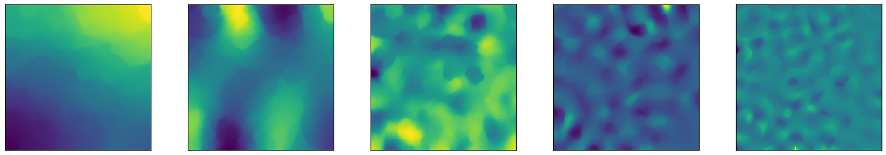

Fig. 2. Eigenvectors of the adjacency matrix used in Figure 1. More dominant parts of the spectrum correspond to low-frequency eigenvectors.

in the RGB feature space. Euclidean distances for each pixel-prototype pair are collected in a distance matrix $D \in \mathbb{R}^{n \times c}$ (here, $n = 100 \cdot 100$, $c = 47$). The point $S_0 = \exp_{\mathbb{1}_{\mathcal{W}}}(-\frac{1}{\rho}D)$ is used to initialize a reference S-flow which we integrate numerically by the geometric Euler method with step-length $h = 0.1$ until a low-entropy assignment state is reached.

To visualize the spectral decomposition of Theorem 2, we first compute approximations to the dominant $k \in \{500, 1000, 1500\}$ eigenvectors of Ω by leveraging sparsity. A spectral approximation of the reference S-flow is obtained by dropping the remaining eigenvalues from (4.24), i.e. $\lambda_j \leftarrow 0$ for $j > k$. Numerical integration with the explicit Euler method (constant step-length $h = 0.1$) yields the integral label assignments shown in Fig. 1. As expected, the approximation becomes progressively more faithful to the reference S-flow if a larger fraction of the spectrum of Ω is considered. Note that Ω has full rank $n = 10^4$ so the approximations in Fig. 1 are obtained by considering at most 15% of its spectrum. Clearly, the described method can lead to improved computational efficiency, if the adjacency Ω can be well approximated by a low-rank matrix. The given example illustrates that informative regularizations may be achieved using relatively low-rank adjacency. We additionally observe that less dominant eigenvalues correspond to high-frequency components (see Fig. 2). This is consistent with the approximation behavior shown in Fig. 1; discarding the influence of less dominant eigenvalues leads to smoothing out high-frequency detail while retaining correct label assignment for largely uniform regions.

6 Conclusion

We have constructed an embedding of the assignment manifold $\mathcal{W}$ into a single probability simplex Δ. Under this embedding, S-assignment flows are pushed to replicator dynamics with linear fitness function. Because Δ has intractably large dimension, numerical integration can not be performed directly. However, we show that the embedding into Δ serves as a valuable tool to transfer structural results on replicator dynamics to the study of assignment flows. Conversely, Theorem 1 also identifies a class of replicator dynamics with linear fitness which can be decomposed into much lower-dimensional S-flows. A systematic study of applications is left for future work.

Acknowledgements

This work is supported by the Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC 2181/1 - 390900948 (the Heidelberg STRUCTURES Excellence Cluster).
The original artwork used in Figure 1 was designed by dgim-studio / Freepik.

References

1. Åström, F., Petra, S., Schmitzer, B., Schnörr, C.: Image Labeling by Assignment. Journal of Mathematical Imaging and Vision **58**(2), 211–238 (2017)
2. Benning, M., Möller, M., Nossek, R.Z., Burger, M., Cremers, D., Gilboa, G., Schönlieb, C.B.: Nonlinear spectral image fusion. In: Lauze, F., Dong, Y., Dahl, A.B. (eds.) Scale Space and Variational Methods in Computer Vision. pp. 41–53. Springer International Publishing, Cham (2017)
3. Buades, A., Coll, B., Morel, J.M.: Image Denoising Methods. A New Nonlocal Principle. SIAM Review **52**(1), 113–147 (2010)
4. Bungert, L., Burger, M., Tenbrinck, D.: Computing nonlinear eigenfunctions via gradient flow extinction. In: Lellmann, J., Burger, M., Modersitzki, J. (eds.) Scale Space and Variational Methods in Computer Vision. pp. 291–302. Springer International Publishing, Cham (2019)
5. Burger, M., Eckardt, L., Gilboa, G., Moeller, M.: Spectral representations of one-homogeneous functionals. In: Aujol, J.F., Nikolova, M., Papadakis, N. (eds.) Scale Space and Variational Methods in Computer Vision. pp. 16–27. Springer International Publishing, Cham (2015)
6. Fisher, R.A.: The Genetical Theory of Natural Selection. Dover, New York (1985)
7. Gilboa, G., Moeller, M., Burger, M.: Nonlinear spectral analysis via one-homogeneous functionals: Overview and future prospects. Journal of Mathematical Imaging and Vision **56**(2), 300–319 (2016)
8. Karev, G.P.: On mathematical theory of selection: continuous time population dynamics. Journal of Mathematical Biology **60**(1), 107–129 (2010)
9. Karev, G.P.: Replicator equations and the principle of minimal production of information. Bulletin of Mathematical Biology **72**(5), 1124–1142 (2010)
10. Page, K.M., Nowak, M.A.: Unifying evolutionary dynamics. Journal of Theoretical Biology **219**(1), 93 – 98 (2002)
11. Savarino, F., Schnörr, C.: Continuous-Domain Assignment Flows. Europ. J. Appl. Math. (2020)
12. 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)
13. Zeilmann, A., Savarino, F., Petra, S., Schnörr, C.: Geometric Numerical Integration of the Assignment Flow. Inverse Problems **36**(3), 034004 (33pp) (2020)
14. Zern, A., Zeilmann, A., Schnörr, C.: Assignment Flows for Data Labeling on Graphs: Convergence and Stability. preprint: arXiv (2020), https://arxiv.org/abs/2002.11571