Inaugural-Dissertation

submitted to the

Combined Faculty for the
Natural Sciences and Mathematics

of

Heidelberg University, Germany

for the degree of
Doctor of Natural Sciences

put forward by
M.Sc. Matthias Zisler
born in Nürnberg

Date of oral examination: 06.08.2020

Non-Convex and Geometric Methods for Tomography and Label Learning

Advisor: Prof. Dr. Christoph Schnörr

Zusammenfassung

Datenlabeling ist ein grundlegendes Problem der mathematischen Datenanalyse, bei dem jedem Datenpunkt genau ein einziges Label (Prototyp) aus einer endlichen vordefinierten Menge zugewiesen wird. In dieser Arbeit werden zwei herausfordernde Erweiterungen untersucht, bei denen entweder die Eingabedaten nicht direkt beobachtet werden können oder die Prototypen als Vorwissen nicht verfügbar sind.

Die Hauptanwendung des ersten Szenarios stellt die diskrete Tomographie dar. Es werden mehrere nicht-konvexe variationelle sowie glatte geometrische Ansätze entwickelt, die aus den indirekten Messungen eine Rekonstruktion ermitteln und gleichzeitig der Lösung die bekannten Prototypen zuweisen. Insbesondere wird, basierend auf der KL-Divergenz, eine räumliche Regularisierung von Labelings realisiert, welche die glatte Geometrie der diskreten Wahrscheinlichkeitsverteilungen berücksichtigt, die durch die Fisher-Rao (Informations-) Metrik gegeben ist, der Assignment Mannigfaltigkeit. Schließlich führt die geometrische Sichtweise zu einem glatten Fluss, der sich auf einer Riemannschen Untermannigfaltigkeit entwickelt und die Nebenbedingungen der tomographischen Projektion direkt mit in die Geometrie der Assignments einbezieht. Darüber hinaus werden entsprechende implizite numerische Schemata untersucht, die darauf hinauslaufen eine Folge von konvexen Problemen zu lösen.

Ebenso wird für das zweite Szenario, wenn die Prototypen nicht gegeben sind, ein glattes dynamisches System für unüberwachtes Datenlabeling eingeführt, welches durch geometrische Integration auf der Assignment Mannigfaltigkeit evolviert. Die rigorose Abstraktion von “Daten-Label” zu “Daten-Daten” Entscheidungen führt zu interpretierbaren Datenrepräsentationen mit niedrigen Rang, welche selbst wiederum durch die Assignments parametrisiert sind. Der daraus resultierende Self-Assignment-Fluss lernt latente Prototypen gleichzeitig, während diese als Labels für die Inferenz benutzt werden, im selben Framework. Darüber hinaus bestimmt ein einziger Parameter, die Skala der Regularisierung bezüglich des räumlichen Kontextes, den gesamten Prozess. Durch glatte geodätische Interpolation zwischen verschiedenen Normierungen von Self-Assignment Matrizen auf der positiv definiten Matrix-Mannigfaltigkeit wird eine Einparameter-Familie von Self-Assignment-Flüssen definiert. Dementsprechend kann der vorgeschlagene Ansatz unter verschiedenen Gesichtspunkten wie diskreter optimaler Transport, normalisierte spektrale Schnitte und kombinatorische Optimierung durch vollständig positive Faktorisierungen charakterisiert werden, jeder mit zusätzlich eingebauter räumlicher Regularisierung.

Abstract

Data labeling is a fundamental problem of mathematical data analysis in which each data point is assigned exactly one single label (prototype) from a finite predefined set. In this thesis we study two challenging extensions, where either the input data cannot be observed directly or prototypes are not available beforehand.

The main application of the first setting is discrete tomography. We propose several non-convex variational as well as smooth geometric approaches to joint image label assignment and reconstruction from indirect measurements with known prototypes. In particular, we consider spatial regularization of assignments, based on the KL-divergence, which takes into account the smooth geometry of discrete probability distributions endowed with the Fisher-Rao (information) metric, i.e. the assignment manifold. Finally, the geometric point of view leads to a smooth flow evolving on a Riemannian submanifold including the tomographic projection constraints directly into the geometry of assignments. Furthermore we investigate corresponding implicit numerical schemes which amount to solving a sequence of convex problems.

Likewise, for the second setting, when the prototypes are absent, we introduce and study a smooth dynamical system for unsupervised data labeling which evolves by geometric integration on the assignment manifold. Rigorously abstracting from "data-label" to "data-data" decisions leads to interpretable low-rank data representations, which themselves are parameterized by label assignments. The resulting self-assignment flow simultaneously performs learning of latent prototypes in the very same framework while they are used for inference. Moreover, a single parameter, the scale of regularization in terms of spatial context, drives the entire process. By smooth geodesic interpolation between different normalizations of self-assignment matrices on the positive definite matrix manifold, a one-parameter family of self-assignment flows is defined. Accordingly, the proposed approach can be characterized from different viewpoints such as discrete optimal transport, normalized spectral cuts and combinatorial optimization by completely positive factorizations, each with additional built-in spatial regularization.

Acknowledgments

I would like to thank many people who have contributed to this thesis in various ways. First of all, I would like to thank my advisor, Prof. Christoph Schnörr, who introduced me to this fascinating topic. I am deeply grateful for his guidance, continuous support and openness, as well as for sharing his comprehensive knowledge and foresight, with which he sees the important long range connections.

Furthermore, I would like to express special thanks to Claudius and Stefania, who also contributed a lot to the success of this work. I am very thankful to my present and former colleagues at the Image and Pattern Analysis group. Special thanks to Artjom for all the fruitful discussions and help, as well as the essential coffee and dinner breaks. Many thanks to Bastian for proofreading this thesis and joining our essential breaks. I would like to mention: Alexander, Andreas, Andreea, Dmitrij, Ecaterina, Fabian, Fabrizio, Florian, Francesco, Frank, Freddy, Johannes, Jan, Jonathan, Jörg, Lukas, Mattia, Robert, Ruben, Tabea and Tobias, many of which contributed to this thesis with helpful comments and made the work a pleasant experience. I warmly thank Barbara and Evelyn for all the assistance in administrative matters and for the nice and cheerful chats in between.

I would like to thank my parents, my mother Monika and my father Harald and my little brother Simon for their constant support and encouragement, you were always there for me. Finally, I want to express my deepest gratitute to Vera for all the love, help and support I have received from her over the last 5 years. In addition to proofreading, she has always encouraged me and believed in me.

Contents

List of Symbols

Basic Conventions

$\mathbb{N}$,$\mathbb{N}_0$	natural numbers, natural numbers including 0
$\mathbb{R}$, $\mathbb{R}_+$, $\mathbb{R}_{++}$	real numbers, nonnegative real numbers, positive real numbers
$[n]$	integers $\{1, 2, \ldots, n\}$
$\lvert S \rvert$	cardinality of a finite set S
$\langle \cdot, \cdot \rangle$	standard inner product in $\mathbb{R}^n$
$\lVert \cdot \rVert$	Euclidean and Frobenius norm for vectors and matrices
I_n	identity matrix of $\mathbb{R}^{n \times n}$
$\mathcal{R}(A)$	range of a matrix, $\mathcal{R}(A) = \operatorname{span}(A) = \{Ax \colon x \in \mathbb{R}^n\}$
$\mathcal{N}(A)$	nullspace of a matrix
$\operatorname{tr}(A)$	trace of a matrix $A \in \mathbb{R}^{n \times n}$
$\operatorname{rank}(A)$	rank of a matrix
$A \otimes B$	Kronecker product
$\operatorname{vec}(A)$	vector $\mathbb{R}^{mn}$ representation of a matrix $A \in \mathbb{R}^{m \times n}$
$A^\dagger$	Moore-Penrose generalized inverse of a matrix $A \in \mathbb{R}^{n \times n}$
∇F	Euclidean gradient of a differentiable function $F \colon \mathbb{R}^n \to \mathbb{R}$
$\nabla^2 F$	Hessian of a twice differentiable function $F \colon \mathbb{R}^n \to \mathbb{R}$
e_i	i-th unit vector of $\mathbb{R}^n$
$\mathbb{1}_n$	constant vector of ones $(1, 1, \ldots, 1)^\top \in \mathbb{R}^n$
Δ_c	c dimensional standard probability simplex
$\mathbb{S}^n$	set of symmetric $n \times n$ matrices
$\mathbb{S}^n_+$	set of symmetric nonnegative $n \times n$ matrices
$\mathbb{R}^{n \times c}_+$	set of nonnegative $n \times c$ matrices
$\mathcal{P}^n$	set of symmetric positive definite $n \times n$ matrices
$\operatorname{St}(c, n)$	set of orthonormal $n \times c$ matrices
$\mathcal{O}(n)$	group of orthogonal $n \times n$ matrices
$\mathbb{R}^{n \times c}_*$	set of full-rank $n \times c$ matrices
$\operatorname{GL}(n)$	general linear group of invertible $n \times n$ matrices

Convex Analysis

$\overline{\mathbb{R}}$	extended real line $\overline{\mathbb{R}} = \mathbb{R} \cup \{\infty\}$
$\operatorname{conv} A$	convex hull of a set A
$\operatorname{int} S$	interior of a set S
$\operatorname{bd} S$	boundary of a set S
$\operatorname{dom} F$	domain of a convex function $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$
Π_C	orthogonal projection onto a closed convex set C
∂F	subdifferential of a convex function $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$
F^*	Fenchel conjugate of a function $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$
$\delta_C(x)$	indicator function w.r.t. a convex set $C \subset \mathbb{R}^n$
$\sigma_C(x)$	support function w.r.t. a convex set $C \subset \mathbb{R}^n$
$D_\phi(x, y)$	Bregman divergence induced by ϕ of Legendre type
$\mathrm{KL}(p, q)$	Kullback-Leibler (KL) divergence for $p, q \in \Delta_c$
$\mathrm{Prox}_F(x)$	proximal mapping of a convex function $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$

Riemannian Geometry

$\mathcal{M}$	smooth manifold
$(\mathcal{M}, g)$	Riemannian manifold
$C^\infty(\mathcal{M})$	set of all smooth real-valued functions $f\colon \mathcal{M} \to \mathbb{R}$
$T_p\mathcal{M}$	tangent space at $p \in \mathcal{M}$
$T\mathcal{M}$	tangent bundle
$T_p^*\mathcal{M}$	cotangent space at $p \in \mathcal{M}$
$T^*\mathcal{M}$	cotangent bundle
$\mathfrak{X}(\mathcal{M})$	set of all smooth vector fields on $\mathcal{M}$
$\mathfrak{X}^*(\mathcal{M})$	set of all smooth covector fields (one-forms) on $\mathcal{M}$
$dF_p(v)$	differential of a smooth function $F\colon \mathcal{M} \to \mathcal{N}$ at $p \in \mathcal{M}$
$g_p(v, u)$	Riemannian metric, alternative notation $\langle v, u \rangle_g$
$\|v\|_g$	induced norm on $T_p\mathcal{M}$ of a Riemannian metric g
$d_g(p, q)$	Riemannian distance between $p, q \in \mathcal{M}$
$\operatorname{grad} f$	Riemannian gradient of $f \in C^\infty(\mathcal{M})$
$\mathrm{Exp}_p(v)$	exponential map at $p \in \mathcal{M}$, $v \in T_p\mathcal{M}$
$\mathrm{Log}_p(q)$	logarithmic map at $p \in \mathcal{M}$, i.e. inverse exponential map

Assignment Framework

$(\mathcal{F}, d_{\mathcal{F}})$	feature space as a metric space
$f_i \in \mathcal{F}_n$	data points and set of given data
$f_j^* \in \mathcal{F}_*$	prototypes and set of prototypes
$\mathcal{G} = (\mathcal{I}, \mathcal{E})$	undirected graph
$\mathcal{I}$	index set of data points
n	number of data points $n = \lvert\mathcal{I}\rvert$
$\mathcal{I}_j$	vertices of j-th cluster
n_j	size of j-th cluster, $n_j = \lvert\mathcal{I}_j\rvert$
$\mathcal{J}$	index set of prototypes
c	number of prototypes $c = \lvert\mathcal{J}\rvert$
$\hat{c}$	number of effective prototypess $\hat{\mathcal{J}} \subset \mathcal{J}$
$\mathcal{N}_i$	neighborhood of a vertex i, $\mathcal{N}_i = \{k \in \mathcal{I} \colon ik \in \mathcal{E}\} \cup \{i\}$
$w_{i,k}$	edge weights for $k \in \mathcal{N}_i$
Ω	weighted adjacency matrix on $\mathcal{G}$ with $\Omega_i = \big(w_{i,k}\big)_{k\in\mathcal{N}_i}^{\top}$
ρ	selectivity parameter, $\rho > 0$
$C(W)$	diagonal matrix of cluster sizes, $C(W) = \operatorname{Diag}(W^{\top}\mathbb{1}_n)$
$A_0(W)$	self-affinity matrix
$A_1(W)$	self-influence matrix
$A_s(W)$	one-parameter family of self-assignment matrices, $s \in [0,1]$
$B(W)$	cluster-confusion matrix
$\gamma_s(W)$	geodesic on $\mathcal{P}^c$ between $C(W)$ and $(W^{\top}W)^{-1}$, $s \in [0,1]$
$K_{\mathcal{F}}$	weighted similarity matrix, $K_{\mathcal{F}} \in \mathbb{S}^n$
$D_{\mathcal{F}}$	field of distances, $D_{\mathcal{F}} \in \mathbb{R}_+^{n\times c}$ with $D_{\mathcal{F};i.j} = d_{\mathcal{F}}(f_i, f_j^*)$
$E_s(W)$	clustering criterion, $s \in [0,1]$
$L(W), L_i(W)$	likelihood map, likelihood vector
$L_s(W)$, $L_{s;i}(W)$	generalized likelihood map, likelihood vector, $s \in [0,1]$
$S(W), S_i(W)$	similarity map, similarity vector
$\mathcal{S}$	interior of the standard simplex, $S = \operatorname{int}\Delta_c$
$\mathcal{W}$	assignment manifold, $\mathcal{W} = \mathcal{S} \times \cdots \times \mathcal{S}$ (n factors)
W	assignment matrix, $W \in \mathcal{W}$
$\mathcal{W}^c$	assignments of rank c, $\operatorname{rank}(W) = c$
$\mathcal{W}_*^c$	integral assignments of rank c, $W \in \{0,1\}^{n\times c}$
$\mathbb{1}_{\mathcal{S}}$, $\mathbb{1}_{\mathcal{W}}$	barycenter of $\mathcal{S}$ and $\mathcal{W}$, $\mathbb{1}_{\mathcal{S}} = \frac{1}{c}\mathbb{1}_c$ and $\mathbb{1}_{\mathcal{S}} = \frac{1}{c}\mathbb{1}_n\mathbb{1}_c^{\top}$
T_0, $\mathcal{T}_0$	tangent space of $\mathcal{S}$ and $\mathcal{W}$
$T\mathcal{S}$, $\mathcal{T}\mathcal{W}$	tangent bundle of $\mathcal{S}$ and $\mathcal{W}$, $\mathcal{T}_0 = T_0 \times \cdots \times T_0$ (n factors)
Π_0	orthogonal projection onto T_0 or $\mathcal{T}_0$
R_p, R_W	replicator map defined on $\mathcal{S}$ and $\mathcal{W}$
$\exp_p$, $\exp_W$	lifting map, $\operatorname{Exp}_p \circ R_p$ and $\operatorname{Exp}_W \circ R_W$

F	Euclidean data collected as $F = (f_1, \ldots, f_n)^\top \in \mathbb{R}^{n \times d}$
F^*	Euclidean prototypes collected as $F^* = (f_1^*, \ldots, f_c^*)^\top \in \mathbb{R}^{c \times d}$
$P(i)$	probability distribution on the vertices, $i \in \mathcal{I}$
$P(j\|i)$	posterior probability of selecting label f_j^* given f_i
$Q(k\|j)$	likelihood probability of observing datum f_k given label f_j^*
$\mathcal{N}_{\mathcal{P},i}$	patch neighborhood centered at pixel i
$n_{\mathcal{P}}$	size of patch, $n_{\mathcal{P}} = \|\mathcal{N}_{\mathcal{P},i}\|$
P_i	patch centered at pixel i
P_j^*	j-th prototypical patch
$\mathcal{P}(\mathcal{F}_n)$	collection of all patches induced by $\mathcal{F}_n$
$\mathcal{D}_4$	dihedral group of order 4
$T_\sigma P_i$	transformed patch by $\sigma \in \mathcal{D}_4$
S_t	total scatter
S_w	within-class scatter
S_b	between-class scatter

Discrete Tomography

A	projection matrix, $A \in \mathbb{R}^{m \times n}$
b	tomographic measurements, $b \in \mathbb{R}^m$
u	reconstructed image, $u \in \mathbb{R}^n$
$f_j^* \in \mathcal{F}_*$	prototypes and set of prototypes, $\mathcal{F}_* \subset [0, 1]$
n	number of pixels
m	number of measurements
c	number of prototypes
$\overline{\mathcal{W}}$	set of $n \times c$ row-stochastic matrices
$\mathcal{W}_{\text{tomo}}$	submanifold of tomographic assignments, $\mathcal{W}_{\text{tomo}} \subset \mathcal{W}$
$\mathrm{P}_{\mathcal{F}_*}(W)$	assignment operator, $u = \mathrm{P}_{\mathcal{F}_*}(W)$
$\|\nabla u\|_0$	Potts regularizer
$\|\nabla u\|_1$	discrete total variation
$R_{\mathcal{W}}(W)$	geometric regularizer of assignments
$D(W)$	data term
$D_{\text{unary}}(W, D_{\mathcal{F}})$	separable data term
$D_{\text{inverse}}(W, A, b)$	non-separable data term
λ	regularization parameter, $\lambda \geq 0$
α	discretization parameter, $\alpha \geq 0$

List of Publications

Parts of this thesis have been published in international journals and have been presented at conferences.

1. M. Zisler, A. Zern, S. Petra, and C. Schnörr, "Self-Assignment Flows for Unsupervised Data Labeling on Graphs," *SIAM Journal on Imaging Sciences (in press)*, vol. 13, no. 3, pp. 1113–1156, 2020

2. M. Zisler, A. Zern, S. Petra, and C. Schnörr, "Unsupervised Labeling by Geometric and Spatially Regularized Self-Assignment," in *Proceedings of the 7th International Conference on Scale Space and Variational Methods in Computer Vision*, vol. 11603 LNCS, Springer, 2019, pp. 432–444

3. A. Zern, M. Zisler, S. Petra, and C. Schnörr, "Unsupervised Assignment Flow: Label Learning on Feature Manifolds by Spatially Regularized Geometric Assignment," *Journal of Mathematical Imaging and Vision*, 2020, ISSN: 1573-7683

4. A. Zern, M. Zisler, S. Petra, and C. Schnörr, "Spatially Regularized Geometric Assignment for Unsupervised Label Learning on Manifolds," *Proceedings in Applied Mathematics and Mechanics*, vol. 19, no. 1, 2019

5. A. Zern, M. Zisler, F. Åström, S. Petra, and C. Schnörr, "Unsupervised Label Learning on Manifolds by Spatially Regularized Geometric Assignment," in *Proceedings of the 40th German Conference on Pattern Recognition*, vol. 11269 LNCS, Springer, 2018, pp. 698–713

6. M. Zisler, F. Savarino, S. Petra, and C. Schnörr, "Gradient Flows on a Riemannian Submanifold for Discrete Tomography," in *Proceedings of the 39th German Conference on Pattern Recognition*, vol. 10496 LNCS, Springer, 2017, pp. 294–305

7. M. Zisler, F. Aström, S. Petra, and C. Schnörr, "Image Reconstruction by Multilabel Propagation," in *Proceedings of the 6th International Conference on Scale Space and Variational Methods in Computer Vision*, vol. 10302 LNCS, Springer, 2017, pp. 247–259

8. M. Zisler, S. Petra, C. Schnörr, and C. Schnörr, "Discrete Tomography by Continuous Multilabeling Subject to Projection Constraints," in *Proceedings of the 38th German Conference on Pattern Recognition*, vol. 9796 LNCS, Springer, 2016, pp. 261–272

9. M. Zisler, J. Kappes, C. Schnörr, S. Petra, and C. Schnörr, "Non-Binary Discrete Tomography by Continuous Non-Convex Optimization," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 335–347, 2016, ISSN: 2573-0436

10. J. Kappes, S. Petra, C. Schnörr, and M. Zisler, "TomoGC: Binary Tomography by Constrained Graph Cuts," in *Proceedings of the 37th German Conference on Pattern Recognition*, vol. 9358 LNCS, Springer, 2015, pp. 262–273

CHAPTER 1

Introduction and Overview

1.1. Motivation

From the beginning of mankind it was necessary to think in categories in order to survive. The ability to assign unknown entities (things, plants, animals, situations) to known ones to quickly classify and analyze them is crucial for making decisions. Nowadays in our modern world we try to teach this ability to machines to advance technologies such as self-driving vehicles, doctorless diagnosis and autonomous robotic surgeries in healthcare as well as robo-advisers for algorithmic stock trading and investment. Input data are required for the decision basis of such systems, these can be of visual (images) or abstract nature. Before the data as a whole can be understood on a semantic level by the system, the individual data points (pixels) must be understood and categorized in terms of labels in a first step.

Data or image labeling is a fundamental problem of low-level data analysis. Let data be given on a graph, the (multi-)labeling problem consists of uniquely assigning to each vertex an optimal label from a finite set of prototypes which is pre-specified in advance and constitutes a strong prior knowledge. The set of edges encodes contextual information (regularization) and therefore decisions depend on each other such that data points related by a neighborhood are likely to get the same label assignment. These dependencies make computing a globally optimal solution a challenging combinatorial problem, which is NP-hard in general.

In practice, relaxations serve as computationally feasible approximations to overcome the combinatorial difficulty. Particularly for labeling problems with image data numerous variational formulations were developed to approximate models defined on continuous or discrete domains. Probably the most influential continuous model is the Mumford-Shah model [MS89], which allows to explicitly describe discontinuities in a spatially continuous framework in which the structure of solutions is mathematically well understood. After spatial discretization, the piecewise constant limit directly corresponds to the image labeling problem. The resulting piecewise constant Mumford-Shah model is highly non-convex as a feature of perimeter regularization, therefore convex relaxations [ZGFN08; LBS09; PCCB09] aim to provide a tight but still tractable convex envelope of the original non-convex

functional. We refer to [NTC13] for a comprehensive survey and evaluation.

On the other hand in physics and in the field of computer vision discrete models are prevalent. In statistical mechanics, the Potts model [Pot52] describes interacting spins and is defined on discrete lattices. This is the discrete equivalent of the piecewise constant Mumford-Shah model. The Potts model strongly influenced the development of (probabilistic) graphical models which may be translated into discrete Markov random fields under certain conditions. Despite the discrete nature of graphical models, inference still remains a difficult combinatorial problem. Hence, the local-polytope relaxation, a particular linear programming relaxation [Wer07] is commonly used as convex relaxation of graphical models inference. In addition, for submodular energies, so called move-making algorithms such as the α-expansion algorithm [BVZ01] yield suboptimal solutions in a greedy way. We refer to [KAH+15] for a comprehensive survey and evaluation.

The success of convex relaxations of the labeling problem is based on the fact that a globally optimal solution can be found in polynomial time. Unfortunately, the global solution only minimizes the relaxed functional and not the original non-convex functional. Hence, an additional rounding step is required to project back the solution of the relaxed problem onto the set of admissible discrete solutions, however the performance is hard to control [LLS12]. Additionally, both continuous and discrete convex relaxations are inherently non-smooth, which is the price to pay for convexity.

Motivated to overcome the above-mentioned shortcomings of convex relaxations, the assignment flow framework [ÅPSS17], a smooth dynamical system for data labeling was recently introduced. See Section 2.3 for a brief review. The flow evolves on an elementary statistical manifold, the state space, which decouples label decisions from the feature space used to model the input data. More precisely, each data point is associated with a probability simplex to encode the label decisions by the simplex vertices. The interior of each simplex is turned into a smooth Riemannian manifold which is endowed with the Fisher-Rao (information) metric [BR82]. In particular the driving vector field of the assignment flow is based on the smooth geometry which gradually enforces unique label decisions. Adopting the e-connection of information geometry [Kas89; AC10] enables integration of the flow with sparse and efficient numerical updates. Finally smoothness together with the compositional modular design of the assignment flow allows flexible extensions to more complicated tasks.

So far we have seen that a variety of approaches have already been developed to approximately solve the combinatorial hard labeling problem. In this thesis, however, we consider two challenging extensions of the labeling problem that considerably increase the difficulty. We assume that either prototypes are not available beforehand or the input data cannot be observed directly. Therefore, during the labeling process the prototypes have to be inferred from the data, which extends the labeling problem from a *supervised scenario* to an *unsupervised scenario*. Likewise, if the input data can only be observed indirectly and therefore have to be reconstructed during the labeling process extends the labeling problem with *direct* input data to the setting with *indirect measurements* as input data.

1.2. Multilabel Methods for Tomography Reconstruction

The first two parts of this thesis are dedicated to the multilabel problem with indirect measurements, where the image to be labeled cannot be observed directly, however indirect

measurements related by a linear operator are available instead. This amounts to solving an inverse problem and labeling the solution simultaneously. As a consequence, the multilabel problem with indirect measurements is essentially more difficult to solve.

In this thesis, we are particularly interested in the application of non-binary *discrete tomography reconstruction* [HK99; NW01], which perfectly fits into this setting. Many situations require to keep the number of measurements as low as possible, which leads to a small number of projections angles and hence to a severely under-sampled (ill-posed) reconstruction problem. The use of discrete tomography ranges from medical imaging [BSLB11], to natural sciences and industrial applications, like non-destructive material testing [HFU08]. The measurements correspond to line integrals that sum up all pixel-values (absorptions) over each ray transmitted through the object from different projection angles. Typically the forward projections are known and given as projection matrix encoding the imaging geometry. To cope with such problems, the under-sampled projection data is compensated by the assumption that the finite levels i.e. a set of prototypes, of the reconstructed function are known beforehand. Hence the reconstruction of piecewise constant images from indirect measurements can be regarded as a generalized supervised labeling problem.

The main difficulty lies in the fact that there is no direct input data available for a localized dataterm, which is based on the distances between each data point and prototype. Rather, in the case of discrete tomography, the linear projection matrix introduces non-local decisions by coupling the values of pixels along rays. This means that if we change a label in just one pixel, we may need to change the labels for all pixels to satisfy the projection constraints. Therefore, a natural approach is to drop the combinatorial constraints completely [SP08; GBB+12]. In the end, however, a rounding step is required to project back the continuous solution with possibly intermediate values to a discrete solution. More sophisticated two-step approaches [MFKI10; BS11; TKS+15] first reconstruct an image without the integral constraints and then apply a standard labeling approach for direct measurements. These two-step approaches suffer from two major drawbacks, though. The reconstruction process is decoupled from the labeling process, which means the reconstruction method cannot exploit the crucial prior knowledge that the solution only takes values in a given finite set. In addition, after the second step the resulting labeling is not necessarily consistent with the linear projection constraints.

Therefore, the first part of this thesis is dedicated to the question of how to effectively exploit the additional prior knowledge directly during the reconstruction process. We propose a non-smooth and non-convex variational formulation which jointly reconstructs and labels a solution of a system of linear equations. The difficulty is to intertwine the reconstruction process as closely as possible with the labeling process. Our approach is to express the projection constraints directly in terms of decision variables of the labeling problem. However, this results in a weak dataterm which introduces additional non-integral solutions. As a remedy we introduce an additional dataterm for pixel-wise independent decisions which depends on the current iterate. Finally, this naive fixed-point iteration motivates a non-convex discretization term to exclude non-integral solutions by using the prior knowledge (prototypes).

Although the first approach encodes label decisions by discrete probability distributions, their underlying geometry is ignored. Hence, the second part of this thesis investigates how to formulate the labeling problem with indirect measurements with geometric concepts from the assignment flow framework [ÅPSS17]. As a first step we focus on the spatial

regularization of assignments. A novel regularization term is based on the KL-divergence, which locally approximates the squared geodesic distance on the probability simplex equipped with the Fisher-Rao metric [Kas89]. Additionally, the resulting optimization algorithm exploits the underlying geometry by using the KL-divergence as a proximity measure for generalized proximal mappings. Consequently, no reprojection step onto the pixel-wise simplex constraints is required in numerical algorithms.

Finally, in a second formulation we extend the geometric setting itself to include the projection constraints directly. We restrict the assignment manifold to a Riemannian submanifold satisfying the affine projection constraints. Thereby, the submanifold is equipped with a Hessian Riemannian metric which naturally extends the Fisher-Rao metric of the assignment manifold. In the end we construct the tomographic assignment flow, a smooth gradient flow on a Riemannian submanifold for joint tomographic reconstruction and label assignment.

1.3. Unsupervised Data Labeling and Prototype Learning

The last part of this thesis is dedicated to unsupervised data labeling. In contrast to the previous extension, the input data are directly available. However prototypes are not available beforehand and have to be inferred from the data. In practice, the availability of prototypes as class representatives is a strong requirement. In many applications either prototypes are not available or it is not clear which prototypes represent the classes properly. Again, two-step approaches are a basic remedy. Clustering the data in a preprocessing step yields prototypes which are used by a supervised labeling approach in a second step. However, when learning prototypes this way, the framework in which they are used is not properly taken into account. In particular, the (spatial) context in terms of regularization is not considered.

Therefore in this work, we pursue the strategy of simultaneously learning (latent) prototypes in the very same framework while they are used for inference, i.e. assignment of labels. Smoothness and compositional design of the assignment flow are key to linking label learning and inference as closely as possible. As a consequence, this motivates to extend the supervised assignment flow to unsupervised scenarios. We introduce the idea of self-assignment by removing the necessity of explicit prototypes and replacing them by a copy of the given data. Hence the given data are assigned onto itself by abstracting from "data-label" to "data-data" decisions, which specifies if two data points belong to the same cluster or not. Consequently, prototypes are no longer involved explicitly but implicitly as latent variables resulting from low-rank data representations, which themselves are parameterized by label assignments. Finally, the proposed self-assignment flow is a single smooth process evolving on the assignment manifold by geometric integration to perform prototype learning and label assignment simultaneously. Thereby, only one component of the supervised assignment flow needs to be generalized to take the self-assignments into account as local decisions, while all subsequent building blocks remain unaltered. Moreover, a single parameter, the neighborhood size of the geometric regularization in terms of the spatial context, drives the entire process. This parameter determines how fine or coarse the resulting partition is, and how many corresponding latent prototypes are used effectively. At the same time, the geometric spatial regularization of assignments introduces no bias to the emergence of latent prototypes. By smooth geodesic interpolation between different

normalizations of self-assignment matrices on the positive definite matrix manifold, a one-parameter family of self-assignment flows is defined. Accordingly, the approach can be characterized from different viewpoints, e.g. as performing spatially regularized, rank-constrained discrete optimal transport, or as computing spatially regularized normalized spectral cuts. Regarding combinatorial optimization, our approach successfully determines completely positive factorizations of self-assignments in large-scale scenarios, subject to spatial regularization. As a challenging application, to demonstrate the flexibility and plug-in-and-play principle of the approach, we consider locally invariant patch dictionary learning and assignment as well as transfer to novel data.

1.4. Contribution and Organization

In this work, we investigate non-convex and geometric approaches for two challenging extensions of the data labeling problem introduced above. Regarding the multilabel problem with indirect measurements and with particular focus on discrete tomography reconstruction, our main contributions are:

- A variational formulation for joint label assignment and reconstruction from only a few projection angles. Contrary to existing work, we exploit the additional prior knowledge, the known intensities (prototypes), already during the reconstruction. We show how the resulting overall non-convex energy can be optimized efficiently by a fixed-point iteration which requires to solve a convex problem at each step.
- A regularization term for spatially coherent label assignments which respects the underlying geometry of discrete probability distributions. Furthermore, we evaluate the generalized proximal mapping of the regularizer by a fixed point iteration rather than solving a large non-linear system of equations. This strategy is numerically efficient even for larger problem instances.
- A smooth geometric approach evolving on a Riemannian submanifold to simultaneously perform reconstruction and label assignment. By restricting the feasible set to a submanifold, equipped with a natural extension of the Fisher-Rao metric, we adopt the assignment manifold to discrete tomography. We show that the resulting smooth gradient flow can be geometrically integrated by an iterated implicit scheme.

Regarding unsupervised data labeling and prototype learning, our main contributions are:

- A smooth dynamical system for unsupervised data labeling as a direct generalization of the supervised assignment flow which evolves by geometric integration on the assignment manifold. We show that prototypes emerge indirectly as latent variables such that they are learned in the very same framework in which they are used for inference.
- A one-parameter family of low-rank matrix factorizations of assignments for data clustering, which is defined by smooth geodesic interpolation between different normalizations of self-assignment matrices on the positive definite matrix manifold. We discuss the rich algebraic properties and relate the extreme points to normalized spectral cuts as well as completely positive optimization and discrete optimal transport.

- A locally invariant patch distance which we use to demonstrate the generality and flexibility of the approach in more advanced scenarios of unsupervised patch dictionary learning and transfer to novel data.

The remainder of this thesis is structured as follows: In **Chapter** 2 we briefly review the mathematical background on convex analysis and differential geometry, including the supervised assignment flow as a basic framework. Furthermore, we collect material on discrete tomography and clustering which is required in subsequent chapters.

In **Chapter** 3 we introduce the multilabel problem with indirect measurements and Potts regularization for simultaneous reconstruction and labeling for discrete tomography. As a first step, we reformulate the non-local projection constraints in terms of probability vectors (label assignments) and we motivate and derive a non-convex discretization term to exclude non-integral solutions. We propose a non-smooth and non-convex variational approximation to the combinatorial multilabel problem with indirect measurements. In order to minimize the proposed non-convex energy we work out an optimization algorithm based on the DC framework. The resulting algorithm recovers the naive fixed-point iteration, which is a sequence of convex problems. Finally, we numerically evaluate and compare the proposed approach with state-of-the-art methods for non-binary tomography reconstruction on standard test-datasets with only a few projection angles.

Chapter 4 is devoted to geometric aspects of the multilabel problem with indirect measurements. At first we introduce a non-convex variational formulation which involves spatial regularization of assignments respecting the underlying geometry of discrete probability distributions. We derive an optimization algorithm based on the DC framework to minimize the variational non-convex energy. Generalized Bregman proximal maps allow to solve the convex subproblems in a geometric way. We suggest a converging fixed-point iteration to evaluate the proximal mapping of the geometric regularizer efficiently.

Finally, the geometric point of view leads to the tomographic assignment flow, defined as a Riemannian gradient flow on a submanifold including the projection constraints. Therefore, we motivate a suitable objective function to simultaneously perform tomographic reconstruction and labeling. We derive an iterated implicit scheme to integrate the tomographic assignment flow which relies on the Bregman proximal point method. We numerically evaluate and compare the proposed approaches with state-of-the-art methods for non-binary tomographic reconstruction on standard test-datasets with only a few projection angles.

Finally **Chapter** 5 deals with unsupervised data labeling and label learning. First we introduce the concept of self-assignment by removing explicit prototypes and replacing them by a copy of the given data to abstract from "pixel-label" to "pixel-pixel" decisions. This allows us to introduce a family of self-assignment matrices as low-rank matrix factorizations which have rich algebraic properties. In addition, we give detailed interpretations of the self-assignment matrices and show how latent prototypes emerge and can be recovered. Based on the self-assignment matrices we propose a family of self-assignment flows by generalizing the likelihood map of the assignment flow and discuss the numerical integration as well as key properties. We point out connections of the approach to related work from three different viewpoints: spectral relaxation, discrete optimal transport and matrix factorization with aspects of combinatorial optimization. Finally, we demonstrate and compare the approach in various numerical experiments, including unsupervised and locally invariant patch learning, assignment and transfer to novel data.

Finally, in **Chapter** 6 we draw a conclusion and give an outlook for possible future work.

1.5. Basic Notation

We set $[n] = \{1, 2, \dots, n\}$ for $n \in \mathbb{N}$ and $\mathbb{1}_n = (1, 1, \dots, 1)^\top \in \mathbb{R}^n$. We denote the cardinality of a finite set S by $|S|$. The following spaces of matrices will be used.

- $\mathbb{S}^n$: symmetric $n \times n$ matrices
- $\mathbb{S}^n_+$: symmetric nonnegative $n \times n$ matrices
- $\mathbb{R}^{n\times c}_+$: nonnegative $n \times c$ matrices
- $\mathcal{P}^n$: symmetric positive definite $n \times n$ matrices (see Section 2.2.5.1)
- $\mathrm{St}(c, n)$: orthonormal $n \times c$ matrices (see Section 2.2.5.2)
- $\mathbb{R}^{n\times c}_*$: full-rank $n \times c$ matrices (see Section 2.2.5.3)

$\|\cdot\|$ denotes the Euclidean norm and the Frobenius norm for vectors and matrices, respectively. All other norms will be indicated by a corresponding subscript. I_n is the identity matrix of $\mathbb{R}^{n\times n}$. For a matrix $A \in \mathbb{R}^{n\times c}$, A_i, $i \in [n]$ denotes the row vectors and A^j, $j \in [c]$ denotes the column vectors, $A^\top \in \mathbb{R}^{c\times n}$ the transpose and $A^\dagger$ the Moore-Penrose generalized inverse of A. $\mathrm{tr}(A) = \sum_{i\in[n]} A_{i,i}$ denotes the trace of a square matrix $A \in \mathbb{R}^{n\times n}$.

$$\Delta_c = \{p \in \mathbb{R}^c \,:\, p_i \geq 0 \,\wedge\, \langle p, \mathbb{1}_c\rangle = 1\} \tag{1.1}$$

denotes the standard probability simplex. The orthogonal projection onto a closed convex set C is denoted by Π_C.

For strictly positive vectors $p > 0$, we efficiently denote componentwise subdivision by $\frac{v}{p}$. Likewise, we set $pv = (p_1 v_1, \dots, p_n v_n)^\top$. The exponential *function* applies componentwise to vectors (and similarly for log) and will always be denoted by $e^v = (e^{v_1}, \dots, e^{v_n})^\top$, in order not to confuse it with the exponential *maps* (2.80).

In this work we use the Einstein summation convention, which means if the same index name appears exactly twice, once as an upper index and once as a lower index, then we sum up all corresponding expressions along this index. For $a \in \mathbb{R}^n$ and $b \in \mathbb{R}^n$ we have

$$a^i b_i := \sum_{i=1}^n a^i b_i, \tag{1.2}$$

Note, that there is an unambiguity of the symbol ∇, which in section 2.2.3 and 2.2.4 exclusively denotes an affine connection. Otherwise, the symbol ∇, for a differentiable function $F\colon \mathbb{R}^n \to \mathbb{R}$ denotes its ordinary Euclidean gradient, i.e. $\nabla F = (\partial_1 F, \dots, \partial_n F)^\top$.

CHAPTER 2

Preliminaries

In this chapter basic material is collected that is required in subsequent chapters of this thesis. Regarding the mathematical background, we briefly review basic concepts of *convex analysis* (Section 2.1) and *Riemannian Geometry* (Section 2.2). Subsequently, we introduce the *assignment manifold* and the geometric *assignment flow* (Section 2.3), which is a smooth dynamical system for supervised data labeling. Furthermore, we state some important methods and tools from *data clustering* (Section 2.4). Finally, we describe the background and setup used for *discrete tomography* (Section 2.5).

2.1. Convex Analysis

In this section we briefly review the key concepts of convex analysis. We refer the reader to the textbook [RW09] for a detailed introduction.

We consider **extended real-valued functionals** $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ mapping into the **extended real line** $\overline{\mathbb{R}} := \mathbb{R} \cup \{\infty\} = (-\infty, +\infty]$. This offers a convenient way to include restrictions to a subset $C \subset \mathbb{R}^n$ into functionals with the help of the **indicator function** $\delta_C : \mathbb{R}^n \to \overline{\mathbb{R}}$ defined by

$$\delta_C(x) = \begin{cases} 0 & \text{if } x \in C, \\ \infty & \text{if } x \in \mathbb{R}^n \setminus C. \end{cases} \tag{2.1}$$

The **effective domain** of F is defined as the set

$$\operatorname{dom} F := \{x \in \mathbb{R}^n : F(x) < \infty\}, \tag{2.2}$$

and a functional $F : \mathbb{R}^n \to \overline{\mathbb{R}}$ is called **proper** if $\operatorname{dom} F \neq \emptyset$, thus is not identical to $+\infty$ and by definition attains nowhere $-\infty$. A functional is **coercive** if $F(x) \to \infty$ as $\|x\| \to \infty$. The **epigraph** of a function F is defined by

$$\operatorname{epi} F := \{(x, \alpha) \in \mathbb{R}^n \times \mathbb{R} : F(x) \leq \alpha\}. \tag{2.3}$$

Moreover, the epigraph of a proper function is nonempty. A functional F is **lower semicontinuous (lsc.)** at x if for every convergent sequence $x^k \to x$

$$F(x) \leq \liminf_{k \to \infty} F(x^k) \tag{2.4}$$

holds and **lower semicontinuous (lsc.)** on $\mathbb{R}^n$ if this holds for every $x \in \mathbb{R}^n$.

Theorem 2.1.1. *For $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ we have, F is lsc. if and only if* epi F *is closed.*

A set $C \subseteq \mathbb{R}^n$ is **convex** if for every pair of points $x, y \in C$, the line segment joining these points is contained in C, i.e.

$$(1-\alpha)x + \alpha y \in C, \qquad \forall\, \alpha \in [0,1]. \tag{2.5}$$

A functional $F\colon C \to \overline{\mathbb{R}}$ is called **convex** relative to $C \subset \mathbb{R}^n$, if C is convex and for every $x, y \in \operatorname{dom} F$

$$f((1-\alpha)x + \alpha y) \leq (1-\alpha)f(x) + \alpha f(y), \qquad \forall\, \alpha \in [0,1], \tag{2.6}$$

holds, and **strictly convex** if the above inequality is strict for all $x \neq y$. We have the following relation between convexity of a function and convexity of a set.

Theorem 2.1.2. *A function $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is convex if and only if* epi F *is a convex set.*

We call a functional $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ **strongly convex** or **μ-convex** if $F(x) - \frac{\mu}{2}\|x\|_2^2$ is convex.

2.1.1. Subdifferential

Moreover if F is a smooth function then convexity is equivalently characterized by

$$F(y) \leq F(x) + \langle \nabla F(x), y - x\rangle, \qquad \forall\, x, y \in \operatorname{dom} F, \tag{2.7}$$

which means that, the first order approximation supports the function at every point y. Motivated by this relation, we generalize the notion of the gradient to non-differentiable functions. The **subdifferential** $\partial F(x)$ at $x \in \operatorname{dom} F$ of $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ is defined by

$$\partial F(x) = \{g \in \mathbb{R}^n \colon F(y) \leq F(x) + \langle g, y - x\rangle, \quad \forall\, y \in \mathbb{R}^n\} \tag{2.8}$$

and $\partial F(x) = \emptyset$ if F is not finite at x. The elements of $\partial F(x)$ are called **subgradients** of F at x. The subdifferential of the indicator function δ_C w.r.t a convex set $C \subset \mathbb{R}^n$ at $x \in \operatorname{dom} F$ is $\partial \delta_C(x) = \{g \in \mathbb{R}^n \colon \langle g, y - x\rangle, \quad \forall\, y \in C\}$, i.e. the normal cone. F is differentiable at x if and only if the subdifferential is a singleton $\{\nabla F(x)\}$.

The notion of the subdifferential allows to generalize the Fermat principle, which is neccesary and sufficient in case of convex functions.

Theorem 2.1.3 (Fermat's Rule)**.** *Let $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex, lsc. and $\hat{x} \in \operatorname{dom} F$. Then $\hat{x}$ is a global minimizer of F if and only if*

$$0 \in \partial F(\hat{x}). \tag{2.9}$$

2.1.2. Duality and Fenchel Conjugate

Let $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, then we define the **Fenchel conjugate** of F as

$$F^*\colon \mathbb{R}^n \to \overline{\mathbb{R}}, \quad F^*(p) = \sup_{x \in \mathbb{R}^n} \langle p, x\rangle - F(x), \tag{2.10}$$

which is always convex and lsc. Moreover, if F is lower bounded by an affine functional then F^* is also proper. Directly from the definition we obtain the **Fenchel-Young inequality**

$$\langle x, p\rangle \leq F(x) + F^*(p), \qquad \forall\, x, p \in \mathbb{R}^n. \tag{2.11}$$

The Fenchel conjugate of the indicator function δ_C w.r.t a convex set $C \subset \mathbb{R}^n$ is the **support function**

$$\sigma_C(p) := \delta_C^*(p) = \sup_{x \in C} \langle x, p\rangle. \tag{2.12}$$

The next result characterizes the relationship between the Fenchel conjugate and the subdifferential.

Lemma 2.1.4 (Subgradient Inversion Rule)**.** *Let $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be proper, convex and lsc. Then for $x, p \in \mathbb{R}^n$ the following statements are equivalent:*

(i) $\langle x, p\rangle = F(x) + F^*(p)$,

(ii) $p \in \partial F(x)$,

(iii) $x \in \partial F^*(p)$.

In addition, Fenchel conjugation enables to associate a dual optimization problem to each (primal) optimization problem. The dual problem can be easier to treat numerically treated and may give insight into the primal problem and its solutions. Assuming a certain composite structure of the functional, the Fenchel duality scheme gives sufficient conditions for interchanging and relating the optimality conditions between the primal and the dual problem.

Theorem 2.1.5 (Fenchel Duality Scheme)**.** *Let $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ and $G\colon \mathbb{R}^m \to \overline{\mathbb{R}}$ be convex, proper and lsc. and let $A \in \mathbb{R}^{m\times n}$, $b \in \mathbb{R}^m$ and $c \in \mathbb{R}^n$. Consider the two problems*

$$\inf_{x\in\mathbb{R}^n} \phi(x), \quad \phi(x) = \langle c, x\rangle + F(x) + G(b - Ax), \qquad \textbf{(primal problem)} \tag{2.13a}$$

$$\sup_{y\in\mathbb{R}^m} \psi(y), \quad \psi(y) = \langle b, y\rangle - G^*(y) - F^*(A^\top y - c), \qquad \textbf{(dual problem)} \tag{2.13b}$$

and suppose that strong duality holds, i.e.

$$b \in \operatorname{int}(A \operatorname{dom} F + \operatorname{dom} G), \quad c \in \operatorname{int}(A^\top \operatorname{dom} G^* - \operatorname{dom} F^*). \tag{2.14}$$

Then the optimal solutions $(\overline{x}, \overline{y})$ are determined by

$$0 \in c + \partial F(\overline{x}) - A^\top \partial G(b - A\overline{x}), \quad 0 \in b - \partial G^*(\overline{y}) - A\partial F^*(A^\top \overline{y} - c) \tag{2.15}$$

and connected through

$$\overline{y} \in \partial G(b - A\overline{x}), \qquad \overline{x} \in \partial F^*(A^\top \overline{y} - c), \tag{2.16a}$$

$$A^\top \overline{y} - c \in \partial F(\overline{x}), \qquad b - A\overline{x} \in \partial G^*(\overline{y}). \tag{2.16b}$$

Furthermore, the duality gap vanishes, i.e. $\phi(\overline{x}) = \psi(\overline{y})$.

2.1.3. Bregman Divergences

We introduce distance-like divergence functions, which are potentially asymmetric. In order to precisely define a divergence we need the notion of a particular subclass of convex functions. A proper convex function $\phi\colon \mathbb{R}^n \to \mathbb{R}$ is of **Legendre type** if the following holds

(i) $\operatorname{int}(\operatorname{dom}\phi) \neq 0$,

(ii) ϕ is differentiable and strictly convex on $\operatorname{int}(\operatorname{dom}\phi)$,

(iii) for any sequence $(x^k) \subset \operatorname{int}(\operatorname{dom}\phi)$ convergenging to some boundary point, we have

$$\|\nabla\phi(x^k)\| \longrightarrow +\infty \quad \text{for} \quad x^k \longrightarrow x \in \operatorname{bd}(\operatorname{dom}\phi). \tag{2.17}$$

We associate to a function ϕ of Legendre type a corresponding **Bregman divergence** $D_\phi \colon \operatorname{dom}\phi \times \operatorname{int}(\operatorname{dom}\phi) \to \mathbb{R}_+$, which is defined by

$$D_\phi(x,y) := \phi(x) - \phi(y) - \langle \nabla\phi(y), x-y\rangle. \tag{2.18}$$

Intuitively, a Bregman divergence measures the difference between $\phi(x)$ and the linearization of ϕ arround y. The properties of a Lengendre function imply $D_\phi(x,y) \geq 0$ and $D_\phi(x,y) = 0 \Leftrightarrow x = y$. $D_\phi(x,y)$ is strictly convex with respect to the first argument and $D_\phi(x,y)$ is jointly continuous in both arguments. Furthermore, the positive definite Hessian w.r.t. the first argument

$$\nabla_x^2 D_\phi(x,y) = \nabla^2\phi(x) \succ 0, \qquad \forall x \in \operatorname{int}(\operatorname{dom}\phi), \tag{2.19}$$

shows that a Bregman divergence induces a Riemannian metric whichs turns an open subset of a Euclidean space into a Riemannian manifold.

In particular, for the standard probability simplex $\Delta_c = \{p \in \mathbb{R}^c : \ p_i \geq 0 \wedge \langle p, \mathbb{1}_c\rangle = 1\}$ together with the Legendre function $\phi(p) = \langle p, \log(p)\rangle$, i.e. the **negative entropy** function, we obtain the **Kullback-Leibler (KL) divergence**

$$\mathrm{KL}(p,q) := D_\phi(p,q) = \left\langle p, \log\left(\tfrac{p}{q}\right)\right\rangle, \qquad p,q \in \Delta_c, \tag{2.20}$$

which is also known as relative entropy or information divergence.

2.1.4. Optimization and Algorithms

From the viewpoint of numerical optimization, convex objective functions may not necessarily be differentiable and hence first-order gradient descent schemes are not directly applicable. Replacing the gradient by an arbitrary subgradient of the convex subdifferential is no sufficient remedy, since it is not guaranteed to yield a descent direction. Therefore, we resort to the optimality condition $0 \in \partial F(x)$ and search for a zero of the set-valued mapping $x \mapsto \partial F(x)$. For a convex, proper and lsc. function $F \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ the **proximal mapping** $\operatorname{Prox}_F \colon \mathbb{R}^n \to \mathbb{R}^n$ is defined as the unique minimizer of

$$\operatorname{Prox}_F(x) := \operatorname*{arg\,min}_{y\in\mathbb{R}^n} \frac{1}{2}\|y-x\|_2^2 + F(y), \tag{2.21}$$

of a strongly convex objective function. In addition, a generalized proximal mapping is given when the squared Euclidean distance is replaced by some Bregman divergence. Hence, the proximal mapping can be interpreted as a generalized Euclidean projection. In case of the indicator function δ_C w.r.t a non-empty, closed and convex set $C \subset \mathbb{R}^n$ we have $\operatorname{Prox}_{\delta_C}(x) = \Pi_C(x)$. Furthermore, the proximal mapping characterizes the explicit set inclusion of Fermat's rule by an implicit fixed-point equation. For arbitrary $\gamma > 0$ we have

$$0 \in \partial F(x) \quad \Leftrightarrow \quad 0 \in \gamma\partial F(x) \quad \Leftrightarrow \quad x \in (Id + \gamma\partial F)(x), \tag{2.22a}$$

$$\Leftrightarrow \quad x \in (Id + \gamma\partial F)^{-1}(x) \quad \Leftrightarrow \quad x = \operatorname{Prox}_{\gamma F}(x), \tag{2.22b}$$

which yields the following Theorem.

Theorem 2.1.6. *Let $F \colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex, proper and lsc. and $\gamma > 0$ arbitrary. Then $\overline{x} \in \operatorname{dom} F$ is a minimizer of F if and only if*

$$\overline{x} = \operatorname{Prox}_{\gamma F}(\overline{x}). \tag{2.23}$$

This immediately motivates to numerically compute a minimizer of F by the fixed-point scheme stated in the next proposition.

Proposition 2.1.7. *Let $F\colon \mathbb{R}^n \to \overline{\mathbb{R}}$ be convex, proper and lsc. and $\gamma > 0$ arbitrary. Then the sequence (x_k) generated by the* ***proximal-point iteration***

$$x^{(k+1)} = \operatorname{Prox}_{\gamma F}(x^{(k)}), \tag{2.24}$$

converges to a fixed-point $\overline{x} = \operatorname{Prox}_{\gamma F}(\overline{x})$, i.e. a minimizer of F.

The proof relies on a weaker notion of a contraction (firmly non-expansive), which holds for the proximal mapping for a convex, proper and lsc. function.

In this work we are interested in minimizing a composite objective, which is a problem of the form

$$\min_{x\in\mathbb{R}^n} F(x) + G(Ax), \tag{2.25}$$

with F and G convex, proper, lsc. and A a linear operator. From the optimality conditions of the Fenchel duality scheme and the subgradient inversion rule, the popular Chambolle-Pock **primal-dual algorithm** [CP11] can be derived. The iteration for the stepsizes $\sigma, \tau > 0$ reads

$$x^{(k+1)} = \operatorname{Prox}_{\tau F}\left(x^{(k)} - \tau A^\top y^{(k)}\right), \qquad \textbf{(primal step)} \tag{2.26a}$$

$$\overline{x}^{(k+1)} = 2x^{(k+1)} - x^{(k)}, \qquad \textbf{(extrapolation step)} \tag{2.26b}$$

$$y^{(k+1)} = \operatorname{Prox}_{\sigma G^*}\left(y^{(k)} + \sigma A\overline{x}^{(k+1)}\right), \qquad \textbf{(dual step)} \tag{2.26c}$$

where the primal step is a proximal gradient descent and the dual step is a proximal gradient ascent. The additional extrapolation step is considerably accelerating the algorithm in practice. The iteration of the primal-dual algorithm itself can be recast as a single proximal point iteration with respect to the block-variable (x, y) and provably converges if $\sigma\tau < \lambda_{\max}(A^\top A)^{-1}$ holds. A suitable termination criterion is the primal-dual gap, which is an upper bound on the distance between the current objective value and the optimal objective value of the primal problem.

2.2. Riemannian Geometry

In this section, we briefly summarize the main definitions and results of the theory of smooth manifolds and Riemannian geometry, that are relevant in the context of this thesis. We assume that the reader is already familiar with these concepts and otherwise refer to the textbooks [Lee13; Lee18].

2.2.1. Tangent Bundle and Vector Fields

We consider n-dimensional **smooth manifolds**, which are topological manifolds endowed with an extra structure that allows to extend differential calculus to manifolds. Furthermore, we denote by $C^\infty(\mathcal{M})$ the set of all smooth real-valued functions $f\colon \mathcal{M} \to \mathbb{R}$ on the smooth manifold $\mathcal{M}$. Let $\mathcal{M}$ be a smooth n-dimensional manifold. Then for every point $p \in \mathcal{M}$, a **tangent vector** at p is a linear functional $v\colon C^\infty(\mathcal{M}) \to \mathbb{R}$ that is a **derivation** at p, i.e. it satisfies the product rule

$$v(fg) = f(p)vg + g(p)vf, \qquad \forall\, f, g \in C^\infty(\mathcal{M}). \tag{2.27}$$

The **tangent space** at p, denoted by $T_p\mathcal{M}$, comprises the set of all tangent vectors at p, which is a vector space. Furthermore, each smooth coordinate chart $\varphi\colon U \to V \subset \mathbb{R}^n$ on

some open subset $U \subseteq \mathcal{M}$ with coordinate functions $\varphi(p) = (x^1(p), \dots, x^n(p))$ induces a basis on $T_p\mathcal{M}$, the local **coordinate vectors** $\frac{\partial}{\partial x^1}\big|_p, \dots, \frac{\partial}{\partial x^n}\big|_p$ given by

$$\left.\frac{\partial}{\partial x^i}\right|_p f = \left.\frac{\partial}{\partial x^i}\right|_{\varphi(p)} (f \circ \varphi^{-1}). \tag{2.28}$$

Therefore, every tangent vector $v \in T_p\mathcal{M}$ can be written uniquely in the form

$$v = v^i \left.\frac{\partial}{\partial x^i}\right|_p \tag{2.29}$$

with $v^i = v(x^i)$. For a finite dimensional vector space V each tangent space T_pV is identified with V itself.

Given a smooth map $F\colon \mathcal{M} \to \mathcal{N}$ and $p \in \mathcal{M}$, we define the **differential** of F at p as the linear map $dF_p\colon T_p\mathcal{M} \to T_{F(p)}\mathcal{N}$ satisfying

$$dF_p(v)f = v(f \circ F), \qquad v \in T_p\mathcal{M}. \tag{2.30}$$

The **cotangent space** $T_p^*\mathcal{M} = (T_p\mathcal{M})^*$ at $p \in \mathcal{M}$ is the dual space of the tangent space, which is the set of all real-valued linear functionals on $T_p\mathcal{M}$ called **covectors** at p. Every smooth real-valued functional $f \in C^\infty(\mathcal{M})$ together with $p \in \mathcal{M}$ induces a covector $df_p \in T_p^*\mathcal{M}$, by definition of the differential of f at p. Finally, the covectors $dx^1\big|_p, \dots, dx^n\big|_p$ form a **dual basis** to the coordinate vectors, given as the differential of the coordinate functions $\varphi(p) = (x^1(p), \dots, x^n(p))$. As a consequence, we can uniquely express any covector $\omega \in T_p^*\mathcal{M}$ as

$$\omega = \omega_i dx^i|_p \in T_p^*\mathcal{M}. \tag{2.31}$$

The disjoint union of all tangent spaces and cotangent spaces yields the **tangent bundle** and **cotangent bundle** respectively

$$T\mathcal{M} = \dot{\bigcup_{p\in\mathcal{M}}} \{p\} \times T_p\mathcal{M}, \qquad T^*\mathcal{M} = \dot{\bigcup_{p\in\mathcal{M}}} \{p\} \times T_p^*\mathcal{M}. \tag{2.32}$$

A smooth **vector field** on $\mathcal{M}$ is a smooth map

$$X\colon \mathcal{M} \to T\mathcal{M}, \quad p \mapsto X_p \in T_p\mathcal{M}, \tag{2.33}$$

which is a smooth **section** of the tangent bundle, i.e. the property $\pi_1 \circ X = Id_{\mathcal{M}}$ holds. We denote by $\mathfrak{X}(\mathcal{M})$ the set of all smooth vector fields on $\mathcal{M}$. Analogously, we denote by $\mathfrak{X}^*(\mathcal{M})$ the set of all smooth covector fields on $\mathcal{M}$ (i.e. one-forms). Once a chart (U, φ) is fixed, the vector field X can be locally written in coordinate vectors

$$X_p = X^i(p) \left.\frac{\partial}{\partial x^i}\right|_p, \qquad p \in U \subseteq \mathcal{M}. \tag{2.34}$$

A mapping $F\colon \mathcal{M} \to \mathcal{N}$ is called a **smooth immersion** if F is a smooth map where the differential dF_p is injective and the rank of dF_p as a linear mapping is constant and equal to the dimension of $\mathcal{N}$. Moreover, we call a smooth injective immersion $F\colon \mathcal{M} \to \mathcal{N}$ a **smooth embedding** if F is also a topological embedding, i.e. a homeomorphism onto its image $F(\mathcal{M}) \subseteq \mathcal{N}$. A subset $\mathcal{S} \subseteq \mathcal{M}$ of a smooth manifold $\mathcal{M}$ is an embedded **submanifold** if S is a manifold in the subspace topology, endowed with a smooth structure such that the inclusion map $i\colon \mathcal{S} \hookrightarrow \mathcal{M}$ is a smooth embedding. Every open subset $\mathcal{S} \subseteq \mathcal{M}$ is an embedded submanifold of the same dimension as $\mathcal{M}$ and the tangent spaces coincide after identification.

In the two subsequent sections, we equip a smooth manifold with two additional (independent) structures, a metric tensor and an affine connection.

2.2.2. Riemannian Metrics

The Riemannian metric (also called metric tensor) turns every tangent space into an inner product space which allows to measure lengths and angles between tangent vectors and define a distance on $\mathcal{M}$.

A **Riemannian metric** on $\mathcal{M}$ is a smooth symmetric covariant 2-tensor field (a smooth section of the 2-covariant tensor bundle)

$$g_p(v,u) = \langle v,u\rangle_g, \qquad v,u \in T_p\mathcal{M}, \tag{2.35}$$

that is positive definite at each point $p \in \mathcal{M}$, i.e. $g_p(v,v) \geq 0$ for all $v \in T_p\mathcal{M}$ and $g_p(v,v) = 0$ iff $v = 0$. A smooth manifold $\mathcal{M}$ that is additionally equipped with a Riemannian metric g is called a **Riemannian manifold** and denoted by the tuple $(\mathcal{M}, g)$. Fixing a local chart (U,φ), the Riemannian metric g can be written in local coordinates $\varphi(p) = (x^1(p),\dots,x^n(p))$ by using the basis of the cotangent space

$$g = g_{ij} dx^i \otimes dx^j, \tag{2.36}$$

with $g_{ij} \in C^\infty(\mathcal{M})$. For $p \in \mathcal{M}$ and $v,u \in T_p\mathcal{M}$ in local coordinates $u = u^k \frac{\partial}{\partial x^k}\big|_p, v = v^l \frac{\partial}{\partial x^l}\big|_p$ we have

$$g_p(u,v) = g_{ij}(p) u^k v^l dx^i|_p\big(\tfrac{\partial}{\partial x^k}\big|_p\big) \otimes dx^j|_p\big(\tfrac{\partial}{\partial x^l}\big|_p\big) = g_{ij}(p) u^k v^l \delta^i_k\big|_p \delta^j_l\big|_p = g_{ij}(p) u^i v^j. \tag{2.37}$$

Furthermore, we have the following geometric concepts on any Riemannian manifold. For a tangent vector $v \in T_p\mathcal{M}$ at point $p \in \mathcal{M}$ the **norm** induced by the Riemannian metric is given by

$$\|v\|_g = \sqrt{\langle v,v\rangle_g}, \tag{2.38}$$

and the **angle** between two tangent vectors $v,u \in T_p\mathcal{M}$ is defined as

$$\cos(\theta) = \frac{\langle u,v\rangle_g}{\|u\|_g \|v\|_g}, \tag{2.39}$$

with $\theta \in [0,\pi)$. Moreover, two tangent vectors $v,u \in T_p\mathcal{M}$ are **orthogonal** if $\langle u,v\rangle_g = 0$.

Let g be a Riemmanian metric on $\mathcal{N}$ and the mapping $F : \mathcal{M} \to \mathcal{N}$ be a smooth immersion. Then, the **pullback metric** F^*g gives a metric on $\mathcal{M}$ defined by

$$(F^*g)(u,v)_p = g\big(dF_p(u), dF_p(v)\big)_{F(p)}, \qquad v,u \in T_p\mathcal{M}. \tag{2.40}$$

A diffeomorphism $F\colon \mathcal{M} \to \mathcal{N}$ between $(\mathcal{M},g)$ and $(\mathcal{N},h)$ is called Riemannian **isometry** if $F^*h = g$ holds.

A connected Riemannian manifold $(\mathcal{M},g)$ can be turned into a metric space by defining a metric on $\mathcal{M}$ which is called **Riemannian distance** and is given by the length of the shortest path in $\mathcal{M}$, that connects $p,q \in \mathcal{M}$

$$d_g(p,q) = \inf_{\gamma\in\Gamma(p,q)} L_g(\gamma), \quad \text{with} \quad L_g(\gamma) = \int_a^b \|\dot\gamma\|_g \, dt, \tag{2.41}$$

where $\Gamma(p,q)$ is the set of all smooth curves $\gamma\colon [a,b] \to \mathcal{M}$ connecting p and q.

Let $\mathcal{S} \subseteq \mathcal{M}$ be an embedded submanifold of $(\mathcal{M},g)$ with the includsion map $i\colon \mathcal{S} \to \mathcal{M}$. Then, the pullback metric i^*g is the **induced metric** on $\mathcal{S}$ and (S, i^*g) is called a **Riemannian submanifold**.

The differential of a smooth real-valued function defines a covector field, however the inner product induced by the Riemannian metric allows to identify covectors with tangent vectors. This motivates to define the **Riemannian gradient** of a function $f \in C^\infty(\mathcal{M})$

as the vector field

$$\operatorname{grad} f \in \mathfrak{X}(\mathcal{M}), \tag{2.42}$$

that is uniquely determined by

$$\langle \operatorname{grad} f, X\rangle_g = df(X) = Xf, \qquad \forall X \in \mathfrak{X}(\mathcal{M}). \tag{2.43}$$

Moreover, with the linear tangent-cotangent isomorphism mapping a vector field to a covector field which is given by

$$\hat{g}\colon \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}^*(\mathcal{M}), \quad X \mapsto g(X, \cdot), \tag{2.44}$$

we have an equivalent characterization of the Riemannian gradient, i.e. $\operatorname{grad} f = \hat{g}^{-1}(df)$.

2.2.3. Affine Connections

An affine connection is a differential operator which defines the covariant derivative and the notion of geodesics. Note, that in this section the symbol ∇ is exclusively used for affine connections and in all other sections it denotes the Euclidean gradient.

An **affine connection** on a smooth manifold $\mathcal{M}$ is a linear map

$$\nabla\colon \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}), \quad (X, Y) \mapsto \nabla(X, Y) =: \nabla_X Y, \tag{2.45}$$

with the following properties

(i) $C^\infty(\mathcal{M})$-linearity in the first argument,

$$\nabla_{(X_1+X_2)}Y = \nabla_{X_1}Y + \nabla_{X_2}Y, \tag{2.46a}$$

$$\nabla_{(fX)}(Y) = f\nabla_X Y, \tag{2.46b}$$

(ii) $\mathbb{R}$-linearity and product rule in the second argument

$$\nabla_X(Y_1 + Y_2) = \nabla_X Y_1 + \nabla_X Y_2, \tag{2.47a}$$

$$\nabla_X(fY) = X(f)Y + f\nabla_X Y, \tag{2.47b}$$

for all $f \in C^\infty(\mathcal{M})$ and $X, X_1, X_2, Y, Y_1, Y_2 \in \mathfrak{X}(\mathcal{M})$. In addition we have $\nabla_X f = df(X) = Xf$ for a real-valued function $f \in C^\infty(\mathcal{M})$ (a 0-tensor).

The resulting vector field $\nabla_X Y$ is called the **covariant derivative** of Y in direction X. Furthermore, the connection can be written in local coordinates by defining the **Christoffel symbols of the second kind** of the connection ∇ as

$$\Gamma_{ij}^k \frac{\partial}{\partial x^k} := \nabla_{\frac{\partial}{\partial x^i}} \frac{\partial}{\partial x^j}, \tag{2.48}$$

and for $X = X^i \frac{\partial}{\partial x^i}, Y = Y^j \frac{\partial}{\partial x^j}$ we have

$$\nabla_X Y = \left(X(Y^k) + X^i Y^j \Gamma_{ij}^k\right) \frac{\partial}{\partial x^k}. \tag{2.49}$$

2.2.4. Geodesics and Exponential Map

Note, that in this section the symbol ∇ is exclusively used for affine connections and in all other sections it denotes the Euclidean gradient.

A **smooth curve** on a smooth manifold $\mathcal{M}$ is a smooth mapping $\gamma\colon I \to \mathcal{M}$ and the **velocity** γ' defined by

$$\gamma'\colon I \to T\mathcal{M}, \quad t \mapsto \gamma'(t) \in T_{\gamma(t)}\mathcal{M}, \tag{2.50}$$

is a vector field along γ. The set of all smooth vector fields along γ is denoted by $\mathfrak{X}(\gamma)$. Furthermore, each connection ∇ on $\mathcal{M}$ induces a covariant derivative $D_t\colon \mathfrak{X}(\gamma) \to \mathfrak{X}(\gamma)$ along a smooth curve γ on $\mathcal{M}$. With this notion we are ready to define the **acceleration** of γ to be the vector field $D_t\gamma'$. A smooth curve γ is called a **geodesic** with respect to the connection ∇ if its acceleration is zero, i.e. $D_t\gamma' = 0$. Fixing a smooth chart (U, φ) arround $\gamma(t_0) \in U$ we can write the component functions of γ as $(\varphi \circ \gamma)(t) = (\gamma^1(t), \dots, \gamma^n(t))$ for $t \in (t_0 - \varepsilon, t_0 + \varepsilon)$. Then a geodesic γ is characterized by the **geodesic equation**

$$D_t\gamma' = \left(\ddot{\gamma}^k + \dot{\gamma}^i\dot{\gamma}^j\Gamma_{ij}^k \circ \gamma\right) \frac{\partial}{\partial x^k}\bigg|_\gamma = 0, \tag{2.51a}$$

$$\Leftrightarrow \quad \ddot{\gamma}^k(t) + \dot{\gamma}^i(t)\dot{\gamma}^j(t)\Gamma_{ij}^k(\gamma(t)) = 0, \tag{2.51b}$$

which is a system of second-order orinary differential equations (ODEs) for the real-valued functions $\gamma^1, \cdots, \gamma^n$. Given suitable initial conditions, $\gamma(0) = p \in \mathcal{M}$ and $\gamma'(0) = v \in T_p\mathcal{M}$, the theory of ODEs ensures existence and uniqueness of a maximal geodesic on an open interval $0 \in I \subseteq \mathbb{R}$. This geodesic is also called the geodesic with initial point p and initial velocity v and denoted by $\gamma_v(t)$.

Let $\mathcal{M}$ be a smooth manifold endowed with an affine connection ∇. Then the mapping

$$\operatorname{Exp}_p\colon V_p \subset T_p\mathcal{M} \to \mathcal{M}, \quad v \mapsto \gamma_v(1), \tag{2.52}$$

is called the **exponential map** at $p \in \mathcal{M}$, which is defined on

$$V_p = \{v \in T_p\mathcal{M}\colon \gamma_v \text{ exists on } [0,1]\}. \tag{2.53}$$

The exponential map is a diffeomorphism between $0 \in U \subset T_p\mathcal{M}$ and $p \in V \subset \mathcal{M}$. The inverse exponential map is called the **logarithmic map** and denoted by $\operatorname{Log}_p(q) = \operatorname{Exp}_p^{-1}(q)$ for $p, q \in \mathcal{M}$. A smooth manifold $\mathcal{M}$ is **geodesically complete** if $V_p = T_p\mathcal{M}$ for all $p \in \mathcal{M}$, i.e. every maximal geodesic is defined for all $t \in \mathbb{R}$. Due to the rescaling property we have $\operatorname{Exp}_p(tv) = \gamma_v(t)$.

In the Riemannian setting, the exponential map is also called Riemannian exponential map, that is with respect to Levi-Civita connection. Note that geodesics induced by the Levi-Civita connection are locally length minimizing, i.e. the shortest path between $\gamma_v(a)$ and $\gamma_v(b)$ for nearby points $a, b \in \mathcal{M}$. This motivates to define a generalization of the arithmetic mean in Euclidean spaces. We define the **weighted Riemannian mean** of a set $\{p_1, \dots, p_r\} \subset \mathcal{M}$ of r points with respect to the weights $w = (w_1, \dots, w_r) \in \Delta_r$ as the point $q \in \mathcal{M}$ satisfying

$$q = \operatorname*{arg\,min}_{p \in \mathcal{M}} J_w(p), \qquad \text{with} \qquad J_w(p) = \frac{1}{2}\sum_{i \in [r]} w_i d_g^2(p_i, p), \tag{2.54}$$

with respect to the Riemannian distance $d_g(q, p)$ on $\mathcal{M}$. The Riemannian gradient of J_w [Jos17, Lemma 6.9.4] at $p \in \mathcal{M}$ reads

$$\operatorname{grad}_p J_w = -\sum_{i \in [r]} w_i \operatorname{Log}_p(p_i) \in T_p\mathcal{M}, \tag{2.55}$$

and therefore the neccesary optimality condition for a critical point q is given by

$$\sum_{i \in [r]} w_i \operatorname{Log}_q(p_i) = 0. \tag{2.56}$$

This equation is typically solved by a fixed point iteration

$$q^{(k+1)} = \operatorname{Exp}_{q^{(k)}}\Big(\sum_{i \in [r]} w_i \operatorname{Log}_{q^{(k)}}(p_i)\Big), \tag{2.57}$$

with a suitable initialization $q^{(0)} \in \mathcal{M}$.

2.2.5. Matrix Manifolds

Here we list important embedded submanifolds of $\mathbb{R}^{n\times c}$, i.e. matrix manifolds, that are used in the subsequent chapters of this work. Furthermore, we parametrize all objects with respect to the embedding space $\mathbb{R}^{n\times c}$ instead of using charts. For a detailed introduction to matrix manifolds we refer to [AMS09].

2.2.5.1. The Positive Definite Matrix Manifold $\mathcal{P}^n$

The following is taken from [Bha06]. The open set

$$\mathcal{P}^n = \{S \in \mathbb{S}^n \colon \lambda_i(S) > 0,\ \forall i \in [n]\} \tag{2.58}$$

of symmetric and positive definite matrices is an embedded submanifold of $\mathbb{S}^n$ which is a linear subspace of $\mathbb{R}^{n\times n}$. The smooth manifold $\mathbb{R}_*^{n\times c}$ is called **positive definite matrix manifold** and the tangent spaces $T_S\mathcal{P}^n \cong \mathbb{S}^n$ are identified with $\mathbb{S}^n$ and hence the tangent bundle $T\mathcal{P}^n$ is trivial. Moreover, the manifold $\mathcal{P}^n$ forms a Riemannian manifold with the metric

$$g_S(S_1, S_2) = \operatorname{tr}(S^{-1}S_1S^{-1}S_2), \qquad S_1, S_2 \in \mathbb{S}^n, \quad S \in \mathcal{P}^n \tag{2.59a}$$

and corresponding norm

$$\|T\|_{g(S)} = \|S^{-1/2}TS^{-1/2}\|, \qquad T \in \mathbb{S}^n, \quad S \in \mathcal{P}^n. \tag{2.59b}$$

For any $A, B \in \mathcal{P}^n$, there exists a unique geodesic joining A and B given by

$$\gamma(s) = A^{1/2}\big(A^{-1/2}BA^{-1/2}\big)^s A^{1/2}, \qquad s \in [0,1]. \tag{2.60}$$

2.2.5.2. Compact Stiefel Manifold $\operatorname{St}(c,n)$

Let $c \leq n$, then the set of all $n \times c$ orthonormal matrices

$$\operatorname{St}(c,n) = \{X \in \mathbb{R}^{n\times c} \colon X^\top X = I_c\}, \tag{2.61}$$

is an embedded submanifold of $\mathbb{R}_*^{n\times c}$. The smooth manifold $\operatorname{St}(c,n)$ is called **compact Stiefel manifold** and the tangent spaces are given by

$$T_X\operatorname{St}(c,n) = \{Z \in \mathbb{R}^{n\times c} \colon X^\top Z + Z^\top X = 0\}. \tag{2.62}$$

The manifold $\mathbb{R}_*^{n\times c}$ is turned into a Riemannian manifold with the Euclidean metric

$$g_X(Z_1, Z_2) = \operatorname{tr}(Z_1^\top Z_2), \qquad Z_1, Z_2 \in \mathbb{R}^{n\times c},\ X \in \operatorname{St}(c,n), \tag{2.63}$$

inherited from the embedding space. In case of $n = 1$, the compact Stiefel manifold reduces to the unit sphere in $\mathbb{R}^n$ and in case of $n = c$ it becomes the orthogonal group $\mathcal{O}(n)$.

2.2.5.3. Non-Compact Stiefel Manifold $\mathbb{R}_*^{n\times c}$

Let $c \leq n$, then the set of all $n \times c$ matrices whose columns are linearly independent

$$\mathbb{R}_*^{n\times c} = \{X \in \mathbb{R}^{n\times c} \colon \det(X^\top X) \neq 0\}, \tag{2.64}$$

is an open subset of $\mathbb{R}^{n\times c}$ and consequently is an embedded submanifold of $\mathbb{R}^{n\times c}$. The smooth manifold $\mathbb{R}_*^{n\times c}$ is called **non-compact Stiefel manifold** and the tangent spaces are given by

$$T_X\mathbb{R}_*^{n\times c} = \mathbb{R}^{n\times c}. \tag{2.65}$$

The manifold $\mathbb{R}_*^{n\times c}$ is turned into a Riemannian manifold with the metric

$$g_X(Z_1, Z_2) = \operatorname{tr}\left((X^\top X)^{-1} Z_1^\top Z_2\right), \qquad Z_1, Z_2 \in \mathbb{R}^{n\times c},\ X \in \mathbb{R}_*^{n\times c}. \tag{2.66}$$

In case of $n = c$, the non-compact Stiefel manifold is the general linear group $\mathrm{GL}(n)$, i.e. the set of all invertible $n \times n$ matrices.

2.3. The Assignment Framework

This section reviews the assignment flow [ÅPSS17], which is a basic dynamical system labeling data given on a graph in supervised scenarios. We refer to [Sch20] for a more elaborate exposition and a review of recent developments. Well-posedness and stability of the supervised assignment flow was established in [ZZS20] under suitable conditions. In the subsequent sections, all objects from differential geometry are parametrized with respect to the embedding space $\mathbb{R}^{n\times c}$ instead of using local coordinates.

2.3.1. Assignment Manifold

Let $(\mathcal{F}, d_\mathcal{F})$ be a metric space and

$$\mathcal{F}_n = \{f_i \in \mathcal{F} \colon i \in \mathcal{I}\}, \qquad |\mathcal{I}| = n, \tag{2.67}$$

be given data. Assume that a predefined set of **prototypes**

$$\mathcal{F}_* = \{f_j^* \in \mathcal{F} \colon j \in \mathcal{J}\}, \qquad |\mathcal{J}| = c, \tag{2.68}$$

is given. *Data labeling* denotes the assignment

$$j \to i, \qquad f_j^* \to f_i \tag{2.69}$$

of a single prototype $f_j^* \in \mathcal{F}_*$ to each data point $f_i \in \mathcal{F}_n$. The set $\mathcal{I}$ is assumed to form the vertex set of an undirected graph $\mathcal{G} = (\mathcal{I}, \mathcal{E})$ which defines a relation $\mathcal{E} \subset \mathcal{I} \times \mathcal{I}$ and neighborhoods

$$\mathcal{N}_i = \{k \in \mathcal{I} \colon ik \in \mathcal{E}\} \cup \{i\}, \tag{2.70}$$

where ik is a shorthand for the unordered pair (edge) $(i, k) = (k, i)$.

The assignments (labeling) (2.69) are represented by matrices in the set

$$\mathcal{W}_*^c = \left\{W \in \{0,1\}^{n\times c} \colon W\mathbb{1}_c = \mathbb{1}_n,\ \operatorname{rank}(W) = c\right\} \tag{2.71}$$

with unit vectors W_i, $i \in \mathcal{I}$, called **assignment vectors**, as row vectors. Moreover, the rank constraint ensures that exactly c labels are assigned. These assignment vectors are computed by numerically integrating the assignment flow, Definition 2.3.1, in the following elementary geometric setting. The integrality constraint and the rank constraint of (2.71) is relaxed and the vectors

$$W_i = (W_{i,1}, \ldots, W_{i,c})^\top \in \mathcal{S}, \quad i \in \mathcal{I}, \tag{2.72}$$

are discrete probability measures on the set of labels indexed by $\mathcal{J}$, but still called **assignment vectors**. The vectors W_i, $i \in \mathcal{I}$ are points on the Riemannian manifold (recall (1.1))

$$(\mathcal{S}, g), \qquad \mathcal{S} = \{p \in \Delta_c \colon p > 0\} \tag{2.73}$$

with

$$\mathbb{1}_\mathcal{S} = \frac{1}{c}\mathbb{1}_c \in \mathcal{S}, \qquad (\textbf{barycenter}) \tag{2.74}$$

tangent space

$$T_0 = \{v \in \mathbb{R}^c \colon \langle \mathbb{1}_c, v\rangle = 0\} \tag{2.75}$$

and trivial tangent bundle $T\mathcal{S} = \mathcal{S} \times T_0$, orthogonal projection

$$\Pi_0 \colon \mathbb{R}^c \to T_0, \qquad \Pi_0 = \Pi_{T_0} = I_c - \mathbb{1}_{\mathcal{S}}\mathbb{1}_c^\top \tag{2.76}$$

and the Fisher-Rao metric

$$g_p(u,v) = \sum_{j\in\mathcal{J}} \frac{u^j v^j}{p^j}, \quad p \in \mathcal{S}, \quad u, v \in T_0. \tag{2.77}$$

Based on the linear map

$$R_p \colon \mathbb{R}^c \to T_0, \qquad R_p = \operatorname{Diag}(p) - pp^\top, \qquad p \in \mathcal{S} \tag{2.78}$$

satisfying

$$R_p = R_p\Pi_0 = \Pi_0 R_p, \tag{2.79}$$

exponential maps and their inverses are defined as

$$\operatorname{Exp} \colon \mathcal{S} \times T_0 \to \mathcal{S}, \qquad (p,v) \mapsto \operatorname{Exp}_p(v) = \frac{pe^{\frac{v}{p}}}{\langle p, e^{\frac{v}{p}}\rangle}, \tag{2.80a}$$

$$\operatorname{Exp}_p^{-1} \colon \mathcal{S} \to T_0, \qquad q \mapsto \operatorname{Exp}_p^{-1}(q) = R_p \log \frac{q}{p}, \tag{2.80b}$$

$$\exp_p \colon T_0 \to \mathcal{S}, \qquad \exp_p = \operatorname{Exp}_p \circ R_p, \tag{2.80c}$$

$$\exp_p^{-1} \colon \mathcal{S} \to T_0, \qquad \exp_p^{-1}(q) = \Pi_0 \log \frac{q}{p}. \tag{2.80d}$$

We call the linear map (2.78) *replicator map* because it yields, for any vector field $F \colon \mathcal{S} \to \mathbb{R}^c$ that represents affinity measures for the set of labels (2.68), a vector field R_pF on $\mathcal{S}$ and in turn the corresponding *replicator equation* [HS03]

$$\dot{p}_j = \big(R_pF(p)\big)_j = p_j\big(F_j(p) - \mathbb{E}_p[F]\big) = p_jF_j(p) - \langle p, F(p)\rangle p_j, \qquad j \in \mathcal{J}. \tag{2.81}$$

If $F = \nabla E$ is derived as Euclidean gradient of a potential E, then $R_pF(p) = \operatorname{grad}_{\mathcal{S}} E$ is the corresponding Riemannian gradient with respect to the Fisher-Rao metric (2.77) [ÅPSS17, Prop. 1].

Furthermore, the map Exp corresponds to the e-connection of information geometry, rather than to the exponential map of the Riemannian connection [AN00]. Accordingly, the geodesics with respect to the affine e-connection (2.80a) are not length-minimizing. But they provide a close approximation [ÅPSS17, Prop. 3] and are more convenient for numerical computations. In particular, all simplex constraints (normalization of assignment vectors as discrete distributions) are smoothly 'built in'. Yet, unlike the geometry induced by traditional barrier functions (see, e.g., [NT02]), the information geometry underlying the assignment flow $W(t)$ entails that it may – and in fact *does* [ZZS20] – evolve arbitrarily close to the boundary of the assignment manifold so as to determine unambigous label assignments for t large enough.

Remark 2.3.1. Applying the map $\exp_p$ to a vector in $\mathbb{R}^c = T_0 \oplus \mathbb{R}\mathbb{1}_c$ does not depend on the constant component of the argument, due to (2.79).

Finally, the **assignment manifold** is defined as

$$(\mathcal{W}, g), \qquad \mathcal{W} = \mathcal{S} \times \cdots \times \mathcal{S}. \qquad (n = |\mathcal{I}| \text{ factors}) \tag{2.82}$$

Points $W \in \mathcal{W}$ are row-stochastic matrices $W \in \mathbb{R}^{n\times c}$ with row vectors $W_i \in \mathcal{S}$, $i \in \mathcal{I}$ that

represent the assignments (2.69) for every $i \in \mathcal{I}$. We set

$$\mathcal{T}_0 = T_0 \times \cdots \times T_0 \qquad (n = |\mathcal{I}| \text{ factors}) \tag{2.83}$$

with tangent vectors $V \in \mathbb{R}^{n\times c}$, $V_i \in T_0$, $i \in \mathcal{I}$. All the mappings defined above factorize in a natural way and apply row-wise, e.g. $\operatorname{Exp}_W = (\operatorname{Exp}_{W_1}, \ldots, \operatorname{Exp}_{W_n})$ etc.

2.3.2. Assignment Flow

Based on (2.67) and (2.68), the distance vector field

$$D_{\mathcal{F};i} = \big(d_{\mathcal{F}}(f_i, f_1^*), \ldots, d_{\mathcal{F}}(f_i, f_c^*)\big)^\top, \qquad i \in \mathcal{I} \tag{2.84}$$

is well-defined. These vectors are collected as row vectors of the **distance matrix**

$$D_{\mathcal{F}} \in \mathbb{R}_+^{n\times c}. \tag{2.85}$$

The **likelihood map** and the **likelihood vectors**, respectively, are defined as

$$L_i \colon \mathcal{S} \to \mathcal{S}, \qquad L_i(W_i) = \exp_{W_i}\Big(-\frac{1}{\rho} D_{\mathcal{F};i}\Big) = \frac{W_i e^{-\frac{1}{\rho} D_{\mathcal{F};i}}}{\langle W_i, e^{-\frac{1}{\rho} D_{\mathcal{F};i}}\rangle}, \qquad i \in \mathcal{I}, \tag{2.86}$$

where the scaling parameter $\rho > 0$ is used for normalizing the a-priori unknown scale of the components of $D_{\mathcal{F};i}$ that depends on the specific application at hand.

A key component of the assignment flow is the interaction of the likelihood vectors through *geometric* averaging within the local neighborhoods (2.70). Specifically, using the weights

$$\Omega_i = \Big\{ w_{i,k} \colon k \in \mathcal{N}_i,\ w_{i,k} > 0,\ \sum_{k\in\mathcal{N}_i} w_{i,k} = 1 \Big\}, \quad i \in \mathcal{I}, \tag{2.87}$$

the **similarity map** and the **similarity vectors**, respectively, are defined as

$$S_i \colon \mathcal{W} \to \mathcal{S}, \qquad S_i(W) = \operatorname{Exp}_{W_i}\Big(\sum_{k\in\mathcal{N}_i} w_{i,k} \operatorname{Exp}_{W_i}^{-1}\big(L_k(W_k)\big)\Big), \qquad i \in \mathcal{I}. \tag{2.88}$$

If Exp_{W_i} were the exponential map of the Riemannian (Levi-Civita) connection, then the argument inside the brackets of the right-hand side would just be the negative Riemannian gradient (2.55) with respect to W_i of the center of mass objective function (2.54) comprising the points L_k, $k \in \mathcal{N}_i$, i.e. the weighted sum of the squared Riemannian distances between W_i and L_k. Informally speaking we have the interpretation: $S_i(W)$ moves W_i towards the geometric mean of the likelihood vectors L_k, $k \in \mathcal{N}_i$. Since $\operatorname{Exp}_{W_i}(0) = W_i$, this mean is equal to W_i if the aforementioned gradient vanishes.

After having described all required building blocks, we can state Definition 2.3.1 of a smooth dynmical system for supervised data labeling.

Definition 2.3.1 (Assignment Flow). Let the similarity map $S(W)$ be based on the likelihood map $L(W)$ defined in (2.88), (2.86) and a field of distances $D_{\mathcal{F}} \in \mathbb{R}_+^{n\times c}$ in (2.84) be given. Then, the **assignment flow** is induced by the system of nonlinear ODEs

$$\dot{W} = R_W S(W), \qquad W(0) = \mathbb{1}_{\mathcal{W}}, \tag{2.89a}$$

$$\dot{W}_i = R_{W_i} S_i(W), \qquad W_i(0) = \mathbb{1}_{\mathcal{S}}, \quad i \in \mathcal{I}, \tag{2.89b}$$

where $\mathbb{1}_{\mathcal{W}} \in \mathcal{W}$ denotes the barycenter of the assignment manifold (2.82).

The system (2.89a) collects all systems (2.89b), for every vertex $i \in \mathcal{I}$. The latter systems are coupled within local neighborhoods $\mathcal{N}_i$ due to the similarity vectors $S_i(W)$ given by (2.88).

2.3.3. Geometric Integration

The solution $W(t) \in \mathcal{W}$ is numerically computed by geometric integration of the system of ODEs (2.89a) and determines a labeling $W(T) \in \mathcal{W}_*^c$ for sufficiently large T after a trivial rounding operation. There exists a variety of geometric integration methods [HLW06]. However, in particular, the Lie group method [IMNZ00] is very suitable for our scenario. The application of a Lie group method presumes a Lie group G together with an action $\Lambda\colon G \times \mathcal{M} \to \mathcal{M}$ defined on a manifold $\mathcal{M}$.

In our setting of the assignment flow we identify $G = T_0$ with the flat tangent space and define the action $\Lambda\colon T_0 \times \mathcal{S} \to \mathcal{S}$ as

$$\Lambda(v, p) = \exp_p(v), \qquad v \in T_0,\ p \in \mathcal{S}, \tag{2.90}$$

which satisfies the following properties

$$\Lambda(0, p) = p \quad \text{and} \quad \Lambda(v_1 + v_2, p) = \Lambda(v_1, \Lambda(v_2, p)). \tag{2.91}$$

Then the Lie group method yields the following **tangent space parametrization** of the assignment flow.

Proposition 2.3.1 (Tangent Space Parametrization [ZSPS20])**.** *The solution $W(t)$ of the assignment flow (Definition 2.3.1) emanating from any $W(0) = W_0$ admits the representation*

$$W(t) = \exp_{W_0}\big(V(t)\big), \tag{2.92}$$

where $V(t) \in \mathcal{T}_0$ solves

$$\dot{V} = \Pi_0 S\big(\exp_{W_0}(V)\big). \tag{2.93}$$

Since (2.93) is a system of ODEs in a vector space, we can apply any integration method from the Euclidean setting. In this thesis we only consider the simplest scheme, the **geometric Euler method**. This explicit scheme with fixed step-size $h > 0$ reads

$$W_i^{(k+1)} = \exp_{W_i^{(k)}}\big(hS(W^{(k)})\big), \quad i \in \mathcal{I}, \tag{2.94}$$

which ensures that the assignment flow, Definition 2.3.1, evolves properly on the assignment manifold $\mathcal{W}$. We refer to [ZSPS20] for a detailed evaluation of more advanced methods including schemes with adaptive stepsize selection.

2.4. Data Clustering

Here, we review basic concepts of supervised and unsupervised data clustering, that are required in the subsequent chapters of this thesis.

2.4.1. Scatter Matrices

We will use the following basic concepts of statistical pattern recognition [DK82] for interpreting self-assignment flows from a corresponding angle in Section 5.4.4. Let

$$\mathcal{F}_n = \{f_i \in \mathbb{R}^d,\ i \in \mathcal{I}\} \tag{2.95}$$

denote given data in terms of feature vectors in a Euclidean space. Suppose these data are classified corresponding to the partition

$$\mathcal{I} = \dot{\bigcup_{j\in[c]}} \mathcal{I}_j, \qquad |\mathcal{I}_j| = n_j, \qquad \sum_{j\in[c]} n_j = n = |\mathcal{I}|, \qquad c \in \mathbb{N}, \tag{2.96}$$

that is, datum f_i belongs to class j iff $i \in \mathcal{I}_j$.

We define the empirical quantities

$$P_j = \frac{n_j}{n}, \qquad j \in [c] \qquad \text{(prior probabilities)} \tag{2.97a}$$

$$m_j = \frac{1}{n_j} \sum_{i \in \mathcal{I}_j} f_i, \qquad j \in [c] \qquad \text{(class-conditional mean vectors)} \tag{2.97b}$$

$$m = \sum_{j \in [c]} P_j m_j = \frac{1}{n} \sum_{i \in [n]} f_i \qquad \text{(mean vector)} \tag{2.97c}$$

and the **scatter matrices** (empirical covariance matrices)

$$S_t = \frac{1}{n} \sum_{i \in [n]} (f_i - m)(f_i - m)^\top, \tag{2.98a}$$

$$S_w = \sum_{j \in [c]} P_j \cdot \frac{1}{n_j} \sum_{i \in \mathcal{I}_j} (f_i - m_j)(f_i - m_j)^\top = \frac{1}{n} \sum_{j \in [c]} \sum_{i \in \mathcal{I}_j} (f_i - m_j)(f_i - m_j)^\top, \tag{2.98b}$$

$$S_b = \sum_{j \in [c]} P_j (m_j - m)(m_j - m)^\top. \tag{2.98c}$$

S_w is called the **within-class scatter matrix**, whereas S_b is called the **between-class scatter matrix**. S_t is called the **total scatter matrix** since

$$S_t = S_w + S_b, \tag{2.99}$$

as an elementary computation shows.

In *supervised* scenarios the class-label assignments $i \in \mathcal{I}_j$ are known and the decomposition (2.99) can be computed. Assuming S_w has full rank, a basic objective for dimension reduction by extracting lower-dimensional features from the data $\mathcal{F}_n$ is then given by the class-separability measure

$$\operatorname{tr}(S_w^{-1} S_b). \tag{2.100}$$

Defining the features by $Y^\top x$, for some matrix $Y \in \mathbb{R}^{d \times c}$ to be determined, transforms (2.100) to $\operatorname{tr}((Y^\top S_w Y)^{-1} Y^\top S_b Y)$. Maximizing this objective with respect to Y *simultaneously* maximizes the between-class variation and minimizes the within-class variation. The column vectors of the optimal Y are given by dominant generalized eigenvectors of the matrix pencil (S_b, S_w). The map $Y^\top x$ to a lower-dimensional space preserves the structure of the data, as represented by the scatter matrices S_w, S_b, as much as possible.

2.4.2. Sketching Large Affinity Matrices

In order to cope with large-scale scenarios, we will have to compress large symmetric and positive semi-definite matrices $K \in \mathbb{S}^n$ with $K \succeq 0$. The problem is to obtain a computationally feasible approximation of the best rank-ℓ approximation

$$K_\ell = U_1 D_\ell(K) U_1^\top, \qquad \ell \ll n, \tag{2.101}$$

where D_ℓ and $U_1 \in \mathbb{R}^{n \times \ell}$ contain the dominant eigenvalues and eigenvectors of the spectral decomposition $K = U D(K) U^\top$. Computing (2.101) directly for large n using the Singular Value Decomposition (SVD) is too expensive. Computationally feasible approximations [GM16] yield the **compressed matrix**

$$\widehat{K}_\ell = C A^\dagger C^\top \tag{2.102a}$$

parametrized by a **sketching matrix** $S \in \mathbb{R}^{n\times\ell}$ with

$$C = K^q S, \qquad A = S^\top K^{2q-1} S, \qquad q \in \mathbb{N} \tag{2.102b}$$

which hence has rank at most ℓ. $A^\dagger$ is the Moore-Penrose generalized inverse of A and $q \in \{1,2,3\}$ is a small integer in practice.

In this work, we confine ourselves to the following computationally cheap version of this method for computing (2.102a), based on *uniform sampling* of ℓ columns directly from K. Assuming w.l.o.g. that they form the first ℓ columns of K, the corresponding partition $[n] = [\ell] \cup \big([n] \setminus [\ell]\big)$ and $S = \left(\begin{smallmatrix} I_\ell \\ 0 \end{smallmatrix}\right)$ yields with $q = 1$

$$K = \begin{pmatrix} A & B_1 \\ B_1 & B_2 \end{pmatrix}, \qquad C = \begin{pmatrix} A \\ B_1 \end{pmatrix}, \tag{2.103}$$

and using $AA^\dagger A = A$,

$$\widehat{K}_\ell = \begin{pmatrix} A \\ B_1 \end{pmatrix} A^\dagger \begin{pmatrix} A & B_1 \end{pmatrix} = \begin{pmatrix} A & AA^\dagger B_1 \\ B_1 A^\dagger A & B_1 A^\dagger B_1 \end{pmatrix}. \tag{2.104}$$

Assuming that A has full rank, we obtain the classical **Nyström extension**

$$\widehat{K}_\ell = \begin{pmatrix} A & B_1 \\ B_1 & B_1 A^{-1} B_1 \end{pmatrix} \tag{2.105}$$

introduced in machine learning by [WS01], studied much earlier in linear algebra – see, e.g., the Schur compression matrix and references in [And79] – and analyzed by [DM05].

Choosing $q > 1$ is more expensive due to the multiplication of the large matrix K of (2.102b) but yields in theory a better approximation of (2.101) by (2.102a) with respect to the spectral norm. Further improvement is possible by randomly mixing the columns of K before sampling a subset of them. Specifically, the sketching matrix S is given by

$$S = \sqrt{\frac{n}{\ell}} D_R H R, \tag{2.106}$$

where $D_R \in \mathbb{R}^{n\times n}$ is a diagonal matrix of Rademacher random variables, i.e. $D_{ii} = +1$ or $D_{ii} = -1$ each with probability $\frac{1}{2}$, H is the normalized orthogonal matrix corresponding to the fast Hadamard transform, and $R \in \{0,1\}^{n\times\ell}$ selects the columns of K: $R_{ij} = 1$ iff the ith column of K is the jth column among the selected ℓ columns.

2.4.3. Greedy k-Center Metric Clustering

In order to handle large-scale scenarios, the following simple but effective algorithm from [Har11] can be employed for data reduction in a preprocessing step. The algorithm approximates the k-center clustering along with a *performance guarantee* (2-approximation) and only requires *linear complexity* $\mathcal{O}(nc)$ with respect to the (large) number of data points n. By using a min-max objective (see (2.108) below), selected data points are evenly spread among all data points and hence do not introduce a bias beforehand.

The task of **k-center clustering** is as follows. Given data points $\mathcal{F}_n$ from a metric space $(\mathcal{F}, d_\mathcal{F})$, determine a subset

$$\mathcal{F}_c = \{f_j : j \in \mathcal{J}\} \subset \mathcal{F}_n, \qquad |\mathcal{J}| = c. \tag{2.107}$$

that solves the combinatorially hard optimization problem

$$E_\infty^* = \min_{\mathcal{F}_c \subset \mathcal{F}_n, |\mathcal{F}_c| = c} E_\infty(\mathcal{F}_c), \qquad E_\infty(\mathcal{F}_c) = \max_{f\in\mathcal{F}_n} d_\mathcal{F}(f, \mathcal{F}_c), \tag{2.108}$$

where $d_\mathcal{F}(f, \mathcal{F}_c) = \min_{f'\in\mathcal{F}_c} d_\mathcal{F}(f, f')$.

A **greedy approximation** is computed as follows. Start with an initial point f_1, e.g. chosen randomly in $\mathcal{F}_n$. Then select the remaining $c-1$ points $f_2, \ldots, f_c$ successively by determining the point that is most distant from the current subset of already selected points, to obtain a set $\mathcal{F}_c$ that is a 2-approximation $E_\infty(\mathcal{F}_c) \leq 2E^*_\infty$ of the optimum (2.108) [Har11, Thm. 4.3]. As a consequence, the subset of c points of $\mathcal{F}_c$ are almost uniformly distributed within $\mathcal{F}_n$, as measured by the metric $d_{\mathcal{F}}$.

2.5. Discrete Tomography

The main goal of **computed tomography (CT)** [NW01] is to visualize the internal structure of an object without damaging it. In this thesis we only consider objects in the plane $\mathbb{R}^2$. However, the method can be extended to higher dimensions, like volumes, in a straightforward way. The aquired measurements are obtained by exploiting the interaction between the object and energetic particles. We assume that the measurements consist of a

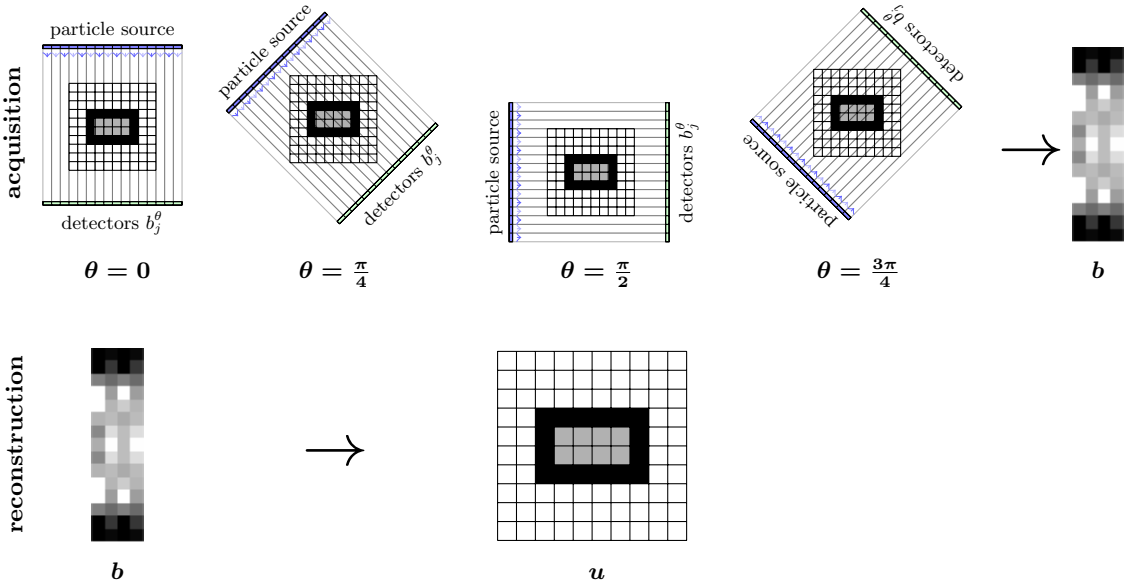

Figure 2.1.: Discrete tomographic setup. The acquisition process of the measurements $b \in \mathbb{R}^m$ is given by different line integrals from different projection angles over the object in investigation. Moreoever, the known projection matrix $A \in \mathbb{R}^{m \times n}$ describes the geometric setup of the acquisition process. As a consequence, the reconstruction problem of discrete tomography amounts in solving the system of linear equations $Au = b$ to recover the unknown image u with intensities restricted to the discrete set $\mathcal{F}_* = \{\square, \blacksquare, \blacksquare\}$.

finite set of parallel and equidistant projection beams such that the whole image domain is covered for each of the equiangular projection angles between 0 and π. All **tomographic measurements** are collected by the vector $b \in \mathbb{R}^m$ where m is the number of measurements. The given **projection matrix** $A \in \mathbb{R}^{m \times n}$ encodes the imaging geometry and each entry a_{ij} corresponds to the length of the line segment of the i-th projection ray passing through the j-th pixel in the image domain. Consequently, **tomographic reconstruction** amounts to solving the linear system of equations $Au = b$ and the reconstructed image is represented

by the vector $u \in \mathbb{R}^n$ where n is the number of pixels.

The field of **discrete tomography** [HK99] assumes that the reconstructed image only take values from a finite set of intensities, which may be given beforehand as prior knowledge. This restriction induces a strong regularization effect, since the search space of feasible solutions is drastically reduced. Hence the reconstructed image u represents a piecewise constant function

$$u \in \mathcal{F}_*^n, \qquad \mathcal{F}_* := \{f_1^*, \dots, f_c^*\} \subset [0,1], \tag{2.109}$$

where u is restricted to a finite range and we assume that $\mathcal{F}_*$ forms an ordered set with values that are normalized between 0 and 1. Therefore, our main focus lies on solving the **discrete reconstruction problem**

$$Au = b \quad \wedge \quad u_i \in \mathcal{F}_*, \quad \forall\, i \in [n], \tag{2.110}$$

and on effectively exploiting additional prior knowledge, given in terms of $\mathcal{F}_*$. Figure 2.1 gives an overview of the discrete tomography setup. Moreoever, in discrete tomography, it is usually assumed to have only a few projection angles given, i.e. there are fewer measurements $m \ll n$ available than pixels n. An important special case is binary discrete tomography, where the reconstructed image is restricted to $\mathcal{F}_* = \{0, 1\}$. Gardner [GGP99] proved that the complexity of the discrete reconstruction problem is NP-complete already in the binary case for more than two projection angles.

CHAPTER 3

Non-Smooth Multilabel Methods for Tomography Reconstruction

3.1. Introduction and Overview

This chapter deals with a challenging generalization of the *multilabel problem*. Given an image, the multilabel problem consists of assigning an optimal label from a finite set of labels (prototypes) to each pixel of a spatial domain. The set of prototypes is pre-specified and constitutes a strong prior knowledge. The optimality of a label configuration is specified by a minimum energy criterion over all admissible label configurations. Usually, discrete label decisions are relaxed to probability vectors (assignments). Additionally, *convex relaxations* [ZGFN08; LBS09; PCCB09] rely on finding a tractable convex approximation of the non-convex energy criterion. Subsequently, approximate label inference in large-scale scenarios can be carried out by efficient convex optimization techniques like the Chambolle Pock primal-dual (PD) algorithm [CP11].

Here, we are considering a generalization of the multilabel problem, where the image to be labeled cannot be observed directly. Instead indirect measurements related by a linear operator are available. Hence, the task is to find a discrete labeling which is a solution of an inverse problem simultaneously. Consequently, the *multilabel problem with indirect measurements* is essentially more difficult than the multilabel problem with direct measurements in general.

One main application in this setting is non-binary *discrete tomography reconstruction* [HK99; NW01] of images $u \in \mathbb{R}^n$ from a small number of noisy measurements given by

$$b = Au + \nu \in \mathbb{R}^m. \tag{3.1}$$

The measurements correspond to line integrals that sum up all pixel-values (absorptions) over each ray transmitted through the object. Thus, the given projection matrix $A \in \mathbb{R}^{m \times n}$ encodes the imaging geometry. Applications range from medical imaging [BSLB11] to natural sciences and industrial applications, like non-destructive material testing [HFU08]. Many situations require to keep the number of measurements as low as possible, which

leads to a small number of projections and hence to a severely under-sampled (ill-posed) reconstruction problem. To cope with such problems, the under-sampled projection data b is compensated by the assumption that the finite range

$$\mathcal{F}_* := \{f_1^*, \dots, f_c^*\} \subset [0, 1], \tag{3.2}$$

of the reconstructed function u is known beforehand. Accordingly, discrete tomography reconstruction amounts to jointly solving a system of linear equations and labeling the solution.

To avoid the combinatorial nature and complexity of the discrete optimization problem, it is common to use convex and non-convex relaxations as approximations. A natural approach is to drop the discrete constraints completely and extend the feasible set $\mathcal{F}_*$ to its convex hull. However, this necessitates a subsequent rounding step which is which is hard to control. The solution of the relaxed problem with possibly intermediate values $x \notin \mathcal{F}_*$ is projected onto the set of admissible discrete solutions. More sophisticated two-step approaches employ a standard multilabel formulation where the already reconstructed image serves as input. However, these two-step approaches suffer from two major drawbacks. The reconstruction process is decoupled from the labeling process, such that the reconstruction method cannot exploit the crucial prior knowledge that the solution only takes values in the finite set $\mathcal{F}_*$. In addition, the resulting labeling after the second step is not necessarily consistent with the linear projection constraints. Therefore, our main concern is to effectively exploit the additional prior knowledge in terms of $\mathcal{F}_*$ already during the reconstruction process. Moreover, we are interested in solving large-scale problem instances by employing convex optimization techniques.

Our starting point are established convex relaxations of the multilabel problem with a Potts prior. First we lift the non-local tomographic projection constraints to probability vectors which encode discrete label decisions. This allows to incorporate information from the indirect measurements into the multilabel problem. However, the back-projection of the indicator variables to gray-values introduces non-integral solutions and results in a very loose relaxation. In other words, the lifted constraints only determine the pixel value but do not indicate how the indicator variables should realize this value (similar to estimating a vector given only its magnitude). A straightforward remedy is to consider an additional unary data term for pixelwise local decisions which depends on the last solution iteratively. This naive fixed-point iteration motivates a non-convex discretization term in order to avoid these unwanted effects and penalize convex combinations of labels. Finally, the proposed variational approximation for joint reconstruction and labeling is non-smooth and non-convex.

The numerical optimization is based on the difference of convex functions (DC) framework. Our proposed energy admits the decomposition into a convex part and a concave part, corresponding to the non-convex discretization term. The resulting algorithm recovers the naive fixed-point iteration, which amounts to solving a sequence of convex optimization problems, provably converging to a stationary point of the proposed non-convex energy.

We also propose a suitable rounding procedure as a post-processing step, since we have relaxed the integrality constraints.

Finally, in a comprehensive numerical evaluation we compare the proposed approach to state-of-the-art approaches from literature and demonstrate the superior reconstruction performance of our approach, even in noisy scenarios with only very few projection angles.

Organization. The remainder of this chapter is organized as follows:

- We review the related work on binary and non-binary discrete tomography (Sect. 3.2).
- We introduce the combinatorial multilabel problem with indirect measurements and Potts regularization for simultaneous reconstruction and labeling for discrete tomography (Sect. 3.3).
- We reformulate the non-local projection constraints in terms of probability vectors (label assignments) as a first step. Next, we motivate and derive a non-convex discretization term to exclude non-integral solutions (Sect. 3.4).
- We propose a non-smooth and non-convex variational approximation to the combinatorial multilabel problem with indirect measurements (Sect. 3.4).
- We work out an optimization algorithm to minimize the proposed non-smooth and non-convex energy based on the DC framework. The resulting algorithm recovers the naive fixed-point iteration, which solves a sequence of convex problems (Sect. 3.5).
- Finally, we numerically evaluate and compare the proposed approach with state-of-the-art methods for non-binary tomography reconstruction on standard test-datasets with only a few projection angles (Sect. 3.6).

3.2. Related Work

In this section, we review the related work on discrete tomography. Discrete tomography considers either *binary* or *non-binary* (multivalued) problems. The latter ones are considerably more involved.

Regarding *binary discrete tomography*, Weber et al. [WSH03; SSWH05] proposed to combine a quadratic program with a non-convex penalty which gradually enforces binary constraints. In our previous work [KPSZ15] we showed how a binary discrete graphical model and a sequence of s-t graph-cuts can be used to take into account the affine projection constraints and to recover high-quality reconstructions.

Regarding *non-binary discrete tomography*, an extension of the latter approach is not straightforward due to the nonlocal projection constraints. Weber [Web09, Chapter 6] proposed a non-convex term for non-binary discrete tomography which we will derive in a natural way in the present work. However, Weber's approach differs with respect to the data term for the projection constraints, regularization and optimization, and additionally requires parameter tuning.

Because u is assumed to be piecewise constant, an obvious approach is to consider sparsity promoting priors. The authors in [SWFU15] proposed a dynamic programming approach for minimizing the ℓ_0-norm of the gradient. However, the set $\mathcal{F}_*$ of feasible intensities is not exploited. In the convex setting, the integrality constraints are dropped and priors like the ℓ_1-norm or the total variation (TV) are used [SP08; GBB+12; DPSS14], with a post-processing step to round the continuous solution to a piecewise constant one.

This approach connects discrete tomography and the fast evolving field of *compressed sensing* with corresponding recovery guarantees [DPSS14]. Again, the prior information of the range of the image to be reconstructed is not involved in the optimization process. Next, we focus on methods that make direct use of the set $\mathcal{F}_*$ during the reconstruction process.

A first step towards incorporating the prior information of discrete labels was done in our preliminary work [ZKS+16], where we combined a total variation formulation for reconstruction with a non-convex coupling term to enforce discrete labels of the solution.

Tuysuzoglu et al. [TKS+15] casted the non-binary discrete reconstruction problem into a series of submodular binary problems within an α-expansion approach by linearizing the ℓ_2-fidelity term around an iteratively updated working point. This local approximation discards a lot of information, and a significantly larger number of projections is required to achieve high-quality reconstructions.

Maeda et al. [MFKI10] suggested a probabilistic formulation which couples a continuous reconstruction with the Potts model. Alternating optimization is applied to maximize the *a posteriori* probability locally. However, there is no guarantee that these alternating continuous and discrete block coordinate steps converge.

Ramlau et al. [KR13] investigated the theoretical regularization properties of the piecewise constant Mumford-Shah functional [MS89] applied to linear ill-posed problems. In earlier work [RR07], they considered discrete tomography reconstruction using this framework. The difficult geometric optimization of the partition is carried out by a level-set approach and the intensities $\mathcal{F}_*$ were additionally estimated in an alternating fashion. By contrast, our approach is based on a convex relaxation of the perimeter regularization and the set $\mathcal{F}_*$ is assumed to be known beforehand.

Varga et al. [VBN12] suggested a heuristic algorithm which is adaptively combing an energy formulation with a non-convex polynomial in order to steer the reconstruction towards the feasible values.

Batenburg et al. [BS11] proposed the *Discrete Algebraic Reconstruction Technique (DART)* algorithm which starts with a continuous reconstruction by a basic algebraic reconstruction method, followed by thresholding to ensure a piecewise constant function. These steps interleaved with smoothing are iteratively repeated to refine the locations where u jumps. This heuristic approach yields good reconstructions in practice but cannot be characterized by an objective function that is optimized.

We consider [BS11; VBN12] as state-of-the-art approaches for experimental comparison.

3.3. Reconstruction by Multilabeling

In this section, we reformulate the discrete tomography reconstruction problem (3.4) as a combinatorial multilabel problem with indirect measurements. This means we cannot directly observe the quantity of interest $u \in \mathbb{R}^n$, the image to be labeled. However, we assume that we are able to observe the quantity $b \in \mathbb{R}^m$, which is related by the linear operator $A\colon \mathbb{R}^n \to \mathbb{R}^m$ to the reconstructed image. In literature the multilabel problem with indirect measurements is a less considered variant since it is essentially more difficult.

In practice, we assume that there are less observations (measurements) than pixels $m \ll n$ and hence that the corresponding equation system is under-determined. We also require full row-rank of the operator A, i.e. $\operatorname{rank}(A) = m$. To cope with such problems, a

common assumption in the field of discrete tomography [HK99] concerns knowledge of the finite range of

$$u \in \mathcal{F}_*^n, \qquad \mathcal{F}_* := \{f_1^*, \dots, f_c^*\} \subset [0,1], \tag{3.3}$$

that is, u represents a piecewise constant function. The discrete set of labels (prototypes) $\mathcal{F}_*$ are c admissible gray-values, which form an ordered set and are normalized between 0 and 1. Therefore, our main focus is to effectively exploit this additional prior knowledge, the set of labels $\mathcal{F}_*$ together with the projection constraints, in order to solve the *discrete reconstruction problem*

$$Au = b \quad \wedge \quad u_i \in \mathcal{F}_*, \quad \forall\, i \in [n]. \tag{3.4}$$

Reversing the forward operator A, even with the additional constraint that the reconstructed solution has to be a piecewise constant function, is severely ill-posed and requires regularization. A common choice is the Potts model [Pot52],

$$R(u) = \|\nabla u\|_0 := |\{i \in [n] \,|\, (\nabla u)_i \neq 0\}|, \tag{3.5}$$

for sparse gradient regularization which favors spatially coherent and piecewise constant images. Therefore regularization is incorporating further prior knowledge and is stabilizing the inversion by further reducing the search space of admissible solutions.

In presence of noise the existence of a solution is no longer guaranteed, since the corrupted measurements b might not be within the range of A. Hence, we use the more general constraints

$$\underline{b}(\epsilon) \leq Au \leq \overline{b}(\epsilon), \tag{3.6}$$

where ϵ is an upper bound of the noise level. As a result, the discrete reconstruction problem can be rewritten as finding a minimizer of the following energy formulation

$$\min_u E(u), \quad E(u) = \lambda \cdot \|\nabla u\|_0 \tag{3.7a}$$

$$\text{s.t.} \quad \underline{b}(\epsilon) \leq Au \leq \overline{b}(\epsilon) \quad \wedge \quad u_i \in \mathcal{F}_*, \quad \forall\, i \in [n]. \tag{3.7b}$$

We refer to problem (3.7a) as a *multilabel problem with indirect measurements* with Potts regularization. The regularization factor $\lambda \geq 0$ is acting as a trade-off between how much we trust the observations (data term) and how much we rely on the model assumption (regularization term).

From the viewpoint of graphical models, the system of affine inequalities induces (very) high-order potentials, a possibly fully connected graph structure. These high-order interactions induced by the non-local constraints result in a non-standard labeling problem which becomes intractable for discrete approaches and larger problem sizes. Therefore, in the next section, we adopt the strategy of solving a sequence of convex relaxations in order to minimize an overall non-convex variational energy, which properly approximates the original problem.

3.4. Non-Convex Variational Approximation

In this section we develop a non-smooth and non-convex variational approximation of the combinatorial multilabel problem with indirect measurements (3.7a). The energy formulation (3.7a) is highly non-convex due to the l_0-norm regularization. In addition, the condition that the solution is restricted to a discrete set of admissible gray-values, constitutes a non-convex constraint. Therefore, our starting point are the established

convex relaxation techniques of the multilabeling problem [ZGFN08; LBS09; PCCB09] with a given direct observation u^0. These relaxations are lifting the discrete label decisions to a higher dimensional space, such that each label corresponds to a unit-vector, which allows to consider the convex hull of these vertices. Consequently, the labeling problem is encoded by probability vectors, which form a convex set. In particular we adapt the most straightforward convex relaxation by Zach et al. [ZGFN08], which is based on layer-wise total variation (TV). Minimizing the energy in (3.8a) below with respect to W over the set of relaxed indicator vectors (3.8b) assigns to each given pixel in u^0 a label from the set $\mathcal{F}_* := \{f_1^*, \dots, f_c^*\}$ in a probabilistic way. The discretized total variation, weighted by λ, and the simplex constraints in the definition of $\overline{\mathcal{W}}$ constitute a basic convex relaxation of the integrality constraints with respect to W, which reads

$$E(W, u^0) = \sum_{i=1}^{n} \sum_{j=1}^{c} W_{i,j} \left(u_i^0 - f_j^*\right)^2 + \lambda \sum_{j=1}^{c} \|\nabla W_j\|_1 \tag{3.8a}$$

$$\text{s.t.} \quad W \in \overline{\mathcal{W}} = \left\{ W \in [0,1]^{n \times c} : \sum_{j=1}^{c} W_{i,j} = 1, \ \forall\, i \in [n] \right\}. \tag{3.8b}$$

Next, we add the tomographic projection constraints $\underline{b} \leq Au \leq \overline{b}$ to the relaxed energy (3.8a) by transforming the indicator variables W back to their corresponding intensities with the linear *assignment operator*

$$\mathrm{P}_{\mathcal{F}_*} : \overline{\mathcal{W}} \to \mathrm{conv}(\mathcal{F}_*^n), \quad W \mapsto \mathrm{P}_{\mathcal{F}_*}(W) = (I_n \otimes {F^*}^\top) \,\mathrm{vec}(W) = WF^*, \tag{3.9}$$

where $F^* := (f_1^*, \dots, f_c^*)^\top$ is the vector of all admissible gray-values. The assignment operator preserves convexity of the resulting energy

$$\begin{aligned} E(W, u^0) = &\sum_{i=1}^{n} \sum_{j=1}^{c} W_{i,j} \left(u_i^0 - f_j^*\right)^2 + \lambda \sum_{j=1}^{c} \|\nabla W_j\|_1 \\ &+ \delta_{\mathbb{R}_+^m}(A\,\mathrm{P}_{\mathcal{F}_*}(W) - \underline{b}) + \delta_{\mathbb{R}_-^m}(A\,\mathrm{P}_{\mathcal{F}_*}(W) - \overline{b}) + \delta_{\overline{\mathcal{W}}}(W), \end{aligned} \tag{3.10}$$

where the constraints $\underline{b} \leq A\,\mathrm{P}_{\mathcal{F}_*}(W) \leq \overline{b}$ and $W \in \overline{\mathcal{W}}$ are implemented by indicator functions $\delta_{\mathbb{R}_+^m}$ and $\delta_{\mathbb{R}_-^m}$.

In tomography, no image u^0 is given beforehand. However, dropping the unary data term in Eq. (3.10) will result in a very loose relaxation, since the constraints are feasible for all convex combinations of prototypes f_j^*. In other words, the constraints only determine the value of a pixel but do not indicate how the indicator variables should realize this value (similar to estimating a vector given only its magnitude). This is reflected in the fact, that the assignment operator is no longer injective for the non-binary case, after relaxing the unit-vectors to probabilities. Hence we are introducing a lot of additional non-integral solutions. Proposition 3.4.1 states, that the assignment operator is bijective for $c \leq 2$.

Proposition 3.4.1 (Bijectivity of the Assignment Operator)**.** *Let $\mathcal{F}_* = \{f_1^*, \dots, f_c^*\} \subset [0,1]$ with $c = |\mathcal{F}_*|$ and the assignment operator $u = \mathrm{P}_{\mathcal{F}_*}(W)$ in (3.9). Then, the linear mapping $\mathrm{P}_{\mathcal{F}_*}$ is bijective if and only if $c \leq 2$.*

Proof. Without loss of generality we can consider the case of a single pixel, $n = 1$. The convex hull $\mathrm{conv}(\mathcal{F}_*)$ is an affine (c-1) simplex if the set of vertices $\mathcal{F}_*$ is affinely independent. Since the vertices $f_c^* \in \mathbb{R}$ are scalars, only $c \leq 2$ are affinely independent. Then, the canonical mapping $\mathrm{P}_{\mathcal{F}_*}$ from the standard simplex to the affine simplex is bijective

since they are homeomorphic if and only if they have the same number of vertices. □

This also explains why the binary case is substantially easier than the non-binary case, since the data term is unambiguous and the convex relaxation is tight. However, despite the drawback of introducing artificial solutions, we still want to employ the linear mixing model $u = \mathrm{P}_{\mathcal{F}_*}(W)$. In order to compensate for these unwanted effects, we have to exclude non-integral solutions in order to stabilize the inversion of the assignment operator in the non-binary case. Consequently, a unary term (local decisions) is required to provide information about how the represented gray-values are to be realized.

A straightforward approach would be to start with some initial reconstruction u^0, e.g. computed using another reconstruction method, followed by iteratively applying the described approach. From a mathematical perspective, this procedure amounts to the following fixed point iteration

$$W^{(k+1)} = \operatorname*{arg\,min}_{W} E\big(W, \mathrm{P}_{\mathcal{F}_*}(W^{(k)})\big), \tag{3.11}$$

where at every iteration a convex problem has to be solved whose solution updates the unary data term. Therefore, in each iteration we are simultaneously reconstructing from the tomographic constraints and labeling the previous solution. This immediately raises the question whether the iteration converges and which overall energy is actually optimized?

To address these questions, we first eliminate u^0 in a principled way. Note that $E(W, u^0)$ in (3.10) is differentiable with respect to the second argument. Hence we invoke Fermat's (first order) optimality condition, which reads

$$\nabla_u E(W, u) = 0 \quad \Leftrightarrow \quad u^* = \mathrm{P}_{\mathcal{F}_*}(W). \tag{3.12}$$

This means that the optimal u^* must be equal to the weighted average of the labels f_j^* which is exactly the assignment operator. Substituting the optimality condition (3.12) back into the energy (3.10) results in the final version of the proposed energy which only depends on W,

$$\begin{aligned} E(W) = &\sum_{i=1}^{n}\sum_{j=1}^{c} W_{i,j}\left((\mathrm{P}_{\mathcal{F}_*}(W))_i - f_j^*\right)^2 + \lambda \sum_{j=1}^{c} \|\nabla W_j\|_1 \\ &+ \delta_{\mathbb{R}^m_+}(A\,\mathrm{P}_{\mathcal{F}_*}(W) - \underline{b}) + \delta_{\mathbb{R}^m_-}(A\,\mathrm{P}_{\mathcal{F}_*}(W) - \overline{b}) + \delta_{\overline{\mathcal{W}}}(W). \end{aligned} \tag{3.13}$$

This new energy formulation, Eq. (3.13), is clearly non-convex because of the products in the first term which measures the discreteness of W. Therefore, we call this term *discretization term* and define it by

$$D(W) := \sum_{i=1}^{n}\sum_{j=1}^{c} W_{i,j}\left((\mathrm{P}_{\mathcal{F}_*}(W))_i - f_j^*\right)^2, \tag{3.14}$$

where the i-th summand of (3.14) reads

$$\sum_{j=1}^{c} W_{i,j}\Big(\sum_{l=1}^{c} W_{i,l} f_l^* - f_j^*\Big)^2 = \sum_{j=1}^{c} W_{i,j}(f_j^*)^2 - \Big(\sum_{j=1}^{c} W_{i,j} f_j^*\Big)^2 \tag{3.15}$$

which is concave with respect to the vector W_i. Consequently, the term $D(W)$ given by (3.14) is concave as well. Figure 3.1 shows a plot of the discretization term $D(W)$ for one pixel (n=1) and the discrete set of values $\mathcal{F}^* = \{0.0, 0.4, 1.0\}$. We can see, that the minima are attained at vertices of the simplex which in turn correspond to unit-vectors. The

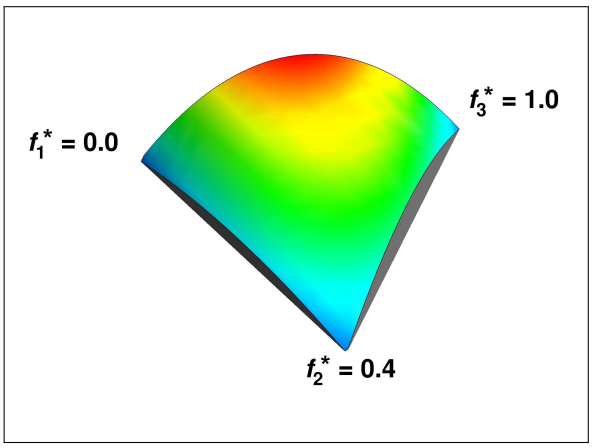

Figure 3.1.: Visualization of the discretization term $D(W)$ for the one-pixel case ($n = 1$), over the probability simplex Δ_c, the vertices correspond to the values $\mathcal{F}^* = \{0.0, 0.4, 1.0\}$. We can see that, the minima are attained at the vertices of the simplex which correspond to unit-vectors. Moreover, the transition between the lowest label $f_1^* = 0.0$ and the highest label $f_3^* = 1.0$ is penalized the most, which is necessary to avoid convex combinations.

proposed discretization term in (3.14) is compensating unwanted effects resulting from the weak data term. Consequently, convex combinations of indicator variables yield a higher energy and hence non-integral solution are excluded (for λ sufficently small). Therefore, the inversion of the assignment operator in (3.9) is regularized. This concave discretization term is acting like a unary data term with pixelwise independent decisions in the setting, where the input image is directly available. Weber [Web09, Chapter 6] proposed this term for discrete tomography which arises here in a natural way, whereas his overall approach differs with respect to the data term for the projection constraints, regularization and optimization. Finally, we have derived a non-convex variational approximation of the original combinatorial problem (3.7a). In the next section, we exploit the special structure of the objective function (3.13) to relate and justify the fixed point iteration (3.11) as a sound and efficient optimization strategy.

3.5. Sequential Convex Optimization

In this section, we aim to find a minimizer of the non-convex energy (3.13) for joint reconstruction and labeling, which was derived in the previous section. The lack of convexity makes finding a global optima in a reasonable time challenging. Hence, we focus on local optimality instead which still is non-trivial. To this end, we construct an iterative algorithm which converges to a critical point. In order to get a highly adapted algorithm, which exploits the special problem structure, we reformulate the objective function (3.13) as a difference of convex functions (DC) program [PE86]. Finally, we justify the straightforward fixed point iteration (3.11), which solves a sequence of convex problems, as a sound and efficient optimization algorithm.

3.5.1. DC Programming

A large subclass of non-convex optimization problems can be written as a difference of convex functions (DC). Such DC programs [PE86] can be solved very efficently by the DC algorithm (DCA). Both together constitute the backbone of non-convex optimization and global optimization techniques. Accordingly, basic concepts of convex optimization like duality and KarushKuhnTucker (KKT) conditions were extended to DC functions [Tol78]. The generic form of a DC program is given by

$$W^* = \arg\min_W g(W) - h(W), \tag{3.16}$$

where $g(W)$ and $h(W)$ are assumed to be proper, lower semicontinuous, convex functions. Based on local optimality conditions and DC duality there exists a simplified version of the DC algorithm [PH97] for minimizing (3.16). This algorithm guarantees convergence to a critical point by starting with $W^0 \in \mathrm{dom(g)}$ and then alternatingly applying the updates

$$V^{(k)} \in \partial h(W^{(k)}) \qquad \text{and} \qquad W^{(k+1)} \in \partial g^*(V^{(k)}) \tag{3.17}$$

until a termination criterion is reached, where g^* denotes the Fenchel conjugate (Section 2.1.2) of g. The DCA is summarized in Algorithm 3.1. In order to apply the DC algorithm

Algorithm 3.1: Generic DCA

Init: choose any $W^0 \in \mathrm{dom(g)}$

while *not converged* **do**

$$V^{(k)} \in \partial h(W^{(k)}) \tag{3.18a}$$

$$W^{(k+1)} \in \partial g^*(V^{(k)}) \tag{3.18b}$$

$k \leftarrow k + 1$

Output: $W^* = W^{(k)}$

to our non-convex energy $E(W)$, defined in Eq. (3.13), we rewrite the energy as a DC function $E(W) = g(W) - h(W)$. It is worth mentioning that DCA only depends on the DC components g and h but not on the objective function itself. Moreover, DC decompositions are not unique and theoretically a DC function has infinitely many decompositions, each leading to different instance of DCA. Hence, choosing an appropriate DC decomposition will have a crucial impact on the speed of convergence, robustness and efficiency of the DC algorithm. It is still an open question how the optimal DC decomposition for a given DC program can be found systematically. Of course, this heavily depends on the problem specific structure and in order to handle large-scale settings it is important that the sub-problems can be explicitly evaluated or at least solved in an inexpensive iterative way.

A reasonable and natural decomposition of the energy (3.13) is to set $h(W) = -D(W)$ since the discretization term (3.14) is concave by (3.15). We denote by $g(W)$ the remaining convex terms from Eq. (3.13).

The second step (3.18b) of Algorithm 3.1 is in general not very applicable in practice since the Fenchel conjugate of g may not admit a closed form expression. In order to make the step (3.18b) explicit, we apply the subgradient inversion rule (Lemma 2.1.4) of convex

analysis to obtain

$$W^{(k+1)} \in \partial g^*(V^{(k)}) \quad \Leftrightarrow \quad V^{(k)} \in \partial g(W^{(k+1)}) \quad \Leftrightarrow \quad 0 \in \partial g(W^{(k+1)}) - V^{(k)} \tag{3.19}$$

which is equivalent to the convex optimization problem

$$W^{(k+1)} = \arg\min_W g(W) - \langle V^{(k)}, W \rangle. \tag{3.20}$$

Furthermore, we rewrite (3.20) by strategically adding terms independent from W, which do not affect the minimizer and get

$$W^{(k+1)} = \arg\min_W g(W) - \left(h(W^{(k)}) + \langle V^{(k)}, W - W^{(k)} \rangle \right). \tag{3.21}$$

This immediately reveals the main idea and obvious interpretation of DCA, to approximate the DC program by a sequence of convex programs. At each iteration k of DCA the concave part is linearly majorized (by a first-order Taylor expansion if h is differentiable) around the current iterate $W^{(k)}$.

Since our defined h is differentiable, the first step of (3.17) reads

$$V^{(k)} \in \partial h(W^{(k)}) \quad \Leftrightarrow \quad V^{(k)} = -\nabla D(W^{(k)}), \tag{3.22}$$

where the Euclidean gradient of D is given by Proposition 3.5.1.

Proposition 3.5.1 (Gradient of Discretization Term)**.** *Let $D(W)$ be defined as*

$$D(W) = \sum_{i=1}^{n} \sum_{j=1}^{c} W_{i,j} \big((\mathrm{P}_{\mathcal{F}_*}(W))_i - f_j^* \big)^2. \tag{3.23}$$

Then the component $\big(\nabla D(W)\big)_{k,l}$ of the gradient of D at W with respect to W for pixel $k \in [n]$ and label $l \in [c]$ is given by

$$(\nabla D(W))_{k,l} = \frac{\partial D(W)}{\partial W_{k,l}} = \big((\mathrm{P}_{\mathcal{F}_*}(W))_k - f_l^* \big)^2. \tag{3.24}$$

Proof. Using $\mathrm{P}_{\mathcal{F}_*}(W)_k = \sum_{j=1}^{c} W_{k,j} f_j^*$, we compute

$$\frac{\partial D(W)}{\partial W_{k,l}} = \big((\mathrm{P}_{\mathcal{F}_*}(W))_k - f_l^* \big)^2 + 2 \sum_{i=1}^{n} \sum_{j=1}^{c} W_{i,j} \big((\mathrm{P}_{\mathcal{F}_*}(W))_i - f_j^* \big) \frac{\partial}{\partial W_{k,l}} \Big(\sum_m W_{i,m} f_m^* - f_j^* \Big) \tag{3.25a}$$

$$= \big((\mathrm{P}_{\mathcal{F}_*}(W))_k - f_l^* \big)^2 + 2 f_l^* \sum_{j=1}^{c} W_{k,j} \big((\mathrm{P}_{\mathcal{F}_*}(W))_k - f^* \big) \tag{3.25b}$$

$$= \big((\mathrm{P}_{\mathcal{F}_*}(W))_k - f_l^* \big)^2 + 2 f_l^* \Big((\mathrm{P}_{\mathcal{F}_*}(W))_k \sum_{j=1}^{c} W_{k,j} - \sum_{j=1}^{c} W_{k,j} f_j^* \Big) \tag{3.25c}$$

$$= \big((\mathrm{P}_{\mathcal{F}_*}(W))_k - f_l^* \big)^2 + 2 f_l^* \big((\mathrm{P}_{\mathcal{F}_*}(W))_k \cdot 1 - (\mathrm{P}_{\mathcal{F}_*}(W))_k \big) \tag{3.25d}$$

$$= \big((\mathrm{P}_{\mathcal{F}_*}(W))_k - f_l^* \big)^2 \tag{3.25e}$$

□

Combining equations (3.22) and (3.24) and inserting into equation (3.20) yields

$$W^{(k+1)} = \arg\min_W E\left(W, \mathrm{P}_{\mathcal{F}_*}(W^{(k)}) \right) \tag{3.26a}$$

$$= \arg\min_W g(W) + \sum_{i=1}^{n} \sum_{j=1}^{c} W_{i,j} \left((\mathrm{P}_{\mathcal{F}_*}(W^{(k)}))_i - f_j^* \right)^2. \tag{3.26b}$$

We notice that the DC algorithm, summarized as Algorithm 3.2 below, agrees with the iteration (3.11), and hence *proves its convergence* to a stationary point of the non-convex energy (3.13). This answers the second part of the question raised in the previous section. We showed the connection between iteratively applying (3.11) (fixed point iteration) and

Algorithm 3.2: DC Fixed Point Algorithm

Init: choose $W^{(0)} = \mathbb{1}_{\mathcal{W}}$

while *not converged* **do**

$$W^{(k+1)} = \arg\min_{W} E\left(W, \mathrm{P}_{\mathcal{F}_*}(W^{(k)})\right) \tag{3.27}$$

$k \leftarrow k + 1$

Output: $W^* = W^{(k)}$

the overall non-convex energy (3.13). Finally, we apply the primal-dual (PD) algorithm proposed by [CP11] to solve the convex subproblems (3.27). We use the barycenter as an unbiased initialization, which is natural for variables defined on the probability simplex.

3.5.2. Rounding Step

Recall that the discretization term $D(W)$ of (3.13) only steers the solution to the finite set of feasible values $\mathcal{F}_*$, and the integrality constraints are relaxed. As a consequence, for vanishing regularization parameter λ, the minimizer W will correspond to indicator vectors W_i that assign a unique label to each pixel i. For larger values of λ which are more common in practice, the minimizing vectors W_i will generally not be integral. Therefore, a post-processing step for rounding the solution is required.

We assume that Algorithm 3.2 returns a minimizer W^* of (3.13) after termination and we propose to select a label for each pixel i by solving the local problems

$$\hat{u}_i^* = \arg\min_{f^* \in \mathcal{F}_*} |(\mathrm{P}_{\mathcal{F}_*}(W^*))_i - f^*|, \qquad i \in [n], \tag{3.28}$$

as a post-processing step. Note that this method differs from the common rounding procedure of multilabeling approaches which select the label f_j^* if $W_{i,j} = \max\{W_{i,1}, \ldots, W_{i,c}\}$.

3.6. Experiments

In this section, we illustrate and compare our proposed method to state-of-the-art approaches which can handle non-binary tomography reconstruction from only a few projection angles. Specifically, we considered the Discrete Algebraic Reconstruction Technique (*DART*) [BS11] and the energy minimization method from Varga et al. [VBN12] (*Varga*). Finally, we denote our proposed joint reconstruction and labeling approach (*LayerTV*), since it is based on the layer-wise total variation, which is directly applied to the indicator variables. Table 3.1 summarizes all considered approaches.

Implementation details. We implemented the proposed approach (LayerTV) described in Section 3.3 with the DC fixed point algorithm 3.2 which minimizes the non-convex energy (3.13) by a sequence of convex problems. Each subproblem of Algorithm 3.2 was approximately solved using the primal-dual (PD) algorithm [CP11] limited to 1000

Shortcut	Reference	Regularization	Implementation
LayerTV (proposed)	Eq. (3.13)	anisotropic TV	ours (MATLAB), Alg. 3.2
DART	[BS11]	Gaussian filter	ASTRA-toolbox [APB+15]
Varga	[VBN12]	Sobolev semi-norm	ours (MATLAB)

Table 3.1.: Overview of the approaches used for a competitive evaluation.

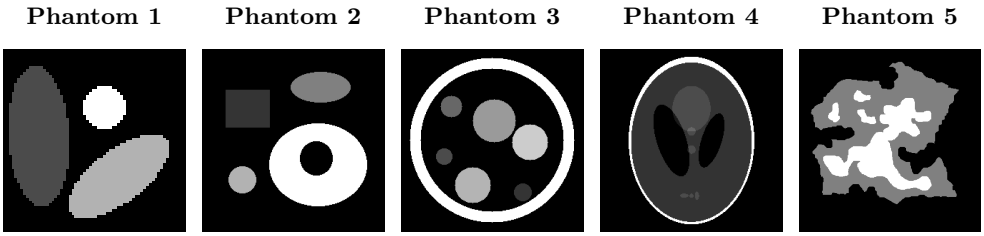

Figure 3.2.: The 5 different phantoms used for the numerical evaluation.

iterations or until the primal-dual gap drops below 0.1. The outer iteration was terminated if the change of the energy between two subsequent iterations, normalized by the number of pixels, was smaller than 10^{-5} in the noiseless case and 10^{-4} in the noisy case. Additionally, we limited the number of outer iterations to 20. The discrete total variation in (3.13) was implemented by its anisotropic version. That means a 4-connected neighborhood system was used and the fully separable l_1-norm was applied to the components of the gradient at each pixel.

For DART we used the publicly available implementation included in the ASTRA-toolbox [APB+15] and for the method of Varga [VBN12] we used our own implementation in MATLAB since no public code was available. We tried to use the default parameters of the competing approaches as proposed by their authors. However, since the test-datasets differ in size, we slightly adjusted the parameters in order to get best results for every algorithm and problem instance. Therefore, the parameters were adjusted for each test-dataset but kept constant over different projection angles.

Data Setup. We adapted the binary phantoms from Weber et al. [WNS+06] to multivalued test-datasets with more labels, shown as phantom 1,2 and 3 by Figure 3.2. Phantom 5 in Figure 3.2 was taken from [BS11] and phantom 4 is the well-known Shepp-Logan phantom [SL74]. We also created noisy scenarios by applying Poisson noise to the measurements b with a signal-to-noise ratio of $SNR = 20$ db. The geometric setup of the projection matrices was created by the ASTRA-toolbox [APB+15], where we used parallel beam projections along equidistant angles between 0 and 180 degrees. Each entry a_{ij} of the projection matrix A corresponds to the length of the line segment of the i-th projection ray passing through the j-th pixel in the image domain. The width of the sensor-array was set to 1.5 times the image size, so that every pixel intersects with a least a single projection ray and each sensor has the size of a pixel.

Performance measure. For the evaluation we measured the *relative pixel error*, that

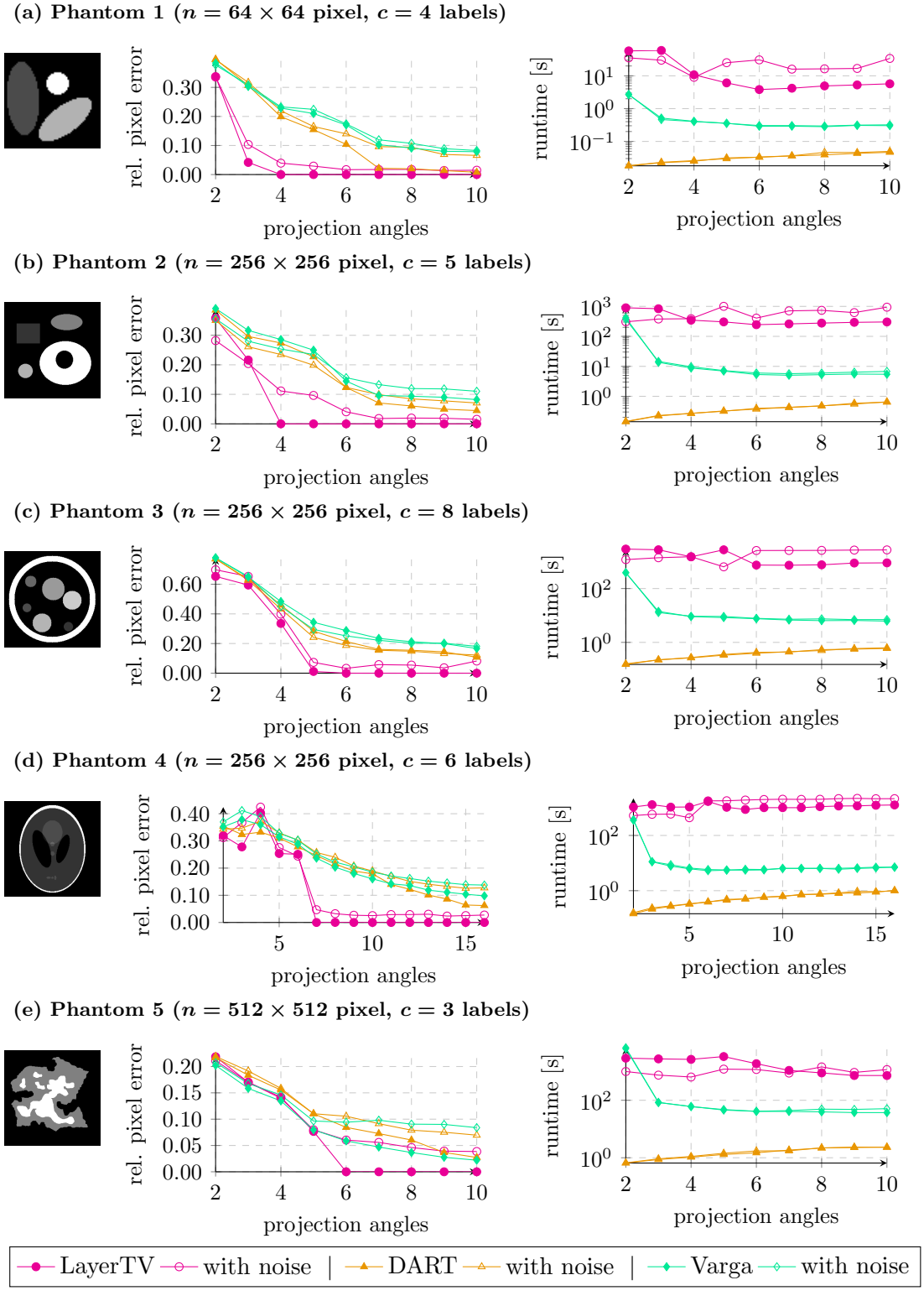

Figure 3.3.: Numerical evaluation of the approaches for the different test-datasets and increasing (but small) numbers of projections, in the noiseless case **(filled markers)** and in the noisy case **(non-filled markers)**, with noise level $SNR = 20$ db. The relative pixel error is shown. The proposed approach (LayerTV) gives perfect reconstructions with the smallest number of projection angles in the noiseless case and also returns high-quality reconstructions in the presence of noise, compared to the other approaches.

is the relative number of erroneously reconstructed pixels compared to the ground truth u^*

$$Err_{\text{pxl}}(u) := \frac{1}{n}\sum_{i}^{n} d(u_i, u_i^*); \quad d(x,y) := \begin{cases} 0 & \text{if } x = y, \\ 1 & \text{if } x \neq y. \end{cases} \tag{3.29}$$

This measure was used for the discrete solutions of the algorithms (after rounding).

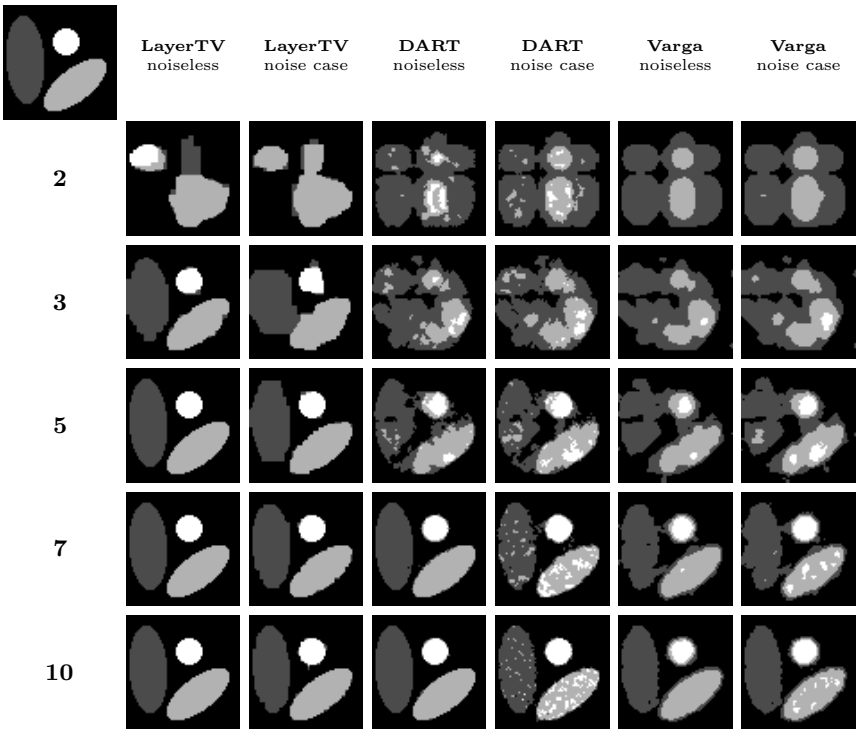

Figure 3.4.: Visual results of experiment phantom 1.

Results. Figure 3.3 shows all results of the numerical evaluation. For each of the test-datasets (phantoms 1 - 5), the left plot displays the relative pixel error (after rounding) for increasing numbers of projection angles. On the right, the corresponding runtime is shown as a log-scaled plot. For each algorithm two curves are drawn: filled markers correspond to the noiseless case and non-filled markers correspond to the noisy case.

The results show that the proposed approach returns a perfect reconstruction with the least number of projection angles in the noiseless case among all approaches. Similar quality improvements are visible for more challenging noisy scenarios by the proposed algorithm. In the noiseless case, phantom 1 can be almost perfectly reconstructed from only 3 projection angles and fully from 4 by the proposed approach whereas DART needs 7 projections to achieve similar reconstruction quality and the method of Varga needs at least 7 projections for a reasonable reconstruction. These visual differences can be seen in Figure 3.4. The relative performance of all approaches is similar overall for phantoms 2, 3 and 5, with the exception that the approach of Varga performed better than DART in the noiseless case of phantom 5. Figure 3.5 shows the results for phantom 4, where our approach is able to perfectly reconstruct from merely 7 projections in the noiseless case and returns a good piecewise-constant result in the noisy case.

Regarding the runtime (right plots from Figure 3.3), DART is the fastest followed by

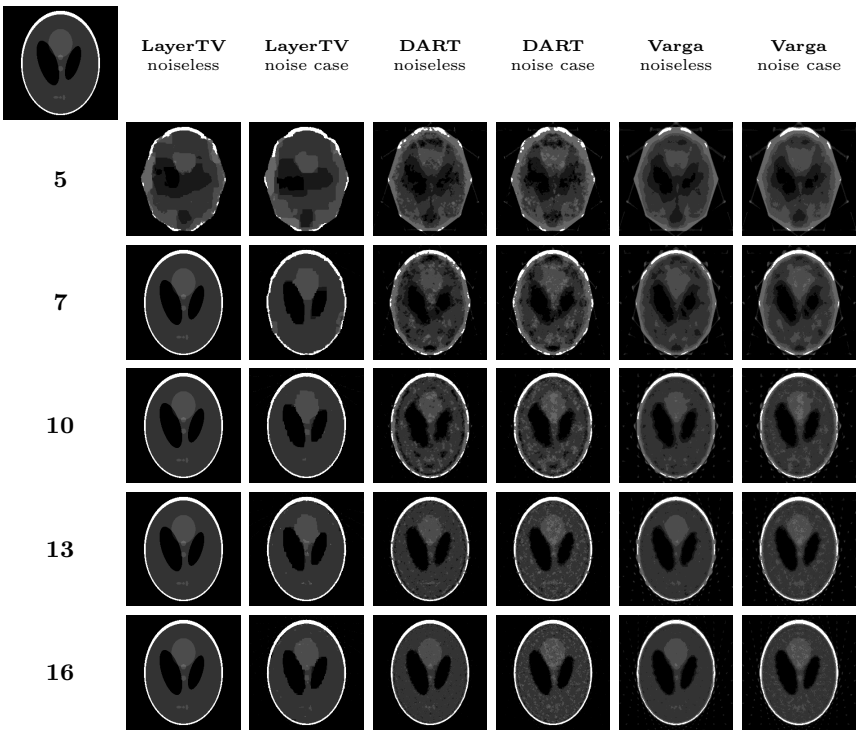

Figure 3.5.: Visual results of experiment phantom 4.

Varga. The proposed approach clearly requires more runtime to return more accurate solutions. However, in settings where runtime is crucial, the proposed approach could easily be parallelized and implemented e.g. in CUDA to run on modern graphics cards.

3.7. Conclusion

In this chapter we have presented a novel non-convex and non-smooth variational approach for solving the discrete tomography reconstruction problem in the general non-binary case. We have combined an established convex relaxation of the multilabeling Potts problem with the non-local tomographic projection constraints. The feasible set of labels is properly taken into account during the reconstruction process by a non-convex discretization term. This term naturally emerges when the function to be reconstructed is represented as a convex combination of gray-values. Moreover, non-integral solutions, introduced by lifting the projection constraints to indicator variables were successfully excluded. The proposed energy for joint reconstruction and labeling was reliably and efficiently minimized by the DC algorithm which provably converges to a stationary point. The resulting algorithm recovers the naive fixed-point iteration, which amounts to solving a sequence of convex optimization problems. The reconstruction performance turned out to be superior to the state of the art.

In the context of this work, an immediate question arises, if the supervised joint reconstruction and labeling approach can be extended to the unsupervised case. For direct observations, the generalization of the multilabel problem to unsupervised scenarios is

investigated in the last chapter of this work, within the smooth and geometric framework of the assignment flow. However, the unsupervised setting with indirect measurements, where neither the prototypes are given beforehand nor the data to be labeled is directly available, poses a challenging task to be considered as future work.

A different direction of future work can be to explore the connection of he proposed discretization term (3.14) to spatially continuous variational formulations, which may indicate potential for further improvement.

Further future work can include improvement of the runtime and focus on theoretical aspects of this approach.

CHAPTER 4

Geometric Multilabel Methods for Tomography Reconstruction

4.1. Introduction and Overview

In the previous chapter we have introduced a non-convex variational approximation to solve a challenging generalization of the multilabel problem. Thereby, the input image to be labeled is not available directly, however indirect measurements related by a linear operator are available instead. One key application of this setting is *discrete tomography* [HK99], which introduces the problem of reconstructing *piecewise constant* functions from projection data, that are taken from few projection angles only. Such extremely ill-posed inverse problems are motivated by industrial applications, like quality inspection [HFU08]. Regularization of such problems essentially rests upon the fact that the functions to be reconstructed only take values in a finite set of labels

$$\mathcal{F}_* := \{f_1^*, \dots, f_c^*\} \subset [0,1]. \tag{4.1}$$

This is similar to the common image labeling problem in computer vision, with the essential difference that the image to be labeled is only *indirectly* observed. Specifically, after discretization the resulting representation $u \in \mathbb{R}^n$ and projection data b given by

$$Au = b \quad \text{s.t.} \quad u_i \in \mathcal{F}_*, \quad \forall\, i \in [n] \tag{4.2}$$

are observed, where the projection matrix A is known but underdetermined (fewer rows than columns). Hence, the task is to simultaneously reconstruct an inverse problem and select for each pixel of the solution one of the labels $\mathcal{F}_*$. This can be considered as a *multilabel problem with indirect measurements* which is consequently considerably more difficult to solve. To avoid the combinatorial nature of such discrete optimization problems, it is common to encode and relax hard label decisions to discrete probability distributions.

In this chapter we focus on *geometric* aspects of the multilabel problem in contrast to Chapter 3. The space of discrete probability distributions over the set of labels for each pixel constitutes a smooth manifold, the assignment manifold, with the smooth geometry induced by the Fisher-Rao metric [BR82] on a single simplex. Moreover, *information geometry* [Kas89; AC10] bridges geometry with convex analysis by introducing Bregman

divergences, distance like functions, which locally respect the underlying geometry. This allows to generalize established convex optimization algorithms to the geometric setting. The assignment flow (2.89) is based on these smooth geometric concepts to perform image labeling with direct observations. Therefore, in this chapter we investigate how to adapt and generalize the geometric building blocks of the assignment flow to handle indirect measurements as input data. In particular we focus on the following aspects:

- geometric spatial regularization of assignments,
- geometric numerical algorithms by Bregman proximal optimization and
- extension of the geometric setting itself to include the projection constraints.

Finally, we propose two different geometric approaches to joint reconstruction and labeling.

The first part of this chapter is concerned with adapting the approach from Chapter 3 to the geometric setting of the assignment manifold. We propose a non-convex variational approximation to the combinatorial multilabel problem with indirect measurements in which the individual building blocks of the energy are motivated by geometric considerations. Therefore, the regularization term measures the spatial coherence with respect to the KL-divergence between discrete probability distributions of neighboring pixels. This allows to incorporate the underlying geometry of relaxed label decisions into the regularization. In addition, we add an entropy term for uniform rounding towards more discrete labelings and hence we only require a basic rounding procedure.

The proposed energy admits the decomposition into a convex part and a concave part. Therefore, we can apply the difference of convex functions (DC) algorithm, which amounts to solving a sequence of convex optimization problems. Moreover, the optimization algorithm to minimize each convex subproblem exploits the underlying geometry by using the KL-divergence as a proximity measure for generalized proximal mappings. The resulting iterated primal-dual algorithm is capable of solving large-scale scenarios efficiently. Finally, we demonstrate the proposed approach in a comprehensive numerical evaluation and compare it with state-of-the-art approaches from literature. We focus on the key application of non-binary discrete tomography as well as deblurring and denoising with joint labeling.

The second part of this chapter is concerned with the geometric setting itself. Here we are going one step further and restrict the assignment manifold to a Riemannian submanifold

$$\mathcal{W}_{\text{tomo}} \subset \mathcal{W}, \tag{4.3}$$

satisfying the affine projection constraints. Therefore, the feasible set $\mathcal{W}_{\text{tomo}}$ is equipped with a Hessian Riemannian metric which naturally extends the Fisher-Rao metric of the assignment manifold. Subsequently, we construct a smooth gradient flow for any smooth objective function restricted to the submanifold. In addition, we define a corresponding objective function for joint tomographic reconstruction and label assignment.

We derive an iterated implicit scheme to numerically integrate the resulting tomographic assignment flow which relies on the Bregman proximal point method. Each convex subproblem is solved approximately by the generalized primal-dual algorithm with KL-divergences as proximity measure. This also allows to efficiently handle large-scale scenarios. Finally, we demonstrate the proposed approach in a comprehensive numerical evaluation and compare it with state-of-the-art approaches from literature.

Organization. The remainder of this chapter is organized as follows:

- We briefly review the related work on binary and non-binary discrete tomography (Sect. 4.2).
- We propose a non-convex variational approximation to the combinatorial multilabel problem with indirect measurements. The spatial regularization is implemented by considering KL-divergences between pairs of pixels (Sect. 4.3.1).
- We derive an optimization algorithm based on the DC framework to minimize the variational non-convex energy. Generalized Bregman proximal maps allow to solve the convex subproblems in a geometric way (Sect. 4.3.2).
- We suggest a converging fixed-point iteration to evaluate the proximal mapping of the geometric regularizer efficiently (Prop. 4.3.1)
- We propose the tomographic assignment flow as a Riemannian gradient flow on a submanifold by including the projection constraints and extending the Fisher-Rao metric in a natural way (Sect. 4.4.1).
- We motivate a suitable objective function to simultaneously perform tomographic reconstruction and labeling (Sect. 4.4.2).
- We derive an iterated implicit scheme to integrate the tomographic assignment flow which relies on the Bregman proximal point method (Sect. 4.4.3).
- Finally, we numerically evaluate and compare the proposed approaches with state-of-the-art methods for non-binary tomography reconstruction on standard test-datasets with only a few projection angles (Sect. 4.3.3 and 4.4.4)

4.2. Related Work

A natural class of approaches is based on minimizing *convex sparsifying functionals* depending on u (e.g. total variation) subject to the affine subject constraints (3.4), but *without* the labeling constraints [SP08; GBB+12; DPSS14]. Unless sufficient conditions for unique recovery are met, in terms of the number of projection measurements relative to the complexity of the discontinuity set of u [DPSS14], the performance of the necessary rounding post-processing step is difficult to control. Likewise, a binary discrete graphical model was adopted by [KPSZ15], and a sequence of s-t graph-cuts was solved to take into account the affine projection constraints. An extension to the non-binary case (multiple labels) seems to be involved. The authors of [SWFU15] minimize the ℓ_0-norm of the gradient directly by a dynamic programming approach, but do not exploit the set $\mathcal{F}_*$ of feasible labels for regularization.

A first step towards incorporating the prior information of discrete labels was done in our preliminary work [ZKS+16], where we combined a total variation formulation for reconstruction with a non-convex coupling term to enforce discrete labels of the solution. Other approaches that aim to enforce the labeling constraints by *continuous non-convex* optimization include [WSH03; SSWH05]. In this work, the approach proposed in Chapter 3 also constitutes a non-convex variational formulation.

Unlike our two geometric approaches proposed below, which limit the degrees of freedom by restricting the feasible set to a Riemannian manifold, these approaches work in the higher-dimensional ambient Euclidean space and hence are more susceptible to poor initializations and local minima.

Further approaches that define the state of the art include [VBN12; BS11]. The authors of [VBN12] proposed a heuristic algorithm that adaptively combines an energy formulation with a non-convex polynomial representation, in order to steer the reconstruction towards the feasible label set. Batenburg et al. [BS11] proposed the *Discrete Algebraic Reconstruction Technique (DART)* algorithm which starts with a continuous reconstruction by a basic algebraic reconstruction method, followed by a thresholding operation. These steps, interleaved with smoothing, are iteratively repeated to refine the locations of boundaries. This heuristic approach yields good reconstructions in practice, but cannot be characterized by an objective function which is optimized.

We regard [BS11; VBN12; ZPSS16] as state-of-the-art approaches in our experimental comparison.

4.3. Geometric Spatial Regularization by Bregman Divergences

This section is concerned with generalizing the approach from Chapter 3 to the geometric setting of discrete probability distributions. The Fisher-Rao metric [BR82] is induced by the Boltzmann-Shannon entropy which is a Legendre function. On the other hand, this Legendre function also induces the KL-divergence, which locally approximates the squared geodesic distance on the probability simplex equipped with the Fisher-Rao metric [Kas89]. Therefore, we propose a non-convex variational approximation where pairwise interactions of the regularization term are based on the KL-divergence. In this sense, spatial regularization naturally respects the information geometric properties of the assignment manifold $\mathcal{W}$. Moreover, the derived algorithm to solve the variational formulation itself relies on the KL-divergence. For this we use optimization methods which involve generalized proximal mappings with respect to Bregman divergences to incorporate the constraint set in a geometric way.

4.3.1. Variational Energy Formulation

In this section we introduce a non-convex variational energy with geometric regularization to approximate the combinatorial multilabel problem with (possibly) indirect measurements (4.2). In Chapter 3 we already derived a successful strategy for incorporating linear affine constraints into a relaxation based on total variation regularization. Here, we adopt the same strategy, however with a different regularization, which respects the underlying smooth geometry of discrete probability distributions. Moreover, in contrast to (3.28), an additional entropy term is already enforcing more discrete solutions during optimization and consequently a basic rounding procedure is sufficient.

In the following we assume, that the reconstructed function $u \in \mathcal{F}^n$ is piecewise-constant with pre-specified levels from a feature space $\mathcal{F}$, denoted by

$$u \in \mathcal{F}_*^n, \qquad \mathcal{F}_* := \{f_1^*, \dots, f_c^*\} \subset \mathcal{F}. \tag{4.4}$$

To encode this prior knowledge in terms of labels, we relax the hard assignments of a label

from $\mathcal{F}_*$ to each pixel $i \in [n]$ to the convex set of row-stochastic matrices, given by

$$\overline{\mathcal{W}} = \left\{ W \in [0,1]^{n \times c} : \sum_{j=1}^{c} W_{i,j} = 1, \ \forall\, i \in [n] \right\}, \tag{4.5}$$

such that each row of W_i is a discrete distribution describing the assignment at pixel i. As a generalization of (3.13), we propose the non-convex variational energy formulation

$$\min_{W \in \overline{\mathcal{W}}} E(W), \qquad E(W) = \lambda R_{\mathcal{W}}(W) - \alpha \langle W, \log(W) \rangle + D(W), \tag{4.6}$$

which consists of three basic building blocks detailed below:

(i) regularization for spatial coherence controlled by parameter $\lambda \geq 0$,

(ii) an entropy term for enforcing a unique decision with weight $\alpha \geq 0$

(iii) a data term D for direct or indirect observations

Generally, convex relaxations [LBS09; LS11; CCP12] yield non-integral solutions after optimization and hence a rounding scheme is required. In our case, however, the concave entropy term promotes integral solutions during optimization.

Data Term. We consider two cases of the data term, the *separable* and *non-separable*. Consequently, we implement the data term $D(W)$ in (4.6) with two different flavors:

- The separable case refers to the setting in which noisy image data u^0 *can* be directly observed. This is the standard scenario of the multilabel problem, which allows to define a data term based on local decisions. Hence, a distance function $d_{\mathcal{F}}(\cdot, \cdot)$ measures the similarity to the priors $f_j^* \in \mathcal{F}_*$ for $j \in [c]$. We define the data term

  $$D_{\text{unary}}(W, D_{\mathcal{F}}) = \langle W, D_{\mathcal{F}} \rangle \quad \text{with} \quad D_{\mathcal{F};i,j} = d_{\mathcal{F}}(u_i^0, f_j^*) \tag{4.7}$$

  for separable labeling problems which measures the correlation between assignments and the distance matrix.

- The non-separable case refers to problems where image data u^0 *cannot* be directly observed since it is the solution of the inverse problem (4.2). In (3.9) we have introduced the assignment operator

  $$u = \mathrm{P}_{\mathcal{F}_*}(W) = WF^* \quad \text{with} \quad F^* := (f_1^*, \ldots, f_c^*)^\top, \tag{4.8}$$

  which assigns to each pixel i a convex combination of labels in terms of the distribution W_i. As a consequence, we instead minimize the distance $d(A\,\mathrm{P}_{\mathcal{F}_*}(W), b)$ in (4.2) between the forward projection by A of the back-projected assignments and the given measurements b. In Chapter 3.4 it was shown that, if the data term $D(\mathrm{P}_{\mathcal{F}_*}(W))$ only depends on the assigned solution $\mathrm{P}_{\mathcal{F}_*}(W)$, will result in a weak relaxation. Therefore, the introduction of a concave discretization term (3.14) is required. The generalization of the discretization term to a general feature space $\mathcal{F}$ with a distance function $d_{\mathcal{F}}(\cdot, \cdot)$ is straightforward. Accordingly, we define the data term

  $$D_{\text{inverse}}(W, A, b) = d(A\,\mathrm{P}_{\mathcal{F}_*}(W), b) + \sum_{i=1}^{n} \sum_{j=1}^{c} W_{i,j} d_{\mathcal{F}} \left((\mathrm{P}_{\mathcal{F}_*}(W))_i, f_j^* \right), \tag{4.9}$$

  for non-separable labeling problems with indirect measurements.

Note, when the discretization term in (4.9) is constrained to the simplex ($n = 1$), the vertices of the simplex are its minima. In this case, the entropy term in (4.6) has the same minimizers as the discretization term, which enforces integral solutions. However, the discretization term in (4.9) constitutes a meaningful descent direction w.r.t. to the labels (pushing the assignment $\mathrm{P}_{\mathcal{F}_*}(W)$ towards the label values f_j^*). Furthermore, linearization of the discretization term using the squared Euclidean distance at a point W^0 resembles D_{unary} with $u^0 = \mathrm{P}_{\mathcal{L}}(W^0)$, see Chapter 3.4 for details.

Geometric Regularizer. To enforce spatial coherence over pixelwise probability distributions Zach et. al. [ZGFN08] regularize each individual layer W^j by total variation in order to get a convex relaxation of the Potts model. In Chapter 3.4, we employed this approach to derive the non-convex variational approximation (3.13). A tighter relaxation is obtained by coupling the regularization across layers as in [CCP12; LBS09]. However, these works employ Euclidean norms that disregard the underlying geometry of the discrete probability distributions and hence necessitate re-projection onto the simplex.

Instead, we propose a regularizer $R_{\mathcal{W}}(W)$ (see (4.6)) which respects the underlying geometry of the probability simplex by coupling probability distributions across layers via the KL-divergence. Our regularizer is defined as

$$R_{\mathcal{W}}(W) := \sum_{i=1}^{n} \sum_{k \in \mathcal{N}_i} \frac{1}{|\mathcal{N}_i|} \mathrm{KL}(W_i, W_k), \tag{4.10}$$

which enforces spatial coherence by pairwise interactions in neighborhoods $\mathcal{N}_i$ induced by the underlying grid-graph of the image.

Remark 4.3.1. It is well-known that the KL-divergence locally approximates the *squared* geodesic distance on the probability simplex equipped with the Fisher-Rao metric [Kas89]. In this sense, (4.10) naturally respects the information geometric properties of the underlying manifold. Furthermore, for this particular manifold, our formulation *without* approximation of the quadratic geodesic distance would correspond to a *non-local* extension of a quadratic regularizer in the framework of [WDS14]. Finally, Proposition 4.3.1 states two basic properties of the geometric regularizer.

Proposition 4.3.1 (Basic Properties)**.** *Let $W \in \overline{\mathcal{W}}$ and $R_{\mathcal{W}}(W)$ be defined by (4.10). Then*

(i) $R_{\mathcal{W}}(W)$ is a convex function

(ii) $R_{\mathcal{W}}(W)$ is the KL*-divergence between W_i and the geometric mean of the vectors W_k indexed by $k \in \mathcal{N}(i)$,*

$$R_{\mathcal{W}}(W) = \sum_{i=1}^{n} \mathrm{KL}\left(W_i, \mathrm{gm}(\{W_k\}_{k \in \mathcal{N}_i})\right), \quad \mathrm{gm}(\{W_k\}_{k \in \mathcal{N}_i}) := \prod_{k \in \mathcal{N}_i} W_k^{\frac{1}{|\mathcal{N}_i|}}. \tag{4.11}$$

Proof. Assertion 1. follows from the joint convexity of the KL-divergence [CT06]. The

second claim is implied by

$$R_{\mathcal{W}}(W) = \sum_{i=1}^{n} \sum_{k \in \mathcal{N}_i} \frac{1}{|\mathcal{N}_i|} \langle W_i, \log\Big(\frac{W_i}{W_k}\Big)\rangle = \sum_{i=1}^{n} \langle W_i, \log\Big(\prod_{k \in \mathcal{N}_i} \Big(\frac{W_i}{W_k}\Big)^{\frac{1}{|\mathcal{N}_i|}}\Big)\rangle \tag{4.12a}$$

$$= \sum_{i=1}^{n} \langle W_i, \log\Big(W_i^{|\mathcal{N}_i|\frac{1}{|\mathcal{N}_i|}} \prod_{k \in \mathcal{N}_i} \big(W_k^{\frac{1}{|\mathcal{N}_i|}}\big)^{-1}\Big)\rangle = \sum_{i=1}^{n} \mathrm{KL}\big(W_i, \prod_{k \in \mathcal{N}_i} W_k^{\frac{1}{|\mathcal{N}_i|}}\big). \tag{4.12b}$$

□

In the following section we reformulate the objective function defined in (4.6) as a difference of convex (DC) program and work out a corresponding optimization algorithm which naturally respects the underlying geometry of probability distributions.

4.3.2. Bregman Proximal Optimization

In this section our aim is to derive a numerical algorithm which efficiently finds a minimizer of the variational formulation (4.6). The proposed energy admits the decomposition into a convex part and a concave part which corresponds to the discretization term together with the entropy. This suggests to use the generic difference of convex functions (DC) framework [PE86]. Moreover, we are interested in solving the subproblems efficiently by using the problem specific structure. In particular, the relaxed solution has to be constrained to the convex set $\overline{\mathcal{W}}$ of pixelwise simplex constraints. In contrast to the preliminary version of this approach, introduced in Chapter 3, we incorporate these constraints in a geometric way with help of a suitable Bregman divergence. Based on these divergences, generalized proximal mappings can realize a non-linear and implicit gradient descent, without the requirement to project onto the feasible set at every iteration. This allows to use a generalized version of the Chambolle Pock primal-dual (PD) algorithm [CP16], to solve the sub-problems efficiently even in large-scale scenarios.

DC Programming. A large subclass of non-convex objective functions are DC functions which can (locally) be minimized by DC programming [PE86]. The basic form of a DC program is given by

$$W^* = \arg\min_{W} \; g(W) - h(W), \tag{4.13}$$

where $g(W)$ and $h(W)$ are proper, lower semicontinuous, convex functions. The DC algorithm [PH97] for minimizing (4.13) guarantees convergence to a critical point. This amounts to starting at $W^{(0)} \in \mathrm{dom(W)}$ and alternatingly applying the updates

$$V^{(k)} \in \partial h(W^{(k)}) \qquad \text{and} \qquad W^{(k+1)} \in \partial g^*(V^{(k)}) \tag{4.14}$$

until a termination criterion is reached. Here, g^* denotes the Fenchel conjugate (Section 2.1.2) of g.

To apply the DC algorithm to our non-convex energy $E(W)$ defined in (4.6) we rewrite $E(W) = g(W) - h(W)$ as a DC function. We identify the concave part of the DC

Algorithm 4.1: Iterated Primal-Dual Algorithm

Init: choose $W^{(0)} = \mathbb{1}_{\mathcal{W}}$ (barycenter), $Q^{(0)} \in \operatorname{dom}(D^*) = \mathbb{R}^m$ and $\tau, \sigma > 0$

while *not converged* **do**

set $\hat{W} = W^{(k)}$

while *not converged* **do**

$$W^{(l+1)} = \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} \lambda R_{\mathcal{W}}(W) + \langle W, \mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top} Q^{(l)}) - \nabla h(\hat{W}) \rangle + \tfrac{1}{\tau} \operatorname{KL}(W, W^{(l)}) \tag{4.19a}$$

$$Q^{(l+1)} = \operatorname*{arg\,min}_{Q} D^*(Q) - \langle Q, A\, \mathrm{P}_{\mathcal{F}_*}(2W^{(l+1)} - W^{(l)}) \rangle + \tfrac{1}{2\sigma} \|Q - Q^{(l)}\|_2^2 \tag{4.19b}$$

$l \leftarrow l + 1$

$k \leftarrow k + 1$; $W^{(k)} = W^{(l+1)}$

Output: $W^* = W^{(k)}$

decomposition and set

$$h_{\text{unary}}(W) := \alpha \langle W, \log(W) \rangle \tag{4.15a}$$

$$h_{\text{inverse}}(W) := \alpha \langle W, \log(W) \rangle - \sum_{i=1}^{n} \sum_{j=1}^{c} W_{i,j} d_{\mathcal{F}} \left((\mathrm{P}_{\mathcal{F}_*}(W))_i, f_j^* \right), \tag{4.15b}$$

for the separable and non-separable case respectively. This is a natural decomposition of the energy (4.6) since both entropy and the discretization term are concave. Furthermore, we collect the remaining convex terms of the energy in (4.6) as

$$g(W) := \lambda R_{\mathcal{W}}(W) + D(W), \tag{4.16}$$

where $D(W)$ corresponds to the convex part of the data term from (4.9). In our case since the concave part h is differentiable, the DC algorithm results in the fixed point iteration

$$W^{(k+1)} = \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} \lambda R_{\mathcal{W}}(W) + D(W) - \langle W, \nabla h(W^{(k)}) \rangle, \tag{4.17}$$

where the Euclidean gradient $\nabla h(W)$ for both cases is given by

$$\nabla h_{\text{unary}}(W) = \alpha(\log(W) + \mathbb{1}_{n \times c}) \tag{4.18a}$$

$$\nabla h_{\text{inverse}}(W) = \alpha(\log(W) + \mathbb{1}_{n \times c}) - \left(\mathrm{P}_{\mathcal{F}_*}(W) \mathbb{1}_c^{\top} - \mathbb{1}_n F^{\top} \right)^2. \tag{4.18b}$$

Regarding the second term of the non-separable case, we refer to Proposition 3.5.1 for the gradient with $\mathcal{F}_* \subset \mathbb{R}$. The vectorial case $\mathcal{F}_* \subset \mathbb{R}^d$ together with $d_{\mathcal{F}}(x, y) = \|x - y\|_2^2$ is straightforward by computing the Euclidean gradient for each channel.

Solving the Fixed Point Iteration. The generic DC algorithm reduces to the fixed point iteration (4.17), which amounts in solving a sequence of convex optimization problems. Each convex subproblem is constrained to the convex set $\overline{\mathcal{W}}$ of assignments matrices. We use Algorithm 4.1 to iteratively solve the subproblems using the generalized primal-dual algorithm [CP16] with Bregman proximal mappings. In particular, we employ the KL-divergence as a Bregman divergence which allows to incorporate the convex constraints in a smooth and geometric way without the requirement of a projection onto the set $\overline{\mathcal{W}}$. See

Remark 4.3.1.

The *primal update step* (4.19a) of Algorithm 4.1 requires evaluating the generalized proximal operator of the regularizer (4.10). Therefore, we rewrite the primal step (4.19a) as

$$W^{(l+1)} = \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} R_{\mathcal{W}}(W) + \frac{1}{\lambda\tau} \operatorname{KL}(W, P), \tag{4.20}$$

where the argument $P \in \overline{\mathcal{W}}$ is given by the non-linear gradient descent step

$$P = \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} \langle W, \mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top}Q^{(l)}) - \nabla h(\hat{W})\rangle + \frac{1}{\tau} \operatorname{KL}(W, W^{(l)}) \tag{4.21a}$$

$$= \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} \frac{1}{\tau} \operatorname{KL}\Big(W, W^{(l)} \exp\big(-\tau(\mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top}Q^{(l)}) - \nabla h(\hat{W}))\big)\Big) \tag{4.21b}$$

$$= \frac{W^{(l)} \exp\big(-\tau(\mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top}Q^{(l)}) - \nabla h(\hat{W}))\big)}{\langle W^{(l)}, \exp\big(-\tau(\mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top}Q^{(l)}) - \nabla h(\hat{W}))\big)\rangle}. \tag{4.21c}$$

Note that the argmin induces normalization of P, thus $P \in \overline{\mathcal{W}}$.

Next, Proposition 4.3.2 below states that evaluating the generalized proximal mapping (4.20) can be done approximately by an efficient fixed point iteration rather then solving a large system of non-linear equations (optimality conditions). Even for larger problem instances, this fixed point iteration evaluates the proximal mapping very efficiently. Specifically, in our numerical experiments we observed convergence within few iterations and we initialized with $W^{(0)} = P$ for warm start. Due to the fact that the variation in (4.11) with respect to geometric averaging is significantly smaller than in the first argument of the KL-divergence (see [ÅPSS17]), we propose the following converging fixed point iteration which is stated in Proposition 4.3.2.

Proposition 4.3.2 (Evaluation of the Proximal Mapping)**.** *Let $P \in \overline{\mathcal{W}}$ be fixed and define $R_{\mathcal{W}}(W)$ by (4.10), then the fixed point iteration*

$$W_i^{(m+1)} = \operatorname*{arg\,min}_{W_i \in \Delta_c} \operatorname{KL}(W_i, \operatorname{gm}(\{W_k^{(m)}\}_{k \in \mathcal{N}_i})) + \frac{1}{\tau\lambda} \operatorname{KL}(W_i, P_i), \quad \forall\, i \in [n], \tag{4.22}$$

converges for every $W^{(0)} \in \overline{\mathcal{W}}$.

Proof. We evaluate the fixed point iteration (4.22) and obtain

$$W_i^{(m+1)} = \frac{(P_i)^{\frac{1}{1+\tau\lambda}} \operatorname{gm}(\{W_k^{(m)}\}_{k \in \mathcal{N}_i})^{\frac{\tau\lambda}{1+\tau\lambda}}}{\langle (P_i)^{\frac{1}{1+\tau\lambda}}, \operatorname{gm}(\{W_k^{(m)}\}_{k \in \mathcal{N}_i})^{\frac{\tau\lambda}{1+\tau\lambda}} \rangle} \qquad \forall\, i \in [n]. \tag{4.23}$$

Without loss of generality we skip the intermediate normalizations and normalize only the *last* iterate since the normalization of the intermediate steps cancels out. This yields the fixed point iteration with $P \in \overline{\mathcal{W}}$ fixed and initial point $W^{(0)} \in \overline{\mathcal{W}}$

$$W_i^{(m+1)} = (P_i)^{\frac{1}{1+\tau\lambda}} \operatorname{gm}(\{W_k^{(m)}\}_{k \in \mathcal{N}_i})^{\frac{\tau\lambda}{1+\tau\lambda}}. \tag{4.24}$$

Taking the logarithm of (4.24), substituting $v^{(m)} = \log(W^{(m)})$ and $r = \log(P)$, gives

$$v_i^{(m+1)} = \frac{1}{1+\tau\lambda} r_i + \frac{\tau\lambda}{1+\tau\lambda} \sum_{k \in \mathcal{N}_i} \frac{1}{|\mathcal{N}_i|} v_k^{(m)}. \tag{4.25a}$$

Rewriting the neighborhood interactions by the associated stochastic matrix Ω, with

$\Omega_{i,k} := 1/|\mathcal{N}_i|$ for $k \in \mathcal{N}_i$ and 0 otherwise, we get the explicit expression

$$v^{(m+1)} = \frac{1}{1+\tau\lambda} r + \underbrace{\frac{\tau\lambda}{1+\tau\lambda}\Omega}_{:=\tilde{\Omega}} v^{(m)} = \frac{1}{1+\tau\lambda} \sum_{l=0}^{m} \tilde{\Omega}^l r + \tilde{\Omega}^{m+1} v^{(0)}. \tag{4.26}$$

Since Ω per definition, is a stochastic matrix and $\tau\lambda(1+\tau\lambda)^{-1} < 1$ it follows that

$$\lim_{m\to\infty} \tilde{\Omega}^m = 0. \tag{4.27}$$

Thus $|\lambda_i| < 1$ holds for all eigenvalues of $\tilde{\Omega}$ and $(I_n - \tilde{\Omega})$ is invertible. This implies that the geometric series of the matrix $\tilde{\Omega}$ converges to

$$v^* = \lim_{m\to\infty} v^{(m)} = \frac{1}{1+\tau\lambda}(I_n - \tilde{\Omega})^{-1} r. \tag{4.28a}$$

Resubstituting the continuous functions $W = \exp(v^*)$, $r = \log(P)$ into (4.28a) and normalization finally gives

$$W_i^* = \frac{\exp\left(\frac{1}{1+\tau\lambda}(I_n - \tilde{\Omega})^{-1}\log(P)\right)_i}{\langle \exp\left(\frac{1}{1+\tau\lambda}(I_n - \tilde{\Omega})^{-1}\log(P)\right)_i, \mathbb{1}_c\rangle}, \tag{4.29a}$$

which yields the limit point W^* independent of $W^{(0)}$. □

Due to the convexity of D and the standard Euclidean proximal mapping, the *dual update step* 4.19b of Algorithm 4.1 can be evaluated in a straightforward manner. We refer to [CP11; CP16] for details.

Parameter Selection. Following the parameter selection of [CP16, Example 7.2] we set $\tau = \sqrt{c}/L_{12}^2$ in the primal update and $\sigma = 1/\sqrt{c}$ in the dual update. Note that this parameter configuration implies that the condition $\sigma\tau \leq \|A\,\mathrm{P}_{\mathcal{F}_*}(\cdot)\|^2$ holds, where the operator norm is given by

$$L_{12} = \|A\,\mathrm{P}_{\mathcal{F}_*}(\cdot)\| = \sup_{\|x\|_1 \leq 1} \|A(I_n \otimes F^{*\top})x\|_2 = \max_{j\in[m]} \|(A(I_n \otimes F^{*\top}))_j\|_2, \tag{4.30}$$

with respect to the mixed ℓ_1-ℓ_2-norm. This stems from the fact that, in the primal, the negative entropy is 1-strongly convex with respect to the ℓ_1-norm when restricted to the simplex, which induces the KL-divergence.

4.3.3. Experiments

In this section we evaluate our proposed variational model (4.6) for separable and non-separable data terms. Regarding indirect observations, we compare the proposed method to state-of-the-art approaches in the setting of discrete tomography reconstruction from only a few projection angles. In all experiments we used a $|\mathcal{N}| = 3 \times 3$ neighborhood system to specify the pairwise relations in the proposed geometric regularization. We avoid numerical issues when evaluating the KL-divergences for almost discrete assignments by adopting the renormalization strategy from [ÅPSS17]. In the end, to guarantee fully discrete solutions we use only a simple pixelwise maximum likelihood (argmax) rounding scheme. Finally, we denote our proposed joint reconstruction and labeling approach (*LayerKL*), since the regularization is based on the KL-divergence which couples the layers (labels) of the assignments in a geometric way.

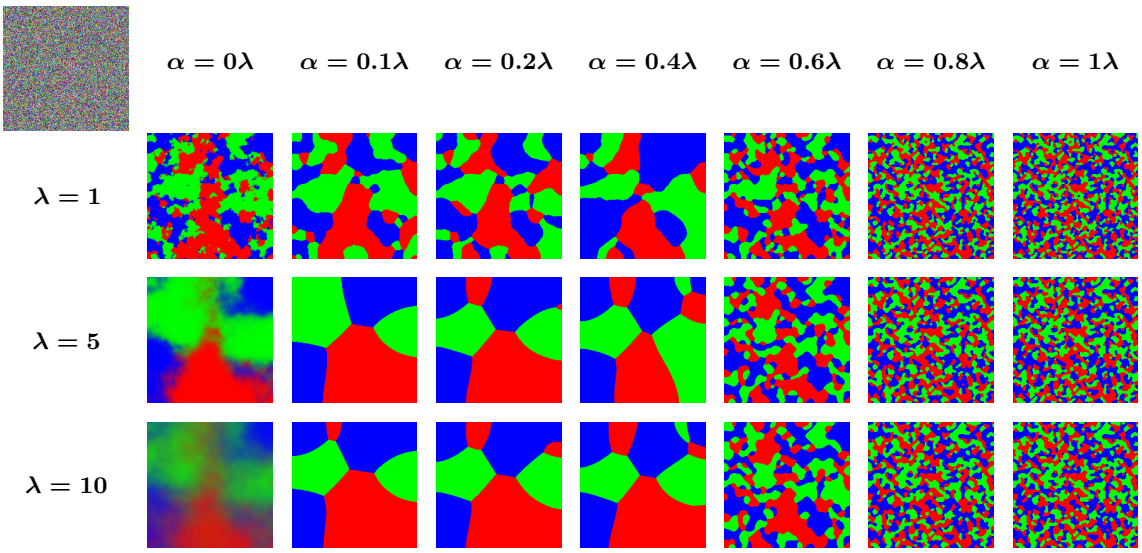

Figure 4.1.: Random color noise (top left corner) serves as direct input data and is labeled with $\mathcal{F}_* = \{$■, ■, ■$\}$ with varying regularization parameter λ and discretization parameter α.

Parameter Influence. This experiment shows the influence of the regularization parameter λ and the discretization parameter α. We generated random color noise u^0 and implemented the data term in (4.7) by $D_{\text{unary}}(W, D_{\mathcal{F}})$ with the distances

$$D_{\mathcal{F};i,j} = d_{\mathcal{F}}(u_i^0, f_j^*) = \|u_i^0 - f_j^*\|_2^2, \quad \text{where} \quad f_j^* \in \mathcal{F}_* = \{■, ■, ■\}, \tag{4.31}$$

between the direct input image and the labels. Figure 4.1 shows that larger values of λ act as a smoothing parameter enforcing larger homogeneous regions, whereas α favors consistency over the discrete label space, which leads to faster discretization and stops the diffusion process. The absence of the entropy term promoting integral solutions is illustrated in the left most-column ($\alpha = 0$). Note, that in this experiment no rounding was applied to show the effect of rounding during optimization.

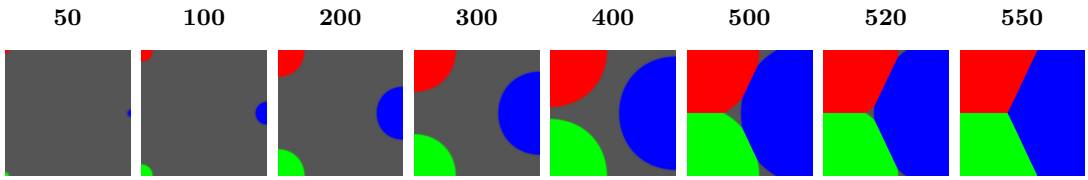

Figure 4.2.: Evolution of interfaces for increasing iteration number with the label set $f_j^* \in \mathcal{F}_* = \{$■, ■, ■$\}$. The interfaces are propagated by the geometric regularization uniformly from the three seed pixels into the uninformative image until they meet in a triple junction.

Interface Propagation. In this example we illustrate the information propagation (diffusion) by the geometric regularization in the case when the data term is uninformative. We use the same model configuration and $\mathcal{F}_*$ as in the previous experiment with $\lambda = 10$ and $\alpha = 1$. We set the input data u^0 to a constant gray image with three seed pixels:

- one ■ -pixel in the top left corner,
- one ■ -pixel in the bottom left corner and
- one ■ -pixel in the middle of the right edge.

Figure 4.2 displays the evolution of interfaces for increasing iterations. We see that the information given by the three seed pixels is uniformly propagated into the image until the interfaces meet in a triple junction which demonstrates uniform propagation speeds. Note that in this experiment, no rounding was applied.

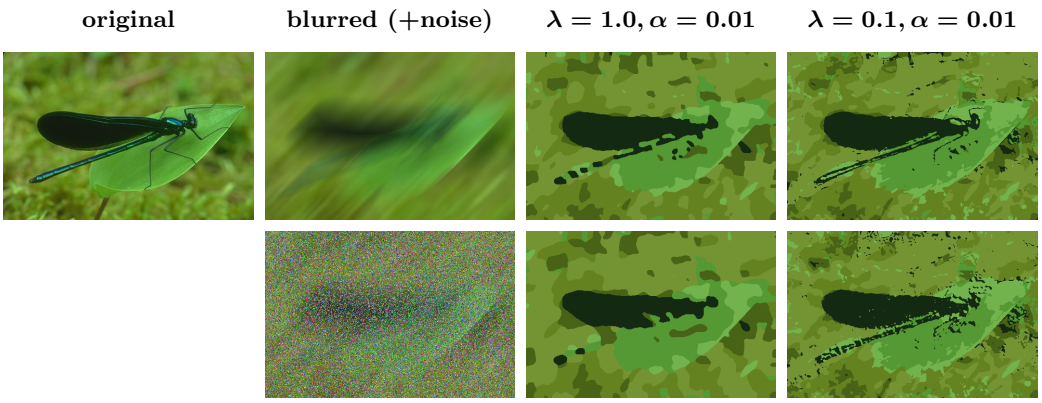

Figure 4.3.: Reconstruction and labeling with the proposed model (4.6) of a severely blurred and noisy image of an insect with discrete label set $\mathcal{F}_* = \{$■, ■, ■, ■, ■, ■$\}$. In both cases the approach is able to reconstruct fine details which are present in the original image.

Joint Deblurring and Labeling. We used the non-separable data term D_{inverse} defined in (4.9) and implemented by the ℓ_1-norm as distance for the indirect measurements. The discretization term was extended component-wise to each color channel corresponding to the quadratic Euclidean norm. The label set $\mathcal{F}_* = \{$■, ■, ■, ■, ■, ■$\}$ was generated by K-means clustering with 6 clusters of the original image, which is displayed in Figure 4.3. The same figure shows the reconstruction of a severely blurred picture of an insect (motion blur of 65 pixel length) and joint labeling with two different parameter configurations: high regularization and low regularization. In a more challenging setting we additionally corrupted 50% the blurred image with random colors drawn from a uniform distribution. In both cases we were able to reconstruct fine details of the original image.

Discrete Tomography Reconstruction. The reconstruction problem in discrete tomography aims to recover an image $u \in \mathbb{R}^n$ from a small number of possibly noisy measurements $b = Au + \nu \in \mathbb{R}^m$. The latter correspond to line integrals that sum up all absorptions over each ray transmitted through the object. A given projection matrix $A \in \mathbb{R}^{m \times n}$ encodes the imaging geometry. Here we used the parallel beam setup. The width of the sensor-array was set to 1.5 times the image size, so that every pixel intersects with at least a single projection ray. We used the non-separable data term D_{inverse} (4.9) implemented by indicator functions to enforce the constraints, see Chapter 3.4 for details.

We compare the proposed model (LayerKL) to state-of-the-art approaches for non-binary discrete tomography reconstruction from only a few projection angles. Specifically, we considered DART [BS11], the energy minimization method of Varga et. al. [VBN12] (Varga) and the preliminary version (LayerTV) introduced in Chapter 3, with a layer-wise total variation regularizer.

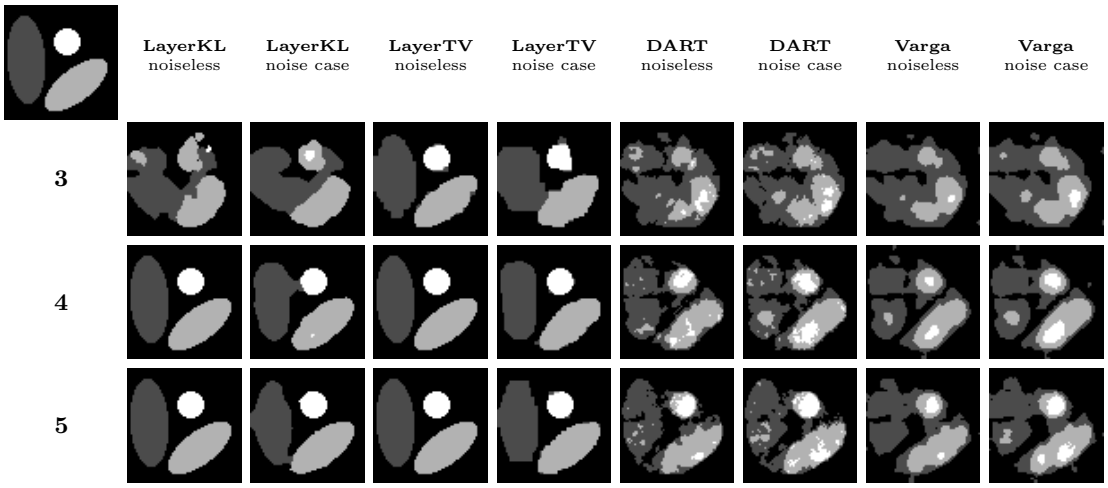

Figure 4.4.: Visual results of experiment with phantom 1 (ellipses, $n = 64 \times 64$).

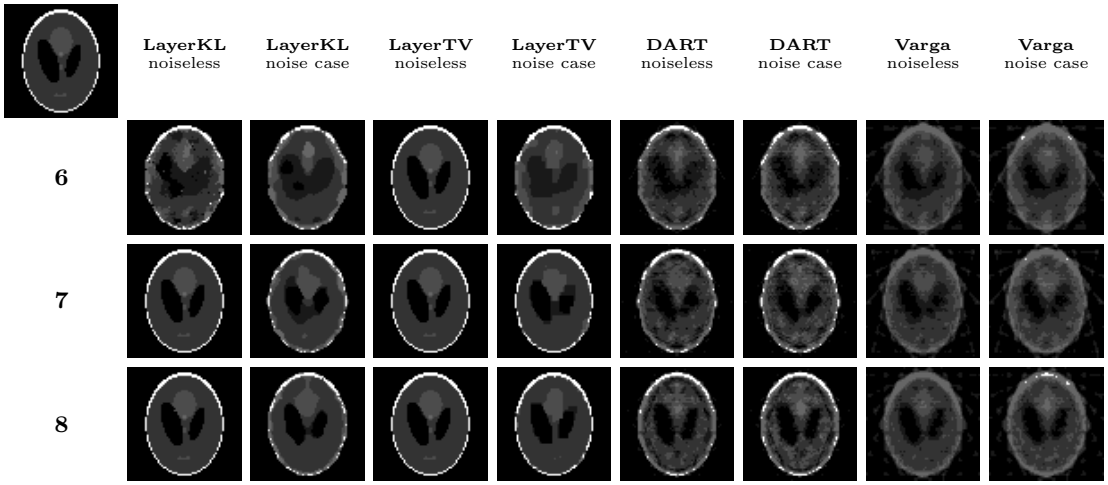

Figure 4.5.: Visual results of experiment with phantom 4 (Shepp-Logan [SL74], $n = 64 \times 64$).

Setup. For the evaluation we measured the relative pixel error, that is the relative number of erroneously reconstructed pixels as compared to the ground truth. We tried to use the default parameters of the competing approaches as proposed by their authors. However, since the test-datasets differ in size, we slightly adjusted the parameters to get best results for every algorithm and problem instance.

Results. In Figure 4.5 and 4.4 the proposed approach gives perfect reconstructions with a low number of projection angles in the noiseless case and also returns high-quality

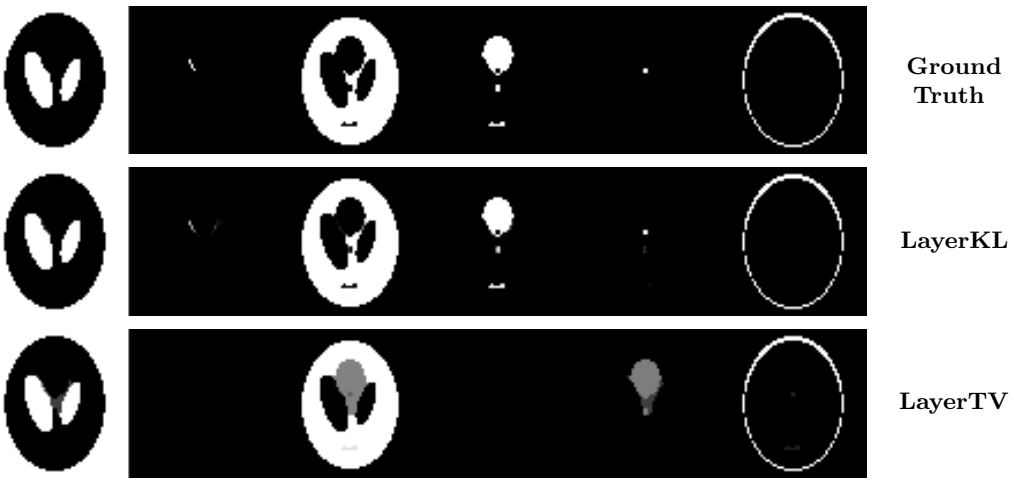

Figure 4.6.: Reconstruction of the Shepp-Logan ($n = 64 \times 64$) phantom [SL74] from only 7 projections, where the indicator variables W^j are shown for each layer $j \in [c]$ from left to right. White denotes the selected label. The LayerTV produces a non-integral solution with a convex combination of the labels - illustrating the need for rounding, whereas our proposed model LayerKL directly gives a meaningful and almost discrete labeling.

reconstructions in the presence of noise. The preliminary approach LayerTV needs one fewer projection, but a non-trivial rounding strategy must be used. This behavior is depicted in Figure 4.6 for the Shepp-Logan phantom [SL74] from 7 projections where LayerTV clearly gives a non-integral solution and requires a special rounding strategy to obtain a meaningful reconstruction. Further details are given in the caption. Figure 4.7 shows the numerical evaluation of the approaches for increasing (but small) numbers of projection angles, in the noiseless case (filled markers) and in the noisy case (non-filled markers), with Poisson noise $SNR = 20$ dB.

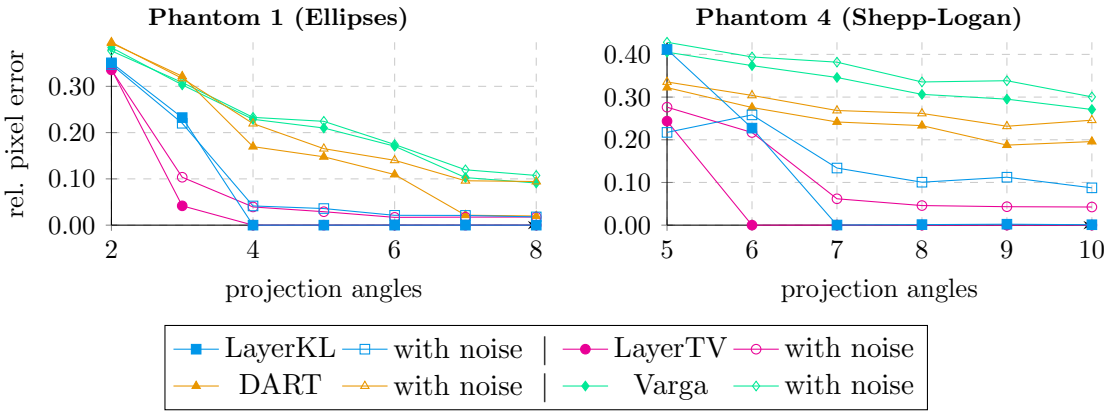

Figure 4.7.: Relative pixel error relative to the number of projections angles. Regarding the Shepp-Logan phantom [SL74] LayerTV can reconstruct the phantom with one fewer projection, but a special rounding strategy is performed to obtain a meaningful solution.

4.4. Reconstruction by Gradient Flows on a Riemannian Submanifold

This section is concerned with a completely geometric approach for approximating the combinatorial multilabel problem with indirect measurements. Previously, we have derived a variational approach (4.6), where regularization and optimization are based on the KL-divergence in order to respect the geometric properties of the assignment manifold $\mathcal{W}$. Here we are going one step further by additionally incorporating the affine constraints from the indirect measurements directly into the geometry. Therefore, we are restricting the assignment manifold to a submanifold which is equipped with a natural extension of the Fisher-Rao metric [BR82]. More precisely, we endow the relative interior of the feasible set with a Hessian Riemannian metric. Then we construct a smooth gradient flow, for any smooth objective function, restricted to the submanifold. Accordingly, we define a corresponding objective function for joint tomographic reconstruction and label assignment. Finally, we show that the resulting flow can be numerically integrated by an iterated implicit scheme based on a Bregman proximal point iteration.

4.4.1. Tomographic Assignment Flow

In the case of direct observations, the assignment flow (2.89) is a smooth geometric approach for labeling where each label $j \in [c]$ and corresponding prototype

$$f_j^* \in \mathcal{F}_* := \{f_1^*, \dots, f_c^*\} \subset \mathbb{R}, \tag{4.32}$$

is represented by a vertex $e_j \in \mathbb{R}^c$ of the probability simplex. Here we restrict to real-valued labels for the purpose of discrete tomographic reconstruction. The relative interior of the set of feasible label assignments to all pixels corresponds to the set of row-stochastic matrices with full support, denoted by $\mathcal{W} \subset \mathbb{R}_{++}^{n \times c}$. Afterwards the feasible set $\mathcal{W}$ is turned into a Riemannian manifold using the Fisher-Rao (information) metric [BR82].

We now consider the situation where the image data are only indirectly observed through the projection constraints (3.4). To this end, we extend the assignment flow approach using techniques developed by [ABB04], in order to restrict the smooth Riemannian flow to assignments that respect the projection constraints.

Our starting point is the observation that the Riemannian metric used in the assignment flow approach (2.89) is induced by the Hessian of a convex Legendre function, the negative Bolzmann-Shanon entropy which reads

$$h(W) := \langle W, \log(W) - \mathbb{1}_{n \times c} \rangle, \tag{4.33}$$

with domain restricted to the relative interior of

$$\overline{\mathcal{W}} = \{W \in \mathbb{R}_+^{n \times c} \colon W \mathbb{1}_c = \mathbb{1}_n\}. \tag{4.34}$$

In order to take into account the projection constraints (3.4), we introduce the assignment operator

$$\mathrm{P}_{\mathcal{F}_*} \colon \overline{\mathcal{W}} \to \mathrm{conv}(\mathcal{F}_*^n), \quad W \mapsto \mathrm{P}_{\mathcal{F}_*}(W) = (I_n \otimes {F^*}^\top) \operatorname{vec}(W) = W F^*, \tag{4.35}$$

where $F^* := (f_1^*, \dots, f_c^*)^\top$ is the vector of all admissible gray-values, that explicitly characterizes the reconstructed function

$$u = \mathrm{P}_{\mathcal{F}_*}(W) = W F^* \in \mathbb{R}^n \tag{4.36}$$

in terms of given labels F^* and assignments W. Based on this correspondence and (3.4),

we restrict the feasible set $\overline{\mathcal{W}}$ to

$$\overline{\mathcal{W}}_{\text{tomo}} = \left\{ W \in \mathbb{R}_+^{n\times c} : \; B\,\mathrm{vec}(W) = y \right\}, \quad B := \begin{pmatrix} A(I_n \otimes F^{*\top}) \\ I_n \otimes \mathbb{1}_c^\top \end{pmatrix} \quad y := \begin{pmatrix} b \\ \mathbb{1}_n \end{pmatrix}. \tag{4.37}$$

The following Proposition 4.4.1 states a non-degeneracy property of the matrix B which is crucial for the smooth geometric construction below. The proof exploits the structure of B and the properties of the Kronecker product.

Proposition 4.4.1 (Rank of B)**.** *The matrix B has full row rank by construction, if the measurement matrix A has full row rank.*

Proof. We rewrite the matrix $B \in \mathbb{R}^{(m+n)\times cn}$ as

$$B = \begin{pmatrix} A(I_n \otimes F^{*\top}) \\ I_n \otimes \mathbb{1}_c^\top \end{pmatrix} = \underbrace{\begin{pmatrix} A & 0 \\ 0 & I_n \end{pmatrix}}_{:=X} \underbrace{\begin{pmatrix} I_n \otimes F^{*\top} \\ I_n \otimes \mathbb{1}_c^\top \end{pmatrix}}_{:=Y}. \tag{4.38}$$

The matrix $X \in \mathbb{R}^{(m+n)\times 2n}$ has full row rank since the matrix A has full row rank by assumption.

Furthermore, the matrix $Y \in \mathbb{R}^{2n\times cn}$ has full row rank, i.e. $\mathrm{rank}(Y) = 2n$ since

$$\mathrm{rank}(I_n \otimes F^{*\top}) = \mathrm{rank}(I_n)\,\mathrm{rank}(F^{*\top}) = n, \tag{4.39a}$$

$$\mathrm{rank}(I_n \otimes \mathbb{1}_c^\top) = \mathrm{rank}(I_n)\,\mathrm{rank}(\mathbb{1}_c^\top) = n, \tag{4.39b}$$

and moreover the two blocks are linearly independent by construction (it is not possible to express a row of the first block of Y with a linear combination of rows from the second block and vice versa, except for $F^* = \lambda\mathbb{1}$ with $\lambda \neq 0$).

Thus, we have

$$\mathrm{rank}(B) = \mathrm{rank}(XY) = \mathrm{rank}(X) = m + n, \tag{4.40}$$

which shows that B has full row rank. □

Our next step is to restrict the manifold $\mathcal{W}$ to a manifold $\mathcal{W}_{\text{tomo}}$, based on the restriction of $\overline{\mathcal{W}}$ to $\overline{\mathcal{W}}_{\text{tomo}}$. We adopt the convex Legendre function $h(W)$, defined in (4.33), and take as its domain the strictly positive orthant $\mathcal{M} = \mathbb{R}_{++}^{n\times c}$. Then the Hessian $\nabla^2 h(W) = \frac{1}{W}$ (*componentwise* inverse) smoothly depends on $W \in \mathcal{M}$ and defines the linear mapping

$$H(W)\colon \mathbb{R}^{n\times c} \to \mathbb{R}^{n\times c}, \quad U \mapsto H(W)U := \big(U_{i,j}/W_{i,j}\big)_{i\in[n],j\in[c]}, \tag{4.41}$$

which is positive definite. Based on the canonical identification of tangent spaces $T_W\mathcal{M} \simeq \mathbb{R}^{n\times c}$ for open subsets of the Euclidean space, the mapping $H(W)$ defines the Riemannian metric

$$\langle U, V\rangle_W^H := \langle H(W)U, V\rangle, \qquad \forall\, W \in \mathcal{M}, \quad U, V \in \mathbb{R}^{n\times c}. \tag{4.42}$$

Given some smooth objective

$$J\colon \mathbb{R}^{n\times c} \to \mathbb{R}, \quad W \mapsto J(W), \tag{4.43}$$

the corresponding Riemannian gradient field restricted to $\mathcal{M}$ is given by

$$\mathrm{grad}\, J_{|\mathcal{M}}(W) := H(W)^{-1}\nabla J(W). \tag{4.44}$$

Next we consider the smooth submanifold of $\mathcal{M}$, given by

$$\mathcal{W}_{\text{tomo}} = \mathrm{rint}(\overline{\mathcal{W}}_{\text{tomo}}) = \mathcal{M} \cap \{W \in \mathbb{R}^{n\times c} : \; B\,\mathrm{vec}(W) = y\} \subset \mathcal{M} \tag{4.45}$$

with tangent space $T_W\,\mathcal{W}_{\text{tomo}} \simeq \mathcal{N}(B)$. The metric on $\mathcal{M}$ induces a metric on $\mathcal{W}_{\text{tomo}}$. Furthermore, the Riemannian gradient field of $J(W)$ restricted to $\mathcal{W}_{\text{tomo}}$ reads

$$\mathrm{grad}\, J_{|\,\mathcal{W}_{\text{tomo}}}(W) = \Pi_W^{\mathcal{N}(B)}\big(H(W)^{-1}\nabla J(W)\big), \tag{4.46}$$

where $\Pi_W^{\mathcal{N}(B)}$ is the $\langle\cdot,\cdot\rangle_W^H$-orthogonal projection onto the nullspace $\mathcal{N}(B)$. Since the matrix B has full rank due to Proposition 4.4.1, this projection reads

$$\Pi_W^{\mathcal{N}(B)}(H(W)^{-1}\nabla J(W)) = \operatorname{vec}^{-1}\left[\left(I - (BD_W^H B^\top)^{-1}BD_W^H\right)(D_W^H)^{-1}\operatorname{vec}[\nabla J(W)]\right], \tag{4.47a}$$

$$\text{with } D_W^H = \operatorname{Diag}[\operatorname{vec}(H(W))]. \tag{4.47b}$$

The vector field $-\operatorname{grad} J_{|\mathcal{W}_{\text{tomo}}}(W)$ for $W \in \mathcal{W}_{\text{tomo}}$ is the steepest descent direction in the corresponding tangent space $\mathcal{N}(B)$. Furthermore, minimization of an objective J on the Riemannian manifold $\left(\mathcal{W}_{\text{tomo}}, \langle\cdot,\cdot\rangle_W^H\right)$ amounts to finding the solution trajectory $W(t)$ of the continuous dynamical system

$$\dot{W}(t) + \operatorname{grad} J_{|\mathcal{W}_{\text{tomo}}}(W(t)) = 0, \quad W(0) = W^{(0)} \in \mathcal{W}_{\text{tomo}}, \tag{4.48}$$

with an initial condition $W^{(0)} \in \mathcal{W}_{\text{tomo}}$.

4.4.2. Objective Function

After having defined the geometric setting in the previous section, we are able to specify a suitable objective function for joint reconstruction and labeling.

We adopt and modify the assignment flow approach (2.89) for our purpose. Assume we have some tomographic reconstruction which is given in terms of the assignments $\hat{W}$. Accordingly, we define a distance matrix

$$D(\hat{W}) := \frac{1}{\rho}\Big(\|\operatorname{P}_{\mathcal{F}_*}(\hat{W})_i - f_j^*\|_2^2\Big)_{i\in[n],\, j\in[c]} \tag{4.49}$$

with the assignment operator $\operatorname{P}_{\mathcal{F}_*}(\hat{W})$ given by (4.35) and a scaling parameter $\rho > 0$. Furthermore, we compute a similarity matrix $S(\hat{W})$ as described in (2.88) to induce geometric spatial regularization.

Based on $S(\hat{W})$, we define the following objective function

$$J(W, \hat{W}) = \operatorname{KL}\left(W, S(\hat{W})^{1+\alpha}\right), \quad \alpha > 0. \tag{4.50}$$

Remark 4.4.1. Minimizing J with respect to W encodes two aspects. First, the discrete assignment distributions comprising W should be consistent with the spatially regularized similarities $S(\hat{W})$ that correspond to the lifted distances $D(\hat{W})$ between the reconstructed function $\operatorname{P}_{\mathcal{F}_*}(\hat{W})$ and the labels $\mathcal{F}_*$. Second, since W appears as first argument of the KL-distance, W matches the prominent *modes* of the discrete distributions comprising $S(\hat{W})$ (cf. [Min05]), and hence enforces unique labelings. The damping parameter α enables to control this "rounding property".

Since the assignment $\hat{W}$ is not given beforehand, we pursue an iterative strategy and set $\hat{W} = W^{(k)}$ to the current iterate, in order to compute $W^{(k+1)}$ by minimizing (4.50). In the next section, we formulate this process in a more principled way as a fixed point iteration, that properly discretizes and solves the continuous flow (4.48).

4.4.3. Iterated Implicit Scheme

In this section we want to find a solution trajectory of the initial value problem (4.48) associated with the steepest Riemannian gradient descent of the convex objective function $J(W)$ in (4.50) on the smooth manifold $\mathcal{W}_{\text{tomo}}$. Following [ABB04], we reformulate (4.48)

as a differential inclusion for a time interval (T_m, T_M) corresponding to the unique maximal solution of (4.48) and obtain

$$\frac{d}{dt}\nabla h(W(t)) + \nabla J(W(t)) \in \mathcal{N}(B)^{\perp}, \quad W(t) \in \mathcal{W}_{\text{tomo}}, \quad W(0) = W^{(0)} \in \mathcal{W}_{\text{tomo}}, \tag{4.51}$$

with h given by (4.33). Since $J(W)$ is convex, the time discretization

$$\frac{d}{dt}\nabla h(W(t)) \approx \frac{\nabla h(W^{(k+1)}) - \nabla h(W^{(k)})}{\mu_k}, \tag{4.52}$$

together with $W^{(0)} \in \mathcal{W}_{\text{tomo}}$ yields the implicit scheme

$$\nabla h(W^{(k+1)}) - \nabla h(W^{(k)}) + \mu_k \nabla J(W^{(k+1)}) \in \mathcal{N}(B)^{\perp}, \quad B\operatorname{vec}(W^{(k+1)}) = y, \tag{4.53}$$

where μ_k is a step-size parameter. These relations are in fact the optimality conditions of the *Bregman proximal point method* with the KL-divergence as a proximity measure

$$W^{(k+1)} \in \operatorname*{arg\,min}_{W \in \mathbb{R}_+^{n\times c}} J(W, \hat{W}) + \frac{1}{\mu_k}\operatorname{KL}(W, W^{(k)}) \quad \text{s.t.} \quad B\operatorname{vec}(W) = y. \tag{4.54}$$

We solve (4.54) for fixed $W^{(k)}$ by an iterative algorithm to perform an implicit integration step on the flow (4.48). In order to update the fixed $\hat{W}$ in $J(W, \hat{W})$ defined by (4.50), we set $\hat{W} = W^{(k)}$. Inserting $\hat{W}$ into (4.54) and combining the KL-divergences as a multiplicative convex combination with respect to the second argument yields the fixed point iteration

$$W^{(k+1)} \in \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} \operatorname{KL}\Big(W, \underbrace{\big(W^{(k)}\big)^{\frac{1}{1+\mu_k}}\big(S(W^{(k)})\big)^{\frac{\mu_k(1+\alpha)}{1+\mu_k}}}_{:=T_{\mu_k}(W^{(k)})}\Big) \quad \text{s.t.} \quad A\,\mathrm{P}_{\mathcal{F}_*}(W) = b, \tag{4.55}$$

where the constraints $W \in \mathbb{R}_+^{n\times c}$ and $B\operatorname{vec}(W) = y$ are rewritten as $W \in \overline{\mathcal{W}}$ and $A\,\mathrm{P}_{\mathcal{F}_*}(W) = b$. Regarding convergence of the fixed point iteration (4.55), we use a non-summable diminishing step-size parameter

$$\mu_k := \frac{1}{0.005 \cdot k \cdot \|A\,\mathrm{P}_{\mathcal{F}_*}(W^{(k)}) - b\|_2} \quad \text{with} \quad \lim_{k\to\infty} \mu_k = 0. \tag{4.56}$$

Consequently, the operator T_{μ_k} becomes $T_{\mu_k} \longrightarrow \mathrm{Id}$ for $k \to \infty$ and the influence of the objective function $J(W)$ vanishes. When the iteration converges, Eq. (4.55) reduces to the KL-projection onto the fixed feasible set $\overline{\mathcal{W}}_{\text{tomo}}$.

Solving the Fixed Point Iteration. Algorithm 4.2 solves equation (4.55) iteratively using the generalized primal-dual algorithm [CP16]. The primal update step (4.57a) can be evaluated in closed form

$$W^{(l+1)} = \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} \operatorname{KL}(W, \tilde{W}) + \langle W, \mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top}Q^{(l)})\rangle + \frac{1}{\tau}\operatorname{KL}(W, W^{(l)}) \tag{4.58a}$$

$$= \frac{(W^{(l)})^{\frac{1}{1+\tau}}(\tilde{W})^{\frac{\tau}{1+\tau}}\exp(-\frac{\tau}{1+\tau}\mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top}Q^{(l)}))}{\langle (W^{(l)})^{\frac{1}{1+\tau}}(\tilde{W})^{\frac{\tau}{1+\tau}}, \exp(-\frac{\tau}{1+\tau}\mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top}Q^{(l)}))\rangle}. \tag{4.58b}$$

The dual update step (4.57b) admits a closed form

$$Q^{(l+1)} = \operatorname*{arg\,min}_{Q} \langle Q, b - A\,\mathrm{P}_{\mathcal{F}_*}(2W^{(l+1)} - W^{(l)})\rangle + \frac{1}{2\sigma}\|Q - Q^{(l)}\|_2^2 \tag{4.59a}$$

$$= Q^{(l)} + \sigma(A\,\mathrm{P}_{\mathcal{F}_*}(2W^{(l+1)} - W^{(l)}) - b), \tag{4.59b}$$

by evaluating the standard proximal operator since the measurements b are in the Euclidean space $\mathbb{R}^m$.

Algorithm 4.2: Iterated Implicit Primal-Dual Scheme

Init: choose $W^{(0)} = \mathbb{1}_{\mathcal{W}}$ (barycenter), dual variable $Q^{(0)} \in \mathbb{R}^m$ and $\tau, \sigma > 0$
Parameters: selectivity $\rho > 0$, discretization $\alpha > 0$, trust region $\mu_k > 0$

while *not converged* **do**
  set $\tilde{W} = T_{\mu_k}(W^k)$
  while *not converged* **do**

$$W^{(l+1)} = \operatorname*{arg\,min}_{W \in \overline{\mathcal{W}}} \mathrm{KL}(W, \tilde{W}) + \langle W, \mathrm{P}_{\mathcal{F}_*}^{\top}(A^{\top} Q^{(l)}) \rangle + \tfrac{1}{\tau} \mathrm{KL}(W, W^{(l)}) \tag{4.57a}$$

$$Q^{(l+1)} = \operatorname*{arg\,min}_{Q} \langle Q, b - A\,\mathrm{P}_{\mathcal{F}_*}(2W^{(l+1)} - W^{(l)}) \rangle + \tfrac{1}{2\sigma} \|Q - Q^{(l)}\|_2^2 \tag{4.57b}$$

    $l \leftarrow l + 1$
  $k \leftarrow k + 1; \quad W^{(k)} \leftarrow W^{(l+1)}$
Output: $W^* = W^{(k)}$

Parameter Selection. For the step-size parameters τ and σ of the iterated primal-dual algorithm, we adopt the parameter values of [CP16, Example 7.2] and set $\tau = \sqrt{c}/L_{12}^2$ for the primal update and $\sigma = 1/\sqrt{c}$ for the dual update. This choice implies that the condition $\sigma\tau \|A\,\mathrm{P}_{\mathcal{F}_*}(\cdot)\|^2 \leq 1$ for convergence holds, with the operator norm

$$L_{12} = \|A\,\mathrm{P}_{\mathcal{F}_*}(\cdot)\| = \sup_{\|x\|_1 \leq 1} \|A(I_n \otimes F^{*\top})x\|_2 = \max_{j \in [m]} \|(A(I_n \otimes F^{*\top}))_j\|_2, \tag{4.60}$$

with respect to the mixed $\ell_1 - \ell_2$-norm. This reflects the fact that the negative entropy is 1-strongly convex with respect to the ℓ_1-norm when restricted to the probability simplex.

4.4.4. Experiments

In this section we evaluate the proposed smooth geometric flow (4.48) for joint tomographic reconstruction and labeling, implemented by the concrete objective function (4.50) which is motivated by the assignment flow approach (2.89). We compared the proposed approach (*TomoFlow*) to state-of-the-art approaches for discrete tomography, including the Discrete Algebraic Reconstruction Technique (*DART*) [BS11], the energy minimization method of Varga et al. [VBN12] (*Varga*), and our variational approximation (LayerTV), which was introduced in Chapter 3, with a layer-wise total variation regularizer.

Setup. We adapted the binary phantoms from Weber et al. [WNS+06] to the multilabel case, shown as phantom 1,2 and 3 in Figure 4.8. Phantom 4 is the well-known Shepp-Logan phantom [SL74]. Furthermore, we simulated noisy scenarios by applying Poisson noise to the measurements b with a signal-to-noise ratio of $SNR = 20$ db. The geometrical projection setup was created by the ASTRA-toolbox [APB+15], where we used parallel projections along equidistant angles between 0 and 180 degrees. The width of the sensor-array was set 1.5 times the image size, such that every pixel is intersecting with at least one single projection ray.

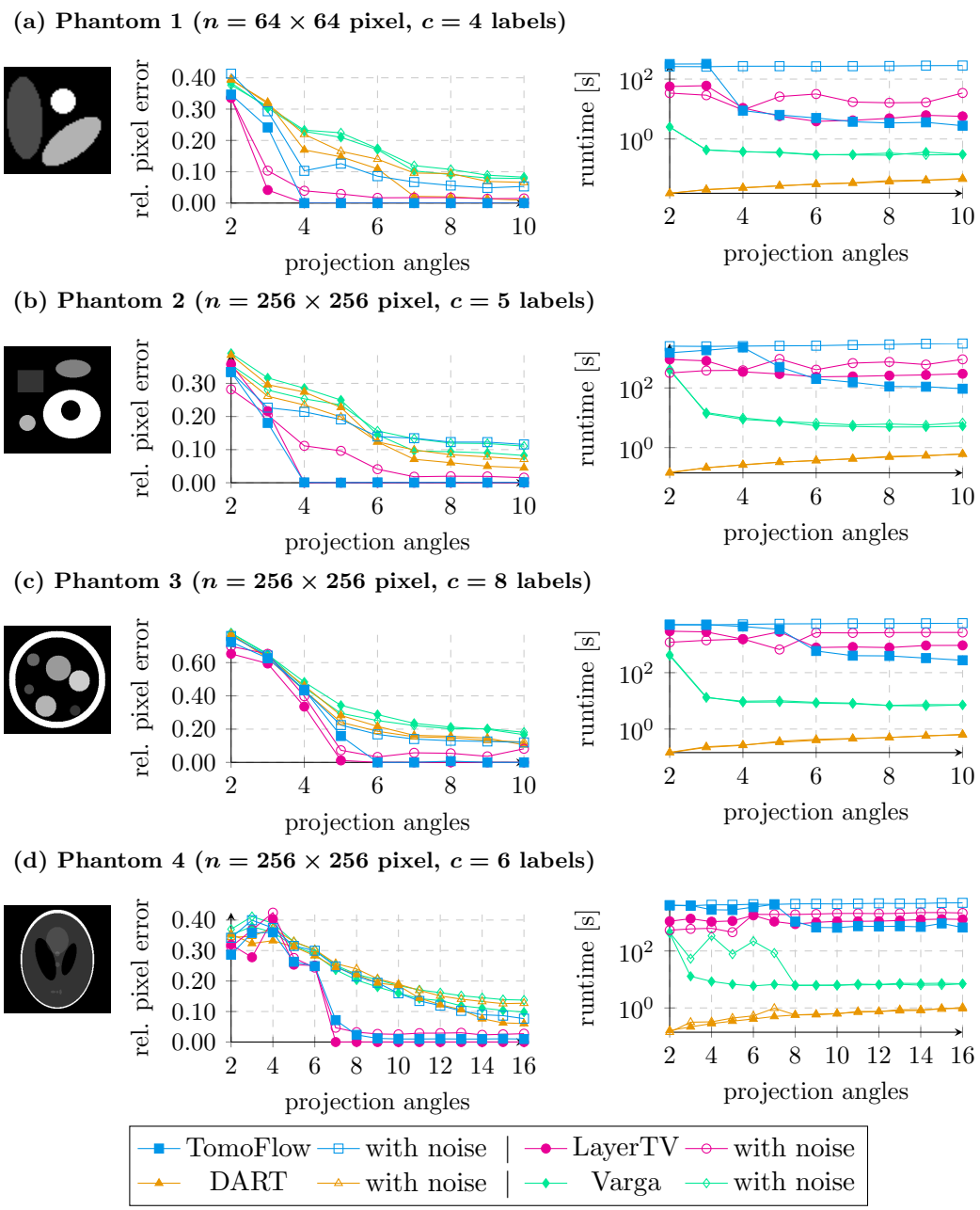

Figure 4.8.: Evaluation of the approaches for different test-datasets and increasing (but small) numbers of projections angles, in the noiseless case **(filled markers)** and in the noisy case **(non-filled markers)**, noise level $SNR = 20$ dB. The relative pixel error and runtime is displayed. The proposed approach TomoFlow gives perfect reconstructions with a small number of projection angles in the noiseless case and also returns good reconstructions in the presence of noise, compared to the other approaches. The single competitive approach, LayerTV, uses a special rounding strategy to obtain meaningful solutions (phantom 3 and 4) and a dedicated data term to cope with Poisson noise.

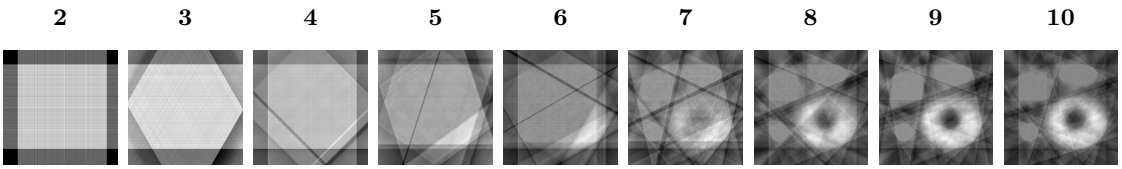

Figure 4.9.: "Implicit data terms" generated by the tomographic constraints, in terms of the reprojected dual variable $A^\top Q$ in algorithm 4.2 (scaled to $[0, 1]$ and inverted for better visualization), after convergence, for phantom 2 and an increasing number of projection angles. The proposed approach achieves a perfect reconstruction from only 4 projection angles. The missing information is effectively compensated by geometric label assignment and spatial coherence due to geometric averaging.

Implementation details. The subproblems of Algorithm 4.2 were approximately solved by the generalized primal-dual algorithm [CP16]. For the multiplicative updates (4.58b), we adopted the renormalization strategy from [ÅPSS17] to avoid numerical issues close to the boundary of the manifold, that correspond to unique label assignments. The outer iteration was terminated when the reprojection error

$$\|A\,\mathrm{P}_{\mathcal{F}_*}(W^{(k)}) - b\|_2 < 0.1, \tag{4.61}$$

dropped below the threshold 0.1. Additionally, we limited the number of total iterations to 20.000. For the geometric averaging (cf. (2.88) and (4.50)), we used a $\mathcal{N} = 3 \times 3$ neighborhood for the smaller phantom 1 and $\mathcal{N} = 5 \times 5$ for all others. In order to reconstruct from *noisy* measurements, we modified the proposed approach by using the squared ℓ_2-reprojection error as a relaxed data term, so that the objective (4.50) reads

$$J(W, \hat{W}) = \mathrm{KL}\left(W, S(\hat{W})^{1+\alpha}\right) + \frac{1}{2}\|A\,\mathrm{P}_{\mathcal{F}_*}(W) - b\|_2^2, \quad \alpha > 0, \tag{4.62}$$

which is smooth and convex in W. In this case, the fixed point iteration (4.55) is applied to the modified objective (4.62) and the dual update step (4.57b) of algorithm 4.2 results in an additional rescaling, i.e.

$$Q^{(l+1)} = \frac{Q^{(l)} + \sigma(A\,\mathrm{P}_{\mathcal{F}_*}(2W^{(l+1)} - W^{(l)}) - b)}{1 + \sigma} \tag{4.63}$$

compared to (4.59b). However, we remark, that the ℓ_2 data term has very limited ability to treat Poisson noise, whereas the inequality projection constraints used by LayerTV are superior.

Regarding DART we used the publicly available implementation included in the ASTRA-toolbox [APB+15], for Varga [VBN12] and LayerTV (Chapter 3) we used our own implementations in MATLAB. We used the default parameters of the competing approaches as proposed by the respective authors. However, since the test-datasets differ in size, we slightly adjusted the parameters in order to get best results for every algorithm and problem instance.

Results. Figure 4.8 summarizes the numerical evaluation of the approaches for increasing (but small) numbers of projections, in the noiseless case (filled markers) and in the noisy case (non-filled markers), with Poisson noise $SNR = 20$ dB. Each test-dataset is depicted in the leftmost column, followed by the relative pixel error and runtime.

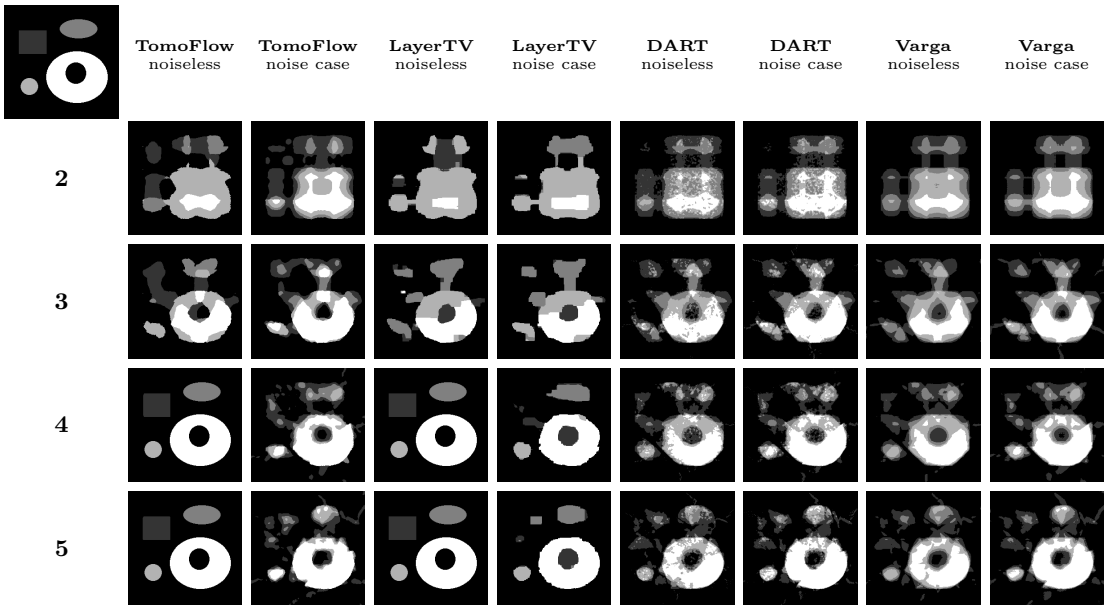

Figure 4.10.: Visual results of experiment phantom 2.

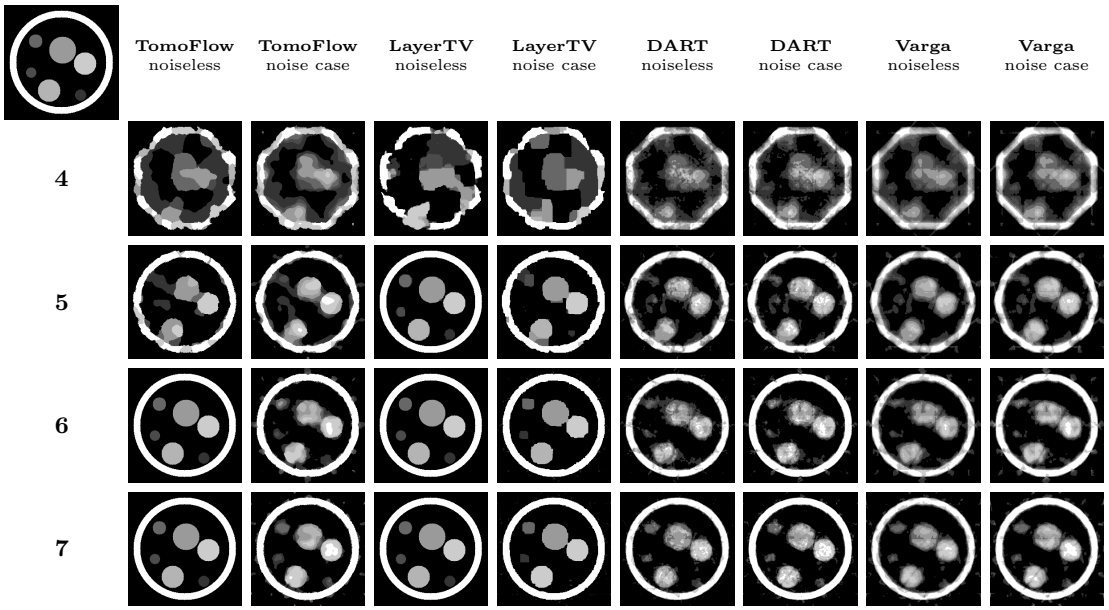

Figure 4.11.: Visual results of experiment phantom 3.

The proposed approach TomoFlow achieved perfect reconstructions with a small number of projection angles in the noiseless case. Only LayerTV needed one fewer projection at phantom 3 and 4. LayerTV however tends to return non-integral solutions when the regularization parameter is large and requires afterwards a special rounding strategy to obtain a meaningful reconstruction. In noisy scenarios, LayerTV performs better due to use of the inequality projection constraints, followed by the proposed method TomoFlow that outperforms both DART and Varga. Figure 4.9 shows the poor "implicit data terms" generated by the tomographic constraints in the case of phantom 2, to illustrate the severe ill-posedness of these inverse problems (see caption for more details).

Considering the runtime (right plots from figure 4.8), DART is the fastest approach followed by Varga. The proposed approach TomoFlow and LayerTV are clearly consuming more runtime to return more accurate solutions. In the noiseless case and with a sufficient number of projection angles, TomoFlow is faster. We point out that the proposed approach could easily be parallelized using graphics cards. In Figure 4.10 and 4.11 the visual results are displayed for the phantoms 2 and 3.

4.5. Conclusion

In this chapter we presented two novel approaches to jointly solving tomographic reconstruction and label assignment with particular focus on the underlying geometric setting of discrete probability distributions.

The first approach was a non-convex variational energy, where the spatial regularization term is implemented by pairwise KL-divergences. This inter-layer coupling of the assignments preserves the information geometric properties of the underlying statistical manifold. Additionally, we showed that evaluating the proximal mapping of the regularization term, relying on the KL-divergence, can be efficiently done by a converging fixed-point iteration. A primal-dual algorithm was applied to solve each subproblem. The numerical evaluation illustrated the competitiveness of the approach compared to state-of-the-art methods for discrete tomography reconstruction as well as deblurring and denoising with joint labeling.

The second approach was a gradient flow which evolves on a Riemannian submanifold. Therefore, we derived a suitable Riemannian structure on the feasible set in order to optimize a smooth objective function on a manifold that respects the projection constraints. The Riemannian gradient flow for our particular objective function combines tomographic reconstruction and labeling in a smooth and mathematically sound way. We derived an iterated implicit scheme to numerically integrate the resulting tomographic assignment flow. Finally, a numerical evaluation illustrated the approach compared to state-of-the-art discrete tomography reconstruction algorithms.

In the context of this work, an immediate question arises, whether the supervised joint reconstruction and labeling approaches can be extended to the unsupervised case. For direct observations, the generalization of the multilabel problem to unsupervised scenarios is investigated in the last chapter of this work within the smooth and geometric framework of the assignment flow. However, the unsupervised setting with indirect measurements, where neither the prototypes are given beforehand nor the data to be labeled is directly available, poses a challenging task to solve for future work.

Possible future work could include a rigorous mathematical analysis of convergence of the fixed-point iteration (4.55) and stability of the corresponding Riemannian gradient descent flow (4.48) that entails iterative updates $\hat{W} = W^{(k)}$ of the objective function $J(W, \hat{W})$. Such issues are not covered by standard convex programming theory.

Furthermore, a promising extension of the tomographic flow concerns the ability to handle inequality constraints, in order to further improve the performance in scenarios with high noise levels.

CHAPTER 5

Self-Assignment Flows for Unsupervised Data Labeling on Graphs

5.1. Introduction and Overview

In the previous chapters we mainly considered a generalization of the data labeling problem, where the data to be labeled cannot be observed directly, however indirect linear measurements are available instead. Therefore the task at hand was to infer an optimal labeling which is a solution of a linear inverse problem. We proposed non-smooth and non-convex variational energies as well as smooth geometric approaches respecting the underlying manifold structure of probabilistic assignments. Hereby, the main focus was discrete tomography reconstruction as a challenging application. Additionally, we assumed a set of prototypes to be given beforehand, a strong prior knowledge, which comprises the *supervised scenario* of image labeling.

In this chapter we are interested in the complementary task, where the input data is given directly but prototypes are not available beforehand and for that have to be inferred from the data, resulting in the *unsupervised scenario* of data labeling. In practice, the availability of prototypes as class representatives is a strong requirement. In many applications either prototypes are not available or it is not clear what prototypes represent the classes properly. A basic remedy is to cluster the data in a preprocessing step. However, such clustering does not properly take into account the framework in which the resulting prototypes are subsequently used for classification. Therefore, we pursue the strategy of simultaneously learning (latent) prototypes in the very same framework while they are used for inference, i.e. assignment of labels.

We have already seen in the previous chapter, that assignment flows (2.89) provide a powerful and flexible framework for supervised image labeling which is fully interpretable from the mathematical viewpoint. Moreover, the compositional structure of the defining vector fields allows modular design derived from the smooth geometric setting of information geometry. This smoothness particularly enables the design of efficient algorithms using

geometric integration. A natural question is therefore, how to generalize the framework of the assignment flow to the unsupervised scenario?

We briefly review the *coupled flow* approach [ZZÅ+18; ZZPS20a] where we took a first step towards extending the supervised assignment flow to the unsupervised setting. The coupled flow combines two geometric flows, the assignment flow with a time varying distance matrix as well as explicit evolution of prototypes to enable the adaption of prespecified prototypes. However, numerical integration on the corresponding feature manifolds can be computationally demanding. Additionally, balancing the relative evolution velocity between the two individual processes defines a free user parameter, and criteria for determining proper values of this balancing parameter for specific applications are not yet available.

In order to alleviate these shortcomings we introduce the concept of *self-assignment* by removing the necessity of explicit prototypes and replacing them by a copy of the given data. Hence we abstract from "pixel-label" decisions to "pixel-pixel" decisions in the absence of prototypes. Instead of directly assigning pixels to labels, we aim to decide for each pair of pixels if they belong to the same cluster or not. These pairwise relationships are encoded by *self-assignment matrices* which are low-rank matrix factorizations. Eventually, a one-parameter family of self-assignment matrices is defined by smooth geodesic interpolation between different normalizations which correspond to two major relaxations of the graph partitioning problem. Specifically, the interpolation marks a trade off between combinatorial aspects (completely positive and doubly stochastic) and spectral aspects (iso-spectral), which also reflects the well-known fact that the positivity and orthogonality constraints are mutually exclusive.

Consequently, we also have to abstract the input data to a pairwise data affinity matrix by considering a similarity measure between all pairs. This includes the basic scenarios of pattern recognition and machine learning: distances between Euclidean feature vectors, Riemannian distances between manifold-valued features and kernel matrices after embedding given feature vectors into a reproducing kernel Hilbert space (RKHS) [HSS08]. Next, we correlate self-assignments with pairwise similarities of the data as a clustering criterion (data term), which has a clear interpretation and allows to exploit the low-rank structure effectively.

Finally, we extend the supervised assignment flow (2.89) and propose the *self-assignment flow* [ZZPS19b; ZZPS20b]. Thereby, only the likelihood map (2.86) is generalized by lifting the gradient of the clustering criterion onto the assignment manifold to achieve maximization of the criterion. Except for this more general definition, the subsequent building blocks of the assignment flow remain unaltered, as do numerical schemes for integration [ZSPS20].

A key parameter is the scale, the size of local neighborhoods in which evolving assignments driven by the flow affect each other. This parameter determines how fine or coarse the resulting partition is, and how many corresponding latent prototypes can be recovered under additional assumptions. At the same time, no bias affects the emergence of latent prototypes. Furthermore, the very same framework is used for both learning these prototypes and subsequent contextual data labeling (classification). In contrast to the earlier coupled approach with explicit prototype evolution, the self-assignment flow is a single process exclusively evolving on the assignment manifold.

In numerical experiments we illustrate the proposed self-assignment flow and compare it to state-of-the art clustering approaches on various basic examples of image analysis. In addition, we demonstrate its plug-in-and-play principle by using a locally invariant patch distance function for unsupervised patch dictionary learning and assignment as well as transfer to novel data.

Organization. The remainder of this chapter is organized as follows:

- We briefly review the preliminary coupled flow to unsupervised data labeling with explicit prototype updates on corresponding feature manifolds (Sect. 5.2).
- We motivate the concept of self-assignment by removing explicit prototypes and replacing them by a copy of the given data to abstract from "pixel-label" decisions to "pixel-pixel" decisions. (Sect. 5.3.1).
- We introduce a family of self-assignment matrices as low-rank matrix factorizations by smooth geodesic interpolation between different normalizations (Sect. 5.3.2). We state basic algebraic properties of the factorizations in (Prop. 5.3.2).
- We give detailed interpretations of the self-assignment matrices and show how latent prototypes emerge and can be recovered (Sect. 5.3.3).
- We state mathematical properties of the set of full-rank assignments (Sect. 5.4.1).
- We propose the family of self-assignment flows (Sect. 5.4.3) by generalising the likelihood map based on the self-assignment matrices (Sect. 5.4.2) and discuss the numerical integration (Sect. 5.4.5) as well as key properties (Sect. 5.4.3 and 5.4.4) of the flows.
- We discuss connections of the approach to related work from three different viewpoints: spectral relaxation (Sect. 5.5.1), discrete optimal transport (Sect. 5.5.2) and matrix factorization with aspects of combinatorial optimization (Sect. 5.5.3).
- Finally, we demonstrate and compare the approach in various numerical experiments (Sect. 5.6), including unsupervised and locally invariant patch learning, assignment and transfer to novel data (Sect. 5.6.4).

5.2. Coupled Flow: Direct Prototype Evolution

This section briefly summarizes our work [ZZÅ+18; ZZPS20a] on the *coupled flow* approach, which extends the assignment flow (2.89) to unsupervised image labeling by learning appropriate features as labels from the data directly. Eventually, the coupled flow has sparked the idea of self-assignment, which is detailed in the consecutive sections by further abstracting from prototypes to pairwise relations between data points. Consequently, prototypes are no longer involved explicitly in the self-assignment flow but implicitly determined simultaneously by a single flow. In contrast, the coupled flow approach couples two geometric flows, the assignment flow and the explicit prototype evolution on the feature manifold, in order to adapt labels to representative feature values during the spatially regularized

assignment process.

Following [ZZPS20a], the generic version of the coupled flow reads

$$\textbf{(CF)} \quad \begin{cases} \dot{W}_i(t) = R_{W_i(t)} S_i(W(t)), & W_i(0) = \mathbb{1}_{\mathcal{S}}, \qquad i \in \mathcal{I}, \\ \dot{f}_j^*(t) = \alpha \sum_{i \in \mathcal{I}} q_{j,i}(W) \operatorname{grad} d_{\mathcal{F}}^2(f_j^*(t), f_i), & f_j^*(0) = f_{0;j}^*, \quad \alpha > 0, \quad j \in \mathcal{J}, \end{cases} \tag{5.1}$$

which is coupling two geometric flows, the assignment flow (first component, $W_i(t)$) and prototype evolution (second component, $f_j^*(t)$), where the user parameter α enables to adjust the relative time scale between the flows. The first component determines spatially regularized assignments $W(t)$ by a slightly generalized version the assignment flow (2.89) which involves a *time-varying* distance matrix depending on the evolving prototypes $f_j^*(t)$ instead of fixed prototypes as the original formulation does. The prototype evolution itself is determined by the second component which represents Riemannian gradient descent flows corresponding to the weighted Riemannian means

$$f_j^* = \operatorname*{arg\,min}_{f \in \mathcal{F}} \sum_{i \in \mathcal{I}} q_{j,i}(W) d_{\mathcal{F}}^2(f, f_i), \qquad j \in \mathcal{J}. \tag{5.2}$$

For each prototype $f_j^*(t)$, the weights q depend on the assignments W, i.e. the evolution of the spatially regularized assignments and prototypes interact. The initial prototypes $f_j^*(0)$ are determined by an approximation of metric clustering which covers the feature space almost uniformly, as described in Section 2.4.3. Since the inference of assignments is separately represented by $W(t) \in \mathcal{W}$, the approach applies to arbitrary manifold-valued data as long as the corresponding Riemannian means are well-defined and computationally feasible.

The weights $q_{j,i}(W)$ determining the Riemannian means (5.2) in turn depend on the assignments $W(t)$. Consequently, each prototype $f_j^*(t)$ evolves towards the average of all data points f_i assigned to it by $W(t)$. More precisely in [ZZPS20a] we proposed a family of weights $q_{j,i}$ which is given by interpolation between two distinguished cases. The first one is motivated by incorporating spatial regularization to the manifold-valued expectation maximization (EM) iteration [MP00; BMDG05] by replacing the mixture coefficients with the pixelwise localized and spatially regularized assignments $W(t)$.

In the second case we consider the column normalized assignments as weights

$$q_{j,i}(W) = \big(C(W)^{-1} W^\top\big)_{j,i}, \qquad j \in \mathcal{J},\, i \in \mathcal{I}, \tag{5.3}$$

where the diagonal matrix $C(W)$ is defined in Eq. (5.8a). This choice is inspired by the so called soft-assignments used in manifold-valued soft-k-means clustering [Teb07] to perform mean-shift fixed point updates of the prototypes, which are induced by the optimality conditions. Interestingly, the very same weights appear in the recovery of latent prototypes derived in Section 5.3.3.2, which are also defined as Riemannian means over all data points weighted by the normalized assignments for each cluster. Hence this version of the coupled flow and the self-assignment flow with family parameter $s = 0$ are related in the formation of (latent) prototypes. However, the numerical integration of the coupled flow (5.1) involves explicit numerical updates on the corresponding feature manifold, which can be computationally demanding. Additionally, how fast the two individual processes should evolve relative to each other defines a free user parameter, and criteria for determining proper values of this balancing parameter for specific applications are not yet available. In

contrast, the self-assignment flow, proposed in the consecutive sections, is a single flow of assignments which allows to perform numerical integration by efficient and sparse updates schemes on the assignment manifold, detailed in Section 5.4.5.

5.3. Self-Assignment

This section prepares the generalization of the assignment flow (2.89) from supervised labeling to completely unsupervised labeling. The prototypes (2.68) no longer are involved but implicitly determined simultaneously in contrast to the coupled flow detailed in the previous Section 5.2 which involves explicit evolution of prototypes on the feature manifold. Here, our approach is (i) to assign given data (2.67) to itself in terms of a self-affinity matrix parameterized by the assignment matrix $W \in \mathbb{R}^{n\times c}$ (Section 5.3.2), which ensures computational feasibility since $c \ll n$, and (ii) to generalize later on the likelihood map (2.86) accordingly (Section 5.4.2). Except for this more general definition of the likelihood map $L(W)$, the subsequent similarity map $S(W)$ given by (2.88) remains unaltered, as do numerical schemes for integrating the flow (2.89) [ZSPS20].

In fact, we define a one-parameter family of self-assignment matrices by geodesic interpolation of two extreme points on the positive definite manifold, that admit natural probabilistic interpretations of the corresponding self-assignments. Properties of these matrices also provide the basis for the interpretation of resulting self-assignment flows (Section 5.3.3) and for pointing out connections to related work (Section 5.5).

5.3.1. From Labeling to Partitioning

Since the prototypes $\mathcal{F}_*$ are unknown, we replace them by the given data $\mathcal{F}_n$, to abstract from "data-label" relations to "data-data" relations. Along with $\mathcal{F}_n$ and the underlying graph $\mathcal{G} = (\mathcal{I}, \mathcal{E})$, we assume a weighted similarity matrix

$$K_{\mathcal{F}} \in \mathbb{S}^n \tag{5.4}$$

to be given with entries

$$K_{\mathcal{F};i,k} = (K_{\mathcal{F}})_{i,k} = k_{\mathcal{F}}(f_i, f_k), \qquad i,k \in \mathcal{I} \tag{5.5}$$

measuring the similarity of data points f_i, f_k in terms of a symmetric function $k_{\mathcal{F}}$. A high similarity value indicates that two data points are likely to be associated with the same cluster. Consequently, defining a suitable similarity is important for practical application, because it encodes aspects of the feature space topolgy which is how the approach sees the data. The matrix $K_{\mathcal{F}}$ is positive definite if $k_{\mathcal{F}}$ evaluates the inner product of a data embedding into a corresponding reproducing kernel Hilbert space (RKHS) space [HSS08]. A basic example is a Euclidean feature space $(\mathcal{F}, d_{\mathcal{F}})$ with norm $d_{\mathcal{F}}(f_i, f_k) = \|f_i - f_k\|$ and

$$k_{\mathcal{F}}(f_i, f_k) = e^{-d_{\mathcal{F}}(f_i,f_k)^2/\sigma^2}. \tag{5.6}$$

Let $W \in \mathcal{W}_*^c$ be a labeling. The column vectors W^j, $j \in \mathcal{J}$, of W indicate which data points f_i are assigned to the j-th cluster $\mathcal{I}_j$ corresponding to the partition

$$\{\mathcal{I}_j\}_{j=1}^c \quad \text{with} \quad \mathcal{I} = \dot{\bigcup_{j\in\mathcal{J}}} \mathcal{I}_j \quad \text{and} \quad \mathcal{I}_j \cap \mathcal{I}_l = \emptyset \tag{5.7}$$

of the vertices $\mathcal{I}$ representing data points $\mathcal{F}_n$. Let us define the diagonal matrix

$$C(W) := \operatorname{Diag}(W^\top \mathbb{1}_n) = \operatorname{Diag}(n_1, \ldots, n_c) \in \mathbb{S}_+^c \tag{5.8a}$$

with the cardinalities of each cluster

$$n_j := |\mathcal{I}_j|, \qquad j \in \mathcal{J} \tag{5.8b}$$

as entries. The quadratic form

$$\frac{1}{2}\langle W^j, K_{\mathcal{F}} W^j\rangle = \frac{1}{2}\sum_{i,k\in\mathcal{I}} k_{\mathcal{F}}(f_i, f_k) W_{i,j} W_{k,j} = \frac{1}{2}\sum_{i\in\mathcal{I}_j} k_{\mathcal{F}}(f_i, f_i) + \sum_{i,k\in\mathcal{I}_j\colon i\neq k} k_{\mathcal{F}}(f_i, f_k) \tag{5.9}$$

measures the *size* of cluster $\mathcal{I}_j$ by the first sum of the right-hand side, which for common kernel functions like (5.6) is proportional to the number n_j of data points assigned to cluster j. The second sum measures the *connectivity* in terms of the weights $k_{\mathcal{F}}(f_i, f_k)$ of all edges $ik \in \mathcal{E}$ connecting points i and k in this cluster. Assuming that all clusters are non-empty, which amounts to the assumption

$$\operatorname{rank}(W) = c, \tag{5.10}$$

we normalize the preceding expression by the cluster cardinality and subsequently sum over all clusters, to obtain

$$\sum_{j\in\mathcal{J}} \frac{1}{2n_j}\langle W^j, K_{\mathcal{F}} W^j\rangle = \frac{1}{2}\sum_{j\in\mathcal{J}} \frac{1}{n_j}\sum_{i\in\mathcal{I}_j} k_{\mathcal{F}}(f_i, f_i) + \sum_{j\in\mathcal{J}} \frac{1}{n_j}\sum_{i,k\in\mathcal{I}_j\colon i\neq k} k_{\mathcal{F}}(f_i, f_k) \tag{5.11a}$$

$$= \frac{1}{2}\sum_{j\in\mathcal{J}} \frac{1}{n_j}(W^\top K_{\mathcal{F}} W)_{j,j} \overset{(5.8a)}{=} \frac{1}{2}\operatorname{tr}\left(C(W)^{-1} W^\top K_{\mathcal{F}} W\right) \tag{5.11b}$$

$$= \frac{1}{2}\operatorname{tr}\left(K_{\mathcal{F}} A_0(W)\right), \tag{5.11c}$$

with

$$A_0(W) = W C(W)^{-1} W^\top, \qquad W \in \mathcal{W}_*^c. \tag{5.11d}$$

For common kernel functions like (5.6), the first sum on the right-hand side of (5.11a) is a constant. Objective (5.11c) therefore essentially measures the normalized similarity weights *not* cut by the partition of the underlying graph.

Thus, the problem of partitioning the data and the underlying graph into c clusters takes the form

$$\max_W \operatorname{tr}\left(K_{\mathcal{F}} A_0(W)\right) \qquad \text{subject to} \qquad W \in \mathcal{W}_*^c. \tag{5.12}$$

We record basic properties of the matrix $A_0(W)$.

Lemma 5.3.1. *Let $W \in \mathcal{W}_*^c$. Then the matrix $A_0(W)$ given by (5.11d) is*

(a) nonnegative and symmetric,

(b) doubly stochastic,

$$A_0(W)\mathbb{1}_n = A_0(W)^\top \mathbb{1}_n = \mathbb{1}_n, \tag{5.13}$$

(c) and completely positive,

$$A_0(W) = YY^\top, \qquad Y \geq 0. \tag{5.14}$$

Proof. (a) is immediate. (b) follows from (5.8a) and the constraint $W \in \mathcal{W}_*^c$ (recall (2.71)). (c) holds with $Y = Y(W) = WC(W)^{-1/2}$. □

Property (c), i.e. a completely positive factorization of the matrix $A_0(W)$ depending on W, reflects the combinatorial difficulty of the optimization problem (5.12) – see, e.g., [BS18;

Bom18] and references therein for more information about completely positive matrix factorization. Therefore, various relaxations of the constraint $W \in \mathcal{W}_*^c$ are discussed next.

5.3.2. Self-Assignment Matrices, Relaxation

We start with the definitions of two basic self-assignment matrices. The first relaxation based on (5.11d) drops both the integrality constraint and the rank constraint.

Definition 5.3.1 (**Self-Affinity Matrix**)**.** The *self-affinity matrix* is defined as the factorization

$$A_0(W) := WC(W)^{-1}W^\top, \qquad W \in \mathcal{W}. \tag{5.15}$$

The second definition is based on the observation that equivalent expressions for the normalizing matrix

$$C(W) = W^\top W \qquad \text{if} \qquad W \in \mathcal{W}_*^c \tag{5.16}$$

differ after relaxing the feasible set $\mathcal{W}_*^c$. The equality (5.16) constitutes a generalized orthogonality constraint which is equivalent to the integrality constraint. Dropping the integrality constraint but keeping the rank constraint yields full-rank assignment matrices

Definition 5.3.2 (**Full-Rank Assignments**)**.** The the set of *full-rank assignment matrices* with column-rank $c \leq n$ is defined as

$$\mathcal{W}^c := \big\{W \in \mathcal{W}\colon\ \operatorname{rank}(W) = c\big\}, \tag{5.17}$$

as admissible assignments such that the normalizing matrix $W^\top W$ is non-singular. This gives the following definition.

Definition 5.3.3 (**Self-Influence Matrix**)**.** The *self-influence matrix* is defined as the factorization

$$A_1(W) := W(W^\top W)^{-1}W^\top, \qquad W \in \mathcal{W}^c. \tag{5.18}$$

Definitions 5.3.1 and 5.3.3 differ by the normalizing matrices $C(W)$ and $W^\top W$, both of which are positive definite. It is then natural to define a one-parameter family of factorized matrices in terms of a geodesic connecting $C(W)$ and $W^\top W$ on the manifold of positive definite matrices $\mathcal{P}^c$, which gives rise to the following definition.

Definition 5.3.4 (**Self-Assignment Matrix**)**.** The *self-assignment matrix* with parameter s is defined as the factorization

$$A_s(W) := W\gamma_s(W)^{-1}W^\top, \qquad s \in [0,1], \qquad W \in \begin{cases} \mathcal{W}, & \text{if } s = 0, \\ \mathcal{W}^c, & \text{if } s > 0, \end{cases} \tag{5.19a}$$

with normalizing matrix

$$\gamma_s(W) = C(W)^{\frac{1}{2}}\big(C(W)^{-\frac{1}{2}}W^\top W C(W)^{-\frac{1}{2}}\big)^s C(W)^{\frac{1}{2}} \ \in\ \mathcal{P}^c. \tag{5.19b}$$

Note that Definition 5.3.4 corresponds to Definition 5.3.1 and 5.3.3 if $s = 0$ and $s = 1$, respectively. Figure 5.1 depicts the relations between members of the family of self-assignment matrices (Definition 5.3.4).

The following proposition collects properties of self-assignment matrices defined above. Property (h) refers to a relation between matrices $A_1\big(W(t)\big)$ and $A_1\big(W(t')\big)$; for any $t, t' \in [0,T]$, they share the same eigenvalues.

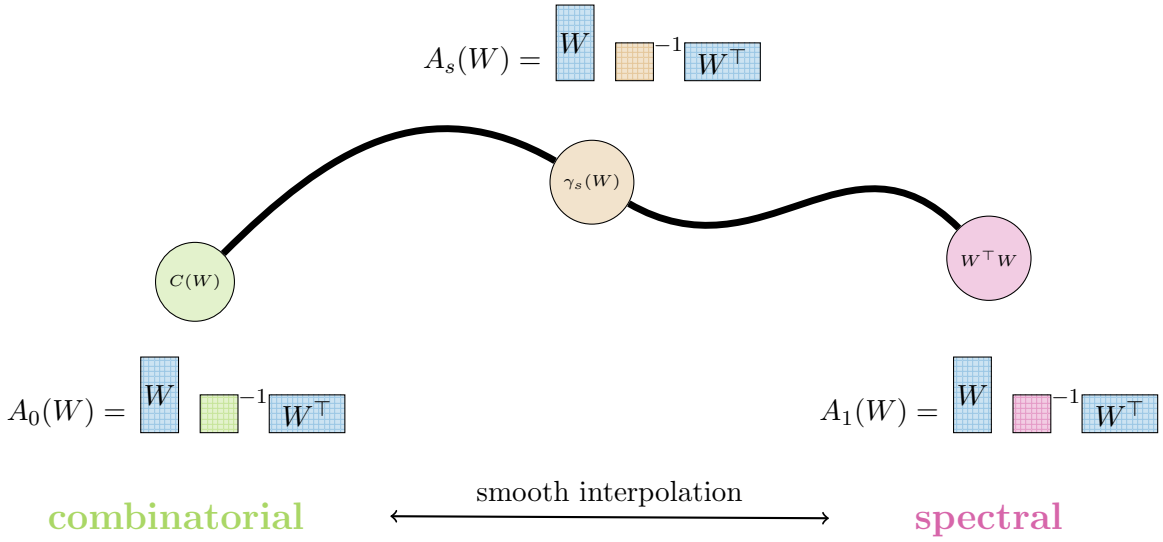

Figure 5.1.: Schematic overview of the one-parameter family of self-assignment matrices $A_s(W)$ (Definition 5.3.4) which is defined by geodesic interpolation on the positive definite matrix manifold between $C(W)$, normalizing the self-affinity matrix $A_0(W)$ and $W^\top W$, normalizing the self-influence matrix $A_1(W)$. The single parameter s parameterizes the smooth interpolation between combinatorial and spectral properties, summarized in Proposition 5.3.2.

Proposition 5.3.2 (Properties of Self-Assignment Matrices)**.** *Let $A_0(W)$ and $A_1(W)$ be given by Definition 5.3.1 and 5.3.3, respectively. Then these matrices have (✓) or do not have (✗) the following properties.*

		self-affinity $A_0(W)$	***self-influence*** $A_1(W)$
	admissible assignments	$W \in \mathcal{W}$	$W \in \mathcal{W}^c$
(a)	*symmetric*	✓	✓
(b)	*positive semi-definite*	✓	✓
(c)	*nonnegative*	✓	✗
(d)	*doubly stochastic*	✓	✗
(e)	*completely positive*	✓	✗
(f)	*rank*	$\leq c$	$= c$
(g)	*orthogonal projection*	✗	$\Pi_{\mathcal{R}(W)}$
(h)	*iso-spectral*	✗	✓
(i)	*eigenvalues* $\in$	$[0,1]$	$\{0,1\}$
(j)	*multiplicity* $(\lambda = 1)$	$= 1$	$= c$
(k)	*multiplicity* $(\lambda = 0)$	$\geq n - c$	$= n - c$
(l)	*eigenvector(s)* $(\lambda = 1)$	$\mathbb{1}_n$	$\big(W(W^\top W)^{-\frac{1}{2}}\big)^j, \quad j \in \mathcal{J}$

Proof. (a)-(f) are clear. We focus on (g)-(l).

(g) One easily checks that $A_1(W) = A_1(W)^2$ is idempotent whereas $A_0(W)$ is not. Taking into account (a) implies the assertion for $s = 1$.

(h) Follows from (i) and (j) for $s = 1$.

(i) Case $s = 0$: The lower eigenvalue bound 0 follows from (a),(b), the upper bound 1 from (d) and [BP94, Thm. 5.3]. Case $s = 1$: This is immediate due to (g).

(j) Case $s = 0$: $W \in \mathcal{W}$ implies that $A_0(W)$ is strictly positive. (i) and [BP94, Thm. 1.4] then imply the assertion. Case $s = 1$: This is immediate due to (f),(g).

(k) Both assertions follow from (f).

(l) Case $s = 0$: Follows from (d) and [BP94, Thm. 5.3]. Case $s = 1$: Setting $Y = W(W^\top W)^{-1/2}$, one directly computes $A_1(W)Y = Y$ and $Y^\top Y = I_c$.

□

The last definition of this section concerns the 'difference' between the normalizing matrices $C(W)$ and $W^\top W$ of Definitions 5.3.1–5.3.4.

Definition 5.3.5 (Cluster-Confusion Matrix). The *cluster-confusion matrix* is defined as the matrix factorization

$$B(W) := C(W)^{-1} W^\top W \in \mathbb{R}_+^{c \times c}, \qquad W \in \mathcal{W}. \tag{5.20}$$

Proposition 5.3.3 (Properties of the Cluster-Confusion Matrix)**.** *The cluster-confusion matrix $B(W)$ has the following properties:*

(a)	*entry-wise positive:*	$B(W) > 0$	
(b)	*row stochastic:*	$B(W)\mathbb{1}_c = \mathbb{1}_c$	
(c)	*pure clusters:*	$B(W) = I_c$ *if and only if*	$W \in \mathcal{W}_*^c$.
(d)	*rank lower bound:*	$0 \leq \operatorname{tr}\big(B(W)\big) \leq \operatorname{rank}(W)$ *with equality if*	$W \in \mathcal{W}_*^c$

Proof. (a)-(c) directly follow from the definitions of $B(W)$ and W_*^c. (d) follows from $\operatorname{tr}(B(W)) = \operatorname{tr}(A_0(W))$ together with Proposition 5.3.2 (c) and (i). □

5.3.3. Relaxations: Interpretation

We take a closer look at the relaxations of problem (5.12).

5.3.3.1. Self-Affinity Matrix

Following [ÅPSS17], we interpret each entry of the assignment matrix $W \in \mathcal{W}$ as posterior probability

$$P(j|i) = W_{i,j}, \qquad j \in \mathcal{J}, \quad i \in \mathcal{I} \tag{5.21}$$

of label j, conditioned on the observation of the data point f_i. In accordance with the present unsupervised scenario, we assume the uniform prior distribution

$$P(i) = \frac{1}{n}, \quad i \in \mathcal{I} \tag{5.22}$$

of the data. Marginalization yields the label distribution

$$P(j) = \sum_{i \in \mathcal{I}} P(j|i)P(i) = \frac{1}{n}\big(W^\top \mathbb{1}_n\big)_j, \tag{5.23}$$

which measures the size of cluster $\mathcal{I}_j$ in terms of the relative mass of assignments. Invoking Bayes' rule, we compute the distribution analogous to (5.21), but with the roles of data and labels reversed, to obtain

$$Q(k|j) = \frac{P(j|k)P(k)}{P(j)} = \frac{W_{k,j}}{\sum_{i \in \mathcal{I}} W_{i,j}} = \big(C(W)^{-1}W^\top\big)_{j,k}. \tag{5.24}$$

The probabilities of self-assignments $f_i \leftrightarrow f_k$, $i, k \in \mathcal{I}$ then result from marginalization over the labels

$$A_{0;i,k}(W) := \sum_{j \in \mathcal{J}} Q(k|j)P(j|i) = \sum_{j \in \mathcal{J}} W_{i,j}\big(C(W)^{-1}W^\top\big)_{j,k} = \big(WC(W)^{-1}W^\top\big)_{i,k}. \tag{5.25}$$

This expression explains the relaxation that is at the basis of Definition 5.3.1. It specifies the probability that two vertices i and k get assigned the same label (no matter which one), i.e. that they belong to the same cluster.

Finally, the derivation of problem (5.12) – cf. (5.11) – showed that optimizing the assignments in order to maximize the correlation (inner product) of $A_0(W)$ and $K_\mathcal{F}$ amounts to cover the most similar data points by the components of the partition (clusters).

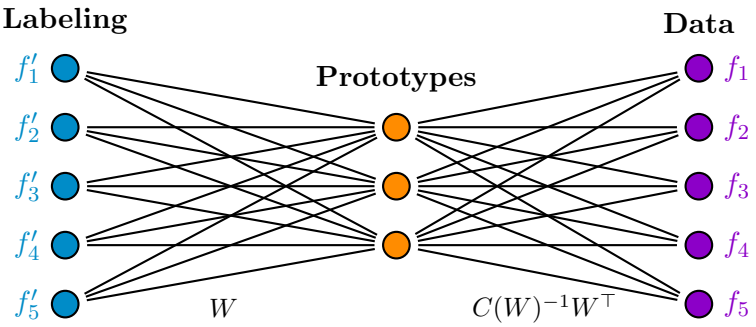

Figure 5.2.: The self-affinity matrix $A_0(W)$ due to Definition 5.3.1 comprises the probabilities for each pair of data points $f_i, f_k \in \mathcal{I}$ to belong to the same cluster. The factorization (5.25) of $A_0(W)$ admits the interpretation that optimizing the assignments W implicitly forms prototypes f^*_j, $j \in \mathcal{J}$ that are assigned to the data themselves so as to maximize the correlation with pairwise similarities given as entries of the matrix $K_\mathcal{F}$.

5.3.3.2. Recovery of Latent Prototypes

Although problem (5.12) does not involve prototypes (2.68), such prototypes can be recovered from the solution W to the problem relaxation discussed in Section 5.3.3.1. Specifically, the probabilities $Q(i|j)$ given by (5.24) indicate the contribution of each data point f_i to cluster j. Consequently, adopting the manifold assumption that the data $\mathcal{F}_n$ are sampled on a Riemannian manifold, prototypes can be recovered as weighted Riemannian means by solving

$$f^*_j = \operatorname*{arg\,min}_{f \in \mathcal{F}} \sum_{i \in \mathcal{I}} \big(C(W)^{-1}W^\top\big)_{j,i} d^2_\mathcal{F}(f, f_i), \qquad j \in \mathcal{J}. \tag{5.26}$$

In the basic case of Euclidean data $\mathcal{F}_n \subset \mathbb{R}^d$, this problem yields the closed form averages

$$f_j^* = \sum_{i \in \mathcal{I}} \big(C(W)^{-1} W^\top\big)_{j,i} f_i, \qquad j \in \mathcal{J}. \tag{5.27}$$

Figure 5.2 illustrates the data self-assignment via the self-affinity matrix and latent prototypes.

Remark 5.3.1. Reconvering latent prototypes from integral assignments $W \in \mathcal{W}_*^c$ results in Riemannian means with uniform weights $Q(i|j) = \frac{1}{n_j}$, normalization by the cluster sizes n_j. That means each data point of a cluster contributes equally to its representative.

Remark 5.3.2 (Relation and Differences to Basic Clustering)**.** Choosing the squared Euclidean norm $d_{\mathcal{F}}^2(f, f_i) = \|f - f_i\|^2$ in Eq. 5.26 determines the prototype f_j^*, like k-means clustering, as arithmetic mean (5.27) of the data f_i assigned to cluster $j \in \mathcal{J}$ by the variables $\big(C(W)^{-1}W^\top\big)_{j,i}$. However, unlike k-means clustering and its variants that alternatingly update prototypes and assignment variables, the prototypes f_j^*, $j \in \mathcal{J}$ are *not explicitly* involved in our self-assignment flow approach. Rather, Eq. 5.27 is evaluated *after* convergence of the self-assignment flow (5.56). Section 5.4.4 reveals that the dependency $f_j^* = f_j^*(W)$ of (5.26) is consistent with graph partitioning through the self-assignment flow $W(t)$, in that prototypes f_j^* that are *implicitly* determined by (5.27) maximize class separability. The usual initialization problem of basic clustering is handled by the self-assignment approach through the initialization $W(0)$ of (5.56a) in terms of the given data. In addition, we point out that basic clustering is lacking the influence of the spatial assignment *regularization* through geometric averaging – cf. (2.88) – on the formation of prototypes.

5.3.3.3. Self-Influence Matrix

Let $W \in \mathcal{W}^c$ be given and temporarily assume that d-dimensional Euclidean feature vectors are given as data $\mathcal{F}_n$ and collected as row vectors in the matrix

$$F = (f_1, \ldots, f_n)^\top \in \mathbb{R}^{n \times d}. \tag{5.28}$$

Let the matrix

$$F^* = (f_1^*, \ldots, f_c^*)^\top \in \mathbb{R}^{c \times d} \tag{5.29}$$

collect the prototypes. Given W and F, a least-squares fit yields

$$F^* = \operatorname*{arg\,min}_{G \in \mathbb{R}^{c \times d}} \frac{1}{2} \|WG - F\|_F^2 = (W^\top W)^{-1} W^\top F, \tag{5.30}$$

which is well-defined since $W \in \mathcal{W}^c$ has full rank. Using these prototypes in turn for predicting data $\hat{F}$ by assignment yields

$$\hat{F} = WF^* = W(W^\top W)^{-1} W^\top F = \Pi_{\mathcal{R}(W)} F = A_1(W) F. \tag{5.31}$$

Finally, optimizing the assignments W in order to obtain the best prediction of the data itself, gives with $A_1(W)^2 = A_1(W)$

$$\operatorname*{arg\,min}_{W \in \mathcal{W}^c} \frac{1}{2} \|A_1(W)F - F\|_F^2 = \operatorname*{arg\,max}_{W \in \mathcal{W}^c} \operatorname{tr}\big(A_1(W) F F^\top\big), \tag{5.32}$$

and the initial assumption of Euclidean data can be dropped by replacing the Euclidean Gram matrix $FF^\top$ with a general inner product matrix $K_{\mathcal{F}}$ corresponding to an embedding of the data into a reproducing kernel Hilbert space.

As a result, the relaxation of problem (5.12) due to Definition 5.3.3 can be interpreted as finding the best c-dimensional subspace $\mathcal{R}(W)$ spanned by the (soft) indicator vectors of the c clusters (column vectors of W) for self-prediction of the given data.

A related spectral interpretation results from rewriting the objective as

$$\operatorname{tr}\big(A_1(W)K_{\mathcal{F}}\big) = \operatorname{tr}\big(W(W^\top W)^{-1}W^\top K_{\mathcal{F}}\big) = \operatorname{tr}\big((W^\top W)^{-\frac{1}{2}}W^\top K_{\mathcal{F}}W(W^\top W)^{-\frac{1}{2}}\big) \tag{5.33a}$$

$$= \operatorname{tr}\big(Y(W)^\top K_{\mathcal{F}}Y(W)\big), \qquad Y(W) = W(W^\top W)^{-\frac{1}{2}}. \tag{5.33b}$$

We conclude from Proposition 5.3.2 that $Y(W)$ varies over the compact Stiefel manifold,

$$Y(W) \in \mathrm{St}(c,n) = \{X \in \mathbb{R}^{n\times c} \colon X^\top X = I_c\}, \tag{5.34}$$

and that the objective (5.33) is the Rayleigh quotient [AMS09, Sec. 4.8] whose maximizer Y spans the subspace of the c dominant eigenvectors of $K_{\mathcal{F}}$ [HM96, Ch. 1]. Note, however, that $Y(W)$ cannot vary freely but is parameterized by $W \in \mathcal{W}^c$.

5.3.3.4. Comparision between $A_0(W)$ and $A_1(W)$.

The Factorization $A_1(W)$ differs from $A_0(W)$ in that the normalizing matrix $C(W)$ of the former self-assignment matrix is replaced by $W^\top W$ in the latter. A consequence due to Proposition 5.3.2 is that $A_1(W)$ is no longer doubly stochastic and may have negative entries. Hence the probabilistic interpretation (5.25) of the factorization of $A_0(W)$ no longer holds for $A_1(W)$. Since $A_0(W)$ has strictly positive entries, it can be seen as an irreducible adjacency matrix and the underlying graph has only one connected component. On the other hand, unlike $A_0(W)$, the matrix $A_1(W)$ has fixed rank c and embeds the data immediately into a corresponding subspace.

Formulas (5.27) and (5.30) for the formation of latent prototypes (Euclidean case) are the same when using $A_0(W)$ or $A_1(W)$, up to the different normalizing matrices. How these prototypes are used to represent the data is made explicit by Figure 5.2 and Eq. (5.31), respectively. In case of integral assignments $W \in \mathcal{W}_*^c$ both matrices $A_0(W)$ and $A_1(W)$ are equivalent. What labelings are computed, however, depends on the self-assignment flow (Section 5.4) and hence on the parameter $s \in [0,1]$

With help of the cluster-confusion matrix $B(W)$, the self-influence matrix $A_1(W)$ can be rewritten as

$$A_1(W) = WB(W)^{-1}C(W)^{-1}W^\top, \tag{5.35}$$

which resembles the structure of $A_0(W)$. In view of (5.27), the latent prototypes are additionally transformed by the inverse of $B(W)$.

5.3.3.5. Cluster-Confusion Matrix

Using (5.21) and (5.24), the entries of the cluster-confusion matrix (5.20) take the form

$$B_{j,l}(W) := \sum_{i\in\mathcal{I}} P(l|i)Q(i|j) = \big(C(W)^{-1}W^\top W\big)_{j,l}, \qquad j,l \in \mathcal{J}, \tag{5.36}$$

which is a marginalization over the vertices. This expression may be interpreted as the probability that clusters $\mathcal{I}_j$ and $\mathcal{I}_l$ are connected (soft partition) as opposed to the case of integral assignments $W \in \mathcal{W}_*^c$, in which case $B(W) = I_c$ and all clusters are disjoint (hard partition).

5.4. Self-Assignment Flows

In this section, we generalize the assignment flow (2.89) to the unsupervised scenario discussed in Section 5.3. We explore mathematical properties of the submanifold of full-rank assignments (Section 5.4.1) to determine a smooth geometric setting that the resulting unsupervised flows are well-defined. Generalizing the likelihood map (2.86) is the major step (Section 5.4.2). The remaining components of the assignment flow remain unchanged, except for starting the flow at the perturbed barycenter of the assignment manifold in order to break the symmetry through the data, in the absence of prototypes (Section 5.4.3). Next, we complement in Section 5.4.4 the interpretations of the relaxations underlying the self-assignment flow (Section 5.3.3) and show that the latent prototypes determined by the flow maximize class separability. Finally, numerical aspects are discussed in Section 5.4.5.

5.4.1. Riemannian Submanifold of Full-Rank Assignments

Here we collect important mathematical aspects of full-rank assignments $\mathcal{W}^c$, Definition 5.3.2 in order to determine a smooth geometric setting, in which the unsupervised generalization of the assignment flow can be described properly. The set of full-rank assignments is equivalently characterized as intersection of the assignment manifold with the non-compact Stiefel manifold

$$\mathcal{W}^c = \mathcal{W} \cap \mathbb{R}_*^{n \times c}. \tag{5.37}$$

Hence $\mathcal{W}^c$ is an open subset of $\mathcal{W}$, since it is the intersection of open sets. Proposition 5.4.1 states the fact, that $\mathcal{W}^c$ is a Riemannian submanifold of the assignment manifold $\mathcal{W}$ endowed with the Fisher-Rao product metric.

Proposition 5.4.1. *The open subset $\mathcal{W}^c \subset \mathcal{W}$ is an (embedded) Riemannian submanifold of the assignment manifold $\mathcal{W}$ equipped with the Fisher-Rao product metric. In particular the tangent bundle is trivial and we have $T_W \mathcal{W}^c = \mathcal{T}_0$.*

Proof. [Lee13, Prop. 5.1] □

The assignment manifold $\mathcal{W}$ is connected, however the full-rank assignment submanifold $\mathcal{W}^c$ is defined by intersection and hence its connectedness is not immediate. Since we are interested in integration over the Riemannian submanifold $\mathcal{W}^c$ we require $\mathcal{W}^c$ to have a single connected component, otherwise some points of $\mathcal{W}^c$ may not be reachable from arbitrary initialization.

In case of strictly less labels c than data points n, which is relevant for our application, Proposition 5.4.2 answers the question of connectedness positively.

Proposition 5.4.2. *The smooth manifold $\mathcal{W}^c$ is path-connected for $c < n$.*

Proof. The proof relies on the basis exchange Lemma A.1.1 and on the fact, that in case of $c < n$ there exists at least one row of $W \in \mathcal{W}^c$ which can be used temporarily to swap two rows by a continuous path with constant rank. For the complete proof see Appendix A.1. □

For the sake of completeness, we show in Proposition 5.4.3 that the submanifold $\mathcal{W}^c$ is not connected for $c = n$.

Proposition 5.4.3. *The smooth manifold $\mathcal{W}^n$ is not connected.*

Proof. We use the fact, that the general linear group $\mathrm{GL}(n)$ has two connected components and construct two disconnected points in $\mathcal{W}^n \subset \mathrm{GL}(n)$. For the complete proof see Appendix A.1. □

Since $\mathcal{W}^c$ is an open submanifold of the assignment manifold, the tangent spaces are coinciding and hence the manifolds are of equal dimension (Proposition 5.4.1). Therefore we can directly use the same geometric setting, the dually flat structure borrowed from information geometry by using the Fisher-Rao product metric together with the e-connection. The corresponding exponential maps and their inverses are defined in Section 2.3.1, eq. (2.80), by restricting to the submanifold. However, we must be aware that geodesics from $W \in \mathcal{W}^c$ in direction $V \in \mathcal{T}_0$, i.e.

$$\gamma_{W,V}(t) = \mathrm{Exp}_W(tV) \in \mathcal{W}^c, \tag{5.38}$$

are only defined in a local vicinity of the starting point W. For larger $t > 0$ we may leave the submanifold, which means that the assignment matrix has rank decay. We will get back to this point when integrating (see Section 5.4.5) the self-assignment flow, which is derived in the following sections.

5.4.2. Generalized Likelihood Map

In the supervised case, for a given distance matrix $D_{\mathcal{F}}$ (2.84), *local* label assignment is simply achieved by determining separately the smallest component of the vectors $D_{\mathcal{F};i}$, for every vertex $i \in \mathcal{I}$. This corresponds to solving

$$\min_{W \in \mathcal{W}} \mathrm{tr}(D_{\mathcal{F}} W^\top) \tag{5.39}$$

and the likelihood map (2.86) lifts the scaled negative *gradient* of this objective function component-wise to $\mathcal{W}$. In view of problem (5.12) and the family of self-assignment matrices due to Definition 5.3.4, a natural approach to generalize this supervised set-up to the unsupervised case is to consider a clustering criterion instead, which is defined in Definition 5.4.1. Then, the local decisions based on $-D_{\mathcal{F}}$ in the likelihood map are replaced by the gradient $\nabla E_s(W)$ of the clustering objective.

Definition 5.4.1 (Clustering Criterion). Let $K_{\mathcal{F}} \in \mathbb{S}^n$ be a weighted similarity matrix and $A_s(W)$ a self-assignment matrix for $s \in [0, 1]$. Then, the *clustering criterion* is defined as the following optimization problem:

$$\max_W \; E_s(W) \quad \text{subject to} \quad W \in \begin{cases} \mathcal{W}, & \text{if } s = 0 \\ \mathcal{W}^c, & \text{if } s \in (0, 1] \end{cases} \tag{5.40a}$$

$$E_s(W) = \mathrm{tr}\left(K_{\mathcal{F}} A_s(W)\right), \tag{5.40b}$$

where $E_s(W)$ is the *clustering objective*.

The corresponding gradients of the clustering objective, Definition 5.4.1, are stated in the following Proposition 5.4.4 for $s = 0$ and $s = 1$.

Proposition 5.4.4 (Euclidean Gradient of $E_s(W)$)**.** *Let $E_s(W)$ be given as in (5.40b). Then, the Euclidean gradient of $E_s(W)$ reads:*

$s = 0$:

$$\nabla E_0(W) \;=\; 2K_{\mathcal{F}}WC(W)^{-1} - \mathbb{1}_n \operatorname{diag}\big(C(W)^{-1}W^\top K_{\mathcal{F}}WC(W)^{-1}\big)^\top, \tag{5.41a}$$

$s = 1$:

$$\nabla E_1(W) \;=\; 2\big(I_n - A_1(W)\big)K_{\mathcal{F}}W(W^\top W)^{-1}. \tag{5.41b}$$

Proof. The calculation of the matrix valued gradients of (5.40b) for $s = 0$ and $s = 1$ is straightforward with help of the directional derivative in direction V, i.e.

$$dE_s(W)[V] = \frac{d}{dt}E_s(W + tV)\big|_{t=0} = \operatorname{tr}(\nabla E_s(W)V^T), \qquad \forall V \in \mathbb{R}^{n\times c}. \tag{5.42}$$

□

In order to substantiate this approach, we interpret the gradients from Proposition 5.4.4 using the concepts from Section 5.3.3. For illustration, let $K_{\mathcal{F}} = FF^\top$ be a Euclidean inner product matrix, with F given by (5.28). Equation (5.27) determining the latent prototypes as averages weighted by the likelihood $Q(i|j)$, defined in Equation (5.24), reads

$$f_j^* = \sum_{i\in\mathcal{I}} \big(C(W)^{-1}W^\top\big)_{j,i} f_i = \big(C(W)^{-1}W^\top F\big)_j, \qquad (F^*)^\top = F^\top WC(W)^{-1}. \tag{5.43}$$

From Proposition 5.4.4 with $s = 0$ we have

$$\nabla E_0(W) = 2FF^\top WC(W)^{-1} - \mathbb{1}_n \operatorname{diag}\big((F^\top WC(W)^{-1})^\top F^\top WC(W)^{-1}\big)^\top \tag{5.44a}$$

$$= 2F(F^*)^\top - \mathbb{1}_n \operatorname{diag}(F^*(F^*)^\top)^\top, \tag{5.44b}$$

$$\big(\nabla E_0(W)\big)_{i,j} = 2\langle f_i, f_j^*\rangle - \|f_j^*\|^2 = -\|f_i - f_j^*\|^2 + \|f_i\|^2, \tag{5.44c}$$

where the prototypes $f_j^* = f_j^*(W)$ depend on W. The last term on the right-hand-side of (5.44c) does not depend on j and hence is factored out – cf. Remark 2.3.1 – when lifting the vector (5.44c) to the assignment manifold as follows. Hence, we ignore this term and generalize the likelihood map (2.86) to

$$L_{0;i}(W_i) = \exp_{W_i}\Big(\frac{1}{\rho}\nabla E_0(W)_i\Big) = \exp_{W_i}\Big(-\frac{1}{\rho}\big(\|f_i - f_j^*\|^2\big)_{j\in\mathcal{J}}\Big), \tag{5.45}$$

which amounts to replacing the distance vectors $D_{\mathcal{F};i}$, for *given* prototypes in the supervised case, by a *varying* squared distance depending on *latent* prototypes, which emerge when the assignments $W(t)$ follow the assignment flow.

A second relation is stated in Proposition 5.4.5 and Corollary 5.4.6 to motivate Definition 5.4.1, where we still assume the Euclidean setting with $K_{\mathcal{F}} = FF^\top$. In the absence of prototypes $\mathcal{F}_*$ we replace them by a copy of the input data $\mathcal{F}$ and construct a quadratic distance matrix $D_{\mathcal{F},\mathcal{F}}$ according to (2.84).

Proposition 5.4.5 (Kernel Matrix and Distance Matrix)**.** *Let $K_{\mathcal{F}} = FF^\top \in \mathbb{S}^n$ be a linear kernel matrix with F given by (5.28) and $A_0(W)$ a self-affinity matrix, defined in Definition 5.3.4. Then, we have the equality:*

$$\frac{1}{2}\operatorname{tr}\big(A_0(W)D_{\mathcal{F},\mathcal{F}}\big) = \operatorname{tr}\big(F^\top(I - A_0(W))F\big) = \operatorname{tr}\big((I - A_0(W))K_{\mathcal{F}}\big) \tag{5.46}$$

Proof. A straightforward calculation shows the

$$\frac{1}{2}\operatorname{tr}\big(A_0(W)D_{\mathcal{F},\mathcal{F}}\big) = \frac{1}{2}\sum_{i,k\in\mathcal{I}} A_0(W)_{i;k}\|F_i - F_k\|_2^2 \tag{5.47a}$$

$$= \frac{1}{2}\sum_{i,k\in\mathcal{I}} A_0(W)_{i;k}\|F_i\|_2^2 - 2A_0(W)_{i;k}\langle F_i, F_k\rangle + A_0(W)_{i;k}\|F_k\|_2^2 \tag{5.47b}$$

$$= \sum_{i\in\mathcal{I}}\|F_i\|_2^2 - \sum_{i\in\mathcal{I}}\langle F_i, \sum_{k\in\mathcal{I}} A_0(W)_{i;k}F_k\rangle \tag{5.47c}$$

$$= \operatorname{tr}\big(FF^\top\big) - \operatorname{tr}\big(F^\top A_0(W)F\big) = \operatorname{tr}\big((I - A_0(W))K_{\mathcal{F}}\big), \tag{5.47d}$$

where the third equality follows from Proposition 5.3.2 (d) with $s = 0$. □

Proposition 5.4.5 together with Definition 5.4.1 immediately imply Corollary 5.4.6.

Corollary 5.4.6. *Let $K_{\mathcal{F}} = FF^\top \in \mathbb{S}^n$ be a linear kernel matrix with F given by (5.28) and $A_0(W)$ a self-affinity matrix, defined in Definition 5.3.4. Then, we have the equivalence:*

$$\operatorname*{arg\,max}_{W\in\mathcal{W}} \operatorname{tr}\big(A_0(W)K_{\mathcal{F}}\big) \quad\Leftrightarrow\quad \operatorname*{arg\,min}_{W\in\mathcal{W}} \operatorname{tr}\big(A_0(W)D_{\mathcal{F},\mathcal{F}}\big). \tag{5.48}$$

This means the clustering criterion for $s = 0$, Definition 5.4.1, can be seen simply as replacing assignments in (5.39) by self-assignments and prototypes by the input data themselves.

Moreover, since this interpretation only depends on the inner product matrix $FF^\top$, it generalizes to data embeddings into a reproducing kernel Hilbert space and a corresponding data affinity matrix $K_{\mathcal{F}}$ with entries (5.5).

Now let $s = 1$. We return to the 'spectral' interpretation in terms of (5.33) and (5.34). Proposition 5.4.7 collects two properties of the clustering objective (5.40b).

Proposition 5.4.7 (Properties of $E_1(W)$)**.** *Let $W \in \mathcal{W}^c$. Then, the clustering objective $E_1(W)$, defined in (5.40b) for $s = 1$, has the following properties:*

(a)	*homogeneity:*	$E_1(WG) = E_1(W)$ *for all* $G \in \mathrm{GL}(c)$
(b)	*stationarity:*	$\nabla E_1(W) = 0$ *if and only if* $\operatorname{span}(W)$ *is an invariant subspace of* $K_{\mathcal{F}}$

Proof. (a) a direct calculation shows

$$E_1(WG) = \operatorname{tr}\big(K_{\mathcal{F}}WG(G^\top W^\top WG)^{-1}G^\top W^\top\big) = E_1(W). \tag{5.49}$$

(b) since $(W^\top W)^{-1}$ is non-singular we have

$$\nabla E_1(W) = 0 \quad\Leftrightarrow\quad 2\big(I_n - A_1(W)\big)K_{\mathcal{F}}W(W^\top W)^{-1} = 0 \tag{5.50a}$$

$$\Leftrightarrow\quad K_{\mathcal{F}}W - W(W^\top W)^{-1}W^\top K_{\mathcal{F}}W = 0 \tag{5.50b}$$

$$\Leftrightarrow\quad \exists\, R = (W^\top W)^{-1}W^\top K_{\mathcal{F}}W \in \mathbb{R}^{c\times c}\colon\ K_{\mathcal{F}}W - WR = 0 \tag{5.50c}$$

$$\Leftrightarrow\quad \operatorname{span}(K_{\mathcal{F}}W) \subset \operatorname{span}(W) \tag{5.50d}$$

□

The Riemannian gradient of the Rayleigh quotient [AMS09, Sec. 4.8]

$$E_1(Y) = \operatorname{tr}(Y^\top K_{\mathcal{F}} Y) \tag{5.51}$$

over the compact Stiefel manifold (5.34) equipped with the standard Euclidean metric reads [AMS09, Sec. 4.8]

$$\operatorname{grad} E_1(Y) = 2(I_n - YY^\top)K_{\mathcal{F}}Y \quad \in \quad T_Y \mathrm{St}(c,n). \tag{5.52}$$

Next, we relate the Euclidean gradient (5.41b) to the Riemannian gradient (5.52) in Proposition 5.4.8, by taking into account the parametrization $Y(W) \in \mathrm{St}(c,n)$ in (5.33).

Proposition 5.4.8 (Relation Between $\nabla E_1(W)$ and $\operatorname{grad} E_1(Y(W))$)**.** *Let $W \in \mathcal{W}^c$. Then, the Euclidean gradient $\nabla E_1(W)$ is related to the Riemannian gradient $\operatorname{grad} E_1(Y)$ by the parametrization $Y(W) = W(W^\top W)^{-\frac{1}{2}} \in \mathrm{St}(c,n)$ in the following way:*

$$\nabla E_1(W) = 0 \quad \Leftrightarrow \quad \operatorname{grad} E_1(Y(W)) = 0. \tag{5.53}$$

$W \in \mathcal{W}^c$ is a stationary point if and only if $Y(W) \in \mathrm{St}(c,n)$ is a stationary point of the Rayleigh quotient (5.51) over the compact Stiefel manifold $\mathrm{St}(c,n)$.

Proof.

$$\nabla E_1(W) = 2\big(I_n - A_1(W)\big)K_{\mathcal{F}}W(W^\top W)^{-1} \tag{5.54a}$$

$$= 2\big(I_n - Y(W)Y(W)^\top\big)K_{\mathcal{F}}Y(W)(W^\top W)^{-\frac{1}{2}} \tag{5.54b}$$

$$= \operatorname{grad} E_1(Y(W))(W^\top W)^{-\frac{1}{2}}. \tag{5.54c}$$

Since the second factor in (5.54c) is non-singular the assertion follows. □

Proposition 5.4.7 and 5.4.8 together with the fact that $K_{\mathcal{F}}$ is symmetric yield the following Corollary 5.4.9.

Corollary 5.4.9. *Let $W \in \mathcal{W}^c$ and $Y(W) = W(W^\top W)^{-\frac{1}{2}} \in \mathrm{St}(c,n)$. Then, $\operatorname{span}(W)$ is an invariant subspace of $K_{\mathcal{F}}$ if and only if the columns of $Y(W)$ are eigenvectors of $K_{\mathcal{F}}$.*

Consequently, the gradient (5.41b) is directly linked to the search direction (5.52) on the compact Stiefel manifold, in order to determine the invariant subspace corresponding to the c dominant eigenvectors of $K_{\mathcal{F}}$.

As a consequence of these considerations, we define for arbitrary s the generalized likelihood map in Definition 5.4.2.

Definition 5.4.2 (Generalized Likelihood Map). Let $s \in [0,1]$ and $E_s(W)$ given by (5.40). Then, the *generalized likelihood map* is defined as

$$L_{s;i}(W_i) = \exp_{W_i}\Big(\frac{1}{\rho}\nabla E_s(W)_i\Big). \tag{5.55}$$

In addition, Definition 5.4.2 allows to fully exploit the low-rank structure of the factorization $A_s(W)$, since storing the full $n \times n$ matrix is not required for any evaluation of the gradients derived in Proposition 5.4.4

The next section introduces the family of self-assignment flows.

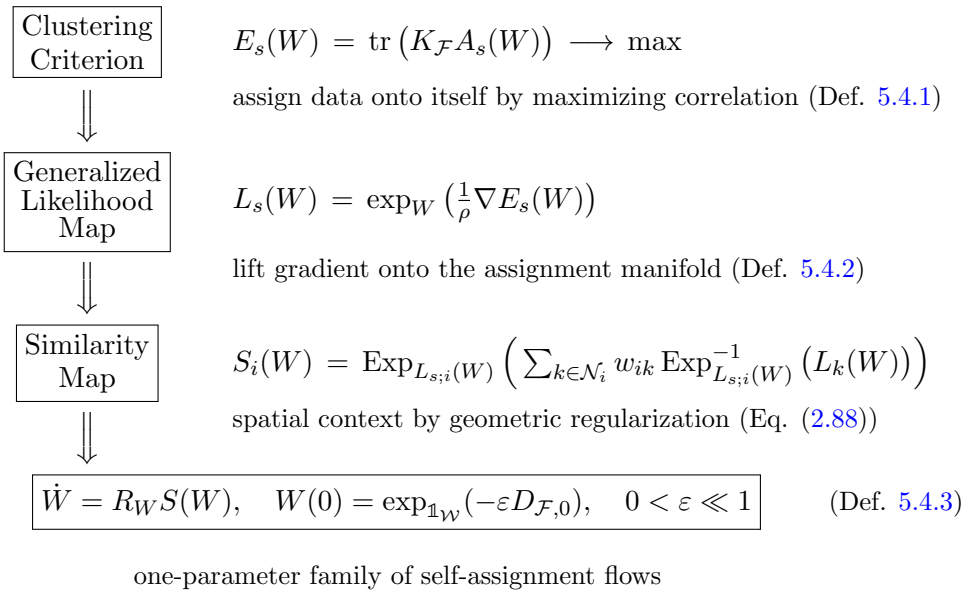

Figure 5.3.: Overview of the essential building blocks composing the one-parameter family of self-assignment flows, Definition 5.4.3. The clustering criterion abstracts "pixel-label" decisions to "pixel-pixel" decisions in the absence of prototypes. Consequently the input data are assigned onto themselves. This fact is encoded by the self-assignment matrix (a low-rank factorization) which encodes for each pair of pixels if they belong to the same cluster or not. The likelihood map (2.86) is generalized by lifting the gradient $\nabla E_s(W)$ with respect to the assignments onto the assignment manifold in order to maximize the correlation between the kernel matrix and the self-assignments (clustering criterion). Except for this more general definition of the likelihood map, the subsequent similarity map (2.88) remains unaltered from the supervised case, as do numerical schemes [ZSPS20] for integrating the resulting self-assignment flows.

5.4.3. The Self-Assignment Flows

Besides replacing the likelihood map (2.86) by the generalized likelihood map (5.55), no further changes are required in order to generalize the assignment flow (2.89) to the unsupervised case, except for the initialization which cannot both start at the barycenter and break the symmetry, without any prior information. This will be achieved by taking a small perturbation of the barycenter as initial point.

Accordingly, we define the one-parameter family of self-assignment flows in Definition 5.4.3.

Definition 5.4.3 (One-Parameter Family of Self-Assignment Flows). Let $s \in [0,1]$ and the similarity map $S(W)$ (2.88) based on the generalized likelihood map $L_s(W)$ (5.55) and $D_{\mathcal{F},0} \in \mathbb{R}_+^{n\times c}$ be given. Then, the one-parameter family of *self-assignment flows* is described by the parametrized dynamical system

$$\dot{W} = R_W S(W), \qquad W(0) = \exp_{\mathbb{1}_{\mathcal{W}}}(-\varepsilon D_{\mathcal{F},0}), \quad 0 < \varepsilon \ll 1 \tag{5.56a}$$

$$W(t) \in \begin{cases} \mathcal{W}, & \text{if } s = 0, \\ \mathcal{W}^c, & \text{if } s \in (0,1]. \end{cases} \tag{5.56b}$$

The matrix $D_{\mathcal{F},0}$ is computed using the given data $\mathcal{F}_n$ as explained in Section 2.4.3. The flow $W(t)$ is restricted to the submanifold of full-rank assignments if $s > 0$.

Figure 5.3 depicts an overview of the building blocks which define the vector field for the one-parameter family of self-assignment flows, defined in Definition 5.4.3.

Proposition 5.3.2 and Eq. (5.16) yield the following.

Corollary 5.4.10. *Let $W(t)$ solve (5.56). Then, for any $t \geq 0$,*

(i) the self-affinity matrix $A_0\big(W(t)\big)$ is doubly stochastic and completely positive, if $s = 0$;

(ii) the self-influence matrix $A_1\big(W(t)\big)$ is iso-spectral, i.e. its eigenvalues satisfy $\lambda_1 = \cdots = \lambda_c = 1$ and $\lambda_{n-c} = \cdots = \lambda_n = 0$, if $s = 1$.

(iii) $A_0\big(W(T)\big) = A_s\big(W(T)\big) = A_1\big(W(T)\big) \quad$ *if* $\quad W(T) \in \mathcal{W}_*^c \quad$ *for* $\quad s \in [0,1]$.

Property *(iii)* relates to the fact that $W(t)$ solving (5.56) approaches a labeling $W(T) \in \mathcal{W}_*^c$ for sufficiently large T after a trivial rounding step. We point out, however, that solving (5.56) generally yields different paths $W(t)$, $t \in [0,T]$ depending on $s \in [0,1]$ and corresponding to the different relaxations discussed in Section 5.3.3. Once a labeling $W(T) \in \mathcal{W}_*^c$ has been computed using any $s \in [0,1]$, the solution is a local optimum of the partitioning problem (5.12). This is what Corollary 5.4.10(iii) says.

Remark 5.4.1 (Parameters of the Self-Assignment Flow)**.** We briefly explain the role of each parameter involved in order to point out, that there is essentially a single user parameter only, that has to be specified.

- Any small positive number $\varepsilon > 0$ determining $W(0)$ by (5.56a) will do in practice.
- The parameter $s \in [0,1]$ of (5.56b) is chosen depending on the application: As Figure 5.5 illustrates, and as a consequence of the interpretations of the self-affinity matrix $A_0(W)$ (Section 5.3.3.1) and the self-influence matrix $A_1(W)$ (Section 5.3.3.3), small values s increase the sensitivity of the self-assignment flow to the spatial structure of the partition of the underlying graph $\mathcal{G}$, whereas large values s make the approach more sensitive with respect to the quantization of the feature space $\mathcal{F}$ in terms of the prototypes, that are implicitly determined by the self-assignment flow (Section 5.3.3.2).
- Parameter ρ of the likelihood map (2.86) merely normalizes the scale of the input similarity matrix $K_{\mathcal{F}}$, that can be small or large depending on the particular data under consideration.
- The fixed stepsize $h > 0$ used in this work for geometric numerical integration (Section 5.4.5) can be determined automatically if a more advanced numerical scheme with adaptive stepsize control is employed, as worked out by [ZSPS20].
- Parameter $c \in \mathbb{N}$ merely specifies an *upper bound* of the number of clusters, whereas the *resulting effective* number of clusters $\hat{c} \leq c$ does *not* need to be specified beforehand (see Definition 5.4.4 below).

As a result, the only parameter that critically influences the result returned by the self-assignment flow is the *size* $|\mathcal{N}_i|$ of the neighborhoods (2.70), that determines the *scale* of geometric spatial regularization (2.88) and, in turn, the number $\hat{c}$ of effective clusters. Section 5.6 provides numerous illustrations.

5.4.4. Self-Assignment Performs Self-Supervision

We interpret the assignment flow from another angle complementing the interpretations discussed in Section 5.3.3.

In Section 5.3.3.2, we showed that following the assignment flow entails learning of latent prototypes that can be explicitly recovered if weighted means in the data space are well-defined and computationally feasible. Let us temporarily adopt the Euclidean situation (5.27). With these recovered prototypes at hand, we get back to Section 2.4.1 and ask how our approach relates to the *supervised* situation where the quality of the clustering can be assessed by objectives like (2.100). Assuming a labeling $W = W(T) \in \mathcal{W}_*^c$ has been determined, let the recovered prototypes $f_j^*(W)$, $j \in \mathcal{J}$ which depend on the assignment, play the role of empirical means m_j, $j \in \mathcal{J}$. We compute the quantities (2.97) in terms of the data matrix F (5.28)

$$P_j = \frac{1}{n}\langle W^j, \mathbb{1}_n \rangle = \frac{1}{n}|\mathcal{I}_j|, \qquad j \in \mathcal{J} \qquad \text{(prior probabilities)} \tag{5.57a}$$

$$f_j^*(W) = F^\top \big(WC(W)^{-1}\big)^j, \qquad j \in \mathcal{J} \qquad \text{(class-conditional mean vectors)} \tag{5.57b}$$

$$f^* = \frac{1}{n} F^\top \mathbb{1}_n, \qquad \text{(mean vector).} \tag{5.57c}$$

Using these expressions, we can evaluate the scatter matrices (2.98) in turn. This is collected in Proposition 5.4.11.

Proposition 5.4.11 (Scatter Matrices for Latent Prototypes)**.** *Let $W \in \mathcal{W}$ and S_w, S_b be defined in (2.98). Then, the following expressions hold:*

$$S_t = \frac{1}{n}\sum_{i \in \mathcal{I}} (f_i - f^*)(f_i - f^*)^\top = \frac{1}{n} F^\top \big(I - \frac{1}{n}\mathbb{1}_n \mathbb{1}_n^\top\big) F, \tag{5.58a}$$

$$S_w(W) = \frac{1}{n}\sum_{j \in \mathcal{J}} \sum_{i \in \mathcal{I}_j} (f_i - f_j^*(W))(f_i - f_j^*(W))^\top = \frac{1}{n} F^\top \big(I - A_0(W)\big) F, \tag{5.58b}$$

$$S_b(W) = \sum_{j \in \mathcal{J}} P_j (f_j^*(W) - f^*)(f_j^*(W) - f^*)^\top = \frac{1}{n} F^\top \big(A_0(W) - \frac{1}{n}\mathbb{1}_n \mathbb{1}_n^\top\big) F. \tag{5.58c}$$

Proof. The assertions follow after straightforward calculations by inserting the expressions (5.57) into (2.98). □

Regarding the dependency on W, we observe that the within-class scatter matrix $S_w(W)$ involves the term $F^\top A_0(W) F$ and the between-class scatter $S_b(W)$ the term $-F^\top A_0(W) F$. Hence, by minimizing the objective (5.12), we *simultaneously* minimize $\operatorname{tr}(S_w)$ and maximize $\operatorname{tr}(S_b)$, which is stated in the following Corollary 5.4.12.

Corollary 5.4.12 (Latent Prototypes Separate Clusters Optimally)**.** *Let $W \in \mathcal{W}$ and $S_w(W)$, $S_b(W)$ given by (5.58). Then, the following equivalence holds:*

$$\operatorname*{arg\,min}_W \ \operatorname{tr}\big(S_w(W)\big) \quad \Leftrightarrow \quad \operatorname*{arg\,max}_W \ \operatorname{tr}\big(S_b(W)\big) \quad \Leftrightarrow \quad \operatorname*{arg\,max}_W \ \operatorname{tr}\big(A_0(W) F F^\top\big). \tag{5.59}$$

Proof. The claim follows by inserting the expressions from Proposition (5.4.11) and removing constant terms. □

We conclude that latent prototypes determined by the self-assignment flow turn a completely unsupervised scenario into a supervised one, in agreement with established measures for class separability like (2.100). This interpretation also remains valid when the relaxation with $s = 1$ and objective (5.32) is used to compute a labeling W, due to Corollary 5.4.10(iii).

Moreover, since the approach only depends on the inner product matrix $FF^\top$, it generalizes to data embeddings into a reproducing kernel Hilbert space and a corresponding data affinity matrix $K_{\mathcal{F}}$ with entries (5.5).

5.4.5. Geometric Numerical Integration

We distinguish the two cases (5.56b).

Case $s = 0$. We directly apply the methods studied in [ZSPS20]. To make this thesis self-contained, we merely state the simplest scheme, the geometric Euler method. This explicit scheme with fixed step-size $h > 0$ reads

$$W_i^{(k+1)} = \operatorname{Exp}_{W_i^{(k)}}\left(hR_{W_i^{(k)}} S(W^{(k)})\right), \quad i \in \mathcal{I}. \tag{5.60}$$

It ensures that the self-assignment flow (5.56a) evolves properly on the assignment manifold $\mathcal{W}$. See [ZSPS20] for more advanced numerical schemes that run 'automatically' through adaptive stepsize control. The iteration (5.60) stops when the average entropy of the assignments $W^{(K)}$ drops at some iteration $k = K$ below the predefined threshold 10^{-3}, which indicates (almost) unique label assignments and hence stationarity of the flow evolution. Then numerical integration is terminated and a labeling $W \in \mathcal{W}_*^{\hat{c}}$, $\hat{c} \leq c$, is determined using $W^{(K)}$ in a trivial postprocessing step by selecting the most likely label for each row $W_i^{(K)}$, $i \in \mathcal{I}$ and removing the $c - \hat{c}$ zero-columns (corresponding to empty clusters) from the resulting labeling.

Definition 5.4.4 (Effective Number $\hat{c}$ of Clusters (Labels)). We call the just described number

$$\hat{c} \leq c \tag{5.61}$$

the *effective number of clusters or labels*, respectively. It is determined by the homogeneity of the data $\mathcal{F}_n$ and by the scale

$$|\mathcal{N}_i|, \quad i \in \mathcal{I} \qquad \textbf{(scale)} \tag{5.62}$$

at which regularization is performed by the assignment flow through the similarity map (2.88). We denote the corresponding index set of labels by

$$\hat{\mathcal{J}} \subset \mathcal{J}, \qquad |\hat{\mathcal{J}}| = \hat{c}. \tag{5.63}$$

Remark 5.4.2. The assertions of Corollary 5.4.10 as well as the considerations in Section 5.4.4 remain valid after replacing the upper bound c of the number of prototypes (labels) and the corresponding index set $\mathcal{J}$ by $\hat{c}$ and $\hat{\mathcal{J}}$, respectively, according to Definition 5.4.4.

Case $s \in (0, 1]$. Integration of the self-assignment flow (5.56a) restricted to the open submanifold $\mathcal{W}^c$ of full-rank assignments (5.17) is more involved. Corresponding geodesics only locally exist on $\mathcal{W}$, i.e. full-rank assignment matrices cannot be guaranteed during the numerical integration process (5.60), see Section 5.4.1 for details. Clearly, if the data affinity matrix $K_{\mathcal{F}}$ has high rank (induced by heterogeneous data) and if the scale (5.62)

for regularization is not chosen too large, a full-rank labeling $W \in \mathcal{W}^c$ may be returned by the self-assignment flow, that is well-defined in view of the relation (5.53).

In order to handle other cases while still using the numerical scheme (5.60) or more sophisticated ones [ZSPS20], we simply replace the inverse normalizing matrix by its Moore-Penrose pseudo-inverse,

$$(W^\top W)^{-1} \quad \longleftarrow \quad (W^\top W)^\dagger. \tag{5.64}$$

Whenever this regularization of the normalizing matrix becomes 'active', we extract the effective number $\hat{c}$ in a postprocessing step, as described above in the case $s = 0$.

5.5. Related Work and Discussion

The literature on clustering is vast. We therefore restrict the discussion to few major methodological directions in the literature: graph cuts and spectral relaxation (Section 5.5.1), discrete regularized optimal transport (Section 5.5.2) and combinatorial optimization for graph partitioning (Section 5.5.3).

5.5.1. Graph Cuts and Spectral Relaxation

Summing up the weights (affinities) of edges that are cut provides a natural quality measure for graph partitioning. To avoid unbalanced partitions, such measures are normalized in various ways, and spectral relaxations of the resulting combinatorial optimization problem renders the computation of good suboptimal solutions feasible. We refer to [vLux07] for a survey.

We focus on two basic balanced cut-criteria that can be expressed by the graph Laplacian

$$L_\mathcal{F} = D_{K,\mathcal{F}} - K_\mathcal{F}, \qquad D_{K,\mathcal{F}} = \operatorname{Diag}(K_\mathcal{F} \mathbb{1}_n) \tag{5.65}$$

and indicator vectors. The *ratio-cut criterion* reads

$$\min_{U \in \mathbb{R}^{n \times c}} \operatorname{tr}(U^\top L_\mathcal{F} U) \quad \text{s.t.} \quad U \geq 0, \quad U^\top U = I_c, \tag{5.66}$$

whereas the *normalized-cut (Ncut) criterion* [SM00] additionally uses the degree matrix $D_{K,\mathcal{F}}$ for normalization,

$$\min_{U \in \mathbb{R}^{n \times c}} \operatorname{tr}(U^\top L_\mathcal{F} U) \quad \text{s.t.} \quad U \geq 0, \quad U^\top D_{K,\mathcal{F}} U = I_c. \tag{5.67}$$

Due to the conjunction of nonnegativity and orthogonality constraints, both problems (5.66) and (5.67) are difficult to optimize globally. *Spectral relaxation* means to drop the element-wise nonnegativity constraint. Then the relaxed problems (5.66) and (5.67) amount to solving an eigenvalue problem and a generalized eigenvalue problem, respectively. The price to pay in either case is that the physical interpretation of U as indicator variables is lost and must be recovered by an additional post-processing step, which is usually done by applying the classical k-means algorithm.

A direct relation to the self-assignment flow is apparant in the case $s = 1$. Substituting $Y = D_{K,\mathcal{F}}^{1/2} U$ in the spectral relaxation of (5.67) results in the problem

$$\max_{Y \in \mathbb{R}^{n \times c}} \operatorname{tr}\left(Y^\top \tilde{K}_\mathcal{F} Y\right) \quad \text{s.t.} \quad Y^\top Y = I_c, \tag{5.68}$$

that is, the Rayleigh quotient of the *normalized* affinity matrix $\tilde{K}_\mathcal{F} = D_{K,\mathcal{F}}^{-1/2} K_\mathcal{F} D_{K,\mathcal{F}}^{-1/2}$ has to be maximized over the compact Stiefel manifold (5.34). As already discussed for $s = 1$ in connection with (5.33), assignments W following the self-assignment flow

parametrize points $Y(W) \in \mathrm{St}(c,n)$ on the compact Stiefel manifold that maximize the Rayleigh quotient: Eq. (5.54c) shows that the driving force of the self-assignment flow (generalized likelihood map) is directly linked to the gradient ascent of the Rayleigh quotient over the compact Stiefel manifold. Finally, when the numerical integration of the self-assignment flow terminates, then the resulting labeling $W \in \mathcal{W}_*^c$ together with (5.16) ensures $Y(W) \geq 0$. Hence, after re-substitution, $U(W) = D_{K,\mathcal{F}}^{-1/2} Y(W)$ is directly feasible for the original problem (5.67) and hence no 'projection' by k-means is required as post-processing.

The common way to take into account *spatial regularization* in spectral clustering is to augment given features by spatial coordinates. However, this strategy suffers from a conceptual shortcoming, since augmentation makes the *same* feature vector differ when it is observed at two different spatial locations. In contrast, the self-assignment flow performs unbiased spatial regularization by smooth geometric averaging and recognizes closeness of features no matter *where* they are observed.

5.5.2. Discrete Regularized Optimal Transport

The theory of optimal transport [Vil09; San15] has become a major modeling framework for data analysis. Here we focus on discrete optimal transport and computational aspects [BCPD99; Pey18].

We consider the case $s = 0$ and the self-affinity matrix $A_0(W)$. Since $A_0(W)$ is doubly stochastic (Prop. 5.3.2), maximizing the objective $E_0(W)$ (5.40b) may be interpreted as a discrete optimal transport problem with cost matrix $K_{\mathcal{F}}$ and uniform marginal measures (5.22), (cf. Corollary 5.4.6). These marginals correspond to the data $\mathcal{F}_n$ and a copy of the data, respectively, resulting in data self-assignment as discussed in Section 5.3.3.1.

For further interpretation, we consider the Euclidean case $K_{\mathcal{F}} = FF^\top$. Inserting the explicit form (5.15) of $A_0(W)$ into the objective $E_0(W)$ and using (5.43), we obtain

$$E_0(W) = \mathrm{tr}(K_{\mathcal{F}} W C(W)^{-1} W^\top) = \mathrm{tr}(W F^* F^\top). \tag{5.69}$$

Maximizing this objective function reveals what this problem relaxation actually means: A linear assignment problem in terms of the assignment matrix W with *varying* inner product matrix $F^*(W)F^\top$ as costs. Moreover, since $W \in \mathcal{W}$, we have a fixed marginal $W\mathbb{1}_c = \mathbb{1}_n$ and a the second marginal $W^\top \mathbb{1}_n = \mathrm{diag}\big(C(W)\big)$ which is *free*. Altogether, a quite difficult problem is solved in terms of W: latent prototypes F^* are formed by *transporting* the uniform prior measure to the support of the respective clusters, so as to maximize the correlation $E_0(W)$ of the assignments W and the inner product matrix $F^*F^\top$.

We point out a key property of the assignment flow that makes this approach work: It is the spatial regularization performed by the similarity map (2.88) that drives the entire process, in addition to the underlying geometry that makes $W(t)$ converge towards hard assignments (labelings). In fact, without spatial regularization, the self-affinity matrix $A_0(W) = I_n$ would maximize $E_0(W)$ assuming the similarity $k_{\mathcal{F}}(f_i, f_k)$ is maximal if $f_i = f_k$, which means that every given data point f_i forms its own cluster. This trivial solution is ruled out, by construction, through the factorization with rank upper bounded by c and through geometric spatial averaging of assignments. The corresponding *scale* in terms of the sizes of the neighborhoods (2.70) determines how coarse or fine the spatial arrangement of the resulting clusters will be.

We informally summarize this discussion: Data self-assignment is defined by uniform

marginal measures and a coupling measure parametrized by the assignment flow. Structure in the data is induced by imposing a low-rank constraint (factorization) on the coupling measure (transport plan) and through spatial regularization of the flow of assignments.

5.5.3. Combinatorial Optimization

Zass and Shashua [ZS05] studied the formulation of the clustering problem

$$\max_{Y \in \mathbb{R}^{n \times c}} \operatorname{tr}(K_{\mathcal{F}} Y Y^\top) \qquad \text{subject to} \tag{5.70a}$$

$$\text{(a) } Y \geq 0, \qquad \text{(b) } \operatorname{rank}(Y) = c, \qquad \text{(c) } Y^\top Y = I_c, \qquad \text{(d) } Y Y^\top \mathbb{1}_n = \mathbb{1}_n \tag{5.70b}$$

in terms of the completely positive factorization $YY^\top$ and the constraints (a)–(d). We notice that the orthogonality constraint (c) with respect to the columns of Y implies (b), and that (a) together with (d) says that $YY^\top$ is doubly stochastic. The authors show that (a)–(d) imply that $W = YC(Y)^{\frac{1}{2}} \in \mathcal{W}_*^c$ is a labeling. This problem formulation differs from more classical conditions ensuring $W \in \mathcal{W}_*^c$ [RW95, Lemma 2.1],

$$W \geq 0, \qquad W\mathbb{1}_c = \mathbb{1}_n, \qquad W^\top \mathbb{1}_n = (n_1, \dots, n_c)^\top, \qquad \operatorname{tr}(W^\top W) = n, \tag{5.71}$$

in that the cluster sizes (third constraint) do not have to be specified beforehand.

Regarding relaxation, the authors of [ZS05] argue that the orthogonality constraint (c) is the weakest one. They propose a two-step procedure after dropping the constraints (b) and (c): approximation of the data similarity matrix $K_{\mathcal{F}}$ by a doubly stochastic matrix using the Sinkhorn iteration, followed by a gradient ascent iteration with stepsize control so as to respect the remaining constraints. The same set-up was proposed by [YC16] except for determining a locally optimal solution by a single iterative process using DC-programming. Likewise, [KYP15] explored symmetric nonnegative factorizations but ignored the constraint enforcing that $YY^\top$ is doubly-stochastic, which is crucial for cluster normalization.

Our approach uses the factorization $A_s(W)$ given by (5.19) instead of $YY^\top$ in (5.70), we can relate the two factorizations by identifying the factor

$$Y(W) = W\gamma_s(W)^{-\frac{1}{2}}, \tag{5.72}$$

which is parametrized by assignments. While the rank constraint (b) and orthogonality constraint (c) are dropped for $s = 0$, the constraints (a) and (d) are 'built in' by construction of

$$Y(W) = WC(W)^{-\frac{1}{2}} \geq 0, \tag{5.73}$$

resulting in a completely positive and double stochastic factorization. Conversely, for $s = 1$, spectral properties are retained (cf. Section 5.5.1). The orthogonality constraint (c) which implies the rank constraint (b), holds for

$$Y(W) = W(W^\top W)^{-\frac{1}{2}} \in \operatorname{St}(c, n), \tag{5.74}$$

whereas constraints (a) and (d) are ignored. This agrees with the observation that the constraints (a) and (c) cause the combinatorial difficulty of formulation (5.70), which renders them to be mutually exclusive ((a) "physical quantity" vs. (c) "exclusive decisions"). However, by Definition 5.3.4 of the one-parameter family of self-assignment matrices, we can smoothly interpolate between combinatorial and spectral properties.

Furthermore, optimization is achieved by a single *smooth and continuous* process, the self-assignment flow (5.56), which enables to apply numerous discrete numerical schemes

[ZSPS20], all of which respect the constraints. Finally, geometric regularization within local neighborhoods of each vertex of the underlying graph through the similarity map (2.88) enforces the formation of 'natural' clusters, whenever assigning the same label to close vertices is more likely to be correct.

5.6. Experiments

In this section, we demonstrate and evaluate the performance of the proposed one-parameter family (5.56) of *self-assignment flows (SAF)* for unsupervised data labeling, using various datasets and feature spaces (Figure 5.4).

After describing specific details of the implementation (Section 5.6.1), we report the study of the two model parameters in Section 5.6.2, and the influence of affinity matrix sketching for data reduction in a preprocessing step, to make learning from large data sets computationally feasible. In Section 5.6.3, we compare our approach to various methods: basic clustering, normalized spectral cuts with spatial regularization, and partitioning using a variational decomposition of the piecewise constant Mumford-Shah model. We focus on an attractive application of our approach in Section 5.6.4: Learning patch dictionaries using the SAF based on a locally invariant distance function. Finally, as a sanity check, we report the application of the SAF to problem data on a graph from a domain that is unrelated to image analysis, to substantiate our claim that our approach applies to any data given on any undirected weighted graph, in principle.

5.6.1. Implementation Details

Throughout this work, the SAF (5.56) was numerically integrated using the geometric explicit Euler scheme (5.60) with step-size $h = 0.1$, as described in Section 5.4.5. For parameter values $s \in (0, 1]$, we applied (5.64) to avoid numerical problems when the effective number of clusters $\hat{c} < c$ (Definition 5.4.4) actually was smaller than the upper bound c, which can be seen as an upper bound for the complexity of the solution. The SAF with $s = 0$ does not encounter any such problems, due to the different normalization involved in (5.15). From Proposition 5.4.4 we have given only gradient expressions for $s = 0$ and $s = 1$ of the clustering criterion, to avoid tedious calculations we have employed automatic differentiation for instances with $s \in (0, 1)$. We adopted from [ÅPSS17] the numerical renormalization step for the assignments with $\varepsilon = 10^{-10}$, to avoid numerical issues for assignments very close to the boundary of the assignment manifold. Numerical integration was terminated when the average entropy of the assignments dropped below the threshold of 10^{-3}, which indicates that the current iterate is very close to an almost unique assignment (labeling) $W^{(k)} \in \mathcal{W}_*^{\hat{c}}$.

Unless specified otherwise, the default value $\rho = 0.1$ (distance normalization in (2.86)) and uniform weights $w_{i,k} = 1/|\mathcal{N}_i|$ (2.87) for assignment regularization were used in all experiments, with neighborhoods $\mathcal{N}_i$ of equal size

$$|\mathcal{N}| := |\mathcal{N}_i|, \quad \forall i \in \hat{\mathcal{I}}, \tag{5.75}$$

for interior pixels $\hat{\mathcal{I}} \subset \mathcal{I}$.

Data $\mathcal{F}_n$ were embedded using the standard Gaussian kernel (5.6) with parameter $\sigma = \sqrt{0.1}$, in order to compute the affinity matrix $K_{\mathcal{F}}$ (5.5). For larger datasets, a sketch of $K_{\mathcal{F}}$ was used as described in Section 2.4.2, with parameters $q = 1$ and $\ell = 100$ random

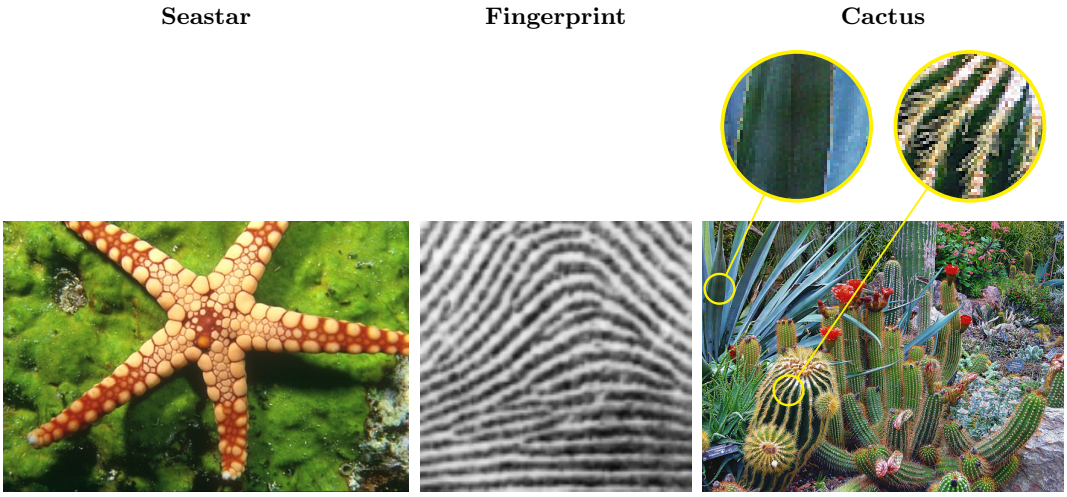

Figure 5.4.: Input image data used in the numerical experiments (Fig. 5.5, 5.8, 5.9, 5.11, 5.12). Close-up views enable to compare the influence of model parameters on local image structure in comparison to alternative approaches from related work. Both the Euclidean RGB-space and locally invariant patch spaces were used as feature spaces. Regarding the latter, additional real image data are processed in Figures 5.14 and 5.15.

samples drawn without replacement; see Section 5.6.2 for a validation. Finally, the initial value $W(0)$ of (5.56a) was chosen as small perturbation of the barycenter (5.56a) with $\varepsilon = 10^{-2}$ and initial distance matrix $D_{\mathcal{F},0}$, computed with the inexpensive greedy k-center clustering algorithm, as explained in Section 2.4.3.

5.6.2. Influence of Model Parameters

The self-assignment flow (SAF) has three model parameters: The parameter s of the self-assignment matrix $A_s(W)$ (5.19a), the neighborhood size $|\mathcal{N}|$ controlling the *scale* of regularization, and the upper bound c on the effective number $\hat{c}$ of labels (5.61).

5.6.2.1. Influence of s, $|\mathcal{N}|$ and c.

Figure 5.5 shows both labelings and recovered prototypes below each panel, depending on s and $|\mathcal{N}|$. We set $c = 16$ which is sufficiently large, since $\hat{c} < c$ quickly happens when lowering s even at the smallest scale of 3×3 pixels. $\hat{c}$ further drops down with larger scale. Regarding the parameter s, we observe:

Small s Spatial regularization is more aggressively enforced, leading to compact codes in terms of smaller numbers $\hat{c}$ of effective prototypes.

Large s Distances in the feature space have more impact. Local image structure is better preserved at the cost of a larger number $\hat{c}$ of effective prototypes.

The second observation underlines the relation of the self-assignment flow, for $s = 1$, to spatially regularized normalized cuts as worked out in Section 5.5.1.

Figure 5.5 illustrates that depending on the application, the properties of the SAF can be continuously controlled by setting the parameter s, thanks to the geodesic interpolation (5.19).

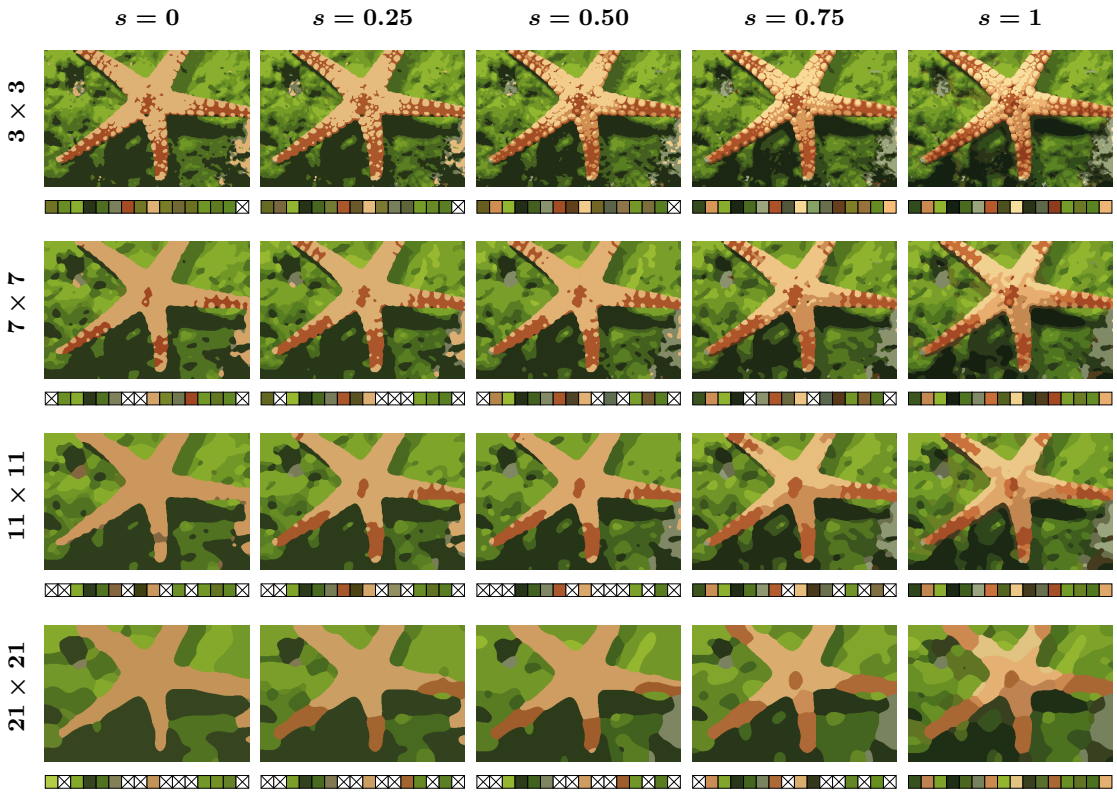

Figure 5.5.: Influence of the model parameters $s \in [0, 1]$ parametrizing the SAF in terms of the self-assignment matrix (5.19), the neighborhood size $|\mathcal{N}|$ controlling the scale of spatial regularization, and the effective number $\hat{c} \leq c = 16$ of labels. Recovered prototypes are displayed below each labeling and aligned to each other (using linear assignment of the clusters) to ease visual comparison. Prototypes that 'died out' are marked by a cross. We observe that due to the geodesic interpolation (5.19), the influence of spatial regularization (small s: compact image codes) relative to the influence of distances in the feature space (large s: preserving local image structure) can be continuously controlled.

5.6.2.2. Evolution of Cluster Sizes, Entropy, and Rank Lower Bound

Figure 5.6 illustrates the *evolution* of the SAF in terms of the following measurements.

Cluster sizes For smaller values of s, more iterations are required for cluster formation. This conforms with the observation in Section 5.6.2.1 that the SAF then promotes spatial regularization. Conversely, larger values of s yield more balanced (uniform) cluster sizes. This is consistent with the observation made in Section 5.6.2.1 that, in this case, the SAF more carefully explores the feature space and preserves local image structure.

Average entropy The panels illustrate that the initial assignment is an ε-perturbation of the barycenter on the assignment manifold, and that the termination criterion was reached in all experiments. In agreement with the preceding point, the SAF converges faster for larger values of s.

Rank lower bound The third row of Figure 5.6 displays the lower bound $\operatorname{tr}\big(B(W^{(k)})\big)$ of $\operatorname{rank}(W^{(k)})$ due to Proposition 5.3.3(d). After termination of the SAF, this lower bound becomes sharp at $W \in \mathcal{W}_*^{\hat{c}}$ and attains the number $\hat{c}$ of effective prototypes.

Cluster-Confusion Matrix The bottom row of Figure 5.6 displays a color coded visualization of the cluster-confusion matrix $B(W)$, which reveals the scattering between clusters for $s < 1$ in accordance with the rank lower bound.

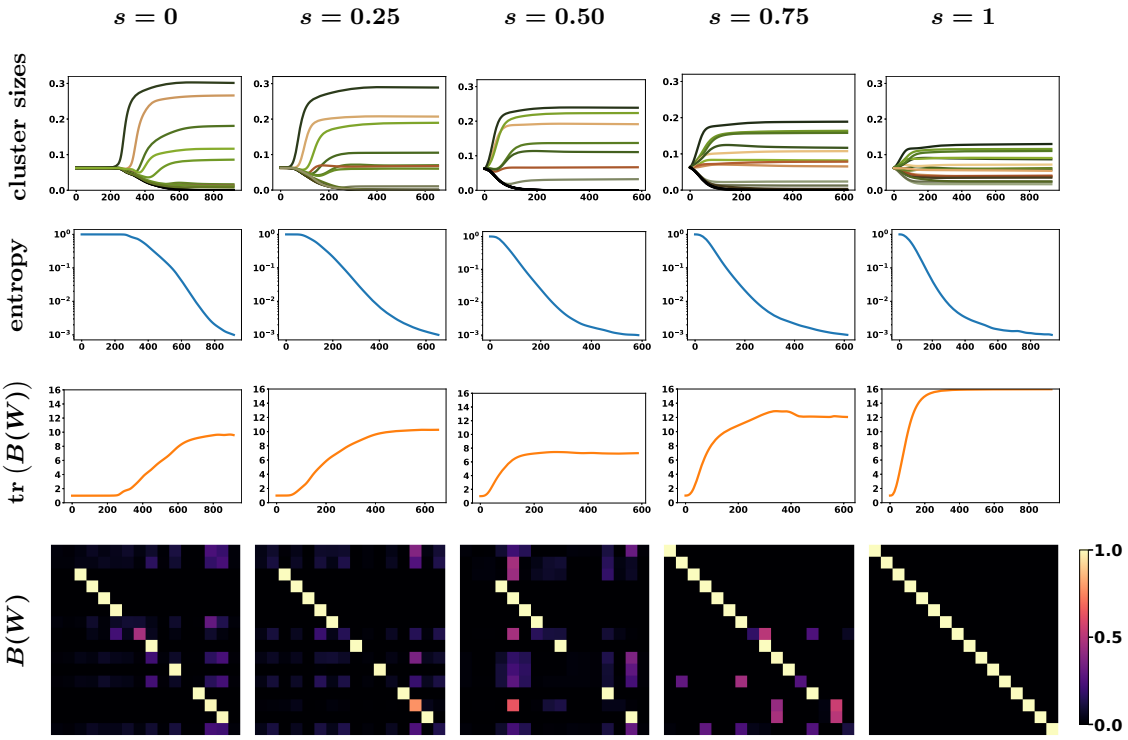

Figure 5.6.: Evolution of relative cluster sizes, average entropy and lower bound of $\operatorname{rank}(W^{(k)})$ as a function of the SAF, depending on the iterations k for the experiment with $|\mathcal{N}| = 11 \times 11$ depicted by Fig. 5.5. TOP: Smaller values of s promote spatial regularization. Hence more iterations are required to form clusters. Larger values of s yield more uniform cluster sizes which reflects the stronger influence of feature similarity and the preservation of local image structure. 2ND ROW: The average entropy illustrates the random initialization ε-close to the barycenter and that the termination criterion is reached in all experiments. The entropy decays faster for larger values of s. 3RD ROW: The lower rank bound due to Proposition 5.3.3(d) becomes sharp when the SAF terminates at some labeling $W \in \mathcal{W}_*^{\hat{c}}$ and attains the number $\hat{c}$ of effective labels. BOTTOM: Visualizing the entries of the cluster-confusion matrix $B(W)$ exhibits scattering between clusters for $s < 1$ which is in accordance with the rank lower bound.

5.6.2.3. Influence of Affinity Matrix Sketching

We evaluate the influence of sketching the data affinity matrix $K_{\mathcal{F}}$ in a preprocessing step, as described in Section 2.4.2, using the parameter value $q = 1$ and varying sample sizes ℓ.

To this end, we focused on the experiment with $s = 0$, $|\mathcal{N}| = 3 \times 3$ depicted by Figure 5.5 and compared the labelings obtained with and without sketching $K_{\mathcal{F}}$. To handle the latter case where $K_{\mathcal{F}}$ requires ≈ 177 GB of memory, we computed the entries for every matrix-vector multiplication on the fly on GPUs using the software library KeOps[1], rather than holding the matrix in memory.

Figure 5.7 displays the relative error of different label assignments after sketching, depending on the sample size ℓ, where 100% corresponds to all $n = 321 \times 481$ columns of $K_{\mathcal{F}}$. For each value ℓ, 100 runs were made using different random seeds. Figure 5.7 displays the *average* error along with the standard deviation. The corresponding curves show that $\ell = 100$ samples, i.e. merely 0.065% of all data points, suffice to eliminate the effect of data reduction by sketching the input affinity matrix. This also underlines the ability of the proposed SAF to abstract data by compact regions (clusters) represented by only a few prototypes.

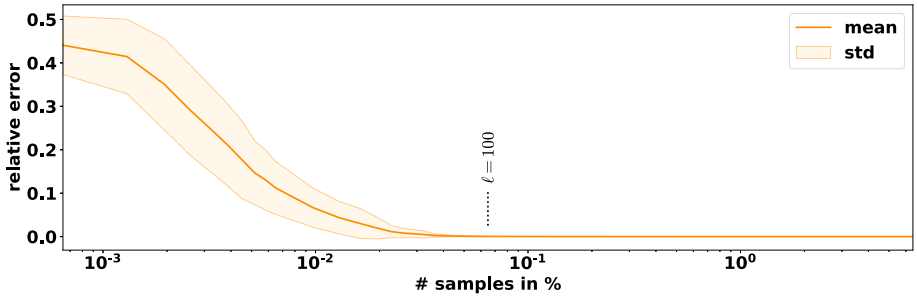

Figure 5.7.: *Average relative labeling error* plotted together with the standard deviation, resulting from data reduction by sketching (see Sec. 2.4.2) the data affinity matrix $K_{\mathcal{F}}$ for SAF in dependency of the relative number of sampled pixels l (represented in %). The curves show that merely 0.065% of all data points (corresponding to $\ell = 100$ randomly sampled columns of $K_{\mathcal{F}}$) suffice to eliminate the effect of data reduction.

5.6.3. Comparison to Other Methods

We compared the SAF to the following methods:

Nearest neighbor clustering **k-means** and **k-center** clustering (no spatial regularization), to show the influence of spatial regularization performed by the SAF on both labeling and prototype formation;

AF *Supervised* assignment flow [ÅPSS17] with spatial regularization, using *fixed* prototypes computed beforehand using nearest neighbor clustering, to highlight that the SAF *simultaneously* performs unsupervised label *learning* and label *assignment*;

[1] B. Charlier, J. Feydy, and J.-A. Glauns, *KeOps Kernel Operations on the GPU*, 2018, https://www.kernel-operations.io/keops/index.html

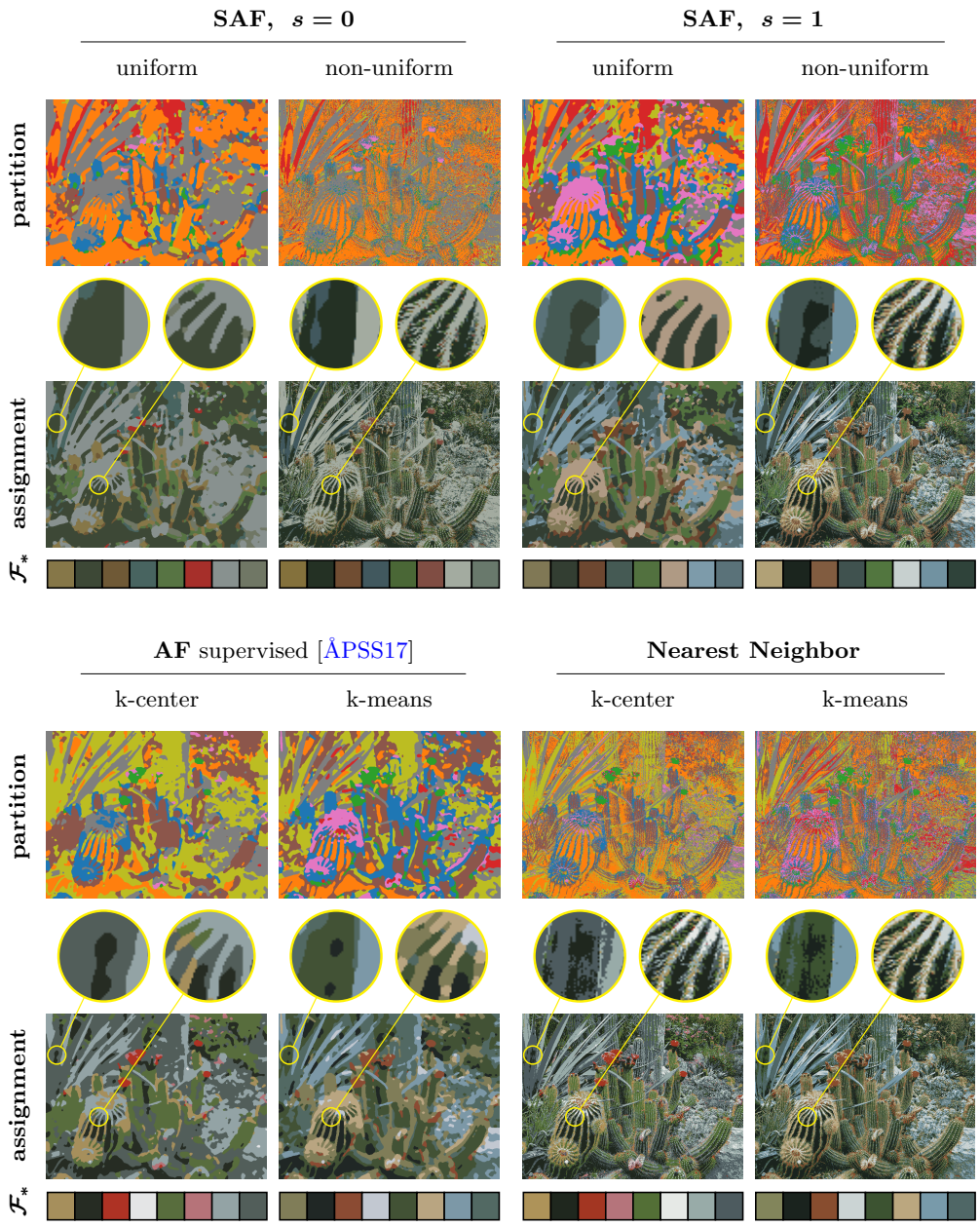

Figure 5.8.: Comparison of the SAF to nearest neighbor clustering and supervised assignment flow (AF). Inspecting the results and the close-up views shows: Nearest neighbor clustering yields noisy label assignments due to the absence of spatial regularization. The AF returns spatially coherent partitions that may locally look unnatural (see close-up views), since the prototypes are fixed and do not adapt to the spatial components of the resulting partition. The unsupervised SAF learns labels adaptively during label assignment. The resulting partitions have a natural spatial structure with increased details if $s = 1$. The latter effect is considerably enhanced, independent of s, when nonuniform weights are used.

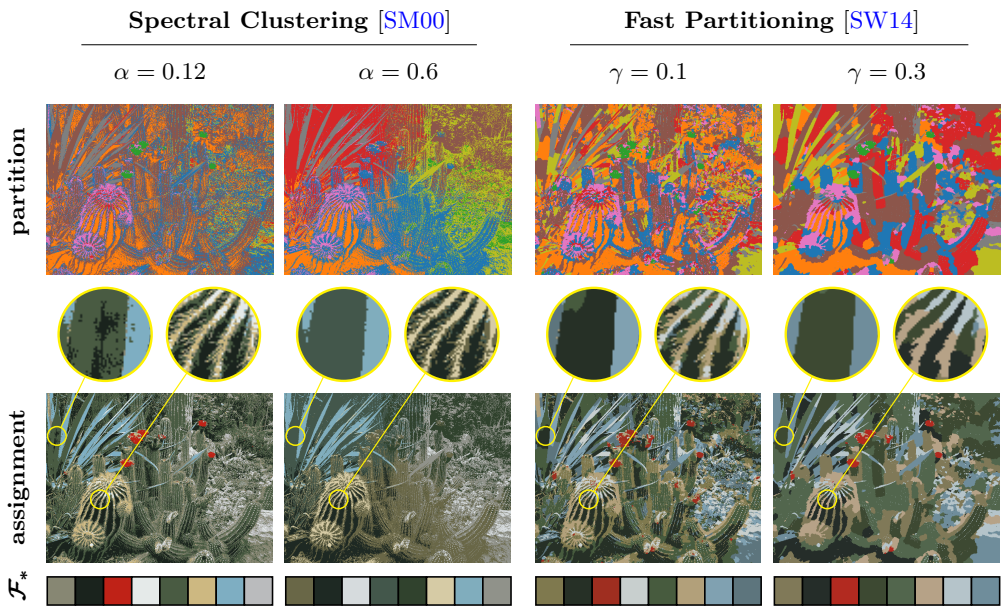

Figure 5.9.: Comparison of the SAF to spectral clustering using feature vectors augmented by spatial coordinates and normalized cuts, and to fast partitioning that approximates the piecewise constant Mumford-Shah model. Spatial regularization as performed by spectral clustering is clearly suboptimal, since weak regularization returns noisy partitions where strong regularization yields biased clusters (e.g. red cluster). See the last paragraph of Section 5.5.1 for an explanation. Fast partitioning yields good labelings but does not consistently enforce the scale of spatial regularization through the choice of γ – see, e.g. the small red clusters in the panel on the right-hand side. This reflects that fast partitioning directly operates on the feature space rather then separating data representation from inference, as does the SAF.

Spectral clustering We computed partitions using normalized spectral cuts [SM00] after augmenting feature vectors by spatial coordinates x_i, $i \in \mathcal{I}$ for spatial regularization. The resulting data affinity matrix reads

$$K_{\mathcal{F}i,k} = \exp\big(-(\tfrac{1}{\sigma^2}\|f_i - f_k\|_2^2 + \alpha\|x_i - x_k\|_2^2)\big), \quad i,k \in \mathcal{I}, \tag{5.76}$$

with parameter $\alpha > 0$ controlling the influence of spatial regularization.

Fast partitioning A variational decomposition of the piecewise-constant Mumford-Shah approach to image partitioning proposed by [SW14], using the publicly available implementation "Pottslab" from the authors. The method operates directly on values in the feature space, instead of using a reformulation with labels. Therefore, the number of clusters can be large. For this reason, we applied an additional k-means clustering step to the (over-segmented) results in order to have a direct comparison in terms of labels and prototypes.

Two variants of the SAF were evaluated for comparison: (i) using *uniform* weights for spatial regularization; (ii) using *nonuniform* weights determined in "non-local means fashion" by

$$w_{i,k} = \frac{\tilde{w}_{i,k}}{\langle \tilde{w}_i, \mathbb{1}_n \rangle} \quad \text{with} \quad \tilde{w}_{i,k} = \begin{cases} \exp\big(-\frac{1}{\rho}\|P_i - P_k\|_F^2\big), & \text{if } k \in \mathcal{N}_i, \\ 0, & \text{else,} \end{cases} \tag{5.77}$$

where P_i denotes the patch centered at pixel i. Throughout, the patch size as well as the neighborhood size $|\mathcal{N}|$ for geometric averaging was chosen to be 5×5 pixels.

The user parameters of all other methods were manually tuned so as to obtain best comparable results.

5.6.3.1. Nearest Neighbor Clustering, Supervised Assignment Flow

Figure 5.8 displays the results. The close-up view of the results of nearest neighbor clustering shows noisy label assignments even in homogeneous regions, due to the absence of spatial regularization. By contrast, the AF returns spatially coherent labelings. However, since the labels (prototypes) are fixed beforehand, their assignments yield partitions that may locally look unnatural (see close-up views). Note that the prototypes displayed for the AF were recomputed after convergence from the resulting partition and, therefore, differ from the nearest neighbor prototypes that were used as input labels for computing the AF.

In comparison with these methods, the SAF yields more natural partitions due to forming the labels *during* label assignment and preserves fine structure for $s = 1$, in agreement with the experiments discussed in Section 5.6.2. This latter effect is considerably enhanced when nonuniform weights (5.77) are used, independently of s, without compromising the quality of the spatial structure of the resulting partitions.

5.6.3.2. Spatial Feature Augmentation and Normalized Spectral Cuts

Figure 5.9 displays the corresponding results for spectral clustering and fast partitioning, respectively, using two parameter values enforcing weak and strong spatial regularization in either case.

We observe that spectral clustering is highly sensitive to the value of α. Small values yield noisy partitions, whereas larger values yield biased partitions (e.g. red cluster). We attribute this strange behavior to the conceptual deficiency of spatial regularization performed by feature augmentation, as discussed in the last paragraph of Section 5.5.1.

Fast partitioning returned the closest labelings to those computed by the SAF. The scale of spatial regularization is not consistently enforced everywhere, however, as e.g. the small red dots on the cactus arms reveal. We attribute this to the above-mentioned fact that fast partitioning directly operates on the feature space, rather than separating data representation from inference using labels and label assignments. In addition, the variational decomposition may be susceptible to getting stuck in suboptimal minima.

5.6.4. Learning and Assignment of Patch Dictionaries

In this section, we base the self-assignment flow (SAF) on more advanced features, viz. feature *patches*, and a corresponding locally invariant distance function.

5.6.4.1. Locally Invariant Patch Distances

Let

$$\mathcal{N}_{\mathcal{P},i}, \quad i \in \hat{\mathcal{I}}, \qquad\qquad n_{\mathcal{P}} := |\mathcal{N}_{\mathcal{P},i}|, \quad \forall i \tag{5.78}$$

denote quadratic sections centered at pixel (vertex) i of the underlying image grid graph, with uniform size $n_{\mathcal{P}} = 2k + 1$ for some $k \in \mathbb{N}$, for every i. We only consider region centers

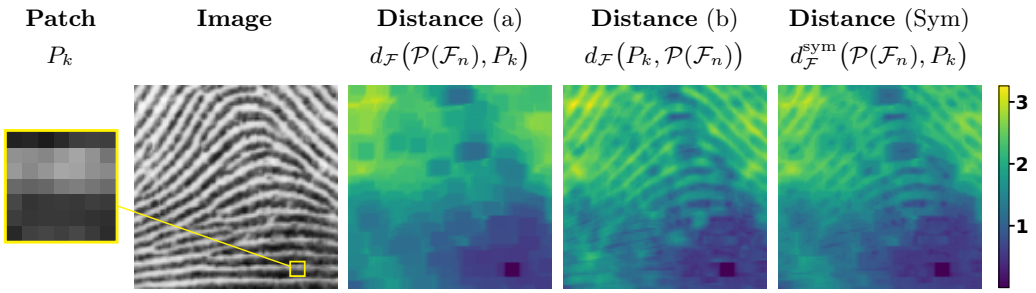

Figure 5.10.: Visualization of the distance functions (5.85) and (5.86) evaluated for a single patch P_k and all patches $\mathcal{P}(\mathcal{F}_n)$ of size $n_{\mathcal{P}} = 7 \times 7$ of the depicted image. The evaluation of distance (a) amounts to determining the minimal distance of P_k to *all* equivalence classes of patches generated by the patches of the entire image. As a consequence, equivalence classes close to P_k generate the 'blocky' graph of the distance function. Conversely, evaluation of distance (b) amounts to comparing the *single* equivalence class generated by P_k to all image patches. As a consequence, the graph of the distance function reflects the original image structure in more detail. The symmetric distance (rightmost panel) is the pointwise minimum of distance (a) and (b). It is apparent that neither distance (a) nor (b) dominates the other distance.

at interior grid points $i \in \hat{\mathcal{I}} \subset \mathcal{I}$ such that no section $\mathcal{N}_{\mathcal{P},i}$ extends beyond the boundary of the graph, which implies

$$\mathcal{N}_{\mathcal{P},i} \subset \mathcal{I}, \quad \forall i \in \hat{\mathcal{I}}. \tag{5.79}$$

We define a *patch centered at pixel* i as the ordered tuple of data points

$$P_i = \big(f_{k_1}, \dots, f_i, \dots, f_{k_{n_{\mathcal{P}}}}\big) \subset \mathcal{F}_n^{n_{\mathcal{P}}}, \qquad k_1, \dots, k_{n_{\mathcal{P}}} \in \mathcal{N}_{\mathcal{P},i}, \quad i \in \hat{\mathcal{I}}, \tag{5.80}$$

where the particular chosen order is arbitrary, but should be fixed for all patches. The individual patch features are denoted by

$$P_{i;m} = f_m, \quad m \in \mathcal{N}_{\mathcal{P},i} \tag{5.81}$$

and the collection of all patches induced by the data $\mathcal{F}_n$ is denoted by

$$\mathcal{P}(\mathcal{F}_n) = \big\{P_i \in \mathcal{F}_n^{n_{\mathcal{P}}} : \, i \in \hat{\mathcal{I}}\big\}. \tag{5.82}$$

In order to define distance functions on the set of patches, which are invariant under certain Euclidean isometries in the 2D image domain, we consider the dihedral group

$$\mathcal{D}_4 = \left\{ \left(\begin{smallmatrix} 1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & -1 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} -1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ -1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} -1 & 0 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 0 & -1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & 1 \\ 1 & 0 \end{smallmatrix}\right), \left(\begin{smallmatrix} 0 & -1 \\ -1 & 0 \end{smallmatrix}\right) \right\} \subset \mathcal{O}(2) \tag{5.83}$$

generated by the following elements of the two-dimensional orthogonal group $\mathcal{O}(2)$: four two-dimensional rotations by $\{0°, 90°, 180°, 270°\}$ and the two reflections with respect to the local coordinate axes, using the center pixel as origin. Since local grid coordinates are mapped onto each other, we can identify each transformation of the group $\mathcal{D}_4$ with a corresponding permutation σ of the pixel locations within the patch domain. Accordingly, writing with abuse of notation $\sigma \in \mathcal{D}_4$, the corresponding transformed patch (5.80) is given and denoted by

$$T_\sigma P_i = \big(f_{\sigma(k_1)}, \dots, f_i, \dots, f_{\sigma(k_{n_{\mathcal{P}}})}\big), \qquad k_1, \dots, k_{n_{\mathcal{P}}} \in \mathcal{N}_{\mathcal{P},i}, \qquad \sigma \in \mathcal{D}_4. \tag{5.84}$$

We point out that no interpolation is required to compute these patch transformations.

In addition to the transformations (5.84), we consider all translations $P_i \mapsto P_k$, $k \in \hat{\mathcal{N}}_{\mathcal{P},i}$

of patch P_i mapping the center location i to some grid location k within its own region $\hat{\mathcal{N}}_{\mathcal{P},i} := \mathcal{N}_{\mathcal{P},i} \cap \hat{\mathcal{I}}$ restricted to interior pixels. We factor out these $|\mathcal{D}_4| \cdot n_{\mathcal{P}} = 8 \cdot n_{\mathcal{P}}$ degrees of freedom by considering all corresponding transformations of patch P_i as *equivalent*. These equivalence classes of patches provide the basis for invariant patch distances as defined next.

We define the *asymmetric patch distance* between two patches centered at pixel $i \in \hat{\mathcal{I}}$ and $k \in \hat{\mathcal{I}}$ by

$$d_{\mathcal{F}}(P_i, P_k) = \min_{\substack{\sigma \in \mathcal{D}_4 \\ j \in \hat{\mathcal{N}}_{\mathcal{P},i}}} \sum_{m \in [n_{\mathcal{P}}]} d_{\mathcal{F}}\big((T_\sigma P_j)_m, P_{k;m}\big) \tag{5.85}$$

and the *symmetric patch distance* by

$$d^{\mathrm{sym}}_{\mathcal{F}}(P_i, P_k) = \min\big\{d_{\mathcal{F}}(P_i, P_k), d_{\mathcal{F}}(P_k, P_i)\big\}. \tag{5.86}$$

Figure 5.10 illustrates these locally invariant distance functions.

5.6.4.2. Recovery of Patch Prototypes and Images

Distance (5.86) defines the affinity matrix (5.5) by (5.6) and in turn the likelihood map (5.55) and the similarity map (2.88). As a consequence, the self-assignment flow can be integrated to obtain the assignment $W(t)$. In this section we focus on the recovery of prototypical patches and on 'explanations' of input images by assigning these prototypical patches. The corresponding results are illustrated by numerical examples in the subsequent Sections 5.6.4.3 and 5.6.4.4.

According to Section 5.3.3.2, prototypical patches representing each cluster are determined as weighted averages

$$P^*_j = \operatorname*{arg\,min}_{P \in \mathcal{P}(\mathcal{F})} \sum_{i \in \hat{\mathcal{I}}} \big(C(W)^{-1} W^\top\big)_{j,i} d^2_{\mathcal{F}}(P_i, P), \qquad j \in \mathcal{J}, \tag{5.87}$$

with respect to the *asymmetric* patch distance (5.85), since the prototypical patch $P \in \mathcal{P}(\mathcal{F})$ is not contained in the set of all image patches $\mathcal{P}(\mathcal{F}_n)$ (5.82).

Using these prototypes, the corresponding image is computed as follows. For each prototypical patch P^*_j, the optimal transformation for the assignment to pixel i is determined as

$$(\sigma^*_{i,j}, l^*_{i,j}) = \operatorname*{arg\,min}_{\substack{\sigma \in \mathcal{D}_4 \\ l \in \hat{\mathcal{N}}_{\mathcal{P},i}}} \sum_{m \in [n_{\mathcal{P}}]} d_{\mathcal{F}}\big((T_\sigma P_l)_m, P^*_{j;m}\big). \tag{5.88}$$

Using these transformations, a prototypical patch is assigned to every pixel $i \in \hat{\mathcal{I}}$. This implies that, for each pixel i, patches assigned to pixels $k \in \hat{\mathcal{N}}_{\mathcal{P},i}$ may assign a corresponding patch entry to pixel i. Averaging these entries, normalized by the number of values contributing to pixel i, defines the restored image value at pixel i.

To make this mathematically more precisely, we assume binary assignments $W \in \mathcal{W}^c_*$ and define the set of contributing pixels (overlapping after translation $l^*_{i,j}$) to pixel $i \in \hat{\mathcal{I}}$ by

$$M_i = \bigcup_{j \in \mathcal{J}} \big\{k \in M_{i,j} : W_{k,j} = 1\big\}, \quad \text{with} \quad M_{i,j} = \big\{k \in \hat{\mathcal{N}}_{\mathcal{P},i} : i \in \hat{\mathcal{N}}_{\mathcal{P},l^*_{k,j}}\big\}, \tag{5.89}$$

the reconstructed value for pixel i is defined as the average

$$f'_i = \frac{1}{|M_i|} \sum_{j \in \mathcal{J}} \sum_{k \in M_{i,j}} W_{k,j} \big(T_{\sigma^*_{i,j}} P^*_j\big)_{m(l^*_{k,j}, i)}, \tag{5.90}$$

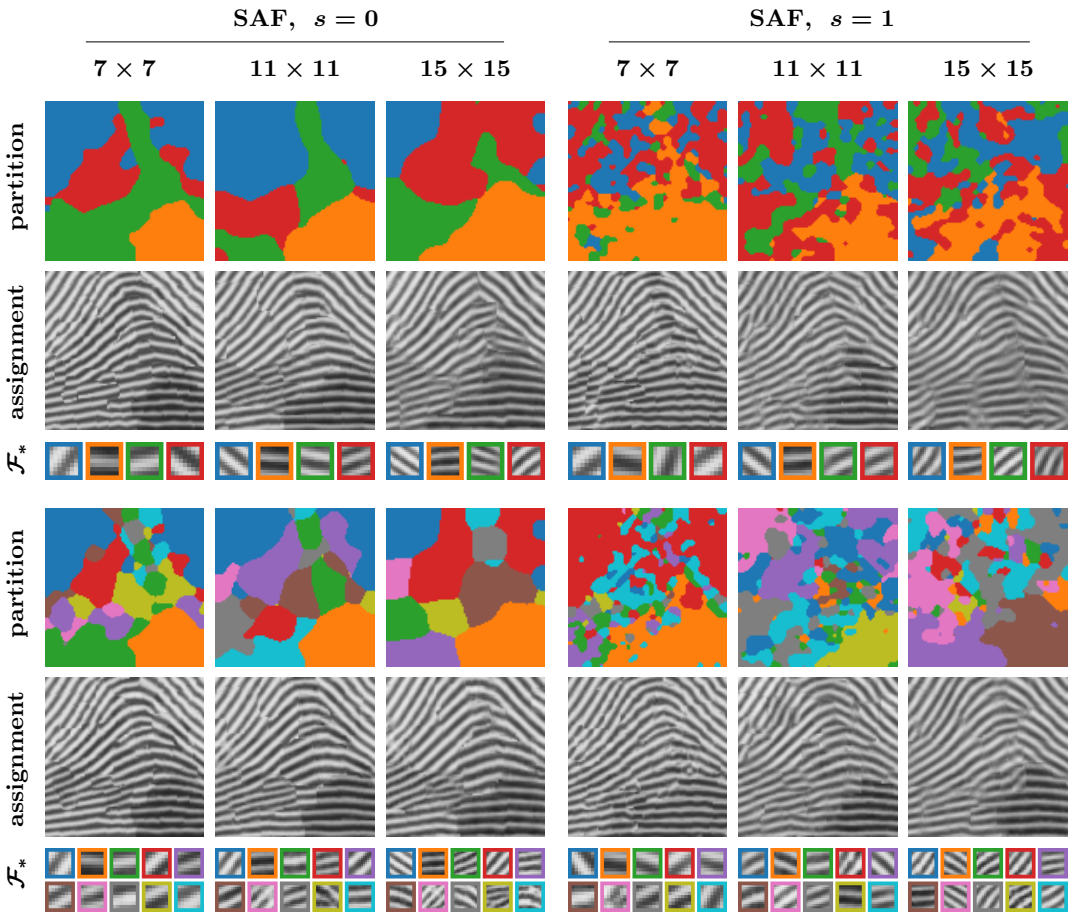

Figure 5.11.: Determination of locally invariant patch prototypes, their assignment to the original image data and the corresponding partitions (depicted with pseudo-colors), using the SAF ($s = 0$ and $s = 1$), different patch sizes (7×7, 11×11, 15×15) and numbers of prototypes ($c = 4$ and $c = 10$). The underlying transformation group enables accurate image representations even with only $c = 4$ patches, provided the patch size is close to the spatial scale of local image structure (here: 7×7 pixels). This performance deteriorates for larger patch sizes. The SAF with $s = 0$ yields partitions that are spatially more regular than the partitions computed with $s = 1$, since the latter tends to cover the feature space more uniformly with patch prototypes, in agreement with the result depicted by by Figure 5.5.

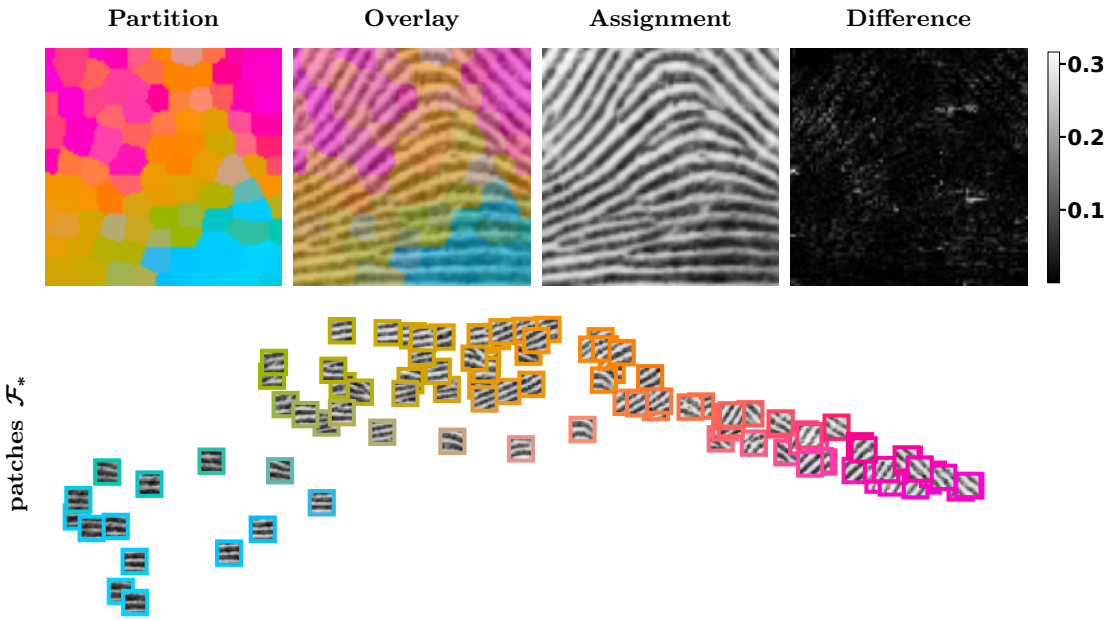

Figure 5.12.: Experiment of Fig. 5.11 repeated with a larger patch dictionary (patch size $n_{\mathcal{P}} = 15 \times 15$) leads to a detailed representation of local image structure. Although overlapping regions of assigned prototypical patches are averaged at each pixel in order to restore an image, the result 'Assignment' is close to the input data 'Image' of Fig. 5.10, due to using the locally invariant patch distance. Panel 'Difference' shows the difference as gray-value plot. The lower panel displays a 2D embedding of the learned prototypical patches. The corresponding colors indicate their assignment in 'Partition' and 'Overlay'. Clusters in the lower panel, e.g. those colored pink and blue, illustrate the invariance under discrete rotations and reflections.

where $m(l^*_{k,j}, i)$ denotes the entry of the prototypical patch j, which overlaps with pixel i after translating its center by $l^*_{k,j}$ from the contributing pixel k.

5.6.4.3. Patch-Based Self-Assignment Flow

Figure 5.11 illustrates image partitions, the corresponding $c = 4$ and $c = 10$ prototypical patches of sizes $n_{\mathcal{P}} \in \{7 \times 7, 11 \times 11, 15 \times 15\}$, as well as their assignment to the input image data as described in the preceding section, based on integrating the SAF with $s = 0$ and $s = 1$ and spatial regularization parameter $|\mathcal{N}| = 3 \times 3$.

In agreement with the discussion of the results depicted by Figure 5.5, we observe that the SAF with $s = 0$ returns partitions with a more regular spatial structure, whereas the SAF with $s = 1$ tends to cover the feature space more uniformly which yields partitions with a irregular spatial structure.

The image recovered by assigning the prototypical patches exhibits relatively sharp spatial structures, despite the small number of prototypes ($c \in \{4, 10\}$) and the pixelwise averaging of grayvalues assigned by multiple patches. This illustrates that the small transformation group defined in Section 5.6.4.1 which does not even require image interpolation, actually is quite powerful. For example, the large blue region of the partition shown in Figure 5.11 resulting from the SAF with $s = 0$ and 7×7 patches, indicates the optimal assignment of

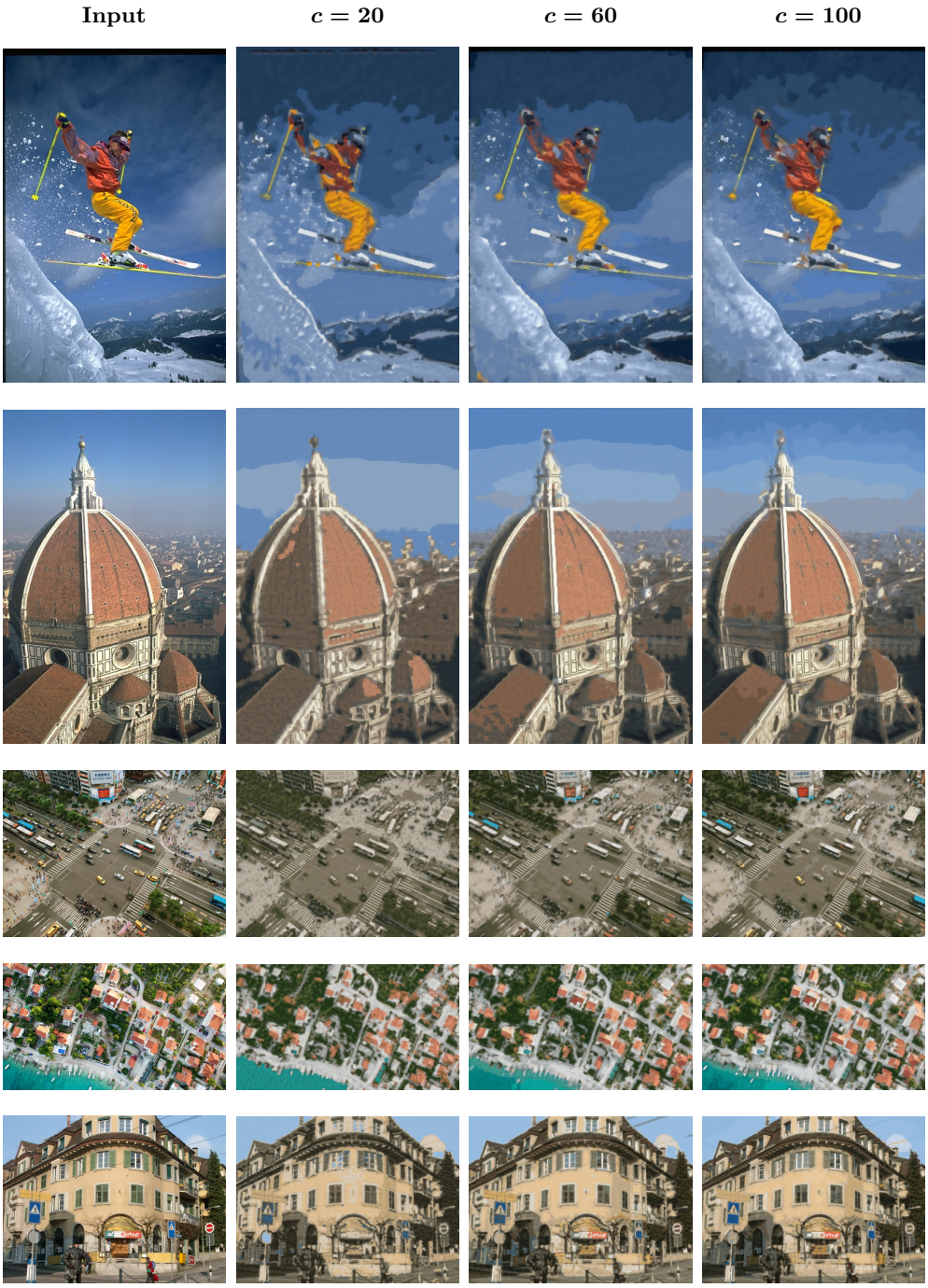

Figure 5.13.: 'Input' images (left-most column) are represented in a compact way by unsupervised patch learning and assignment using the SAF with $s = 1$, $|\mathcal{N}| = 3 \times 3$ for geometric regularization, and with increasing dictionary sizes $c \in \{20, 60, 100\}$ of locally invariant patches of size 7×7, as described in Section 5.6.4. The recovered images are shown in the remaining three columns. We observe that for more complex real-world scenarios, a larger number of patches is required for representing all local details (e.g., see the arcs of the dome in the second row). This suggests to extend the local patch invariance towards affine transformations with arbitrary rotations and scalings, which requires more expensive interpolation of the pixel-grid, however.

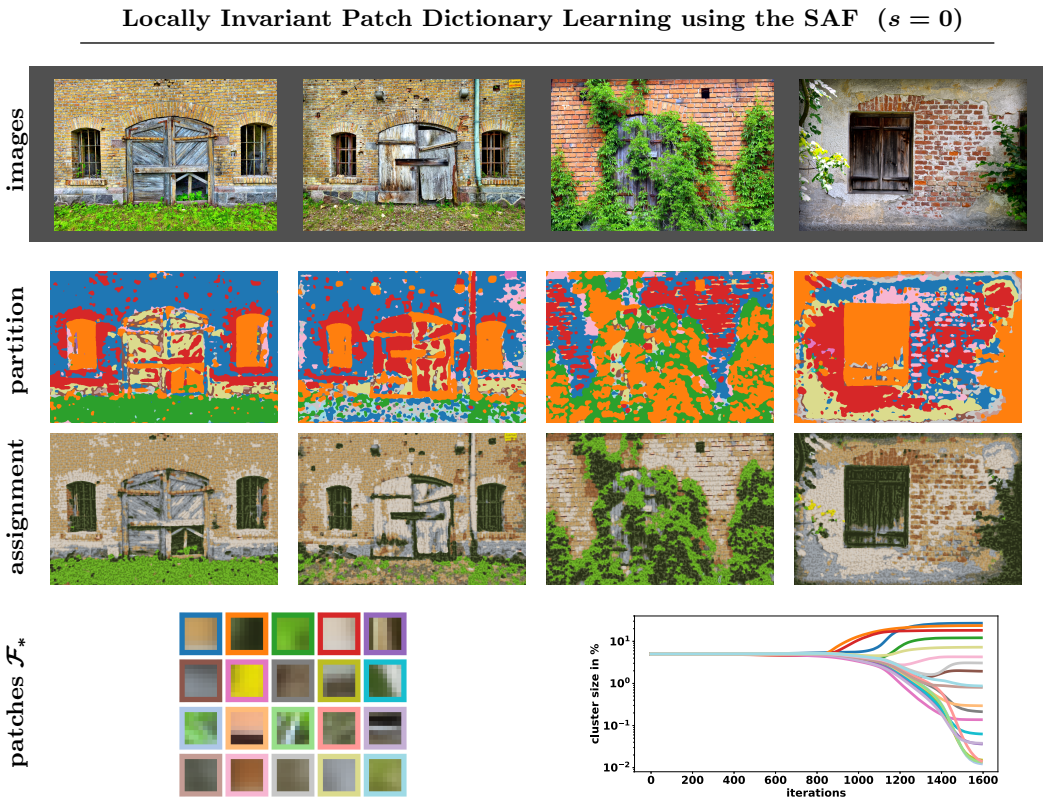

Figure 5.14.: The bottom row shows a dictionary of $c = 20$ locally invariant patches of size 7×7 pixels, respectively learned from the four images shown in the top row using the SAF with $s = 0$ and $|\mathcal{N}| = 3 \times 3$ pixels. The second and third row illustrate the patch assignments with pseudo-colors and the recovered image data, respectively. Closeness of the restored images to the input data, despite the small size of the patch dictionary, demonstrates the effectiveness of the underlying discrete transformation group. The evolution of cluster sizes (bottom row, right panel) illustrates the ability of the SAF to resolve 'conflicting' assignments due to mutually overlapping patches successfully, along with the formation of invariant patch prototypes, in a completely unsupervised way.

patches from a *single* equivalence class only. These patches fit quite accurately to image structures with different orientations and local edge profiles. This effect deteriorates when using patch sizes that are much larger than the typical variations of local image structure, as the results for the patch size 15×15 with $c = 4$ and $c = 10$ show.

As a comparison, Figure 5.12 shows the result for a larger number $c = 100$ of prototypes with patch size $n_{\mathcal{P}} = 15 \times 15$, which leads to a detailed representation of local image structure. The lower panel displays a two-dimensional embedding of the weighted graph with prototypes as patches and the similarities (5.6) as weights. Representatives of equivalence classes of patches that are close to each other, are grouped together. Factoring out the group of transformations effectively copes with different edge profiles and orientations. Panel 'Difference' shows the absolute difference between the input image and the reconstruction from patch labeling.

We additionally evaluated the unsupervised patch-based SAF approach using various

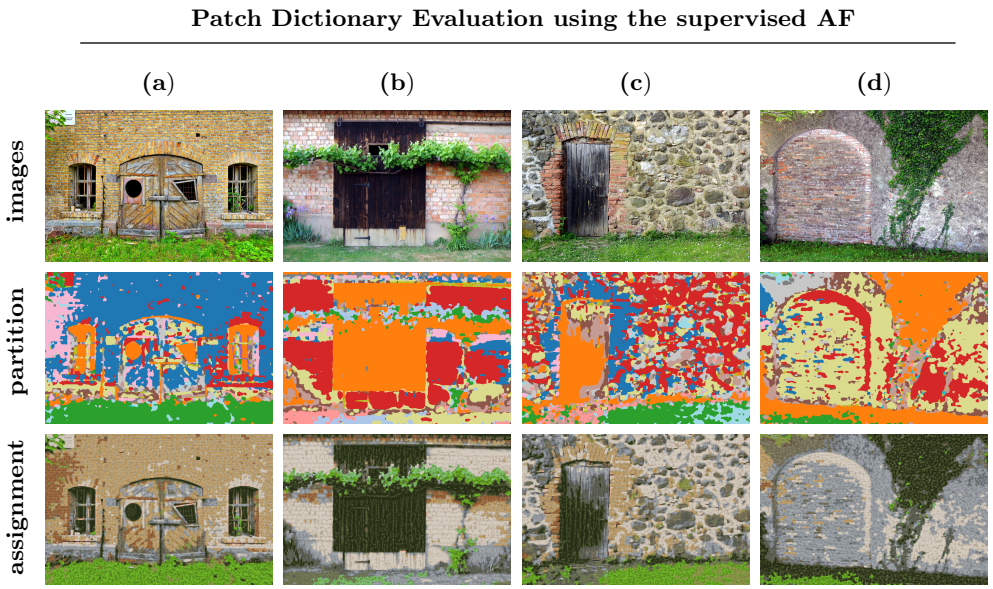

Figure 5.15.: Supervised regularized assignment of the locally invariant patch dictionary from Figure 5.14 using the AF, to four *novel* images (top row). Since these images are semantically similar to the training data from Figure 5.14, the restored images are close to the input data, except for image (c) whose stone wall texture is not present in the training data.

real-world images. Figure 5.13 depicts the input data as well as the resulting patch assignments for an increasing number of labels $c \in \{20, 60, 100\}$.

5.6.4.4. Patch Assignment to Novel Data

We repeated the experiment illustrated by Figure 5.11 using the data shown in Figure 5.14. $c = 20$ locally invariant prototypical patches of size 7×7 pixels were learned from 4 images using the SAF with $s = 0$ and $|\mathcal{N}| = 3 \times 3$ pixels. The restored images shown in the third row are remarkably close to the input data (first row), despite the small size $c = 20$ of the patch dictionary. This demonstrates again the effectiveness of the underlying discrete transformation group.

Figure 5.15 shows in the top row *novel* image data that are semantically similar to the training images of Figure 5.14 regarding the local image structure and texture (brick/stone, door/window, grass/ivy). The corresponding partitions and recovered images solely resulted from assigning the patch dictionary depicted by Figure 5.14 to the data by the supervised assignment flow. Again, the quality of image represention using this small dictionary is remarkable, except for the stone wall texture shown in column (c) of Figure 5.15, which is not present in the training data.

5.6.5. Regularized Clustering of Weighted Graph Data

Our approach can be applied to any data given on any undirected weighted graph. For illustration, we include an additional experiment using data not related to image analysis.

Figure 5.16 shows data in terms of a weighted graph $(\mathcal{I}, \mathcal{E}, K_{\mathcal{E}})$ adopted from [GN02]. It

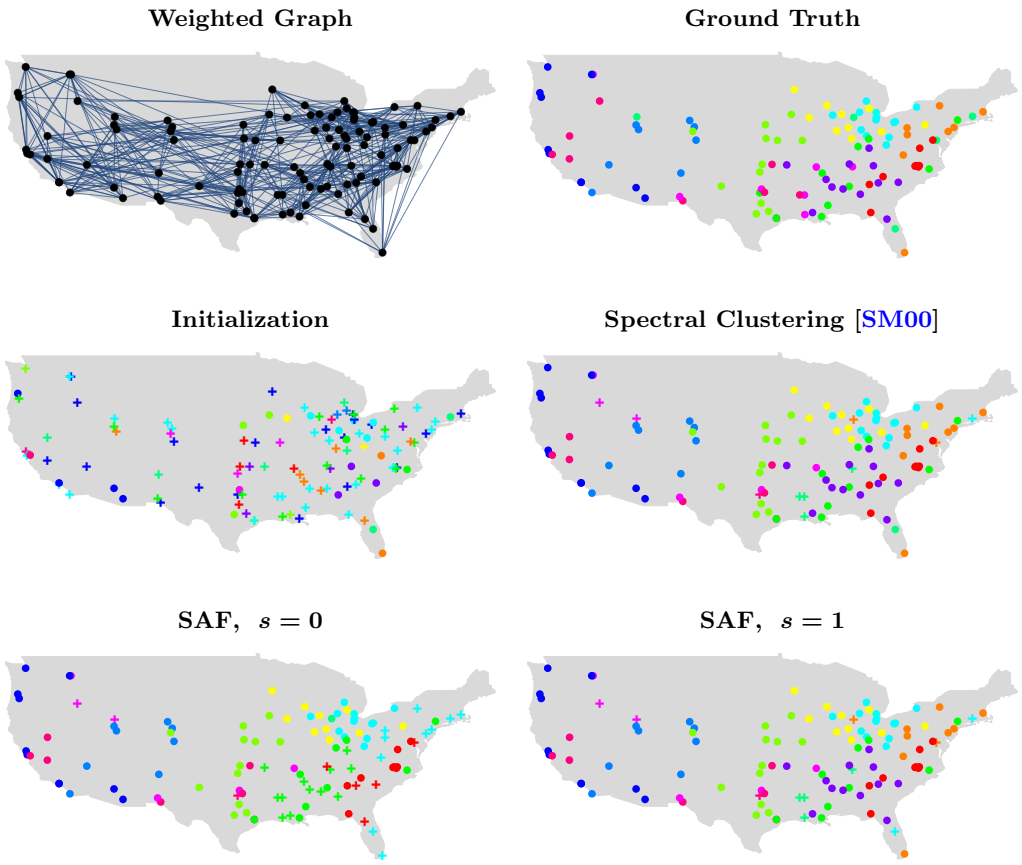

Figure 5.16.: Weighted graph data of American football games between Division IA colleges during the regular season fall 2000 are clustered. Each node represents a team and edge weights indicate the number of games played between two teams. The colored nodes in 'Ground Truth' show the subdivision of the teams into 12 conferences (clusters), that primarily play against each other in a first period. Graph partitioning with $c = 12$ was performed using the SAF with $s = 0$ and $s = 1$, and with weights defined by (5.91). Markers ✚ indicate labels assigned to nodes that differ from ground truth. Starting from the initialization (2nd row, left panel) which is noisy, the SAF with $s = 1$ returns almost the ground truth labeling and is also close to the result of directly applying spectral clustering to $K_{\mathcal{E}}$. The SAF with $s = 0$ enforces label assignments that are spatially more regular, and with empty clusters orange and purple.

represents the network of American football games between Division IA colleges during the regular season fall 2000. Teams are subdivided into 12 conferences, mainly based on geographical distance, that primarily play against each other in a first period. Afterwards, the conference champions play against each other in the final games. Each node of the network represents a team. Edge weights $K_{\mathcal{E}i,k}$ represent the number of games played between two teams. Labels for each vertex indicate the conference to which a team belongs, displayed by a corresponding color in Figure 5.16 (ground truth). We considered this labeling as ground truth for the task to partition the graph into $c = 12$ classes. The initial perturbation of the barycenter (5.56a) in terms of a distance matrix $D_{\mathcal{F},0}$ was computed by assigning feature vectors to each node based on the c dominant eigenvectors of $K_{\mathcal{E}}$, followed by greedy k-center clustering (Section 2.4.3). Markers ✚ indicate nodes that were assigned to a conference different from ground truth. Weights were defined as

$$w_{i,k} = \frac{\tilde{w}_{i,k}}{\langle \tilde{w}_i, \mathbb{1}_n \rangle} \quad \text{with} \quad \tilde{w}_{i,k} = K_{\mathcal{E}i,k} + \operatorname{Diag}(K_{\mathcal{E}} \mathbb{1}_n), \tag{5.91}$$

i.e. by adding the total number of games played by each team to the diagonal.

The nearest neighbor assignment of the initial distance matrix contains many erroneous assignments (Figure 5.16, initialization). The results of the SAF with $s = 1$ reproduces almost the ground truth labeling and is also close to the result of applying spectral clustering [SM00] directly to $K_{\mathcal{E}}$. The SAF with $s = 0$ enforces assignments with a more regular spatial structure. Both findings agree with observations made in preceding experiments; see e.g. Figure 5.5.

5.7. Conclusion

In this chapter we introduced a smooth dynamical system for unsupervised contextual data labeling on arbitrary undirected weighted graphs. To this end, we extended the supervised assignment flow to unsupervised scenarios, where no labels are available. To remove the necessity of prototypes we replaced them with a copy of the input data by abstracting from "pixel-label" decisions to "pixel-pixel" decisions. The resulting family of self-assignment flows takes a pairwise affinity matrix as input data and maximizes the correlation with a self-assignment matrix, corresponding to a low-rank factorization parameterized by variables of the assignment flow. This allows the recovery of latent prototypes, which are emerging during the evolution.

A single parameter $s \in [0, 1]$ determines the self-assignment matrix as smooth geodesic interpolation between the self-affinity matrix ($s = 0$) and the self-influence matrix ($s = 1$). It enables to control the relative influence of spatial regularization and the preservation of local image structures. In addition, a second parameter, the neighborhood size $|\mathcal{N}|$ for geometric averaging of assignments, controls the scale of the resulting partition and the resulting number of clusters. In particular, the self-assignment flow is a single process exclusively evolving on the assignment manifold, such that numerical techniques developed by [ZSPS20] for integrating the assignment flow directly apply.

We related our approach mathematically to different relevant viewpoints: rank-constrained discrete optimal transport, normalized spectral cuts with unbiased spatial regularization and relaxed combinatorial optimization in large-scale scenarios subject to spatial regularization. We demonstrated the properties of our approach using various experiments including the unsupervised learning of patch dictionaries using a locally invariant distance function.

In the context of this work, an immediate question is, if the unsupervised self-assignment flow can be extended to labeling problems with indirect measurements as input, which were considered in the supervised case in the first two chapters for the application of discrete tomographic reconstruction. This generalized setting in which neither the prototypes are given beforehand nor the data to be labeled is directly available poses a challenging task to solve for future work.

Another possible extension of the self-assignment flow concers problems defined on directed graphs. This requires an asymmetric generalization of the self-assignment matrix factorization involving two coupled assignment flows as factors. Moreover connections to co-clustering have to be investigated, where the task is to simultaneously partition two different data sets and associate the clusters between both sets

A different open problem concerns the incorporation of parameter estimation of weights for application-specific adaptive regularization [HSPS19] in the unsupervised self-assignment flow.

Promising directions of further research also include application-dependent extensions of the invariance group in order to learn compact patch dictionaries using the self-assignment flow in various scenarios.

Finally a rigorous mathematical exploration of the proposed smooth dynamical system in terms of stability and convergence, analogous to the study [ZZS20] for the supervised assignment flow, is left for future work.

CHAPTER 6

Conclusion

Summary

With the work presented in this thesis we hope to have provided general, but also practically applicable approaches for two challenging extensions of the data labeling problem. In particular, we assumed that either the input data cannot be observed directly or the prototypes are not available beforehand. In both cases, the missing information had to be inferred during the assignment of labels. Therefore, our strategy has been to consistently link the recovery of missing information as closely as possible to the labeling process. As a result, it turned out that smoothness together with the compositional modular design of the assignment flow framework are the key to intertwining both processes tightly. Finally, a geometric view of the data labeling problem allowed the extension to more complicated tasks.

Chapters 3 and 4 of this work were concerned with the scenario where no direct input data are available. Instead, indirect measurements are available. The main application we considered was discrete tomography reconstruction from only a few projection angles with known intensities. Our strategy was to express the projection constraints directly in terms of decision variables of the labeling problem. However, non-integral solutions are introduced as convex combinations of prototypes. Motivated by the naive fixed-point iteration, that labels the previous solution, we derived a non-convex discretization term for pixelwise independent decisions which excludes non-integral solutions by using the prior knowledge (prototypes). The resulting overall non-convex energy for joint reconstruction and labeling was reliably and efficiently minimized by the difference of convex functions algorithm which provably converges to a stationary point.

Afterwards we focused on the underlying geometric aspects of label decisions encoded by discrete probability distributions. As a first step we derived a regularization term for spatially coherent assignments based on the Bregman divergence which locally approximates the squared geodesic distance induced by the Fisher-Rao information metric on the open probability simplex. Finally, the geometric point of view led to a smooth Riemannian gradient flow evolving on a submanifold including the tomographic projection constraints

directly into the geometry of assignments. Therefore, the feasible set was equipped with a Hessian Riemannian metric which naturally extends the Fisher-Rao metric of the assignment manifold. Furthermore, we investigated corresponding implicit numerical schemes which rely on Bregman proximal mappings for geometric integration. The reconstruction performance turned out to be superior to the state of the art.

Chapter 5 was devoted to unsupervised scenarios of the data labeling problem. We have introduced and studied a smooth dynamical system for unsupervised contextual data labeling on arbitrary undirected weighted graphs. To this end, we extended the supervised assignment flow to the unsupervised setting, where no labels are available. In the absence of prototypes, we rigorously abstracted from "data-label" to "data-data" decisions by introducing interpretable low-rank data representations, which themselves are parameterized by label assignments. In particular, we have defined a one-parameter family of self-assignment matrices by smooth geodesic interpolation between the normalization factors of different low-rank matrix factorizations. This enabled interpolation between combinatorial and spectral aspects. Finally, based on these factorizations we generalized the likelihood map of the supervised assignment flow in a natural way. The resulting unsupervised self-assignment flow simultaneously performs learning of latent prototypes in the very same framework in which they are used for inference, i.e. label assignment.

We related the approach mathematically to different relevant viewpoints: rank-constrained discrete optimal transport, normalized spectral cuts with unbiased spatial regularization and relaxed combinatorial optimization in large-scale scenarios subject to spatial regularization. We demonstrated the properties of our approach using various experiments including the unsupervised learning of patch dictionaries using a locally invariant distance function.

Future Work

Finally, we would like to point out promising directions for further research that seem most central to us:

- Combining the unsupervised scenario with indirect measurements as input. This generalized setting in which neither the prototypes are given beforehand nor the data to be labeled is directly available poses a challenging task to solve.
- Integrating the parameter estimation of weights for application-specific adaptive regularization [HSPS19] into the unsupervised self-assignment flow framework.
- Extending the self-assignment flow to directed graphs, which requires an asymmetric formulation of the self-assignment matrix factorization resulting in two coupled assignment flows as factors. Moreover connections to co-clustering have to be investigated, where the task is to simultaneously partition two different data sets and associate the clusters between both sets.
- Analyzing in a rigorous mathematical way the proposed smooth dynamical systems, the tomographic assignment flow and the self-assignment flow, in terms of stability and convergence. Such issues are not covered by standard theory for dynamical systems.

The author hopes that this these stimulates corresponding future research work.

APPENDIX A

Supplementing Proofs

A.1. Supplementing Proofs of Chapter 5

The following lemma is required to prove the next proposition.

Lemma A.1.1 (Basis Exchange Lemma [Fis14])**.** *Let V be a finite dimensional vector space over K and $B = \{b_1, b_2, \ldots, b_n\}$ a basis of V. Moreover let $w \in V$ be represented as a linear combination $w = \sum_{i=1}^{n} \lambda_i b_i$ with $\lambda_i \in K$. Then for $k \in [n]$ such that $\lambda_k \neq 0$ the basis vector b_k can be exchanged by w and $\tilde{B} = \{b_1, \ldots, b_{k-1}, w, b_{k+1}, \ldots, b_n\}$ is also a basis of V.*

The following proposition refers to the manifold $\mathcal{W}^c$ of full-rank assignments that is defined by Definition 5.3.2 and is discussed in Section 5.4.1.

Proposition A.1.2. *The smooth manifold $\mathcal{W}^c$ is path-connected for $c < n$.*

Proof. First we observe that permuting the rows of the matrix $W \in \mathcal{W}^c$ with $c < n$ can be realized by a continuous path in the set $\mathcal{W}^c$. Let σ be a permutation of n elements and $P_\sigma \in \mathbb{R}^{n \times n}$ the associated permutation matrix. The row permuted matrix $P_\sigma W$ again belongs to the set $\mathcal{W}^c$, since positivity and normalization are row-wise constraints and the global rank constraint is invariant under permutation of rows. Then a continuous path from W to $P_\sigma W$ can be constructed by decomposing the permutation into successive transpositions. Therefore it is sufficient to consider a path which swaps row k and l of the matrix W. Since $W \in \mathbb{R}^{n \times c}$ and $\operatorname{rank}(W) = c$ with $c < n$ we can select c linearly independent rows from W which are indexed by the set I and we denote the indices of the remaining rows by $\tilde{I} = [n] \setminus I$. It is important that $\tilde{I}$ is non-empty since $|\tilde{I}| = n - c > 0$. Now we distinguish three cases:

1. swap rows $k, l \in I$: let $p \in \tilde{I}$

$$W^{(0)} = \sum_{i \in I} e_i W_i + \sum_{j \in \tilde{I}} e_j W_j \tag{A.1a}$$

$$W^{(1)} = \sum_{i \in I} e_i W_i + \sum_{j \in \tilde{I} \setminus \{p\}} e_j W_j + e_p W_k \tag{A.1b}$$

$$W^{(2)} = \sum_{i \in I \setminus \{k\}} e_i W_i + \sum_{j \in \tilde{I} \setminus \{p\}} e_j W_j + e_p W_k + e_k W_l \tag{A.1c}$$

$$W^{(3)} = \sum_{i \in I \setminus \{k,l\}} e_i W_i + \sum_{j \in \tilde{I} \setminus \{p\}} e_j W_j + e_p W_k + e_k W_l + e_l W_k \tag{A.1d}$$

$$W^{(4)} = \sum_{i \in I \setminus \{k,l\}} e_i W_i + \sum_{j \in \tilde{I}} e_j W_j + e_k W_l + e_l W_k, \tag{A.1e}$$

   where between each $W^{(i)}$ and $W^{(i+1)}$ for $i = \{0, 1, 2, 3\}$ we define the path

$$\gamma_{i,i+1} : [0, 1] \longrightarrow \mathcal{W}^c, \quad t \longmapsto (1 - t)W^{(i)} + tW^{(i+1)}, \tag{A.2}$$

   which is well-defined since the row-wise positivity constraints and normalization are preserved under convex combinations (simplex is a convex set). In addition, for each transition $\gamma_{i,i+1}(t)$ there always exists a subset of c linear independent rows which are constant with respect to t, hence $\operatorname{rank}(\gamma_{i,i+1}(t)) = c$ for $t \in [0, 1]$. In particular, by construction we have I for $\gamma_{0,1}$, $I \setminus \{k\} \cup \{p\}$ for $\gamma_{1,2}$, $I \setminus \{l\} \cup \{p\}$ for $\gamma_{2,3}$ and I for $\gamma_{3,4}$ as index sets for c constant and linear independent rows. Finally, by joining all $\gamma_{i,i+1}$ we get a continuous path for swapping rows k and l.

2. swap rows $k \in I$ and $l \in \tilde{I}$:

$$W^{(0)} = \sum_{i \in I} e_i W_i + \sum_{j \in \tilde{I}} e_j W_j \tag{A.3a}$$

$$W^{(1)} = \sum_{i \in I} e_i W_i + \sum_{j \in \tilde{I} \setminus \{l\}} e_j W_j + e_l W_k \tag{A.3b}$$

$$W^{(2)} = \sum_{i \in I \setminus \{k\}} e_i W_i + \sum_{j \in \tilde{I} \setminus \{l\}} e_j W_j + e_l W_k + e_k W_l, \tag{A.3c}$$

$$\tag{A.3d}$$

   analogously to the first case we can construct a continuous path in the set $\mathcal{W}^c$ for swapping rows k and l.

3. swap rows $k, l \in \tilde{I}$: we can define a single continuous path

$$\gamma : [0, 1] \longrightarrow \mathcal{W}^c, \quad t \longmapsto (1 - t)W + t(W - e_k W_k - e_l W_l + e_l W_k + e_k W_l), \tag{A.4}$$

   in the set $\mathcal{W}^c$, since I is an index set for c linear independent rows which are constant with respect to t, hence $\operatorname{rank}(\gamma(t)) = c$.

Finally, we can always construct a continuous path from W to $P_\sigma W$ by concatenation of continuous paths of successive transpositions.

Next, let X and Y be two arbitrary points of the set $\mathcal{W}^c$ with fixed $c < n$. Then we can find two permutations P_{σ_W} and P_{σ_V} such that the first c rows of $W = P_{\sigma_W} X$ and $V = P_{\sigma_V} Y$ are linear independent, i.e. the first rows of each matrix form a basis of the vector space $\mathbb{R}^c$. The basic idea is to apply the basis exchange Lemma A.1.1 successively.

Let $p \in [c]$ be a row-index such that W_p can be exchanged with V_1, i.e. $\lambda_p \neq 0$ for the linear combination $V_1 = \sum_{i=1}^{c} \lambda_i W_i$. This exchange can be realized by a continuous path in $\mathcal{W}^c$ through the following points

$$Z^{(0)} = W = \sum_{i=1}^{c} e_i W_i + \sum_{j=c+1}^{n} e_j W_j \tag{A.5a}$$

$$Z^{(1)} = \sum_{i\in[c]} e_i W_i + \sum_{j\in[n]\setminus[c+1]} e_j W_j + e_{c+1} W_p \tag{A.5b}$$

$$Z^{(2)} = \sum_{i\in[c]\setminus\{p\}} e_i W_i + \sum_{j\in[n]\setminus[c+1]} e_j W_j + e_{c+1} W_p + e_p V_1 \tag{A.5c}$$

$$Z^{(3)} = \sum_{i\in[c]\setminus\{p\}} e_i W_i + \sum_{j\in[n]\setminus[c]} e_j W_j + e_p V_1, \tag{A.5d}$$

$$\tag{A.5e}$$

where between each $Z^{(i)}$ and $Z^{(i+1)}$ for $i = \{0, 1, 2\}$ we define the path

$$\gamma_{i,i+1} : [0,1] \longrightarrow \mathcal{W}^c, \quad t \longmapsto (1-t)Z^{(i)} + tZ^{(i+1)}, \tag{A.6}$$

which is well-defined since the row-wise constraints positivity and normalization are preserved under convex combinations (simplex is a convex set). In addition, for each transition $\gamma_{i,i+1}(t)$ there always exists a subset of c linear independent rows which are constant with respect to t, hence $\operatorname{rank}(\gamma_{i,i+1}(t)) = c$ for $t \in [0,1]$. In particular, by construction we have $[c]$ for $\gamma_{0,1}$, $[c] \setminus \{p\} \cup \{c+1\}$ for $\gamma_{1,2}$ and $[c]$ for $\gamma_{2,3}$ as index sets for c constant and linear independent rows. Note that, for the last segment $\gamma_{2,3}$ we have used, that $\{W_1, \ldots, W_{p-1}, V_1, b_{p+1}, \ldots, W_c\}$ is a basis by the basis exchange Lemma A.1.1. Finally, by joining all $\gamma_{i,i+1}$ we find a continuous path for exchanging rows W_p and V_1.
To exchange the V_k, $k \in [c]$ with some W_p we have the starting point

$$Z = \sum_{i\in[k-1]} e_i V_i + \sum_{i\in[c]\setminus[k-1]} e_i W_i + \sum_{i\in[n]\setminus[c]} e_i W_i, \tag{A.7}$$

where we assume that the first c rows are rearranged by some permutation such that the $k-1$ first rows are the already exchanged $\{V_1, \ldots, V_{k-1}\}$ and the remaining $c-k+1$ are (possibly renumbered) $\{W_k, \ldots, W_c\}$ row-vectors of W. We can always find an index $p \in [c]$ with $p > k-1$ with $\lambda_p \neq 0$ for the linear combination $V_k = \sum_{i\in[k-1]} \lambda_i V_i + \sum_{i\in[c]\setminus[k-1]} \lambda_i W_i$, otherwise it would contradict the linear independence of V_i for $i \in [c]$. Hence we can construct a continuous path in $\mathcal{W}^c$ through the following points

$$Z^{(0)} = Z = \sum_{i\in[k-1]} e_i V_i + \sum_{i\in[c]\setminus[k-1]} e_i W_i + \sum_{i\in[n]\setminus[c]} e_i W_i \tag{A.8a}$$

$$Z^{(1)} = \sum_{i\in[k-1]} e_i V_i + \sum_{i\in[c]\setminus[k-1]} e_i W_i + \sum_{j\in[n]\setminus[c+1]} e_j W_j + e_{c+1} W_p \tag{A.8b}$$

$$Z^{(2)} = \sum_{i\in[k-1]} e_i V_i + \sum_{i\in[c]\setminus\{[k-1]\cup\{p\}\}} e_i W_i + \sum_{j\in[n]\setminus[c+1]} e_j W_j + e_{c+1} W_p + e_p V_k \tag{A.8c}$$

$$Z^{(3)} = \sum_{i\in[k-1]} e_i V_i + \sum_{i\in[c]\setminus\{[k-1]\cup\{p\}\}} e_i W_i + \sum_{j\in[n]\setminus[c]} e_j W_j + e_p V_k, \tag{A.8d}$$

$$\tag{A.8e}$$

which is similar to the first exchange step, hence we constructed a continuous path in $\mathcal{W}^c$

to exchange V_k by some W_p with $p > k-1$. Thus we can concatenate all paths required to exchange the first c rows of W and V to a single continuous path. Subsequently we can find a suitable permutation P_{σ_Z} of the resulting Z to rearrange the first c rows such that they coincide with the first c rows of V. Next we can find a path between the remaining $(P_{\sigma_Z}Z)_i = W_i$ and V_i for $i \in [n] \setminus [c]$, which are linear dependent rows (since the first c already form a basis). Therefore, we can define the path in $\mathcal{W}^c$

$$\gamma : [0,1] \longrightarrow \mathcal{W}^c, \quad t \longmapsto (1-t)P_{\sigma_Z}Z + tV, \tag{A.9}$$

where the first c rows are constant with respect to t and linear independent, hence $\operatorname{rank}(\gamma(t)) = c$. Finally we have to undo the permutation from Y to V by using $Y = P_{\sigma_V^{-1}}V$, which again can be realized by some continuous path. The following diagram summarizes all stages to construct a continuous path from X to Y

$$X \xrightarrow{P_{\sigma_W}} W \xrightarrow{V_1 \to W_p} \cdots \xrightarrow{V_k \to W_p} \cdots \xrightarrow{V_c \to W_p} Z \xrightarrow{P_{\sigma_Z}} P_{\sigma_Z}Z \xrightarrow{V_i \to W_i,\, i>c} V \xrightarrow{P_{\sigma_V^{-1}}} Y, \tag{A.10}$$

which proofs that the set $\mathcal{W}^c$ is path-connected if $c < n$. □

The following proposition refers to the manifold $\mathcal{W}^c$ of full-rank assignments that is defined by Definition 5.3.2 and is discussed in Section 5.4.1.

Proposition A.1.3. *The smooth manifold $\mathcal{W}^n$ is not connected.*

Proof. We use the fact, that the general linear group $\mathrm{GL}(n)$ of invertible real matrices has two connected components: matrices with positive determinant and matrices with negative determinant.

After intersecting the general linear group with the assignment constraints, i.e. $\mathcal{W}^n = \mathcal{W} \cap \mathrm{GL}(n) \subset \mathrm{GL}(n)$ we still find two elements with a negative and positive determinant:

for a positive determinant we can choose the identity matrix perturbed by some small ε

$$W^+ = (1-\varepsilon)I_n + \frac{\varepsilon}{n}\mathbb{1}_{n\times n} \in \mathcal{W}^n, \tag{A.11}$$

for a negative determinant we can choose the identity matrix with two rows i and j swapped by a permutation matrix $P_{\pi_{ij}}$ and again perturbed by some small ε

$$W^- = (1-\varepsilon)P_{\pi_{ij}}I_n + \frac{\varepsilon}{n}\mathbb{1}_{n\times n} \in \mathcal{W}^n. \tag{A.12}$$

Since the determinant function is continuous there exists such an ε.

Finally the two elements W^+ and W^- cannot be connected in $\mathcal{W}^n$ since they cannot be connected in the superset $\mathrm{GL}(n)$.

□

List of Figures

List of Tables

List of Algorithms

References

[ABB04] F. Alvarez, J. Bolte, and O. Brahic, "Hessian Riemannian Gradient Flows in Convex Programming," *SIAM Journal on Control and Optimization*, vol. 43, no. 2, pp. 477–501, 2004.

[AC10] S.-I. Amari and A. Cichocki, "Information Geometry of Divergence Functions," *Bulletin of the Polish Academy of Sciences: Technical Sciences*, vol. 58, no. 1, pp. 183–195, 2010.

[AMS09] P.-A. Absil, R. Mahony, and R. Sepulchre, *Optimization Algorithms on Matrix Manifolds*. Princeton University Press, 2009.

[AN00] S.-I. Amari and H. Nagaoka, *Methods of Information Geometry*. Amer. Math. Soc. and Oxford Univ. Press, 2000.

[And79] T. Ando, "Generalized Schur Complements," *Lin. Algebra Appl.*, vol. 27, pp. 173–186, 1979.

[APB+15] W. Aarle, W. Palenstijn, J. Beenhouwer, T. Altantzis, S. Bals, K.-J. Batenburg, and J. Sijbers, "The ASTRA Toolbox: A Platform for Advanced Algorithm Development in Electron Tomography," *Ultramicroscopy*, vol. 157, pp. 35–47, 2015, ISSN: 0304-3991.

[ÅPSS17] F. Åström, S. Petra, B. Schmitzer, and C. Schnörr, "Image Labeling by Assignment," *Journal of Mathematical Imaging and Vision*, vol. 58, no. 2, pp. 211–238, 2017.

[BCPD99] R. E. Burkhard, E. Cela, P. M. Pardalos, and D.-Z. Du, "Linear Assignment Problems and Extensions," in *Handbook Combin. Optimiz. Suppl. Vol. A*. Kluwer Acad. Publ., 1999, pp. 75–149.

[Bha06] R. Bhatia, *Positive Definite Matrices*. Princeton Univ. Press, 2006.

[BMDG05] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh, "Clustering with Bregman Divergences," *J. Mach. Learn. Res.*, vol. 6, pp. 1705–1749, 2005.

[Bom18] I. M. Bomze, "Building a Completely Positive Factorization," *Central Europ. J. Oper. Res.*, vol. 26, no. 2, pp. 287–305, 2018.

[BP94] A. Berman and R. J. Plemmons, *Nonnegative Matrices in the Mathematical Sciences*. SIAM, 1994.

[BR82] J. Burbea and C. R. Rao, "Entropy Differential Metric, Distance and Divergence Measures in Probability Spaces: A Unified Approach," *Journal of Multivariate Analysis*, vol. 12, no. 4, pp. 575–596, 1982.

[BS11] K. Batenburg and J. Sijbers, "DART: A Practical Reconstruction Algorithm for Discrete Tomography," *Image Processing, IEEE Transactions on*, vol. 20, no. 9, pp. 2542–2553, Sep. 2011, ISSN: 1057-7149.

[BS18] A. Berman and N. Shaked-Monderer, "Completely Positive Matrices: Real, Rational, and Integral," *Acta Math. Vietnam*, vol. 43, no. 4, pp. 629–639, 2018.

[BSLB11] J. Bushberg, J. Seibert, E. Leidholdt, and J. Boone, *The Essential Physics of Medical Imaging*, 3rd. Wolters Kluwer, 2011.

[BVZ01] Y. Boykov, O. Veksler, and R. Zabih, "Fast Approximate Energy Minimization via Graph Cuts," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 23, no. 11, pp. 1222–1239, 2001.

[CCP12] A. Chambolle, D. Cremers, and T. Pock, "A Convex Approach to Minimal Partitions," *SIAM Journal on Imaging Sciences*, vol. 5, no. 4, pp. 1113–1158, 2012.

[CP11] A. Chambolle and T. Pock, "A First-Order Primal-Dual Algorithm for Convex Problems with Applications to Imaging," *Journal of Mathematical Imaging and Vision*, vol. 40, no. 1, pp. 120–145, May 2011, ISSN: 0924-9907.

[CP16] ——, "On the Ergodic Convergence Rates of a First-Order Primal-Dual Algorithm," *Mathematical Programming*, vol. 159, no. 1, pp. 253–287, 2016, ISSN: 1436-4646.

[CT06] T. Cover and J. Thomas, *Elements of Information Theory*, 2nd. John Wiley & Sons, 2006.

[DK82] P. A. Devijver and J. Kittler, *Pattern Recognition: A Statistical Approach*. Prentice Hall, 1982.

[DM05] P. Drineas and M. W. Mahoney, "On the Nyström Method for Approximating a Gram Matrix for Improved Kernel-Based Learning," *J. Mach. Learning Res.*, vol. 6, pp. 2153–2175, 2005.

[DPSS14] A. Denitiu, S. Petra, C. Schnörr, and C. Schnörr, "Phase Transitions and Cosparse Tomographic Recovery of Compound Solid Bodies from Few Projections," *Fundamenta Informaticae*, vol. 135, pp. 73–102, 2014.

[Fis14] G. Fischer, *Lineare Algebra, Eine Einführung für Studienanfänger*, 18th revised edition, ser. Grundkurs Mathematik. Wiesbaden: Springer Spektrum, 2014, ISBN: 978-365-80394-5-5.

[GBB+12] B. Goris, W. Broek, K. Batenburg, H. Mezerji, and S. Bals, "Electron Tomography Based on a Total Variation Minimization Reconstruction Technique," *Ultramicroscopy*, vol. 113, pp. 120–130, 2012, ISSN: 0304-3991.

[GGP99] R. J. Gardner, P. Gritzmann, and D. Prangenberg, "On the Computational Complexity of Reconstructing Lattice Sets from Their X-Rays," *Discrete Mathematics*, vol. 202, no. 1-3, pp. 45–71, 1999.

[GM16] A. Gittens and M. W. Mahoney, "Revisiting the Nyström Method for Improved Large-Scale Machine Learning," *J. Mach. Learning Res.*, vol. 17, no. 1, pp. 3977–4041, 2016.

[GN02] M. Girvan and M. E.-J. Newman, "Community Structure in Social and Biological Networks," *Proceedings of the National Academy of Sciences*, vol. 99, no. 12, pp. 7821–7826, 2002.

[Har11] S. Har-Peled, *Geometric Approximation Algorithms*. AMS, 2011.

[HFU08] R. Hanke, T. Fuchs, and N. Uhlmann, "X-ray Based Methods for Non-Destructive Testing and Material Characterization," *Nuclear Instruments and Methods in Physics Research Section A: Accelerators, Spectrometers, Detectors and Associated Equipment*, vol. 591, no. 1, pp. 14–18, 2008, ISSN: 0168-9002.

[HK99] G. Herman and A. Kuba, *Discrete Tomography: Foundations, Algorithms and Applications*. Birkhäuser, 1999.

[HLW06] E. Hairer, C. Lubich, and G. Wanner, *Geometric Numerical Integration: Structure-Preserving Algorithms for Ordinary Differential Equations*. Springer Science & Business Media, 2006, vol. 31.

[HM96] U. Helmke and J. B. Moore, *Optimization and Dynamical Systems*, 2nd. Springer, 1996.

[HS03] J. Hofbauer and K. Siegmund, "Evolutionary Game Dynamics," *Bull. Amer. Math. Soc.*, vol. 40, no. 4, pp. 479–519, 2003.

[HSPS19] R. Hühnerbein, F. Savarino, S. Petra, and C. Schnörr, "Learning Adaptive Regularization for Image Labeling Using Geometric Assignment," *CoRR abs/1910.09976*, Oct. 2019.

[HSS08] T. Hofmann, B. Schölkopf, and A. J. Smola, "Kernel Methods in Machine Learning," *Ann. Statistics*, vol. 36, no. 3, pp. 1171–1220, 2008.

[IMNZ00] A. Iserles, H. Z. Munthe-Kaas, S. P. Nørsett, and A. Zanna, "Lie-Group Methods," *Acta numerica*, vol. 9, pp. 215–365, 2000.

[Jos17] J. Jost, *Riemannian Geometry and Geometric Analysis*, 7th. Springer-Verlag Berlin Heidelberg, 2017.

[KAH+15] J. Kappes, B. Andres, F. Hamprecht, C. Schnörr, S. Nowozin, D. Batra, S. Kim, B. Kausler, T. Kröger, J. Lellmann, N. Komodakis, B. Savchynskyy, and C. Rother, "A Comparative Study of Modern Inference Techniques for Structured Discrete Energy Minimization Problems," *Int. J. Comp. Vision*, vol. 115, no. 2, pp. 155–184, 2015.

[Kas89] R. E. Kass, "The Geometry of Asymptotic Inference," *Statistical Science*, vol. 4, no. 3, pp. 188–234, 1989.

[KPSZ15] J. Kappes, S. Petra, C. Schnörr, and M. Zisler, "TomoGC: Binary Tomography by Constrained Graph Cuts," in *Proceedings of the 37th German Conference on Pattern Recognition*, vol. 9358 LNCS, Springer, 2015, pp. 262–273.

[KR13] E. Klann and R. Ramlau, "Regularization Properties of Mumford–Shah-Type Functionals with Perimeter and Norm Constraints for Linear Ill-Posed Problems," *SIAM Journal on Imaging Sciences*, vol. 6, no. 1, pp. 413–436, 2013.

[KYP15] D. Kuang, S. Yun, and H. Park, "SymNMF: Nonnegative Low-Rank Approximation of a Similarity Matrix for Graph Clustering," *Journal of Global Optimization*, vol. 62, no. 3, pp. 545–574, 2015.

[LBS09] J. Lellmann, F. Becker, and C. Schnörr, "Convex Optimization for Multi-Class Image Labeling with a Novel Family of Total Variation Based Regularizers," in *Proceedings of the IEEE Conference on Computer Vision (ICCV 09) Kyoto, Japan*, 1, 2009, pp. 646–653.

[Lee13] J. M. Lee, *Introduction to Smooth Manifolds*, 2. ed. New York; Heidelberg: Springer, 2013, ISBN: 978-1-4419-9981-8 and 1-4419-9981-7.

[Lee18] J. M. Lee, *Introduction to Riemannian Manifolds*. Springer, 2018, vol. 2.

[LLS12] J. Lellmann, F. Lenzen, and C. Schnörr, "Optimality Bounds for a Variational Relaxation of the Image Partitioning Problem," *Journal of Mathematical Imaging and Vision*, vol. 47, no. 3, pp. 239–257, 2012.

[LS11] J. Lellmann and C. Schnörr, "Continuous Multiclass Labeling Approaches and Algorithms," *SIAM J. Imag. Sci.*, vol. 4, no. 4, pp. 1049–1096, 2011.

[MFKI10] S. Maeda, W. Fukuda, A. Kanemura, and S. Ishii, "Maximum a Posteriori X-Ray Computed Tomography Using Graph Cuts," *Journal of Physics: Conference Series*, vol. 233, 2010.

[Min05] T. Minka, "Divergence Measures and Message Passing," Microsoft Research Ltd., Cambridge, UK, Tech. Rep. MSR-TR-2005-173, 2005.

[MP00] G. McLachlan and D. Peel, *Finite Mixture Models*. Wiley, 2000.

[MS89] D. Mumford and J. Shah, "Optimal Approximations by Piecewise Smooth Functions and Associated Variational Problems," *Communications on Pure and Applied Mathematics*, vol. 42, no. 5, pp. 577–685, 1989, ISSN: 1097-0312.

[NT02] Y. E. Nesterov and M. J. Todd, "On the Riemannian Geometry Defined by Self-Concordant Barriers and Interior-Point Methods," *Found. Comp. Math.*, vol. 2, pp. 333–361, 2002.

[NTC13] C. Nieuwenhuis, E. Toeppe, and E. Cremers, "A Survey and Comparison of Discrete and Continuous Multi-label Optimization Approaches for the Potts Model," *International Journal of Computer Vision*, vol. 104, no. 3, pp. 223–240, 2013.

[NW01] F. Natterer and F. Wübbeling, *Mathematical Methods in Image Reconstruction*. SIAM, 2001.

[PCCB09] T. Pock, A. Chambolle, D. Cremers, and H. Bischof, "A Convex Relaxation Approach for Computing Minimal Partitions," in *Computer Vision and Pattern Recognition, 2009. CVPR 2009. IEEE Conference on*, Jun. 2009, pp. 810–817.

[PE86] T. Pham Dinh and S. El Bernoussi, "Algorithms for Solving a Class of Nonconvex Optimization Problems. Methods of Subgradients," in *Fermat Days 85: Mathematics for Optimization*, ser. North-Holland Mathematics Studies, J.-B. Hiriart-Urruty, Ed., vol. 129, North-Holland, 1986, pp. 249–271.

[Pey18] M. Peyré G.and Cuturi, *Computational Optimal Transport*. CNRS, 2018.

[PH97] T. Pham Dinh and L. Hoai An, "Convex Analysis Approach to D.C. Programming: Theory, Algorithms and Applications," *Acta Math. Vietnamica*, vol. 22, no. 1, pp. 289–355, 1997.

[Pot52] R. B. Potts, "Some Generalized Order-Disorder Transformations," *Mathematical Proceedings of the Cambridge Philosophical Society*, vol. 48, pp. 106–109, 01 Jan. 1952, ISSN: 1469-8064.

[RR07] R. Ramlau and W. Ring, "A Mumford–Shah Level-Set Approach for the Inversion and Segmentation of X-ray Tomography Data," *Journal of Computational Physics*, vol. 221, no. 2, pp. 539–557, 2007.

[RW09] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*. Springer Science & Business Media, 2009, vol. 317.

[RW95] F. Rendl and H. Wolkowicz, "A Projection Technique for Partitioning the Nodes of a Graph," *Ann. Operations Res.*, vol. 58, pp. 155–179, 1995.

[San15] F. Santambrogio, *Optimal Transport for Applied Mathematicians*. Birkhäuser, 2015.

[Sch20] C. Schnörr, "Assignment Flows," in *Handbook of Variational Methods for Nonlinear Geometric Data*, P. Grohs, M. Holler, and A. Weinmann, Eds., Springer, 2020, pp. 235–260.

[SL74] L. Shepp and B. Logan, "The Fourier Reconstruction of a Head Section," *Nuclear Science, IEEE Transactions on*, vol. 21, no. 3, pp. 21–43, Jun. 1974, ISSN: 0018-9499.

[SM00] J. Shi and J. Malik, "Normalized Cuts and Image Segmentation," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 22, no. 8, pp. 888–905, Aug. 2000.

[SP08] E. Y. Sidky and X. Pan, "Image Reconstruction in Circular Cone-Beam Computed Tomography by Constrained, Total-Variation Minimization," *Physics in Medicine and Biology*, vol. 53, no. 17, p. 4777, 2008.

[SSWH05] T. Schüle, C. Schnörr, S. Weber, and J. Hornegger, "Discrete Tomography by Convex-Concave Regularization and D.C. Programming," *Discrete Applied Mathematics*, vol. 151, no. 1–3, pp. 229–243, 2005, ISSN: 0166-218X.

[SW14] M. Storath and A. Weinmann, "Fast Partitioning of Vector-Valued Images," *SIAM Journal on Imaging Sciences*, vol. 7, no. 3, pp. 1826–1852, 2014.

[SWFU15] M. Storath, A. Weinmann, J. Frikel, and M. Unser, "Joint Image Reconstruction and Segmentation Using the Potts Model," *Inverse Problems*, vol. 31, no. 2, p. 025 003, 2015.

[Teb07] M. Teboulle, "A Unified Continuous Optimization Framework for Center-Based Clustering Methods," *J. Mach. Learning Res.*, vol. 8, pp. 65–102, 2007.

[TKS+15] A. Tuysuzoglu, W. Karl, I. Stojanovic, D. Castanon, and M. Unlu, "Graph-Cut Based Discrete-Valued Image Reconstruction," *Image Processing, IEEE Transactions on*, vol. 24, no. 5, pp. 1614–1627, May 2015, ISSN: 1057-7149.

[Tol78] J. F. Toland, "Duality in Nonconvex Optimization," *Journal of Mathematical Analysis and Applications*, vol. 66, no. 2, pp. 399–415, 1978.

[VBN12] L. Varga, P. Balázs, and A. Nagy, "An Energy Minimization Reconstruction Algorithm for Multivalued Discrete Tomography," in *3rd International Symposium on Computational Modeling of Objects Represented in Images, Rome, Italy, Proceedings (Taylor & Francis)*, 2012, pp. 179–185.

[Vil09] C. Villani, *Optimal Transport: Old and New*. Springer, 2009.

[vLux07] U. von Luxburg, "A Tutorial on Spectral Clustering," *Statistics and Computing*, vol. 17, no. 4, pp. 395–416, 2007.

[WDS14] A. Weinmann, L. Demaret, and M. Storath, "Total Variation Regularization for Manifold-Valued Data," *SIAM Journal on Imaging Sciences*, vol. 7, no. 4, pp. 2226–2257, 2014.

[Web09] S. Weber, "Discrete Tomography by Convex-Concave Regularization Using Linear and Quadratic Optimization," *PhD thesis, Ruprecht-Karls-Universität, Heidelberg, Germany*, 2009.

[Wer07] T. Werner, "A Linear Programming Approach to Max-Sum Problem: A Review," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 29, no. 7, pp. 1165–1179, 2007.

[WNS+06] S. Weber, A. Nagy, T. Schüle, C. Schnörr, and A. Kuba, "A Benchmark Evaluation of Large-Scale Optimization Approaches to Binary Tomography," in *Discrete Geometry for Computer Imagery (DGCI 2006)*, ser. LNCS, vol. 4245, Springer, 2006, pp. 146–156.

[WS01] C. K. I. Williams and M. Seeger, "Using the Nyström Method to Speed up Kernel Machines," in *Proc. NIPS*, 2001, pp. 682–688.

[WSH03] S. Weber, C. Schnörr, and J. Hornegger, "A Linear Programming Relaxation for Binary Tomography with Smoothness Priors," *Electronic Notes in Discrete Mathematics*, vol. 12, pp. 243–254, 2003.

[YC16] Z. Yang and E. Corander J.and Oja, "Low-Rank Doubly Stochastic Matrix Decomposition for Cluster Analysis," *Journal of Machine Learning Research*, vol. 17, no. 1, pp. 6454–6478, 2016.

[ZAPS17] M. Zisler, F. Aström, S. Petra, and C. Schnörr, "Image Reconstruction by Multilabel Propagation," in *Proceedings of the 6th International Conference on Scale Space and Variational Methods in Computer Vision*, vol. 10302 LNCS, Springer, 2017, pp. 247–259.

[ZGFN08] C. Zach, D. Gallup, J. Frahm, and M. Niethammer, "Fast Global Labeling for Real-Time Stereo Using Multiple Plane Sweeps," in *Proceedings of the Vision, Modeling, and Visualization Conference 2008, VMV 2008, Konstanz, Germany, October 8-10, 2008*, 2008, pp. 243–252.

[ZKS+16] M. Zisler, J. Kappes, C. Schnörr, S. Petra, and C. Schnörr, "Non-Binary Discrete Tomography by Continuous Non-Convex Optimization," *IEEE Transactions on Computational Imaging*, vol. 2, no. 3, pp. 335–347, 2016, ISSN: 2573-0436.

[ZPSS16] M. Zisler, S. Petra, C. Schnörr, and C. Schnörr, "Discrete Tomography by Continuous Multilabeling Subject to Projection Constraints," in *Proceedings of the 38th German Conference on Pattern Recognition*, vol. 9796 LNCS, Springer, 2016, pp. 261–272.

[ZS05] R. Zass and A. Shashua, "A Unifying Approach to Hard and Probabilistic Clustering," in *Proc. ICCV*, 2005.

[ZSPS17] M. Zisler, F. Savarino, S. Petra, and C. Schnörr, "Gradient Flows on a Riemannian Submanifold for Discrete Tomography," in *Proceedings of the 39th German Conference on Pattern Recognition*, vol. 10496 LNCS, Springer, 2017, pp. 294–305.

[ZSPS20] A. Zeilmann, F. Savarino, S. Petra, and C. Schnörr, "Geometric Numerical Integration of the Assignment Flow," *Inverse Problems*, vol. 36, no. 3, 034004 (33pp), 2020.

[ZZÅ+18] A. Zern, M. Zisler, F. Åström, S. Petra, and C. Schnörr, "Unsupervised Label Learning on Manifolds by Spatially Regularized Geometric Assignment," in *Proceedings of the 40th German Conference on Pattern Recognition*, vol. 11269 LNCS, Springer, 2018, pp. 698–713.

[ZZPS19a] A. Zern, M. Zisler, S. Petra, and C. Schnörr, "Spatially Regularized Geometric Assignment for Unsupervised Label Learning on Manifolds," *Proceedings in Applied Mathematics and Mechanics*, vol. 19, no. 1, 2019.

[ZZPS19b] M. Zisler, A. Zern, S. Petra, and C. Schnörr, "Unsupervised Labeling by Geometric and Spatially Regularized Self-Assignment," in *Proceedings of the 7th International Conference on Scale Space and Variational Methods in Computer Vision*, vol. 11603 LNCS, Springer, 2019, pp. 432–444.

[ZZPS20a] A. Zern, M. Zisler, S. Petra, and C. Schnörr, "Unsupervised Assignment Flow: Label Learning on Feature Manifolds by Spatially Regularized Geometric Assignment," *Journal of Mathematical Imaging and Vision*, 2020, ISSN: 1573-7683.

[ZZPS20b] M. Zisler, A. Zern, S. Petra, and C. Schnörr, "Self-Assignment Flows for Unsupervised Data Labeling on Graphs," *SIAM Journal on Imaging Sciences (in press)*, vol. 13, no. 3, pp. 1113–1156, 2020.

[ZZS20] A. Zern, A. Zeilmann, and C. Schnörr, "Assignment Flows for Data Labeling on Graphs: Convergence and Stability," *CoRR abs/2002.11571*, Feb. 2020.