Assignment Flows and Nonlocal PDEs on Graphs

Dmitrij Sitenko(✉), Bastian Boll, and Christoph Schnörr

Image and Pattern Analysis Group (IPA) and Heidelberg Collaboratory for Image Processing (HCI), Heidelberg University, Heidelberg, Germany
dmitrij.sitenko@iwr.uni-heidelberg.de

Abstract. This paper employs nonlocal operators and the corresponding calculus in order to show that assignment flows for image labeling can be represented by a nonlocal PDEs on the underlying graph. In addition, for the homogeneous Dirichlet condition, a tangent space parametrization and geometric integration can be used to solve the PDE numerically. The PDE reveals a nonlocal balance law that governs the spatially distributed dynamic mass assignment to labels. Numerical experiments illustrate the theoretical results.

Keywords: Image labeling · Assignment flows · Nonlocal PDEs

1 Introduction

Overview, Motivation. As in most research areas of computer vision, state of the art approaches to image segmentation are based on *deep networks*. A recent survey [18] reviews a vast number of different network *architectures* and their empirical performance on various benchmark data sets. Among the challenges discussed in [18, Section 6], the authors write: "... a concrete study of the underlying behavior/dynamics of these models is lacking. A *better understanding* of the *theoretical aspects* of these models can enable the development of better models curated toward various segmentation scenarios."

Among the various approaches towards a better understanding of the *mathematics* of deep networks, the connection between general continuous-times ODEs and deep networks, in terms of so-called *neural ODEs*, has been picked out as a central them [8,17]. A particular class of neural ODEs derived from first principles of information geometry, so-called *assignment flows*, were introduced recently [3,20]. The connection to deep networks becomes evident by applying the simplest *geometric* numerical integration scheme (cf. [22]) to the system of ODEs (23), which results in the *discrete-time* dynamical system (see Sect. 3 for definitions of the mappings involved)

$$W^{(t+1)}(x) = \mathrm{Exp}_{W^{(t)}(x)} \circ R_{W^{(t)}(x)}\big(h_{(t)} S(W^{(t)})(x)\big), \quad x \in \mathcal{V}. \tag{1}$$

Here, $t \in \mathbb{N}$ is the discrete time index (iteration counter), x is any vertex of an underlying graph, and $W^{(t)}(x)$ is the assignment vector that converges for

C. Bauckhage et al. (Eds.): DAGM GCPR 2021, LNCS 13024, pp. 498–512, 2021.
https://doi.org/10.1007/978-3-030-92659-5_32

$t \to \infty$ to some unit vector e_j and thus assigns label j to the data observed at x. The key observation to be made here is that the right-hand side of (1) comprises two major ingredients of any deep network: (i) a *context-sensitive interaction* of the variables over the underlying graph in terms of the mapping $S(\cdot)$, and (ii) a *pointwise nonlinearity* in terms of the exponential map Exp_W corresponding to the e-connection of information geometry [1]. Thus, implementing each iteration of (1) as layer of a networks yields a *deep network*.

The **aim of this paper** is to contribute to the mathematics of deep networks by *representing assignment flows* through a *nonlocal PDE* on the underlying graph. This *establishes a link* between traditional *local* PDE-based image *processing* [21] to modern advanced *nonlocal* schemes of image *analysis*. Regarding the latter nonlocal approaches, we mention the seminal work of Gilboa and Osher [16], a PDE [6] resulting as zero-scale limit of the nonlocal means neighborhood filter [7], nonlocal Laplacians on graphs for image denoising, enhancement and for point cloud processing [14,15], and variational phase-field models [5] motivated by both total-variation based image denoising and the classical variational relaxation of the Mumford-Shah approach to image segmentation [2]. However, regarding image *segmentation*, while some of the afore-mentioned approaches apply to binary fore-/background separation, none of them was *specifically* designed for image segmentation and labeling, with an arbitrary number of labels.

Contribution. Our *theoretical* paper introduces a *nonlocal PDE for image labeling* related to assignment flows on arbitrary graphs. Starting point is the so-called S-parametrization of assignment flows introduced by [19]. The mathematical basis is provided by the nonlocal calculus developed in [12,13], see also [11], whose operators include as *special cases* the mappings employed in the above-mentioned works (graph Laplacians etc.). We show, in particular, that the geometric integration schemes of [22] can be used to *solve* the novel nonlocal PDE, which additionally *generalizes* the *local* PDE derived in [19] corresponding to the zero-scale limit of the assignment flow. In addition, the novel PDE reveals in terms of *nonlocal balance laws* the flow of information between label assignments across the graph.

Organization. Section 2 presents required concepts of nonlocal calculus. Section 3 summarizes the assignment flow approach. Our contribution is presented and discussed in Sect. 4, and illustrated in Sect. 5. We conclude in Sect. 6.

2 Nonlocal Calculus

In this section, following [13], we collect some basic notions and nonlocal operators which will be used throughout this paper. See [11] for a more comprehensive exposition.

Let $(\mathcal{V}, \mathcal{E}, \Omega)$ be a weighted connected graph consisting of $|\mathcal{V}| = n$ nodes with no self-loops, where $\mathcal{E} \subset \mathcal{V} \times \mathcal{V}$ denotes the edge set. In what follows we focus on *undirected* graphs with connectivity given by the neighborhoods $\mathcal{N}(x) = \{y \in$

$\mathcal{V}\colon \Omega(x,y) > 0\}$ of each node $x \in \mathcal{V}$ through nonnegative symmetric weights $\Omega(x,y)$ satisfying

$$0 \leq \Omega(x,y) \leq 1, \qquad \Omega(x,y) = \Omega(y,x). \tag{2}$$

The weighting function $\Omega(x,y)$ serves as the similarity measure between two vertices x and y on the graph. We define the function spaces

$$\mathcal{F}_{\mathcal{V}} := \{f\colon \mathcal{V} \to \mathbb{R}\}, \qquad \mathcal{F}_{\mathcal{V}\times\mathcal{V}} := \{F\colon \mathcal{V}\times\mathcal{V} \to \mathbb{R}\}, \tag{3a}$$
$$\mathcal{F}_{\mathcal{V},E} := \{F\colon \mathcal{V} \to E\}, \qquad \mathcal{F}_{\mathcal{V}\times\mathcal{V},E} := \{F\colon \mathcal{V}\times\mathcal{V} \to E\}, \tag{3b}$$

where E denotes any subset of an Euclidean space. The spaces $\mathcal{F}_{\mathcal{V}}$ and $\mathcal{F}_{\mathcal{V}\times\mathcal{V}}$ respectively are equipped with the inner products

$$\langle f,g\rangle_{\mathcal{V}} = \sum_{x,y\in\mathcal{V}} f(x)g(y), \qquad \langle F,G\rangle_{\mathcal{V}\times\mathcal{V}} = \sum_{x,y\in\mathcal{V}\times\mathcal{V}} F(x,y)G(x,y). \tag{4}$$

Given an *antisymmetric* mapping $\alpha \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ with $\alpha(x,y) = -\alpha(y,x)$ and $\overline{\mathcal{V}}$ defined by (5) and (6), and assuming the vertices $x \in \mathcal{V}$ correspond to points in the Euclidean space $\mathbb{R}^d$, the *nonlocal interaction domain with respect to* α is defined as

$$\mathcal{V}_{\mathcal{I}}^{\alpha} := \{x \in \mathbb{Z}^d \setminus \mathcal{V} : \alpha(x,y) > 0 \text{ for some } y \in \mathcal{V}\}. \tag{5}$$

The set (5) serves as discrete counterpart of the nonlocal boundary on a Euclidean underlying domain with positive measure, as opposed to the traditional local formulation of a boundary that has measure zero. See Figure 1 for a schematic illustration of possible nonlocal boundary configuration. Introducing the abbreviation

$$\overline{\mathcal{V}} = \mathcal{V}\dot{\cup}\mathcal{V}_{\mathcal{I}}^{\alpha} \qquad \text{(disjoint union)}, \tag{6}$$

we state the following identity

$$\sum_{x,y\in\overline{\mathcal{V}}} F(x,y)\alpha(x,y) - F(y,x)\alpha(y,x) = 0, \qquad \forall F \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}. \tag{7}$$

Based on (5), the *nonlocal divergence operator* $\mathcal{D}^{\alpha} : \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}} \to \mathcal{F}_{\overline{\mathcal{V}}}$ and the *nonlocal interaction operator* $\mathcal{N}^{\alpha}\colon \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}} \to \mathcal{F}_{\mathcal{V}_{\mathcal{I}}^{\alpha}}$ are defined by

$$\begin{aligned} \mathcal{D}^{\alpha}(F)(x) &:= \sum_{y\in\overline{\mathcal{V}}} \big(F(x,y)\alpha(x,y) - F(y,x)\alpha(y,x)\big), \qquad x \in \overline{\mathcal{V}}, \\ \mathcal{N}^{\alpha}(G)(x) &:= -\sum_{y\in\overline{\mathcal{V}}} \big(G(x,y)\alpha(x,y) - G(y,x)\alpha(y,x)\big), \quad x \in \mathcal{V}_{\mathcal{I}}^{\alpha}. \end{aligned} \tag{8}$$

Due to the identity (7), these operators (8) satisfy the *nonlocal Gauss theorem*

$$\sum_{x\in\mathcal{V}} \mathcal{D}^{\alpha}(F)(x) = \sum_{y\in\mathcal{V}_{\mathcal{I}}^{\alpha}} \mathcal{N}^{\alpha}(F)(y). \tag{9}$$

The operator $\mathcal{D}^\alpha$ maps two-point functions $F(x,y)$ to $\mathcal{D}^\alpha(F) \in \mathcal{F}_{\overline{\mathcal{V}}}$, whereas $\mathcal{N}^\alpha$ is defined on the domain $\mathcal{V}_\mathcal{I}^\alpha$ given by (5) where nonlocal boundary conditions are imposed. In view of (4), the adjoint mapping $(\mathcal{D}^\alpha)^*$ is determined by the relation

$$\langle f, \mathcal{D}^\alpha(F)\rangle_{\overline{\mathcal{V}}} = \langle (\mathcal{D}^\alpha)^*(f), F\rangle_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}, \quad \forall f \in \mathcal{F}_{\overline{\mathcal{V}}}, \quad F \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}, \tag{10}$$

which yields the operator $(\mathcal{D}^\alpha)^* \colon \mathcal{F}_{\overline{\mathcal{V}}} \to \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ acting on $\mathcal{F}_{\overline{\mathcal{V}}}$ by

$$(\mathcal{D}^\alpha)^*(f)(x,y) := -(f(y) - f(x))\alpha(x,y), \qquad x, y \in \overline{\mathcal{V}}. \tag{11}$$

The *nonlocal gradient operator* is defined as

$$\mathcal{G}^\alpha(x,y) := -(\mathcal{D}^\alpha)^*(x,y). \tag{12}$$

For *vector-valued* mappings, the operators (8) and (11) naturally extend to $\mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}},E}$ and $\mathcal{F}_{\overline{\mathcal{V}},E}$, respectively, by acting *componentwise.*

Using definitions (11), (12), the nonlocal Gauss theorem (9) implies for $u \in \mathcal{F}_{\overline{\mathcal{V}}}$ *Greens nonlocal first identity*

$$\sum_{x\in\mathcal{V}} u(x)\mathcal{D}^\alpha(F)(x) - \sum_{x\in\overline{\mathcal{V}}}\sum_{y\in\overline{\mathcal{V}}} \mathcal{G}^\alpha(u)(x,y)F(x,y) = \sum_{x\in\mathcal{V}_\mathcal{I}^\alpha} u(x)\mathcal{N}^\alpha(F)(x). \tag{13}$$

Given a function $f \in \mathcal{F}_{\overline{\mathcal{V}}}$ and a mapping $\Theta \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ with $\Theta(x,y) = \Theta(y,x)$, we define the linear *nonlocal diffusion operator* as the composition of (8) and $\mathcal{G}^\alpha$,

$$\mathcal{D}^\alpha\big(\Theta\mathcal{G}^\alpha(f)\big)(x) = 2\sum_{y\in\overline{\mathcal{V}}} (\mathcal{G}^\alpha)(f)(x,y)\big(\Theta(x,y)\alpha(x,y)\big). \tag{14}$$

For the particular case with no interactions, i.e. $\alpha(x,y) = 0$ if $x \in \mathcal{V}$ and $y \in \mathcal{V}_\mathcal{I}^\alpha$, expression (14) reduces to

$$\mathcal{L}_\omega f(x) = \sum_{y\in\mathcal{N}(x)} \omega(x,y)\big(f(y) - f(x)\big), \qquad \omega(x,y) = 2\alpha(x,y)^2, \tag{15}$$

which coincides with the *combinatorial Laplacian* [9,10] after reversing the sign.

3 Assignment Flows

We summarize the assignment flow approach introduced by [3] and refer to the recent survey [20] for more background and a review of related work.

Assignment Manifold. Let $(\mathcal{F}, d_\mathcal{F})$ be a metric space and $\mathcal{F}_n = \{f(x) \in \mathcal{F} \colon x \in \mathcal{V}\}$ be given data on a graph $(\mathcal{V}, \mathcal{E}, \Omega)$ with $|\mathcal{V}| = n$ nodes and associated neighborhoods $\mathcal{N}(x)$ as specified in Sect. 2. We encode assignment of nodes $x \in \mathcal{V}$ to a set $\mathcal{F}_* = \{f_j^* \in \mathcal{F}, j \in \mathcal{J}\}$, $|\mathcal{J}| = c$ of predefined prototypes by *assignment vectors*

$$W(x) = (W_1(x), \ldots, W_c(x))^\top \in \mathcal{S}, \tag{16}$$

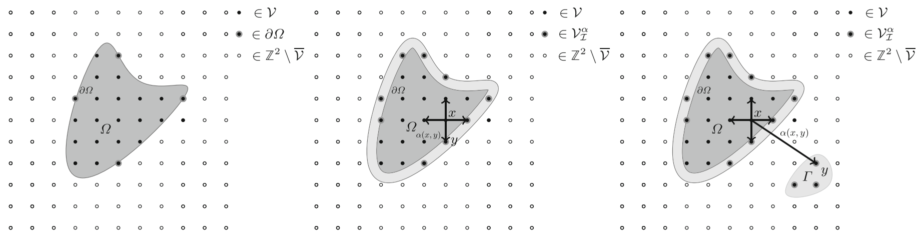

Fig. 1. *Schematic visualization of a nonlocal boundary*: *From left to right:* A bounded open domain $\Omega \subset \mathbb{R}^2$ with continuous *local* boundary $\partial\Omega$ overlaid by the grid $\mathbb{Z}^2$. A bounded open domain Ω with nonlocal boundary depicted by light gray color with nodes • and ● representing the vertices of the graph $\mathcal{V}$ and the interaction domain $\mathcal{V}_{\mathcal{I}}^{\alpha}$, respectively. Possible decomposition of $\mathcal{V}_{\mathcal{I}}^{\alpha}$ satisfying (6). In contrast to the center configuration, this configuration enables nonlocal interactions between nodes ● $\in \Gamma$.

where $\mathcal{S} = \operatorname{rint} \Delta_c$ denotes the relative interior of the probability simplex $\Delta_c \subset \mathbb{R}^c_+$ that we turn into a Riemannian manifold $(\mathcal{S}, g)$ with the Fisher-Rao metric g from information geometry [1,4] at each $p \in \mathcal{S}$

$$g_p(u, v) = \sum_{j \in \mathcal{J}} \frac{u^j v^j}{p^j}, \quad u, v \in T_0 = \{v \in \mathbb{R}^c \colon \langle \mathbb{1}_c, v\rangle = 0\}. \tag{17}$$

The *assignment manifold* $(\mathcal{W}, g)$ is defined as the product space $\mathcal{W} = \mathcal{S} \times \cdots \times \mathcal{S}$ of $n = |\mathcal{V}|$ such manifolds. Points on the assignment manifold have the form

$$W = (\dots, W(x)^\top, \dots)^\top \in \mathcal{W} \subset \mathbb{R}^{n \times c}, \qquad x \in \mathcal{V}. \tag{18}$$

The assignment manifold has the trivial tangent bundle $T\mathcal{W} = \mathcal{W} \times \mathcal{T}_0$ with tangent space

$$\mathcal{T}_0 = T_0 \times \cdots \times T_0. \tag{19}$$

The orthogonal projection onto T_0 is given by $\Pi_0 \colon \mathbb{R}^c \to T_0$, $u \mapsto u - \langle \mathbb{1}_{\mathcal{S}}, u\rangle \mathbb{1}_c$, $\mathbb{1}_{\mathcal{S}} := \frac{1}{c}\mathbb{1}_c$. We denote also by Π_0 the orthogonal projection onto $\mathcal{T}_0$, for notational simplicity.

Assignment Flows. Based on the given data and prototypes, we define the distance vector field on $\mathcal{V}$ by $D_{\mathcal{F}}(x) = \big(d_{\mathcal{F}}(f(x), f_1^*), \dots, d_{\mathcal{F}}(f(x), f_c^*)\big)^\top$, $x \in \mathcal{V}$. These data are lifted to $\mathcal{W}$ to obtain the *likelihood vectors*

$$L(x) \colon \mathcal{S} \to \mathcal{S}, \quad L(W)(x) = \frac{W(x) e^{-\frac{1}{\rho} D_{\mathcal{F}}(x)}}{\langle W(x), e^{-\frac{1}{\rho} D_{\mathcal{F}}(x)}\rangle}, \quad x \in \mathcal{V}, \quad \rho > 0, \tag{20}$$

where the exponential function applies componentwise. This map is based on the affine e-connection of information geometry, and the scaling parameter $\rho > 0$ normalizes the a priori unknown scale of the components of $D_{\mathcal{F}}(x)$. Likelihood vectors are spatially regularized by the *similarity map* and the *similarity vectors*, respectively, given for each $x \in \mathcal{V}$ by

$$S(x)\colon \mathcal{W} \to \mathcal{S}, \quad S(W)(x) = \mathrm{Exp}_{W(x)}\Big(\sum_{y\in\mathcal{N}(x)} \Omega(x,y)\mathrm{Exp}^{-1}_{W(x)}\big(L(W)(y)\big)\Big), \qquad (21)$$

where $\mathrm{Exp}\colon \mathcal{S}\times T_0 \to \mathcal{S}$, $\mathrm{Exp}_p(v) = \frac{pe^{v/p}}{\langle p, e^{v/p}\rangle}$ is the exponential map corresponding to the e-connection. Hereby, the weights $\Omega(x,y)$ determine the regularization of the dynamic label assignment process (23) and satisfy (2) with the additional constraint

$$\sum_{y\in\mathcal{N}(x)} \Omega(x,y) = 1, \quad \forall x \in \mathcal{V}. \qquad (22)$$

The *assignment flow* is induced on the assignment manifold $\mathcal{W}$ by solutions $W(t,x) = W(x)(t)$ of the system of nonlinear ODEs

$$\dot{W}(x) = R_{W(x)}S(W)(x), \qquad W(0,x) = W(x)(0) \in \mathbb{1}_{\mathcal{S}}, \quad x \in \mathcal{V}, \qquad (23)$$

where the map $R_p = \mathrm{Diag}(p) - pp^\top$, $p \in \mathcal{S}$ turns the right-hand side into the tangent vector field $\mathcal{V} \ni x \mapsto \mathrm{Diag}(W(x)) - \langle W(x), S(W)(x)\rangle W(x) \in T_0$.

S-Flow Parametrization. In the following, it will be convenient to adopt from [19, Prop. 3.6] the *S-parametrization* of the assignment flow (23)

$$\dot{S} = R_S(\Omega S), \qquad S(0) = \exp_{\mathbb{1}_{\mathcal{W}}}(-\Omega D), \qquad (24a)$$

$$\dot{W} = R_W(S), \qquad W(0) = \mathbb{1}_{\mathcal{W}}, \quad \mathbb{1}_{\mathcal{W}}(x) = \mathbb{1}_{\mathcal{S}}, \quad x \in \mathcal{V}, \qquad (24b)$$

where both S and W are points on $\mathcal{W}$ and hence have the format (18) and

$$R_S(\Omega S)(x) = R_{S(x)}\big((\Omega S)(x)\big), \quad (\Omega S)(x) = \sum_{y\in\mathcal{N}(x)} \Omega(x,y)S(y), \qquad (25)$$

$$\exp_{\mathbb{1}_{\mathcal{W}}}(-\Omega D)(x) = \big(\dots, \mathrm{Exp}_{\mathbb{1}_{\mathcal{S}}} \circ R_{\mathbb{1}_{\mathcal{S}}}(-(\Omega D)(x)), \dots\big)^\top \in \mathcal{W}, \qquad (26)$$

with the mappings Exp_p, R_p, $p \in \mathcal{S}$ defined after (21) and (23), respectively.

Parametrization (24) has the advantage that $W(t)$ depends on $S(t)$, but not vice versa. As a consequence, it suffices to focus on (24a) since its solution $S(t)$ determines the solution to (24b) by [23, Prop. 2.1.3] $W(t) = \exp_{\mathbb{1}_{\mathcal{W}}}\big(\int_0^t \Pi_0 S(\tau)\mathrm{d}\tau\big)$. In addition, (24a) was shown in [19] to be the Riemannian gradient flow with respect to the potential $J\colon \mathcal{W} \to \mathbb{R}$ given by

$$J(S) = -\frac{1}{2}\langle S, \Omega S\rangle = \frac{1}{4}\sum_{x\in\mathcal{V}}\sum_{y\in\mathcal{N}(x)} \Omega(x,y)||S(x) - S(y)||^2 - \frac{1}{2}||S||^2. \qquad (27)$$

Convergence and stability results for this gradient flow were established by [23].

4 Relating Assignment Flows and Nonlocal PDEs

Using the nonlocal concepts from Sect. 2, we show that the assignment flow introduced in Sect. 3 corresponds to a particular nonlocal diffusion process. This provides an equivalent formulation of the Riemannian flow (24a) in terms of a composition of operators of the form (14).

4.1 *S*-Flow: Non-local PDE Formulation

We begin by first specifying a general class of parameter matrices Ω satisfying (2) and (22), in terms of an anti-symmetric mapping $\alpha \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$.

Lemma 1. *Let* $\alpha, \Theta \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ *be nonnegative anti-symmetric and symmetric mappings, respectively, i.e.* $\alpha(y,x) = -\alpha(x,y)$ *and* $\Theta(x,y) = \Theta(y,x)$, *and assume* α, Θ *satisfy*

$$\lambda(x) = \sum_{y\in\mathcal{N}(x)} \Theta(x,y)\alpha^2(x,y) + \Theta(x,x) \leq 1, \qquad x \in \mathcal{V}, \tag{28a}$$

$$\text{and} \quad \alpha(x,y) = 0, \qquad x,y \in \mathcal{V}^{\alpha}_{\mathcal{I}}. \tag{28b}$$

Then, for arbitrary neighborhoods $\mathcal{N}(x)$, *the parameter matrix* Ω *given by*

$$\Omega(x,y) = \begin{cases} \Theta(x,y)\alpha^2(x,y), & \text{if } x \neq y, \\ \Theta(x,x), & \text{if } x = y. \end{cases} \qquad x,y \in \mathcal{V}, \tag{29}$$

satisfies property (2) *and* (22) *and achieves equality in (28). Additionally, for each function* $f \in \mathcal{F}_{\overline{\mathcal{V}}}$ *with* $f|_{\mathcal{V}^{\alpha}_{\mathcal{I}}} = 0$, *the identity*

$$\sum_{y\in\mathcal{V}} \Omega(x,y)f(y) = \frac{1}{2}\mathcal{D}^{\alpha}(\Theta\mathcal{G}^{\alpha}(f))(x) + \lambda(x)f(x), \qquad x \in \mathcal{V}, \tag{30}$$

holds with $\mathcal{D}^{\alpha}, \mathcal{G}^{\alpha}$ *given by* (8), (12) *and* $\lambda(x)$ *by* (28a).

Proof. For any $x \in \mathcal{V}$, we directly compute using assumption (28) and definition (29),

$$\sum_{y\in\mathcal{V}} \Omega(x,y)f(y) = \sum_{y\in\overline{\mathcal{V}}} \Theta(x,y)\alpha(x,y)^2\big(f(y)\big) + \Theta(x,x)f(x) \tag{31a}$$

$$= -\sum_{y\in\overline{\mathcal{V}}} \Theta(x,y)\alpha(x,y)^2\Big(-\big(f(y)-f(x)\big)\Big) + \lambda(x)f(x) \tag{31b}$$

$$\overset{(11)}{=} -\sum_{y\in\overline{\mathcal{V}}} \Theta(x,y)\big((\mathcal{D}^{\alpha})^*(f)(x,y)\big)\alpha(x,y) + \lambda(x)f(x) \tag{31c}$$

$$= \sum_{y\in\overline{\mathcal{V}}} \frac{1}{2}\Big(\Theta(x,y)\big(-2(\mathcal{D}^{\alpha})^*(f)(x,y)\alpha(x,y)\big)\Big) + \lambda(x)f(x). \tag{31d}$$

Using (12) and (14) yields Eq. (30). □

Next, we generalize the common *local* boundary conditions for PDEs to *volume constraints* for *nonlocal* PDEs on discrete domains. Following [13], given an antisymmetric mapping α as in (5) and Lemma 1, the natural domains $\mathcal{V}^{\alpha}_{\mathcal{I}_N}, \mathcal{V}^{\alpha}_{\mathcal{I}_D}$ for imposing nonlocal *Neumann*- and *Dirichlet* volume constraints are given by a disjoint decomposition $\mathcal{V}^{\alpha}_{\mathcal{I}} = \mathcal{V}^{\alpha}_{\mathcal{I}_N} \dot\cup \mathcal{V}^{\alpha}_{\mathcal{I}_D}$ of the interaction domain (5). The following proposition reveals how the flow (24a) with Ω satisfying the assumptions of Lemma 1 can be reformulated as a nonlocal partial difference equation with *Dirichlet* boundary condition imposed on the entire interaction domain, i.e. $\mathcal{V}^{\alpha}_{\mathcal{I}} = \mathcal{V}^{\alpha}_{\mathcal{I}_D}$. Recall the definition of $\mathcal{S}$ in connection with Eq. (16).

Proposition 1 ***(S-Flow as nonlocal PDE).*** *Let $\alpha, \Theta \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ be as in (28). Then the flow (24a) with Ω given through (29) admits the representation*

$$\partial_t S(x,t) = R_{S(x,t)}\Big(\frac{1}{2}\mathcal{D}^{\alpha}\big(\Theta\mathcal{G}^{\alpha}(S)\big) + \lambda S\Big)(x,t), \qquad on\ \mathcal{V}\times\mathbb{R}_+, \tag{32a}$$

$$S(x,t) = 0, \qquad on\ \mathcal{V}^{\alpha}_{\mathcal{I}}\times\mathbb{R}_+, \tag{32b}$$

$$S(x,0) = \overline{S}(x)(0), \qquad on\ \mathcal{V}\dot\cup\mathcal{V}^{\alpha}_{\mathcal{I}}\times\mathbb{R}_+, \tag{32c}$$

where $\lambda = \lambda(x)$ is given by (28) and $\overline{S}(x)(0)$ denotes the zero extension of the $\mathcal{S}$-valued vector field $S \in \mathcal{F}_{\mathcal{V},\mathcal{S}}$ to $\mathcal{F}_{\overline{\mathcal{V}},\mathcal{S}}$.

Proof. The system (32) follows directly from applying Lemma 1 and

$$\sum_{y\in\mathcal{N}(x)\setminus\mathcal{V}^{\alpha}_{\mathcal{I}}} S(y)\Theta(x,y)\alpha^2(x,y) + \big(\lambda(x) - \sum_{y\in\mathcal{N}(x)}\Theta(x,y)\alpha(x,y)^2\big)S(x) = \Omega S(x). \tag{33}$$

□

Proposition 1 clarifies the connection of the potential flow (24a) with Ω satisfying (22) and the nonlocal diffusion process (32). Specifically, for a nonnegative, symmetric and row-stochastic matrix Ω as in Sect. 3, let the mappings $\widetilde{\alpha}, \widetilde{\Theta} \in \mathcal{F}_{\mathcal{V}\times\mathcal{V}}$ be defined on $\mathcal{V}\times\mathcal{V}$ by

$$\widetilde{\Theta}(x,y) = \begin{cases}\Omega(x,y) & \text{if } y\in\mathcal{N}(x),\\ 0 & \text{else}\end{cases}, \quad \widetilde{\alpha}^2(x,y) = 1, \quad x,y\in\mathcal{V}. \tag{34}$$

Further, denote by $\Theta, \alpha \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ the extensions of $\widetilde{\alpha}, \widetilde{\Theta}$ to $\overline{\mathcal{V}}\times\overline{\mathcal{V}}$ by 0, that is

$$\Theta(x,y) = \big(\delta_{\mathcal{V}\times\mathcal{V}}(\widetilde{\Theta})\big)(x,y), \quad \alpha(x,y) := \big(\delta_{\mathcal{V}\times\mathcal{V}}(\widetilde{\alpha})\big)(x,y) \quad x,y\in\overline{\mathcal{V}}, \tag{35}$$

where $\delta_{\mathcal{V}\times\mathcal{V}}\colon \mathbb{Z}^d\times\mathbb{Z}^d \to \{0,1\}$ is the indicator function of the set $\mathcal{V}\times\mathcal{V}\subset\mathbb{Z}^d\times\mathbb{Z}^d$. Then, for α, Θ as above, the potential flow (24a) is equivalently represented by the system (32). In particular, Proposition 1 shows that the assignment flow introduced in Sect. 3 is a special case of the system (32) with an empty interaction domain (5).

Now we focus on the connection of (32) and the *continuous local* formulation of (24a) on an open, simply connected bounded domain $\mathcal{M} \subset \mathbb{R}^2$, as introduced by [19], that characterizes solutions $S^* = \lim_{t\to\infty} S(t) \in \overline{\mathcal{W}}$ by the PDE

$$R_{S^*(x)}\big(-\Delta S^*(x) - S^*(x)\big) = 0, \quad x \in \mathcal{M}. \tag{36}$$

On the discrete Cartesian mesh $\mathcal{M}^h$ with boundary $\partial\mathcal{M}^h$ specified by a small spatial scale parameter $h > 0$, identify each *interior* grid point $(hk, hl) \in \mathcal{M}^h$ (grid graph) with a node $x = (k, l) \in \mathcal{V}$ of the graph. As in [19], assume (36) is complemented by zero local boundary conditions imposed on S^* on $\partial\mathcal{M}^h$ and adopt the sign convention $L^h = -\Delta^h$ for the basic discretization of the *continuous negative Laplacian* on $\mathcal{M}^h$ by the five-point stencil which leads to strictly positive entries $L^h(x,x) > 0$ on the diagonal. Further, introduce the interaction domain (5) as the local discrete boundary, i.e. $\mathcal{V}^{\alpha}_{\mathcal{I}} = \partial\mathcal{M}^h$ and neighborhoods $\widetilde{\mathcal{N}}(x) = \mathcal{N}(x) \setminus \{x\}$. Finally, let the parameter matrix Ω be defined by (29) in terms of the mappings $\alpha, \Theta \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ given by

$$\alpha^2(x,y) = \begin{cases} 1, & y \in \widetilde{\mathcal{N}}(x), \\ 0, & \text{else}, \end{cases}, \quad \Theta(x,y) = \begin{cases} -L^h(x,y), & y \in \widetilde{\mathcal{N}}(x), \\ 1 - L^h(x,x), & x = y, \\ 0, & \text{else}. \end{cases} \tag{37}$$

Then, for each $x \in \mathcal{V}$, the action of Ω on S reads

$$\sum_{y\in\widetilde{\mathcal{N}}(x)} -L^h(x,y)S(y) + \big(1 - L^h(x,x)\big)S(x) = -\big(-\Delta^h(S) - S\big)(x), \tag{38}$$

which is the discretization of (36) by L^h multiplied by the minus sign. In particular due to relation $R_S(-W) = -R_S(W)$ for $W \in \mathcal{W}$, the solution to the *nonlocal* formulation (32) satisfies Eq. (36). We conclude that the novel approach (32) includes the *local* PDE (36) as special case and hence provides a *natural nonlocal extension.*

4.2 Tangent-Space Parametrization of the S-Flow PDE

Because $S(x,t)$ solving (32) evolves on the non-Euclidean space $\mathcal{S}$, applying some standard discretization to (32) directly will not work. We therefore work out a parametrization of (32) on the *flat* tangent space (19) by means of the equation

$$S(t) = \exp_{S_0}(V(t)) \in \mathcal{W}, \quad V(t) \in \mathcal{T}_0, \qquad S_0 = S(0) \in \mathcal{W}. \tag{39}$$

Applying $\frac{d}{dt}$ to both sides and using the expression of the differential of the mapping $\exp_{S_0}$ due to [19, Lemma 3.1], we get $\dot{S}(t) = R_{\exp_{S_0}(V(t))}\dot{V}(t) \overset{(39)}{=} R_{S(t)}\dot{V}(t)$. Comparing this equation and (24a) shows that $V(t)$ solving the nonlinear ODE

$$\dot{V}(t) = \Omega \exp_{S_0}(V(t)), \qquad V(0) = 0 \tag{40}$$

determines by (39) $S(t)$ solving (24a). Hence, it suffices to focus on (40) which evolves on the flat space $\mathcal{T}_0$. Repeating the derivation above that resulted in the PDE representation (32) of the S-flow (24a), yields the nonlinear PDE representation of (40)

$$\begin{aligned}
&\partial_t V(x,t) = \Big(\frac{1}{2}\mathcal{D}^{\alpha}\big(\Theta\mathcal{G}^{\alpha}(\exp_{S_0}(V))\big) + \exp_{S_0}(V)\Big)(x,t) && \text{on } \mathcal{V}\times\mathbb{R}_+, && \text{(41a)}\\
&V(x,t) = 0 && \text{on } \mathcal{V}^{\alpha}_{\mathcal{I}}\times\mathbb{R}_+, && \text{(41b)}\\
&V(x,0) = \overline{V}(x)(0) && \text{on } \mathcal{V}\dot{\cup}\mathcal{V}^{\alpha}_{\mathcal{I}}\times\mathbb{R}_+. && \text{(41c)}
\end{aligned}$$

From the numerical point of view, this new formulation (41) has the following expedient properties. Firstly, using a parameter matrix as specified by (29) and (35) enables to define the entire system (41) on $\mathcal{V}$. Secondly, since $V(x,t)$ evolves on the flat space T_0, numerical techniques of geometric integration as studied by [22] can here be applied as well. We omit details due to lack of space.

4.3 Non-local Balance Law

A key property of PDE-based models are balance laws implied by the model; see [12, Section 7] for a discussion of various scenarios. The following proposition reveals a *nonlocal* balance law of the assignment flow based on the novel PDE-based parametrization (41), that we express for this purpose in the form

$$\partial_t V(x,t) + \mathcal{D}^{\alpha}(F(V))(x,t) = b(x,t) \quad x\in\mathcal{V}, \tag{42a}$$

$$F(V(t))(x,y) = -\tfrac{1}{2}\Big(\Theta\mathcal{G}^{\alpha}\big(\exp_{S_0}(V(t))\big)\Big)(x,y), \quad b(x,t) = \lambda(x)S(x,t), \tag{42b}$$

where $S(x,t) = \exp_{S_0}(V(x,t))$ by (39).

Proposition 2 (nonlocal balance law of S-flows). *Under the assumptions of Lemma 1, let $V(t)$ solve (41). Then, for each component $S_j(t) = \{S_j(x,t)\colon x\in\mathcal{V}\}$, $j\in[c]$, of $S(t) = \exp_{S_0}(V(t))$, the following identity holds*

$$\begin{aligned}
\frac{1}{2}\frac{d}{dt}\langle S_j,\mathbb{1}\rangle_{\mathcal{V}} + \frac{1}{2}\langle\mathcal{G}^{\alpha}(S_j),\Theta\mathcal{G}^{\alpha}(S_j)\rangle_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}} + \langle S_j,\phi_S-\lambda S_j\rangle_{\mathcal{V}}&\\
+\langle S_j,\mathcal{N}^{\alpha}(\Theta\mathcal{G}^{\alpha}(S_j))\rangle_{\mathcal{V}_{\mathcal{I}^{\alpha}}} = 0&,
\end{aligned}\tag{43}$$

where $\phi_S(\cdot)\in\mathcal{F}_{\mathcal{V}}$ is defined in terms of $S(t)\in\mathcal{W}$ by

$$\phi_S : \mathcal{V}\to\mathbb{R}, \qquad x\mapsto \big\langle S(x), \Pi_0\Omega\exp_{S_0}\big(S(x)\big)\big\rangle. \tag{44}$$

Proof. For brevity, we omit the argument t and simply write $S = S(t), V = V(t)$. In the following, $\odot$ denotes the componentwise multiplication of vectors, e.g. $(S\odot V)_j(x) = S_j(x)V_j(x)$ for $j\in[c]$, and $S^2(x) = (S\odot S)(x)$.

Multiplying both sides of (42a) with $S(x) = \exp_{S_0}(V(x))$ and summing over $x\in\mathcal{V}$ yields

$$\sum_{x\in\mathcal{V}}\big(S\odot\dot{V}\big)_j(x) - \sum_{x\in\mathcal{V}}\frac{1}{2}\Big(S\odot\mathcal{D}^{\alpha}\big(\Theta\mathcal{G}^{\alpha}(S)\big)\Big)_j(x) = \sum_{x\in\mathcal{V}}\big(\lambda S^2\big)_j(x). \tag{45}$$

Applying Greens nonlocal first identity (13) with $u(x) = S_j(x)$ to the second term on the left-hand side yields with (6)

$$\sum_{x\in\mathcal{V}} \big(S \odot \dot{V}\big)_j(x) + \frac{1}{2} \sum_{x\in\mathcal{V}\cup\mathcal{V}_I^\alpha} \sum_{y\in\mathcal{V}\cup\mathcal{V}_I^\alpha} \big(\mathcal{G}^\alpha(S) \odot (\Theta\mathcal{G}^\alpha(S))\big)_j(x,y) \tag{46a}$$

$$+ \sum_{y\in\mathcal{V}_I^\alpha} S_j(y)\mathcal{N}^\alpha\big(\Theta\mathcal{G}^\alpha(S_j)\big)(y) = \sum_{x\in\mathcal{V}} \big(\lambda S^2\big)_j(x). \tag{46b}$$

Now, using the parametrization (39) of S, with the right-hand side defined analogous to (26) and componentwise application of the exponential function to the row vectors of V, we compute at each $x \in \mathcal{V}$: $\dot{S} = S \odot \dot{V} - \langle S, \dot{V}\rangle S$. Substitution into (46) gives for each $S_j = \{S_j(x) \colon x \in \mathcal{V}\}$, $j \in [c]$

$$\frac{1}{2}\frac{d}{dt}\big(\sum_{x\in\mathcal{V}} S_j(x)\big) + \frac{1}{2}\langle\mathcal{G}^\alpha(S_j), \Theta\mathcal{G}^\alpha(S_j)\rangle_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}} + \Big(\sum_{x\in\mathcal{V}}\langle S(x), \dot{V}(x)\rangle\Big)S_j(x) \tag{47a}$$

$$+ \sum_{y\in\mathcal{V}_{\mathcal{I}^\alpha}} S_j\mathcal{N}^\alpha\big(\Theta\mathcal{G}^\alpha(S_j)\big)(y) = \sum_{x\in\mathcal{V}} \big(\lambda S_j^2\big)(x) \tag{47b}$$

which after rearranging the terms yields (43). □

We briefly inspect the *nonlocal balance law* (43) that comprises *four terms*. Since $\sum_{j\in[c]} S_j(x) = 1$ at each vertex $x \in \mathcal{V}$, the *first term* of (43) measures the *rate* of 'mass' assigned to label j over the entire image. This rate is governed by two interacting processes corresponding to the *three remaining terms*:

(i) *spatial* propagation of assignment mass through the nonlocal diffusion process including nonlocal boundary conditions: *second and fourth term*;
(ii) *exchange* of assignment mass with the remaining labels $\{l \in [c] \colon l \neq j\}$: *third term* comprising the function ϕ_S (44).

We are not aware of any other approach, including Markov random fields and deep networks, that makes explicit the flow of information during inference in such an *explicit* manner.

5 Numerical Experiments

In this section, we report numerical results in order to demonstrate two aspects of the mathematical results presented above:

(1) Using geometric integration for numerically solving nonlocal PDEs with appropriate boundary conditions;
(2) zero vs. nonzero interaction domain and the affect of corresponding nonlocal boundary conditions on image labeling.

Numerically Solving Nonlocal PDEs By Geometric Integration. According to Sect. 4.2, imposing the homogeneous Dirichlet condition via the interaction domain (5) makes the right-hand side of (41a) equivalent to (40). Applying a simple explicit time discretization with stepsize h to (41a) results in the iterative update formula

$$V(x,t+h) \approx V(x,t) + h\Pi_0 \exp_{S_0(x)}(\Omega V(x,t)), \qquad h > 0. \tag{48}$$

By virtue of the parametrization (39), one recovers with any nonnegative symmetric mapping Ω as in Lemma 1 the *explicit geometric Euler* scheme on $\mathcal{W}$

$$S(t+h) \approx \exp_{S_0}\big(V(t) + h\dot{V}(t)\big) = \exp_{S(t)}\big(h\Omega S(t)\big), \tag{49}$$

where the last equality is due to property $\exp_S(V_1+V_2) = \exp_{\exp_S(V_1)}(V_2)$ of the lifting map $\exp_S$ and due to the equation $\dot{V} = \Omega S$ implied by (39) and (40). Higher order geometric integration methods [22] generalizing (49) can be applied in a similar way. This provides new perspective on solving a certain class of nonlocal PDEs numerically, conforming to the underlying geometry.

Influence of the Nonlocal Interaction Domain. We considered two different scenarios and compared corresponding image labelings obtained by solving (32a) using the scheme (49), uniform weight parameters but different boundary conditions: (i) *zero-extension* to the interaction domain according to (35) which makes (32) equivalent to (24a) according to Proposition 1; (ii) *uniform extension* in terms of uniform mappings $\Theta, \alpha \in \mathcal{F}_{\overline{V}\times\overline{V}}$ given for each $x \in \mathcal{V}$ by fixed neighborhood sizes $|\mathcal{N}(x)| = 7 \times 7$ and $\alpha^2(x,y) = \frac{1}{7^2}$ if $y \in \mathcal{N}(x)$ and 0 otherwise; $\Theta(x,y) = \frac{1}{7^2}$ if $x = y$ and 1 otherwise. We iterated (49) with step size $h = 1$ until assignment states (24b) of low average entropy 10^{-3} were reached. To ensure a fair comparison and to assess solely the effects of the boundary conditions through nonlocal regularization, we initialized (32) in the same way as (24a) and adopted an uniform encoding of the 31 labels as described by [3, Figure 6].

Figure 2 depicts the results. Closely inspecting panels (c) (*zero extension*) and (d) (*uniform extension*) shows that using a nonempty interaction domain may improve the labeling near the boundary (cf. close-up views), and almost equal performance in the interior domain.

Figure 3 shows the decreasing average entropy values for each iteration (left panel) and the number of iterations required to converge (right panel), for different neighborhood sizes. We observe, in particular, that integral label assignments corresponding to zero entropy are achieved in both scenarios, at comparable computational costs.

A more general system (32) with *nonuniform* interaction is defined through mappings $\alpha, \Theta \in \mathcal{F}_{\overline{\mathcal{V}}\times\overline{\mathcal{V}}}$ measuring similarity of pixel patches analogous to [7]

$$\Theta(x,y) = e^{-G_{\sigma_p} * ||S(x+\cdot) - S(y+\cdot)||^2}, \quad \alpha^2(x,y) = e^{\frac{-|x-y|^2}{2\sigma_s^2}} \quad \sigma_p, \sigma_s > 0, \tag{50}$$

where G_{σ_p} denotes the Gaussian kernel. Decomposition (34) yields a symmetric parameter matrix Ω which not necessarily satisfies property (22). Iterating (49) with step size $h = 0.1$ and $\sigma_s = 1, \sigma_p = 5$ in (50) Fig. 4 visualizes labeling results for different patch sizes. In particular defining Θ by (50) implies a non-zero interaction domain $\mathcal{V}^{\alpha}_{\mathcal{I}}$ as depicted by the right image in Fig. 4.

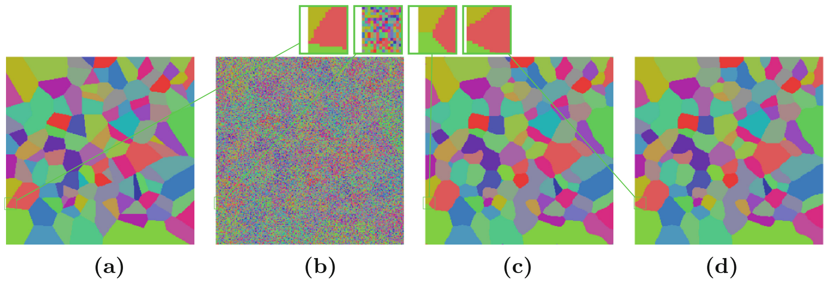

Fig. 2. Labeling through nonlocal geometric flows. **(a)** Ground truth with 31 labels. **(b)** Noisy input data used to evaluate (24a) and (32). **(c)** Labeling returned by (24a) corresponding to a zero extension to the interaction domain. **(d)** Labeling returned by (41) with a uniform extension to the interaction domain in terms of Θ, α specified above. The close-up view show differences close to the boundary, whereas the results in the interior domain are almost equal.

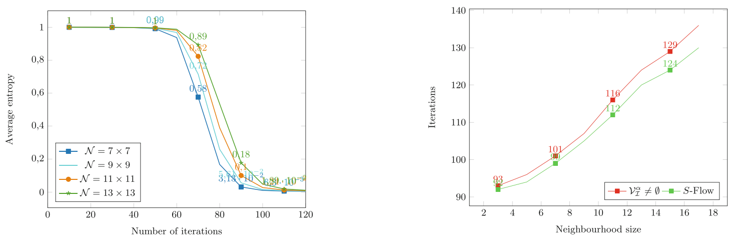

Fig. 3. Left: Convergence rates of the scheme (49) solving (32) with nonempty interaction domain specified by Θ, α above. The convergence behavior is rather insensitive with respect to the neighborhood size. **Right:** Number of iterations until convergence for (32) (●) and (24a) (●). This result shows that different nonlocal boundary conditions have only a minor influence on the convergence of the flow to labelings.

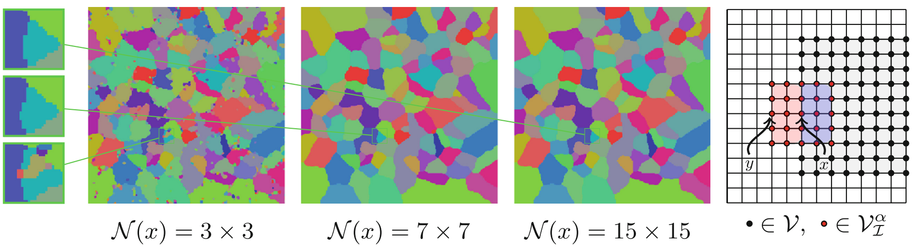

Fig. 4. *From left to right:* Labeling results using (32) for nonuniform interaction domains of size $\mathcal{N}(x) = 3 \times 3$, 7×7 and 15×15, with close up views indicating the regularization properties of the nonlocal PDE (32) with zero Dirichlet conditions. Schematic illustration of the nonlocal interaction domain $y \in \mathcal{V}_{\mathcal{I}}^{\alpha}$ (red area) induced by nodes (blue area) according to (50) with a Gaussian window of size 5×5 centered at $x \in \mathcal{V}$. (Color figure online)

6 Conclusion

We introduced a novel nonlocal PDE motivated by the assignment flow approach. Our results extend established PDE approaches from denoising and image enhancement to image labeling and segmentation, and to a class of *nonlocal* PDEs with nonlocal boundary conditions.

Our future work will study the nonlocal balance law in connection with parameter learning and image structure recognition.

Acknowledgement. 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).

References

1. Amari, S.I., Nagaoka, H.: Methods of Information Geometry. American Mathematical Society and Oxford University Press, Oxford (2000)
2. Ambrosio, L., Tortorelli, V.M.: Approximation of functional depending on jumps by elliptic functional via Γ-convergence. Commun. Pure Appl. Math. **43**(8), 999–1036 (1990)
3. Åström, F., Petra, S., Schmitzer, B., Schnörr, C.: Image labeling by assignment. J. Math. Imaging Vis. **58**(2), 211–238 (2017)
4. Ay, N., Jost, J., Lê, H.V., Schwachhöfer, L.: Information Geometry. Springer, Heidelberg (2017)
5. Bertozzi, A., Flenner, A.: Diffuse interface models on graphs for classification of high dimensional data. SIAM Rev. **58**(2), 293–328 (2016)
6. Buades, A., Coll, B., Morel, J.M.: Neighborhood filter and PDEs. Numer. Math. **105**, 1–34 (2006)
7. Buades, A., Coll, B., Morel, J.M.: Image denoising methods. A new nonlocal principle. SIAM Rev. **52**(1), 113–147 (2010)
8. Chen, R.T.Q., Rubanova, Y., Bettencourt, J., Duvenaud, D.: Neural ordinary differential equations. In: Proceedings of the NeurIPS (2018)
9. Chung, F.: Spectral Graph Theory. American Mathematical Society, Ann Arbor (1997)
10. Chung, F., Langlands, R.P.: A combinatorial Laplacian with vertex weights. J. Comb. Theory Seri. A **5**(2), 316–327 (1996)
11. Du, Q.: Nonlocal Modeling, Analysis, and Computation. SIAM, Philadelphia (2019)
12. Du, Q., Gunzburger, M., Lehoucq, R.B., Zhou, K.: A nonlocal vector calculus, nonlocal volume-constrained problems, and nonlocal balance laws. Math. Models Meth. Appl. Sci. **23**(3), 493–540 (2013)
13. Du, Q., Gunzburger, M., Lehoucq, R.B., Zhou, K.: Analysis and approximation of nonlocal diffusion problems with volume constraints. SIAM Rev. **54**(4), 667–696 (2012)
14. Elmoataz, A., Lezoray, O., Bougleux, S.: Nonlocal discrete regularization on weighted graphs: a framework for image and manifold processing. IEEE Trans. Image Proc. **17**(7), 1047–1059 (2008)

15. Elmoataz, A., Toutain, M., Tenbrinck, D.: On the p-Laplacian and ∞-Laplacian on graphs with applications in image and data processing. SIAM J. Imag. Sci. **8**(4), 2412–2451 (2015)
16. Gilboa, G., Osher, S.: Nonlocal operators with applications to image processing. Multiscale Model. Simul. **7**(3), 1005–1028 (2009)
17. Haber, E., Ruthotto, L.: Stable architectures for deep neural networks. Inverse Prob. **34**(1), 014004 (2017)
18. Minaee, S., Boykov, Y.Y., Porikli, F., Plaza, A.J., Kehtarnavaz, N., Terzopoulos, D.: Image segmentation using deep learning: a survey. IEEE Trans. Pattern Anal. Mach. Intell. (2021). https://ieeexplore.ieee.org/document/9356353
19. Savarino, F., Schnörr, C.: Continuous-domain assignment flows. Eur. J. Appl. Math. **32**(3), 570–597 (2021)
20. Schnörr, C.: Assignment flows. In: Grohs, P., Holler, M., Weinmann, A. (eds.) Handbook of Variational Methods for Nonlinear Geometric Data, pp. 235–260. Springer, Cham (2020). https://doi.org/10.1007/978-3-030-31351-7_8
21. Weickert, J.: Anisotropic Diffusion in Image Processing. B.G. Teubner, Stuttgart (1998)
22. Zeilmann, A., Savarino, F., Petra, S., Schnörr, C.: Geometric numerical integration of the assignment flow. Inverse Probl. **36**(3), 034004 (2020). (33pp)
23. Zern, A., Zeilmann, A., Schnörr, C.: Assignment Flows for Data Labeling on Graphs: Convergence and Stability. CoRR abs/2002.11571 (2020)