On the Geometric Mechanics of Assignment Flows for Metric Data Labeling

Fabrizio Savarino[1], Peter Albers[2], and Christoph Schnörr[1]

[1] Institute of Applied Mathematics, Heidelberg University, Germany
[2] Mathematical Institute, Heidelberg University, Germany
fabrizio.savarino@iwr.uni-heidelberg.de

Abstract. Assignment flows are a general class of dynamical models for context dependent data classification on graphs. These flows evolve on the product manifold of probability simplices, called assignment manifold, and are governed by a system of coupled replicator equations. In this paper, we adopt the general viewpoint of Lagrangian mechanics on manifolds and show that assignment flows satisfy the Euler-Lagrange equations associated with an action functional. Besides providing a novel interpretation of assignment flows, our result rectifies the analogous statement of a recent paper devoted to uncoupled replicator equations evolving on a single simplex, and generalizes it to coupled replicator equations and assignment flows.

Keywords: action functional · assignment flows · image labeling · replicator equation · evolutionary game dynamics

1 Introduction

Assignment flows, originally introduced by [4], are a general class of dynamical models evolving on a statistical manifold $\mathcal{W}$, called *assignment manifold*, for context dependent data classification on graphs. We refer to [13] for a recent survey on assignment flows and related work.

This approach is formulated for a general graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ and can be summarized as follows. Assume for every node $i \in \mathcal{V}$ some data point f_i in a metric space $(\mathcal{F}, d_{\mathcal{F}})$ to be given, together with a set $\mathcal{F}_* = \{f_1^*, \dots, f_n^*\} \subset \mathcal{F}$ of predefined prototypes, also called *labels*. Context based *metric data labeling* refers to the task of assigning to each node $i \in \mathcal{V}$ a suitable label in $\mathcal{F}_*$ based on the metric distance to the given data f_i and the relation between data points encoded by the edge set $\mathcal{E}$.

In order to derive a geometric representation of this problem, the discrete label choice at each node $i \in \mathcal{V}$ is relaxed to a probability distribution over the label space $\mathcal{F}_*$ with full support, represented as a point on the manifold

$$\mathcal{S} := \{p \in \mathbb{R}^n : p > 0 \text{ and } \langle p, \mathbb{1}_n \rangle = 1\}. \tag{1.1}$$

Accordingly, all probabilistic label choices on the graph are encoded as a single point $W \in \mathcal{W}$ on the assignment manifold

$$\mathcal{W} := \mathcal{S} \times \ldots \times \mathcal{S} \qquad (m := |\mathcal{V}| \text{ factors}), \tag{1.2}$$

where the i-th component of $W = (W_k)_{k\in\mathcal{V}}$ represents the probability distribution of label assignments $W_i = (W_i^1, \dots, W_i^n)^\top \in \mathcal{S}$ for the node $i \in \mathcal{V}$. Assignment flows are dynamical systems on $\mathcal{W}$ for inferring probabilistic label assignments that take the form of coupled replicator equations (see Section 4)

$$\dot{W}(t) = \mathcal{R}_{W(t)}[F(W(t))], \quad \text{with} \quad W(t) \in \mathcal{W}, \tag{1.3}$$

where the initial condition $W(0) \in \mathcal{W}$ contains information about the given data points $f_i \in \mathcal{F}$, $i \in \mathcal{V}$. These flows are derived by information geometric principles and usually consist of two interacting processes: non-local regularization of probabilistic label assignments and gradually enforcing unambiguous local decisions at every node $i \in \mathcal{V}$.

In [10, Thm. 2.1], the authors claim that all *uncoupled* replicator equations, i.e. $\dot{p} = R_p F(p)$, on a single simplex, $p(t) \in \mathcal{S}$, satisfy the Euler-Lagrange equation associated with the cost functional (again, see Section 4 for more details)

$$\mathcal{L}(p) := \int_{t_0}^{t_1} \frac{1}{2}\|\dot{p}(t)\|_g^2 + \frac{1}{2}\|R_{p(t)}F(p(t))\|_g^2 dt \quad \text{for curves } p\colon [t_0, t_1] \to \mathcal{S}. \tag{1.4}$$

In this paper, we (i) generalize this result to assignment flows and (ii) show that, in contrast to the claim of [10], the mentioned relation to extremal points of (1.4) holds if and only if condition (1.7) is fulfilled. Unlike the approach taken in [10], we derive this generalization from the more general viewpoint of Lagrangian mechanics on manifolds. This results in a better interpretable version of the Euler-Lagrange equation and leads to a characterization of critical points of the functional in terms of the function F governing the coupled replicator dynamics (1.3). Our main result is summarized in the following theorem.

Theorem 1. *Suppose $F\colon U \to \mathbb{R}^{m\times n}$ is a fitness function defined on an open set $U \subset \mathbb{R}^{m\times n}$ containing $\mathcal{W}$. If $W\colon I = [t_0, t_1] \to \mathcal{W}$ is a solution of the assignment flow (1.3), then $W(t)$ is a critical point of the action functional*

$$\mathcal{L}(W) = \int_{t_0}^{t_1} \frac{1}{2}\|\dot{W}(t)\|_g^2 + \frac{1}{2}\sum_{i\in\mathcal{V}} \mathrm{Var}_{W_i(t)}\big(F_i(W(t))\big) dt, \tag{1.5}$$

that is, $W(t)$ fulfills the Euler-Lagrange equation

$$D_t^g \dot{W}(t) = \frac{1}{2}\sum_{i\in\mathcal{V}} \mathrm{grad}^g \mathrm{Var}_{W_i(t)}\big(F_i(W(t))\big) \quad \textit{for } t \in I = [t_0, t_1], \tag{1.6}$$

if and only if the fitness function F fulfills the condition

$$0 = \mathcal{R}_{W(t)} \circ \big(dF|_{W(t)} - (dF|_{W(t)})^*\big) \circ \mathcal{R}_{W(t)}[F(W(t))], \ \textit{for } t \in I = [t_0, t_1], \tag{1.7}$$

where $(dF|_{W(t)})^*$ is the adjoint linear operator of $dF|_{W(t)}\colon \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ with respect to the Frobenius inner product and $\mathcal{R}_{W(t)}$ is defined in (4.7).

The paper is organized as follows. In Section 2, we introduce our notation and list the necessary ingredients from differential geometry. Section 3 summarizes the required theory of Lagrangian systems on manifolds. Basic properties of assignment manifolds and flows are presented in Section 4, followed by the proof of Theorem 1 together with a counter example for the general claims of [10].

2 Preliminaries

Basic Notation. In accordance with the standard notation in differential geometry, coordinates of vectors have upper indices. For any $k \in \mathbb{N}$, we define $[k] := \{1,\dots,k\} \subset \mathbb{N}$. The standard basis of $\mathbb{R}^d$ is denoted by $\{e_1,\dots,e_d\}$ and we set $\mathbb{1}_d := (1,\dots,1)^\top \in \mathbb{R}^d$. The notation $\langle\cdot,\cdot\rangle$ is used for both, the standard and Frobenius inner product between vectors and matrices respectively. The identity matrix is denoted by $I_d \in \mathbb{R}^{d\times d}$ and the i-th row vector of any matrix A by A_i. The linear dependence of a function F on its argument x is indicated by square brackets $F[x]$. If x is a vector and F a matrix, then Fx is used instead of $F[x]$. For $a,b \in \mathbb{R}^d$, we denote componentwise multiplication (Hadamard product) by $a \diamond b := \mathrm{Diag}(a)b = (a^1b^1,\dots,a^db^d)^\top$ and division, for $b > 0$, simply by $\frac{a}{b} = (\frac{a^1}{b^1},\dots,\frac{a^d}{b^d})^\top$. Similarly, inequalities between vectors or matrices are to be understood componentwise. We further set $a^{\diamond k} := a^{\diamond(k-1)} \diamond a$ with $a^{\diamond 0} := \mathbb{1}_d$. For later reference, we record the following statement here.

Lemma 1. *Assume for each $i \in [k]$ a matrix $Q^i \in \mathbb{R}^{d\times d}$ is given and let $\mathcal{Q}\colon \mathbb{R}^{k\times d} \to \mathbb{R}^{k\times d}$ be the linear map defined by $(\mathcal{Q}[X])_i := Q^iX_i$ for all rows $i \in [k]$. Then, the adjoint linear map $\mathcal{Q}^*$ with respect to the Frobenius inner product is given by $(\mathcal{Q}^*[Y])_i = Q^{i\top}Y_i$ for all $i \in [k]$.*

Proof. This is a direct consequence of $\langle X, \mathcal{Q}^*[Y]\rangle = \sum_{i\in[k]}\langle X_i, (\mathcal{Q}^*[Y])_i\rangle$ and $\langle X, \mathcal{Q}^*[Y]\rangle = \langle \mathcal{Q}[X], Y\rangle = \sum_{i\in[k]}\langle Q^iX_i, Y_i\rangle = \sum_{i\in[k]}\langle X_i, Q^{i\top}Y_i\rangle$ for arbitrary matrices $X, Y \in \mathbb{R}^{k\times d}$. □

Differential Geometry. We assume the reader is familiar with the basic concepts of Riemannian and symplectic manifolds as introduced in standard textbooks, e.g. [8], [9] or [7]. The term "manifold" always means smooth manifold. The tangent and cotangent bundles of a d-dimensional manifold M are $TM = \cup_{x\in M}\{x\} \times T_xM$ and $T^*M = \cup_{x\in M}\{x\} \times T_x^*M$, together with their natural projections $\pi\colon TM \to M$ and $\pi^*\colon T^*M \to M$, sending $(x,v) \in TM$ and $(x,\alpha) \in T^*M$ to x. For local coordinates $(x^1,\dots,x^d)$ on M, a tangent vector $v \in T_xM$ in these coordinates takes the form $v = \sum_{i\in[d]} v^i \frac{\partial}{\partial x^i}|_x$. The differential of a smooth map between manifolds $F\colon M \to N$ at $x \in M$ applied to a vector $v \in T_xM$ is denoted by $dF|_x[v]$. As usual (see e.g. [2, Sec. 3.5.7]), if $M \subset V$ is an embedded submanifold of a vector space V, such as $\mathbb{R}^d$ or $\mathbb{R}^{k\times d}$, then the tangent space at $x \in M$ is identified with the set of velocities of curves through x and, by abuse of notation, we again use T_xM to denote this space

$$T_xM = \{\dot\gamma(0) \in V\colon \gamma \text{ curve in } M \text{ with } \gamma(0) = x\}. \tag{2.1}$$

If N is another submanifold of a vector space V', then the differential $dF|_x[v]$ of a map $F\colon M \to N$ at $x \in M$ can be calculated via a curve $\eta\colon (-\varepsilon,\varepsilon) \to M$, with $\eta(0) = x$ and $\dot\eta(0) = v \in T_xM$, by $dF|_x[v] = \frac{d}{dt}F(\eta(t))|_{t=0}$. Let $I \subset \mathbb{R}$ be an interval. If $\gamma\colon I \to TM$ is an integral curve of a vector field X on TM (or T^*M), i.e. $\dot\gamma(t) = X(\gamma(t))$, then $\pi \circ \gamma\colon I \to M$ (or $\pi^* \circ \gamma$) is called the *base* *integral curve*. For a Riemannian metric h (and similarly for a symplectic form ω) on M, there is a canonical isomorphisms $h^\flat\colon TM \to T^*M$, given by sending a tangent vector $v \in T_xM$ to the one-form $h_x(v,\cdot)\colon T_x^*M \to \mathbb{R}$. Its inverse $h^\sharp\colon T^*M \to TM$ sends a one-form $\alpha \in T_x^*M$ to a unique vector $v_\alpha \in T_xM$ such that $\alpha = h^\flat(v_\alpha) = h(v_\alpha,\cdot)$ holds. In particular, the Riemannian gradient of a function $f\colon M \to \mathbb{R}$ is defined as $\operatorname{grad}^h f := h^\sharp(df)$, where $df = \sum_i \frac{\partial f}{\partial x^i} dx^i$ is the differential of f. Thus, for all $x \in M$, $\operatorname{grad}^h f(x)$ is the unique vector with

$$df|_x[v] = h_x(\operatorname{grad}^h f(x), v), \quad \text{for all } v \in T_xM. \tag{2.2}$$

Furthermore, the Riemannian norm for $v \in T_xM$ is denoted by $\|v\|_h := \sqrt{h_x(v,v)}$ and the covariant derivative (with respect to the Riemannian metric h) of vector fields along curves by D_t^h.

3 Lagrangian Systems on Manifolds

In this section, we give a brief summary of Lagrangian systems on manifolds and mechanics on Riemannian manifolds from [1, Ch. 3].

Suppose M is a smooth manifold. Similar to Hamiltonian systems on momentum phase space T^*M, there is a related concept on the tangent bundle TM, interpreted as *velocity phase space*. In this context, a smooth function $L\colon TM \to \mathbb{R}$ is called *Lagrangian*. For a given point $x \in M$, denote the restriction of L to the fiber T_xM by $L_x := L|_{T_xM}\colon T_xM \to \mathbb{R}$. The *fiber derivative* of L is defined as

$$\mathbb{F}L\colon TM \to T^*M, \quad (x,v) \mapsto \mathbb{F}L(x,v) := dL_x|_v, \tag{3.1}$$

where $dL_x|_v\colon T_xM \to \mathbb{R}$ is the differential of L_x at $v \in T_xM$. The function L is called a *regular Lagrangian* if $\mathbb{F}L$ is regular at all points (meaning that $\mathbb{F}L$ is a submersion), which is equivalent to $\mathbb{F}L\colon TM \to T^*M$ being a local diffeomorphism by [1, Prop. 3.5.9]. Furthermore, L is called *hyperregular Lagrangian* if $\mathbb{F}L\colon TM \to T^*M$ is a diffeomorphism. A class of hpyerregular Lagrangians, including the Lagrangian from Theorem 1, is given in (3.6) below.

The *Lagrange two-form* is defined as the pullback $\omega_L := (\mathbb{F}L)^*\omega^{\text{can}}$ of the canonical symplectic form ω^{can} on the cotangent bundle T^*M under the fiber derivative $\mathbb{F}L$. According to [1, Prop. 3.5.9], ω_L is a symplectic form on T^*M if and only if L is a regular Lagrangian. In the following, we only consider regular Lagrangians. The *action* associated to the Lagrangian $L\colon TM \to \mathbb{R}$ is defined by

$$A\colon TM \to \mathbb{R}, \quad (x,v) \mapsto \mathbb{F}L(x,v)[v] = dL_x|_v[v], \tag{3.2}$$

and the *energy function* by $E := A - L$, having the form

$$E\colon TM \to \mathbb{R}, \quad (x,v) \mapsto \mathbb{F}L(x,v)[v] - L(x,v) = dL_x|_v[v] - L(x,v). \tag{3.3}$$

The *Lagrangian vector field for* L is the unique vector field X_E on TM satisfying

$$dE|_{(x,v)}[u] = \omega_{L,(x,v)}(X_E, u) \quad \text{for all } (x,v) \in T_xM \text{ and } u \in T_{(x,v)}TM, \tag{3.4}$$

that is $X_E = \omega_L^\sharp(dE)$. A curve $\gamma(t) = (x(t), v(t))$ on TM is an integral curve of X_E if $v(t) = \dot{x}(t)$ and the classical Euler-Lagrange equations in local coordinates are satisfied

$$\frac{d}{dt}\left(\frac{\partial L}{\partial v^i}\big(x(t), \dot{x}(t)\big)\right) = \frac{\partial L}{\partial x^i}\big(x(t), \dot{x}(t)\big) \quad \text{for all } i \in [n]. \tag{3.5}$$

Let $\gamma\colon I \to TM$ be any integral curve of X_E. Because of $\frac{d}{dt}E(\gamma) = dE|_\gamma[\dot{\gamma}] = dE|_\gamma[X_E(\gamma)] = \omega_{L,\gamma}(X_E(\gamma), X_E(\gamma)) = 0$, the energy E is constant along γ.

Now, assume (M, h) is a Riemannian manifold. Suppose a smooth function $G\colon M \to \mathbb{R}$, called *potential*, is given and consider the Lagrangian

$$L(x, v) := \frac{1}{2}\|v\|_h^2 - G(x), \quad \forall (x, v) \in TM. \tag{3.6}$$

It then follows (see [1, Sec. 3.7] or by direct computation) that the fiber derivative of L is the canonical isomorphism $\mathbb{F}L = h^\flat\colon TM \to T^*M$. Hence, the Lagrangian L is hyperregular with action A and energy $E = A - L$ given by

$$A(x, v) = \|v\|_h^2 \quad \text{and} \quad E(x, v) = \frac{1}{2}\|v\|_h^2 + G(x) \quad \text{for all } (x, v) \in TM. \tag{3.7}$$

Proposition 1. ***([1, Prop. 3.7.4]).*** *With L as defined in (3.6) on the Riemannian manifold (M, h), the curve $\gamma\colon I \to TM$ with $\gamma(t) = (x(t), v(t))$ is an integral curve of the Lagrangian vector field X_E, i.e. satisfies the Euler-Lagrange equation, if and only if the base integral curve $\pi \circ \gamma = x\colon I \to M$ satisfies*

$$D_t^h \dot{x}(t) = -\operatorname{grad}^h G(x(t)). \tag{3.8}$$

4 Mechanics of Assignment Flows

We now return to the metric data labeling task on a graph $\mathcal{G} = (\mathcal{V}, \mathcal{E})$ from the beginning of this paper. In this section, we consistently use the notation $m = |\mathcal{V}|$ and $n = |\mathcal{F}_*|$ for the number of nodes and labels respectively.

We first give a brief summary of the most important properties of the statistical manifold $\mathcal{W}$, followed by a short description of assignment flows. A more detailed overview can be found in the original work [4] or the recent survey [13]. After this, we apply the general theory of Lagrangian Systems from Section 3 to prove our main result stated as Theorem 1.

4.1 Assignment Manifold and Flows

Assignment Manifold. In the following, we always identify the manifold $\mathcal{W}$ from (1.2) with its matrix embedding

$$\mathcal{W} = \{W \in \mathbb{R}^{m\times n}\colon W > 0 \text{ and } W\mathbb{1}_n = \mathbb{1}_m\}, \tag{4.1}$$

by sending the i-th component W_i of $W = (W_k)_{k\in\mathcal{V}} \in \mathcal{W}$ to the i-th row of a matrix in $\mathbb{R}^{m\times n}$. Therefore, points $W \in \mathcal{W}$ are row stochastic matrices with full support, called *assignment matrices*, with row vectors $W_i = (W_i^1, \dots, W_i^n)^\top \in \mathcal{S}$ representing the relaxed label assignment for every $i \in [m]$. With the identification from (2.1), the tangent space of $\mathcal{S} \subset \mathbb{R}^n$ from (1.1) at any point $p \in \mathcal{S}$ is identified as

$$T_p\mathcal{S} = \{v \in \mathbb{R}^n \colon \langle v, \mathbb{1}_n \rangle = 0\} =: T. \tag{4.2}$$

Hence, $T_p\mathcal{S}$ is represented by the same vector space T for all $p \in \mathcal{S}$. In particular, the tangent bundle is trivial $T\mathcal{S} = \mathcal{S} \times T$. Viewing $\mathcal{W}$ as an embedded submanifold of $\mathbb{R}^{m\times n}$ by (4.1) and using the identification (2.1) for the tangent space, we identify

$$T_W\mathcal{W} = \{V \in \mathbb{R}^{m\times n} \colon V\mathbb{1}_n = 0\} =: \mathcal{T}, \quad \text{for all } W \in \mathcal{W} \subset \mathbb{R}^{m\times n}. \tag{4.3}$$

With this identification, the tangent bundle is also trivial $T\mathcal{W} = \mathcal{W} \times \mathcal{T}$.

From an information geometric viewpoint, e.g. [3] or [5], the Fisher-Rao (information) metric is a "canonical" Riemannian structure on $\mathcal{S}$, given by

$$g_p(u, v) := \big\langle u, \mathrm{Diag}\left(\tfrac{1}{p}\right) v\big\rangle, \quad \text{for all } p \in \mathcal{S}, u, v \in T = T_p\mathcal{S}. \tag{4.4}$$

Next, we define two important matrices, the orthogonal projection of $\mathbb{R}^n$ onto T with respect to the Euclidean inner product

$$P_T := I_n - \tfrac{1}{n}\mathbb{1}_n\mathbb{1}_n \in \mathbb{R}^{n\times n} \quad \text{viewed as} \quad P_T \colon \mathbb{R}^n \to T \tag{4.5}$$

and for every $p \in \mathcal{S}$ the *replicator matrix*

$$R_p := \mathrm{Diag}(p) - pp^\top \in \mathbb{R}^{n\times n} \quad \text{viewed as} \quad R_p \colon \mathbb{R}^n \to T. \tag{4.6}$$

A simple calculation shows that $R_p = R_pP_T = P_TR_p$ as well as $\ker(R_p) = \mathbb{R}\mathbb{1}_n$ hold for all $p \in \mathcal{S}$. Furthermore, if R_p is restricted to the linear subspace $T \subset \mathbb{R}^n$, then $R_p|_T \colon T \to T$ is a linear isomorphism with inverse given by [12, Lem. 3.1]

$$(R_p|_T)^{-1}(u) = P_T\,\mathrm{Diag}\left(\tfrac{1}{p}\right)u, \quad \text{for all } u \in T = T_p\mathcal{S}. \tag{4.7}$$

Now, suppose $J \colon \mathcal{S} \to \mathbb{R}$ is a smooth function defined on some open neighborhood U of $\mathcal{S}$, e.g. $U = \mathbb{R}^n_{>0}$. Then, according to [5, Prop. 2.2], the Riemannian gradient is given by $\mathrm{grad}^g J(p) = R_p\nabla J(p), \quad$ for all $p \in \mathbb{R}^n$, where ∇J is the usual gradient of J on $U \subset \mathbb{R}^n$.

The product metric, again denoted by g, defined by

$$g_W(U, V) := \Sigma_{i\in[m]} g_{W_i}(U_i, V_i), \quad \text{for all } W \in \mathcal{W}, U, V \in \mathcal{T} = T_W\mathcal{W} \tag{4.8}$$

turns $\mathcal{W}$ into a Riemannian manifold. The orthogonal projection $\mathcal{P}_\mathcal{T} \colon \mathbb{R}^{m\times n} \to \mathcal{T}$, $X \mapsto \mathcal{P}_\mathcal{T}[X]$, with respect to the Frobenius inner product of matrices and, for each $W \in \mathcal{W}$, the replicator operator $\mathcal{R}_W \colon \mathbb{R}^{m\times n} \to \mathcal{T}$, $X \mapsto \mathcal{R}_W[X]$, are defined row-wise by

$$(\mathcal{P}_\mathcal{T}[X])_i := P_TX_i \quad \text{and} \quad (\mathcal{R}_W[X])_i := R_{W_i}X_i \quad \text{for all } X \in \mathcal{T}, i \in [m]. \tag{4.9}$$

As a consequence, if a smooth function $J \colon \mathcal{W} \to \mathbb{R}$ is defined on some open neighborhood of $\mathcal{W}$, then the Riemannian gradient is given by

$$\mathrm{grad}^g J(W) = \mathcal{R}_W[\nabla J(W)] \in T_W\mathcal{W} = \mathcal{T}, \quad \text{for all } W \in \mathcal{W}, \tag{4.10}$$

where $\nabla J(W) \in \mathbb{R}^{m\times n}$ is the unique matrix fulfilling $dJ|_W[V] = \langle \nabla J(W), V\rangle$ for all $V \in \mathbb{R}^{m\times n}$. Therefore, $(\nabla J(W))_{ij} = \partial J/\partial W_i^j$, for all $i \in [m]$, $j \in [n]$.

Assignment Flows. The replicator equation is a well known differential equation for modeling various processes in fields such as biology, economy and evolutionary game dynamics, see [6] or [11]. In a typical game dynamics scenario, as described in [6], the labels correspond to different strategies of an agent playing a game and $p = (p^1, \dots, p^n)^\top \in \mathcal{S}$ are the probabilities p^j of playing the j-th strategy, $j \in [n]$. The *fitness function* $F\colon \mathcal{S} \to \mathbb{R}^n$, also called *affinity measure*, represents the payoff $F^j(p)$ for each strategy j depending on the state $p \in \mathcal{S}$. The replicator equation is a consequence of the assumption that the growth rate $\dot{p}^j/p^j$ is given by the difference between the payoff $F^j(p)$ for strategy j and the average payoff $\sum_{k\in[n]} p^k F^k(p) = \langle F(p), p\rangle$, resulting in $\dot{p}^j = p^j(F^j(p) - \langle F(p), p\rangle)$. In vector notation, this can be written using the replicator matrix R_p from (4.6) as

$$\dot{p} = p \diamond F(p) - \langle F(p), p\rangle p = R_p F(p), \quad \text{for all } p \in \mathcal{S}.$$

The replicator dynamics therefore describes a selection process: over time, the agent selects successful strategies more often.

From this game dynamics perspective, assignment flows for data labeling can be seen as a game of interacting agents, where each node $i \in \mathcal{V}$ in the graph represents one agent and the strategies are the labels in $\mathcal{F}_*$. The fitness function (payoff) for node $i \in \mathcal{V}$ is a function $F_i\colon \mathcal{W} \to \mathbb{R}^n$ depending on the global label assignments $W \in \mathcal{W}$ and thereby coupling the label decisions between different nodes. Thus, for each $i \in \mathcal{V}$ the process of label selection on the corresponding simplex $\mathcal{S}$ is described by the replicator equation

$$\dot{W}_i = R_{W_i} F_i(W), \quad W_i(t) \in \mathcal{S},$$

coupled through the $F_i(W)$. In order to express this system of coupled replicator equations in a more compact way, we define the matrix valued fitness function $F\colon \mathcal{W} \to \mathbb{R}^{m\times n}$ with the i-th row given by $(F(W))_i := F_i(W)$. Together with the replicator operator $\mathcal{R}_W$ on $\mathcal{W}$ from (4.9), the coupled replicator equations are compactly expressed through (1.3). We again refer the reader to the survey [13] for applications of this framework to data labeling and related work.

4.2 Proof of Theorem 1

Let $I := [t_0, t_1]$ and suppose $F\colon U \to \mathbb{R}^{m\times n}$ is a fitness function defined on an open set $U \subset \mathbb{R}^{m\times n}$ containing $\mathcal{W}$. Since the squared Riemannian norm and the replicator operator are also defined on $\mathcal{W}$, the functional (1.4) from [10] can be easily extended to curves $W\colon I \to \mathcal{W}$ by simply replacing every occurrence of $p(t)$ with $W(t)$, resulting in

$$\mathcal{L}(W) := \int_{t_0}^{t_1} \frac{1}{2}\|\dot{W}(t)\|_g^2 + \frac{1}{2}\|\mathcal{R}_{W(t)}[F(W(t))]\|_g^2 dt. \tag{4.11}$$

The term $\|\mathcal{R}_{W(t)}[F(W(t))]\|_g^2$ can be rewritten in a slightly more interpretable way. For this, we view the inner product between a vector $x \in \mathbb{R}^n$ and a point

$p \in \mathcal{S}$ as the expected value $\langle x, p\rangle = \mathbb{E}_p[x]$ and similarly $\langle x^{\diamond 2}, p\rangle = \mathbb{E}_p[x^2]$. Thus, it is reasonable to talk about the variance of x with respect to p, given by

$$\mathrm{Var}_p(x) = \mathbb{E}_p[x^2] - (\mathbb{E}_p[x])^2 = \langle x^{\diamond 2}, p\rangle - \langle x, p\rangle^2. \tag{4.12}$$

Lemma 2. *Let $p \in \mathcal{S}$ and $x \in \mathbb{R}^n$, then $\|R_p x\|_g^2 = \langle x, R_p x\rangle = \mathrm{Var}_p(x)$. Thus, for $W \in \mathcal{W}$ and $X \in \mathbb{R}^{m\times n}$, we have $\|\mathcal{R}_W[X]\|_g^2 = \sum_{i\in\mathcal{V}} \mathrm{Var}_{W_i}(X_i) = \langle X, \mathcal{R}_W[X]\rangle$.*

Proof. Since P_T is the orthogonal projection and $R_p x \in T$, the squared norm of the Fisher-Rao metric (4.4) is given by $\|R_p x\|_g^2 = \langle R_p x, P_T \operatorname{Diag}(1/p) R_p x\rangle$. As a result of $R_p = R_p|_T P_T$ and the formula for the inverse of $R_p|_T$ from (4.7), we have $P_T \operatorname{Diag}\left(\frac{1}{p}\right) R_p x = P_T \operatorname{Diag}\left(\frac{1}{p}\right) R_p|_T P_T x = P_T x$. Therefore, $\|R_p\|_g^2 = \langle R_p x, P_T x\rangle = \langle R_p x, x\rangle$ follows. As a consequence of $R_p x = p \diamond x - \langle p, x\rangle p$ we also directly get $\langle x, R_p x\rangle = \langle x, p\diamond x - \langle p, x\rangle p\rangle = \langle x^{\diamond 2}, p\rangle - \langle x, p\rangle^2 = \mathrm{Var}_p(x)$. The statement for $\|\mathcal{R}_W[X]\|_g^2$ is a consequence of the product Riemannian metric (4.8) on $\mathcal{W}$ and the definition of $\mathcal{R}_W$ in (4.9) as a product map. □

The result of the previous lemma explains the expression for $\mathcal{L}$ in Theorem 1. With this, we are in the regime of Lagrangian mechanics on Riemannian manifolds from Section 3 with $M = \mathcal{W}$, Riemannian metric $h = g$ and potential

$$G\colon \mathcal{W} \to \mathbb{R}, \quad G(W) := -\frac{1}{2}\|\mathcal{R}_W[F(W)]\|_g^2 = -\frac{1}{2}\sum_{k\in\mathcal{V}} \mathrm{Var}_{W_k}(F_k(W)). \tag{4.13}$$

For $(W, V) \in T\mathcal{W} = \mathcal{W}\times\mathcal{T}$, the corresponding Lagrangian (3.6) takes the form

$$L(W, V) = \frac{1}{2}\|V\|_g^2 - G(W) = \frac{1}{2}\|V\|_g^2 + \frac{1}{2}\sum_{k\in\mathcal{V}} \mathrm{Var}_{W_k}(F_k(W)).$$

Therefore, the Euler-Lagrange equation (1.6) in Theorem 1 is a direct consequence of Proposition 1. The corresponding energy function (3.7) takes the form $E(W(t), \dot W(t)) = \frac{1}{2}\|\dot W(t)\|_g^2 - \frac{1}{2}\|\mathcal{R}_{W(t)}[F(W(t))]\|_g^2$ and is constant along curves $W\colon I \to \mathcal{W}$ fulfilling the Euler-Lagrange equation (1.6). However, due to this specific form of the energy, it follows that $E(W(t), \dot W(t)) = 0$ holds for all assignment flows (1.3), irrespective of whether or not the Euler-Lagrange equation is satisfied. This fact was also reported in [10] for the uncoupled replicator dynamics on a single simplex.

In the remaining part, we derive the characterization (1.7) for which F the assignment flow fulfills the Euler-Lagrange equation (1.6). We start by considering $\mathcal{R}_W[F(W)]$ as a function of $W \in \mathcal{W}$, denoted by

$$\mathcal{R}[F]\colon \mathcal{W} \to \mathcal{T}, \quad W \mapsto \mathcal{R}[F](W) := \mathcal{R}_W[F(W)].$$

In order to calculate the differential of $\mathcal{R}[F]$, we define the $n\times n$-matrix

$$B(p, x) := \operatorname{Diag}(x) - \langle p, x\rangle I_n - p x^\top, \quad \text{for } p \in \mathcal{S}, x \in \mathbb{R}^n \tag{4.14}$$

and the linear map $\mathcal{B}(W, X)\colon \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ with i-th row

$$(\mathcal{B}(W, X)[V])_i := B(W_i, X_i) V_i, \quad \text{for } W \in \mathcal{W}, X \in \mathbb{R}^{m\times n} \tag{4.15}$$

Lemma 3. *With the identifications $T_W\mathcal{W} = \mathcal{T}$ and $T_{\mathcal{R}_W[F(W)]}\mathcal{T} = \mathcal{T}$, the differential of $\mathcal{R}[F]$ is a linear map $d\mathcal{R}[F]|_W : \mathcal{T} \to \mathcal{T}$, given by*

$$d\mathcal{R}[F]\big|_W[V] = \mathcal{R}_W \circ dF|_W[V] + \mathcal{B}(W, F(W))[V], \quad \text{for } V \in \mathcal{T}.$$

Proof. A short calculation shows $\langle B(W_i, F_i(W))V_i, \mathbb{1}_n\rangle = 0$ for all $i \in \mathcal{V}$, proving that $\mathcal{B}(W,X)[V] \in \mathcal{T}$ holds. Let $\eta\colon (-\varepsilon,\varepsilon) \to \mathcal{W}$ be a curve with $\eta(0) = W$ and $\dot\eta(0) = V$. Keeping in mind $R_p = \mathrm{Diag}(p) - pp^\top$, we obtain for all rows $i \in \mathcal{V}$

$$\begin{aligned}\big(d\mathcal{R}[F]|_W[V]\big)_i &= \tfrac{d}{dt} R_{\eta_i(t)} F_i(\eta(t))\big|_{t=0} = \tfrac{d}{dt} R_{\eta_i(t)}\big|_{t=0} F_i(W) + R_{W_i} \tfrac{d}{dt} F_i(\eta(t))\big|_{t=0} \\ &= \big(\mathrm{Diag}(V_i) - V_i W_i^\top - W_i V_i^\top\big) F_i(W) + \big(\mathcal{R}_W\big[\tfrac{d}{dt} F(\eta(t))\big|_{t=0}\big]\big)_i \\ &= \big(\mathcal{B}(W, F(W))[V]\big)_i + \big(\mathcal{R}_W \circ dF|_W[V]\big)_i,\end{aligned}$$

where $\mathrm{Diag}(V_i)F_i(W) = \mathrm{Diag}(F_i(W))V_i$ and $V_i^\top F_i(W) = F_i(W)^\top V_i$ was used for the last equality. □

Next, we consider the acceleration of curves on $\mathcal{S}$ and $\mathcal{W}$ with respect to the Riemannian metric g, that is the covariant derivative D_t^g of their velocities. Due to $T\mathcal{S} = \mathcal{S} \times T$, we can view the velocity of a curve $p\colon I \to \mathcal{S}$ as a map $\dot p\colon I \to T$. As T is a vector space, we can also consider its second derivative $\ddot p\colon I \to T$. Using the expression from [5, Eq. (2.60)] (with α set to 0), the acceleration $D_t^g \dot p$ of p is related to $\ddot p$ by

$$D_t^g \dot p(t) = \ddot p(t) - \frac{1}{2}\frac{(\dot p(t))^{\diamond 2}}{p(t)} + \frac{1}{2}\|\dot p(t)\|_g^2 p(t) = \ddot p(t) - \frac{1}{2} A(p(t), \dot p(t)),$$

with $A\colon \mathcal{S} \times T \to T$ defined as $A(p,v) := \frac{1}{p} v^{\diamond 2} - \|v\|_g^2 p$. Similarly, as a consequence of $T\mathcal{W} = \mathcal{W} \times \mathcal{T}$, the velocity of a curve $W\colon I \to \mathcal{W}$ can be viewed as a map $\dot W\colon I \to \mathcal{T}$, allowing for the second derivative $\ddot W$. Since the covariant derivative on a product manifold equipped with a product metric is the componentwise application of the individual covariant derivatives, the acceleration of $W(t)$ on $\mathcal{W}$ has the form

$$D_t^g \dot W(t) = \ddot W(t) - \frac{1}{2}\mathcal{A}(W(t), \dot W(t)), \tag{4.16}$$

with i-th row of $\mathcal{A}\colon \mathcal{W} \times \mathcal{T} \to \mathcal{T}$ given by $(\mathcal{A}(W,X))_i := A(W_i, X_i)$ from above.

Lemma 4. *Suppose $W\colon I \to \mathcal{S}$ is a solution of the assignment flow (1.3). Then, the acceleration of $W(t)$, that is the covariant derivative of $\dot W(t)$, takes the form $D_t^g \dot W(t) = \mathcal{R}_{W(t)} \circ dF|_{W(t)} \circ \mathcal{R}_{W(t)}[F(W(t))] + \frac{1}{2}\mathcal{A}\big(W(t), \mathcal{R}_{W(t)}[F(W(t))]\big)$.*

Proof. Since $W(t)$ is a solution of $\dot W(t) = \mathcal{R}_{W(t)}[F(W(t))]$, the second derivative $\ddot W = \frac{d}{dt}\dot W(t)$ takes the form (to simplify notation we drop the dependence on t)

$$\ddot W = \frac{d}{dt}\mathcal{R}_W[F(W)] = d\mathcal{R}[F]|_W[\dot W] \overset{\text{Lem. 3}}{=} \mathcal{R}_W \circ dF|_W[\dot W] + \mathcal{B}(W, F(W))[\dot W]$$

The first term on the right-hand side equals $\mathcal{R}_W \circ dF|_W \circ \mathcal{R}_W[F(W)]$ and the second term $\mathcal{B}(W, F(W))[\mathcal{R}_W[F(W)]]$, where $\mathcal{B}$ is defined in terms of the matrix

B from (4.14). Thus, consider $B(p,x)R_p x$, for $p \in \mathcal{S}$ and $x \in \mathbb{R}^n$. The relations $\langle x, R_p x\rangle = \|\mathbb{R}_p x\|_g^2$ from Lemma 2 and $R_p x = p \diamond (x - \langle p, x\rangle \mathbb{1}_n)$ give $B(p,x)R_p x = (x - \langle p,x\rangle \mathbb{1}_n) \diamond R_p x - \langle x, R_p x\rangle p = \frac{1}{p}(R_p x)^{\diamond 2} - \|R_p x\|_g^2 p = A(p, R_p X)$. This implies $\mathcal{B}(W, F(W))[\mathcal{R}_W[F(W)]] = \mathcal{A}(W, \mathcal{R}_W[F(W)])$ and results in the identity $\ddot{W} = \mathcal{R}_W \circ dF|_W \circ \mathcal{R}_W[F(W)] + \mathcal{A}\big(W, \mathcal{R}_W[F(W)]\big)$. Plugging this expression for $\ddot{W}$ into the one for $D_t^g \dot{W}$ in (4.16) finishes the proof. ◻

In the final step, we calculate the Riemannian gradient for the potential G from (4.13). Since F is defined on an open set $U \subset \mathbb{R}^{m\times n}$, with $\mathcal{W} \subset U$, we identify $T_X U = \mathbb{R}^{m\times n}$ and $T_{F(X)}\mathbb{R}^{m\times n} = \mathbb{R}^{m\times n}$ for all $X \in U$. Accordingly, the differential of F at X is a linear map $dF|_X \colon \mathbb{R}^{m\times n} \to \mathbb{R}^{m\times n}$ and its adjoint with respect to the Frobenius inner product on $\mathbb{R}^{m\times n}$ are denoted by $(dF|_X)^*$.

Lemma 5. *The Riemannian gradient of the potential G from (4.13) is given by* $\operatorname{grad}^g G(W) = -\mathcal{R}_W \circ (dF|_W)^* \circ \mathcal{R}_W[F(W)] - \frac{1}{2}\mathcal{A}(W, \mathcal{R}_W[F(W)])$*, for $W \in \mathcal{W}$*

Proof. Let $W \in \mathcal{W}$. Since the i-th row of $\mathcal{R}_W$ is given by symmetric matrices $R_{W_i} = \operatorname{Diag}(W_i) - W_i W_i^\top$, Lemma 1 implies $\mathcal{R}_W^* = \mathcal{R}_W$. Next, we calculate an expression for $\nabla G(W)$. For this, assume $V \in \mathbb{R}^{m\times n}$ is arbitrary and let $\eta \colon (-\varepsilon, \varepsilon) \to \mathcal{W}$ be a curve with $\eta(0) = W$ and $\dot{\eta}(0) = V$. Then

$$\begin{aligned} dG|_W[V] &= \frac{d}{dt} G(\eta(t))|_{t=0} \overset{\text{Lem. 2}}{=} -\frac{1}{2}\frac{d}{dt}\big\langle F(\eta(t)), \mathcal{R}_{\eta(t)}[F(\eta(t))]\big\rangle\Big|_{t=0} \\ &= -\frac{1}{2}\big\langle dF|_W[V], \mathcal{R}_W[F(W)]\big\rangle - \frac{1}{2}\big\langle F(W), d\mathcal{R}[F]|_W[V]\big\rangle. \end{aligned}$$

With the expression for $d\mathcal{R}[F]|_W$ from Lemma 3 together with $\mathcal{R}_W^* = \mathcal{R}_W$, the second inner product takes the form

$$\begin{aligned} \big\langle F(W), d\mathcal{R}[F]|_W[V]\big\rangle &= \big\langle F(W), \mathcal{R}_W \circ dF|_W[V]\big\rangle + \big\langle F(W), \mathcal{B}(W, F(W))[V]\big\rangle \\ &= \big\langle (dF|_W)^* \circ \mathcal{R}_W[F(W)], V\big\rangle + \big\langle \mathcal{B}^*(W, F(W))[F(W)], V\big\rangle. \end{aligned}$$

Substituting this formula back into the above expression for $dG|_W$ together with $\big\langle dF|_W[V], \mathcal{R}_W[F(W)]\big\rangle = \big\langle V, (dF|_W)^* \circ \mathcal{R}_W[F(W)]\big\rangle$ for the first inner product, results in $dG|_W[V] = \big\langle -(dF|_W)^* \circ \mathcal{R}_W[F(W)] - \frac{1}{2}\big\langle \mathcal{B}^*(W, F(W))[F(W)], V\big\rangle$. Since V is arbitrary, $\nabla G(W) = -(dF|_W)^* \circ \mathcal{R}_W[F(W)] - \frac{1}{2}\mathcal{B}^*(W, F(W))[F(W)]$ follows. Due to (4.10), the Riemannian gradient is given by

$$\operatorname{grad}^g G(W) = -\mathcal{R}_W \circ (dF|_W)^* \circ \mathcal{R}_W[F(W)] - \frac{1}{2}\mathcal{R}_W[\mathcal{B}^*(W, F(W))[F(W)]].$$

Because $\mathcal{B}$ is defined in terms of the matrix B from (4.14), the adjoint $\mathcal{B}^*$ is determined by $B^\top$ through Lemma 1. For $p \in \mathcal{S}$ and $x \in \mathbb{R}^n$, we have

$$\begin{aligned} R_p B^\top(p,x)x &= R_p\big(\operatorname{Diag}(x) - \langle p,x\rangle I_n - xp^\top\big)x = R_p\big(x^{\diamond 2} - 2\langle p,x\rangle x\big) \\ &= p \diamond x^{\diamond 2} - \langle x^{\diamond 2}, p\rangle p - 2\langle p,x\rangle x \diamond p + 2\langle p,x\rangle^2 p \\ &= \big(p \diamond x^{\diamond 2} - 2\langle p,x\rangle x \diamond p + \langle p,x\rangle p\big) - \big(\langle x^{\diamond 2}, p\rangle - \langle p,x\rangle\big)p \\ &= \frac{1}{p}\big(p \diamond x - \langle p,x\rangle p\big)^{\diamond 2} - \|R_p x\|_g^2 p = A(p, R_p x), \end{aligned}$$

where the relation $\langle p, x^{\diamond 2}\rangle - \langle p, x\rangle^2 = \mathrm{Var}_p(x) = \|R_p x\|_g^2$ from (4.12) and Lemma 2 was used in the last line. Therefore, $\mathcal{R}_W[\mathcal{B}^*(W, F(W))[F(W)]] = \mathcal{A}(W, \mathcal{R}_W[F(W)])$ holds which proves the statement. □

Proof (Theorem 1). Suppose $W(t)$ is a solution of the assignment flow (1.3). Due to Lemma 4 and 5, the expression for the acceleration of $W(t)$ and the Riemannian gradient of G at $W(t)$ both contain the term $\frac{1}{2}\mathcal{A}(W(t), \mathcal{R}_{W(t)}[F(W(t))])$ which yields the relation

$$\begin{aligned}
&D_t^g \dot{W}(t) - \frac{1}{2}\sum_{k\in\mathcal{V}} \mathrm{grad}^g \mathrm{Var}_{W_k}(F_k(W)) \overset{(4.13)}{=} D_t^g \dot{W}(t) + \mathrm{grad}^g G(W) \\
&= \mathcal{R}_{W(t)} \circ dF|_{W(t)} \circ \mathcal{R}_{W(t)}[F(W(t))] - \mathcal{R}_{W(t)} \circ (dF|_{W(t)})^* \circ \mathcal{R}_{W(t)}[F(W(t))] \\
&= \mathcal{R}_{W(t)} \circ \big(dF|_{W(t)} - (dF|_{W(t)})^*\big) \circ \mathcal{R}_{W(t)} F(W(t)).
\end{aligned}$$

As a consequence, the characterization of F in (1.7) is equivalent to the Euler-Lagrange equation (1.6). □

Remark 1. As can be seen from the expression of $D_t^g W(t)$ in (4.16), the Euler-Lagrange equation is a second-order differential equation. The reason why all second- and first-order terms disappear in the condition (1.7) for F is due to the fact that any solution of the assignment flow satisfies $\dot{W}(t) = \mathcal{R}_{W(t)}[F(W(t)]$, allowing to replace any occurrences of $\ddot{W}$ and $\dot{W}$ by alternative expressions in terms of the replicator operator. This basically is the statement of Lemma 4.

4.3 Counterexample

It can be shown that in the case of $n = 2$ labels any fitness function F fulfills condition (1.7) and therefore also the Euler-Lagrange equation. However, for $n > 2$ labels this is no longer true in general, as the example below demonstrates. Nevertheless, a large class of fitness functions always fulfilling condition (1.7) is given by those defined as the gradient $F = \nabla\beta$ of an objective function β. Since the corresponding derivative $dF|_x = \mathrm{Hess}\,\beta(x)$ is self-adjoint, the condition is trivially fulfilled.

For the counterexample, assume $n > 2$. We first consider the case of $m = |\mathcal{V}| = 1$ nodes, that is an uncoupled replicator equation on a single simplex. Define the matrix $F := e_2 e_1^\top$, where e_i are the standard basis vectors of $\mathbb{R}^n$. Thus, the fitness is a linear map $p = (p^1, \dots, p^n)^\top \mapsto Fp = p^1 e_2$, fulfilling $dF|_p = F$ and $(dF|_p)^* = F^\top$. After a short calculation, using the relation $R_p e_i = p^i(e_i - p)$ (Einstein summation convention is *not* used), the first coordinate of condition (1.7) takes the form

$$\big(R_p(F - F^\top)R_p F p\big)^1 = -(p^1)^2 p^2 (1 - p^1 - p^2) \neq 0, \quad \text{for all } p \in \mathcal{S}.$$

In the more general case $m > 1$, define the i-th row of the linear fitness $\mathcal{F}[W]$ by $(\mathcal{F}[W])_i := FW_i$. Since $(\mathcal{F}^*[W])_i = F^\top W_i$ by Lemma 1, the counterexample also extends to general coupled replicator equations on $\mathcal{W}$.

5 Conclusion

Starting from the viewpoint of Lagrangian mechanics on manifolds, we showed that assignment flows solve the Euler-Lagrange equations associated with an action functional. We further characterized those solutions in terms of the fitness function F, which allowed to rectify the result of [10] for uncoupled replicator equations on a single simplex.

Regarding future work, there is a relation to Hamiltonian mechanics via the Legendre transformation, which enables to analyze assignment flows as systems of interacting particles from a physics point of view. There also exists a connection to geodesic motion for a modified Riemannian metric on $\mathcal{W}$, the so called *Jacobi metric*, that provides yet another way of characterizing assignment flows.

Acknowledgments. This work is supported by Deutsche Forschungsgemeinschaft (DFG, German Research Foundation) under Germany's Excellence Strategy EXC-2181/1 - 390900948 (the Heidelberg STRUCTURES Cluster of Excellence). PA was also supported by the Transregional Colloborative Research Center CRC/TRR 191 (281071066).

References

1. Abraham, R., Marsden, J.: Foundations of Mechanics. Addison-Wesley Publishing Company, Inc. , Redwood City, 2 edn. (1987)
2. Absil, P.A., Mahony, R., Sepulchre, R.: Optimization Algorithms on Matrix Manifolds. Princeton University Press (2008)
3. Amari, S.i., Nagaoka, H.: Methods of Information Geometry, vol. 191. American Mathematical Soc. (2007)
4. Aström, F., Petra, S., Schmitzer, B., Schnörr, C.: Image Labeling by Assignment. J. Math. Imag. Vision **58**(2), 211–238 (2017)
5. Ay, N., Jost, J., Vân Lê, H., Schwachhöfer, L.: Information Geometry, vol. 64. Springer (2017)
6. Hofbauer, J., Sigmund, K.: Evolutionary Game Dynamics. Bull. Amer. Math. Soc. **40**(4), 479–519 (2003)
7. Jost, J.: Riemannian Geometry and Geometric Analysis. Springer, 7th edn. (2017)
8. Lee, J.M.: Smooth Manifolds. Springer (2013)
9. Lee, J.M.: Introduction to Riemannian Manifolds, vol. 2. Springer (2018)
10. Raju, V., Krishnaprasad, P.: A Variational Problem on the Probability Simplex. In: 2018 IEEE Conf. on Decision and Control (CDC). pp. 3522–3528. IEEE (2018)
11. Sandholm, W.H.: Population Games and Evolutionary Dynamics. MIT press (2010)
12. Savarino, F., Schnörr, C.: Continuous-Domain Assignment Flows. Europ. J. Appl. Math. pp. 1–28 (2020)
13. Schnörr, C.: Assignment Flows. In: Grohs, P., Holler, M., Weinmann, A. (eds.) Handbook of Variational Methods for Nonlinear Geometric Data, pp. 235—260. Springer (2020)