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. Artjom Zern
born in Novosibirsk, Russia

Date of oral examination: .....

Riemannian Flows for Supervised and Unsupervised Geometric Image Labeling

Advisor: Prof. Dr. Christoph Schnörr
PD Dr. Karl Rohr

Zusammenfassung

In dieser Dissertation beschäftigen wir uns mit dem Image Labeling Problem, welches als Unterroutine in zahlreichen Anwendungen in der Bildverarbeitung gebraucht wird. Unsere Arbeit baut auf dem Assignment Flow auf, welcher kürzlich als neuer geometrischer Ansatz zum Image Labeling Problem eingeführt wurde. Dieser Fluss entwickelt sich zeitlich auf der Mannigfaltigkeit der zeilenstochastischen Matrizen, dessen Elemente die Wahrscheinlichkeiten der Zuweisungen von Klassenlabels repräsentieren. Die strikte Trennung von Zuweisungen und Daten ermöglicht es, dass die Daten aus jedem beliebigen metrischen Raum stammen dürfen. Zudem führt eine Glättungsoperation im Raum der Zuweisungen zu einem räumlich regularisierten Labeling, das ohne systematischen Fehler auf dem zugrundeliegenden Graphen realisiert wird.

Der erste Teil der Arbeit ist theoretischen Aussagen zum asymptotischen Verhalten des Assignment Flows gewidmet. Wir zeigen unter schwachen Voraussetzungen an die Parameter, dass der Assignment Flow für Daten in allgemeiner Lage gegen ganzzahlige Wahrscheinlichkeiten und somit eindeutige Zuweisungsentscheidungen konvergiert. Zudem untersuchen wir die Stabilität der möglichen Grenzwerte in Abhängigkeit der Eingabedaten und Parameter. Für stabile Grenzwerte leiten wir Bedingungen her, welche das frühzeitige Erkennen der Konvergenz zu diesen Werten ermöglichen und somit eine Konvergenzgarantie liefern.

Im zweiten Teil der Arbeit diskutieren wir, wie der ursprüngliche Ansatz des Assignment Flows erweitert werden kann, um einfache Label-Statistiken als Vorwissen für eine zusätzliche Regularisierung einzuführen. Diese Label-Statistiken ergeben sich durch globale konvexe Einschränkungen auf lineare Filterstatistiken des resultierenden Labelings. Die zugehörigen Filter können anhand von Beispielen mittels einer Eigenwertzerlegung gelernt werden. Die Wirksamkeit des neuen Ansatzes wird anhand akademischer Beispiele demonstriert.

Im letzten Teil der Arbeit betrachten wir die Situation, in der keine Labels gegeben sind und somit diese prototypischen Elemente aus den Daten mitbestimmt werden müssen. Hierfür führen wir einen zusätzlichen Fluss im Raum der Daten ein, der an den Assignment Flow gekoppelt wird. Der resultierende Fluss adaptiert die Prototypen zeitlich an die Zuweisungswahrscheinlichkeiten. Die simultane Adaption und Zuweisung der Prototypen liefert nicht nur passende Prototypen, sondern verbessert auch die resultierende Bildsegmentierung, was durch Experimente belegt wird. Für diesen Ansatz wird angenommen, dass die Daten auf einer Riemannschen Mannigfaltigkeit liegen. Wir konkretisieren den Ansatz für eine Reihe von Mannigfaltigkeiten aus der Anwendung und evaluieren die resultierenden Verfahren anhand experimenteller Beispiele.

Abstract

In this thesis we focus on the image labeling problem, which is used as a subroutine in many image processing applications. Our work is based on the assignment flow which was recently introduced as a novel geometric approach to the image labeling problem. This flow evolves over time on the manifold of row-stochastic matrices, whose elements represent label assignments as assignment probabilities. The strict separation of assignment manifold and feature space enables the data to lie in any metric space, while a smoothing operation on the assignment manifold results in an unbiased and spatially regularized labeling.

The first part of this work focuses on theoretical statements about the asymptotic behavior of the assignment flow. We show under weak assumptions on the parameters that the assignment flow for data in general position converges towards integral probabilities and thus ensures unique assignment decisions. Furthermore, we investigate the stability of possible limit points depending on the input data and parameters. For stable limits, we derive conditions that allow early evidence of convergence towards these limits and thus provide convergence guarantees.

In the second part, we extend the assignment flow approach in order to impose global convex constraints on the labeling results based on linear filter statistics of the assignments. The corresponding filters are learned from examples using an eigendecomposition. The effectiveness of the approach is numerically demonstrated in several academic labeling scenarios.

In the last part of this thesis we consider the situation in which no labels are given and therefore these prototypical elements have to be determined from the data as well. To this end we introduce an additional flow on the feature manifold, which is coupled to the assignment flow. The resulting flow adapts the prototypes in time to the assignment probabilities. The simultaneous adaptation and assignment of prototypes not only provides suitable prototypes, but also improves the resulting image segmentation, which is demonstrated by experiments. For this approach it is assumed that the data lie on a Riemannian manifold. We elaborate the approach for a range of manifolds that occur in applications and evaluate the resulting approaches in numerical experiments.

Acknowledgements

I would like to thank many people who supported me in writing this thesis. First of all, I would like to thank my supervisor Prof. Dr. Christoph Schnörr at the Faculty of Mathematics and Computer Science at Heidelberg University. He introduced me to this fascinating topic and guided me through my doctorate. I appreciate that he provided interesting ideas and useful literature and shared his foresight. I also value his quick responses.

I would also like to thank my second supervisor PD Dr. Karl Rohr for his support and that he always took time for meetings where we could discuss current issues.

Next, I would like to thank the secretaries Evelyn Wilhelm, Barbara Werner and Marion Münster for their administrative support.

I am very grateful to my former and current colleagues at the *Image and Pattern Analysis* group and the *Biomedical Computer Vision* group for the pleasant working atmosphere and the time outside of work. A special thanks goes to Matthias Zisler for the many productive discussions and the pleasant company during the coffee and meal breaks. I would also like to especially thank Bastian Boll for his nice company, fruitful discussions and for proofreading this thesis. A special thanks also goes to Alexander Zeilmann and his taking care of the essential coffee machine. My friends Lukas Kiefer, Jan Plier and Francesco Silvestri, I would like to thank for the time spent together, especially during the board game nights and for helping me with some tasks. I would like to thank my colleagues Leonid Kostrykin, Christian Ritter and Thomas Wollmann both for the insights I gained in meetings and for the time spent together at social events.

For insightful discussions and pleasant company I would also like to mention Sharib Ali, Freddie Åström, Florian Becker, Johannes Berger, Ecatarina Bodnariuc, Robert Breckner, Tobias Dencker, Andreea Denitiu, Mattia Desana, Felix Draxler, Simon Eck, Qi Gao, Ruben Hühnerbein, Jörg Kappes, Pakpoon Khamchuai, Frank Lenzen, Ran Li, Wei Liao, Timo Milbich, Andreas Neufeld, Stefania Petra, Fabian Radtke, Svenja Reith, Stefan Richter, Fabrizio Savarino, Bogdan Savchynskyy, Bernhard Schmitzer, Jonathan Schwarz, Luis Vale Silva, Dmitrij Sitenko, Roman Spilger, Paul Swoboda, Vera Trajkovska, Nikolai Ufer, Steffen Wolf, Stefan Wörz, and Tabea Zuber.

Many thanks to my friends Judith Stein and Benjamin Kupferer for the relaxing company at lunch and other social events.

Finally, I would like to thank my family — my mother Tatjana and my father Alexander, my little brother Wenzel and his wife Veronique, and my nephew Leo — for their constant support and encouragement.

Funding by the Deutsche Forschungsgemeinschaft via the research training group 1653 Spatio / Temporal Graphical Models and Applications in Image Analysis is also gratefully acknowledged.

Contents

List of Symbols

Basic Conventions

$\mathbb{N}$, $\mathbb{N}_0$	set of natural numbers, set of natural numbers including 0
$\mathbb{Z}$	set of integers
$\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{R}_{>0}$	set of real numbers, nonnegative real numbers and positive real numbers
$\mathbb{C}$	set of complex numbers
$[n]$	finite set $\{1, \dots, n\}$
$\lvert S \rvert$	cardinality of a finite set S
$\mathrm{Re}(z)$, $\mathrm{Im}(z)$	real and imaginary part of a complex number or vector z
$\overline{U}$	closure of a subset $U \subset \mathbb{R}^n$
$\mathrm{int}(U)$	interior of a set U
$\mathrm{rint}(U)$	relative interior of a set U
$\mathrm{C}^k(U, V)$	set of k-times continuously differentiable functions $f \colon U \to V$
∇f	Euclidean gradient of a function f
$\langle \cdot, \cdot \rangle$	standard inner product in $\mathbb{R}^n$
$\lVert \cdot \rVert$	Euclidean norm of a vector and Frobenius norm of a matrix
$\lVert \cdot \rVert_p$	ℓ^p-norm of a vector, i. e. $\lVert v \rVert_p^p = \sum_j \lvert v_j \rvert^p$
e_j	j-th standard basis vector of $\mathbb{R}^n$
$\mathbb{1}_n$	constant vector of ones $(1, \dots, 1)^\top \in \mathbb{R}^n$
I_n	identity matrix in $\mathbb{R}^{n \times n}$
$\mathrm{Diag}(v)$	diagonal matrix with diagonal v
$\ker(A)$, $\mathrm{im}(A)$	kernel and image of a matrix A
$\mathrm{tr}(A)$	trace of a matrix A
$\det(A)$	determinant of a matrix A
$A \odot B$	Hadamard (entrywise) product of matrices A and B
$A \otimes B$	Kronecker product of matrices A and B
$\exp(v)$, $\log(v)$	entrywise exponentation and logarithm of a vector $v \in \mathbb{R}^n$
$\mathrm{expm}(A)$	matrix exponential of a matrix A, i. e. $\mathrm{expm}(A) = \sum_{n \in \mathbb{N}_0} \frac{1}{n!} A^n$
$\mathrm{logm}(A)$	matrix logarithm of a matrix A
$\mathcal{P}_d$	set of positive definite $d \times d$ matrices
Δ_n	probability simplex in $\mathbb{R}^n$, i. e. $\Delta_n = \{p \in \mathbb{R}^n_{\geq 0} \colon \langle p, \mathbb{1}_n \rangle = 1\}$
$D_{\mathrm{KL}}(p, q)$	Kullback-Leibler (KL) divergence for $p, q \in \Delta_n$

Differential Geometry

$(\mathcal{M}, g)$	Riemannian manifold
$T_p\mathcal{M}$	tangent space at $p \in \mathcal{M}$
$T_p^*\mathcal{M}$	cotangent space at $p \in \mathcal{M}$
$T\mathcal{M}$	tangent bundle
$T^*\mathcal{M}$	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 on $\mathcal{M}$
$\widehat{g}$	linear tangent-cotangent isomorphism $\mathfrak{X}(\mathcal{M}) \to \mathfrak{X}^*(\mathcal{M})$
$\mathrm{d}F_p$	differential of a smooth function $F\colon \mathcal{M} \to \mathcal{N}$ at $p \in \mathcal{M}$
$\|v\|_g$	norm on $T_p\mathcal{M}$ induced by the Riemannian metric g
d_g	Riemannian distance
$\operatorname{grad} f$	Riemannian gradient of a function $f \in \mathrm{C}^\infty(\mathcal{M}, \mathbb{R})$
$\operatorname{Exp}_p(v)$	exponential map at $p \in \mathcal{M}$ with $v \in T_p\mathcal{M}$
$\operatorname{Log}_p(q)$	logarithmic map at $p \in \mathcal{M}$ with $q \in \mathcal{M}$

Assignment Framework

(I, E)	undirected graph with vertex set I and edge set E
I, J	index set of data points, index set of labels
m, n	number of data points, number of labels
$\mathcal{N}_i$	neighborhood of $i \in I$ – cf. (2.71)
ω_{ik}	weights for $k \in \mathcal{N}_i$
$\Omega = (\omega_{ik})_{ik}$	weighted adjacency matrix of the graph (I, E)
ρ	selectivity parameter $\rho > 0$
$\mathcal{S}$	open probability simplex in $\mathbb{R}^n$
$\mathcal{W}$	assignment manifold (2.85)
$\mathcal{W}^*$	set of integral assignments
$C_\mathcal{S}$, $C_\mathcal{W}$	barycenter of $\mathcal{S}$ and $\mathcal{W}$, i. e. $C_\mathcal{S} = \frac{1}{n}\mathbb{1}_n$ and $C_\mathcal{W} = \frac{1}{n}\mathbb{1}_m\mathbb{1}_n^\top$
T_0, $\mathcal{T}_0$	tangent space of $\mathcal{S}$, tangent space of $\mathcal{W}$
W	$m \times n$ assignment matrix in $\overline{\mathcal{W}}$
$\exp_p$, $\exp_W$	lifting map (2.90), (2.100a)
$L(W)$	likelihood map (2.102)
$S(W)$	similarity map (2.104)
R_p, R_W	replicator operator (2.91), (2.100b)
Π_0	orthogonal projection onto T_0 or $\mathcal{T}_0$ – cf. (2.81)

List of Publications

[ZRS18] **A. Zern**, K. Rohr, and C. Schnörr. "Geometric Image Labeling with Global Convex Labeling Constraints". In: *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition.* Vol. 10746. LNCS. 2018, pp. 533–547

[Zer+18] **A. Zern**, M. Zisler, F. Åström, S. Petra, and C. Schnörr. "Unsupervised Label Learning on Manifolds by Spatially Regularized Geometric Assignment". In: *German Conference on Pattern Recognition.* 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". In: *PAMM* 19.1 (2019), e201900258. DOI: 10.1002/pamm.201900258

[ZZPS19b] M. Zisler, **A. Zern**, S. Petra, and C. Schnörr. "Unsupervised Labeling by Geometric and Spatially Regularized Self-Assignment". In: *International Conference on Scale Space and Variational Methods in Computer Vision.* 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". In: *Journal of Mathematical Imaging and Vision* (2020). DOI: 10.1007/s10851-019-00935-7. (in press)

[ZZPS20b] M. Zisler, **A. Zern**, S. Petra, and C. Schnörr. "Self-Assignment Flows for Unsupervised Data Labeling on Graphs". In: *SIAM Journal on Imaging Sciences* (2020). (in press)

[ZZS20] **A. Zern**, A. Zeilmann, and C. Schnörr. "Assignment Flows for Data Labeling on Graphs: Convergence and Stability". In: *arXiv preprint arXiv:2002.11571* (2020)

1 Introduction

1.1 Motivation

Image segmentation is a fundamental task needed in numerous applications in computer vision. For video surveillance and autonomous driving, segmentation techniques are used to detect pedestrians and other objects in the street scene [TPM08]. In biological research, cell segmentation is used to analyze cell morphology and cell movement [KSR19]. Additionally, segmentation techniques are used in medical imaging to assist physicians in diagnosing and treating diseases. This includes cancer detection [WER18] as well as blood vessel segmentation to identify pathologies such as aneurysms and dissections [Bie+12, Bie+15]. While segmentation of biomedical images can be done by hand, manual segmentation is time consuming and error-prone. Therefore automatic segmentation methods are needed.

There are numerous segmentation methods and they come in many different forms such as graph partitioning or model-based segmentation. In this thesis we deal with the image labeling problem, which is a subclass of the image segmentation problem. It consists of assigning each pixel a single element from a finite set of labels, so that equally labelled pixels share certain characteristics. By labels we mean prototypical elements from the feature space. The following subproblems arise in the labeling problem, which affect the segmentation result.

1. **Similarity measure.** A dissimilarity function is needed to compare data and labels. This function is used for pixelwise labeling decisions.

2. **Regularization.** To cope with noisy data, prior knowledge of (inter-pixel) label statistics should be incorporated into the image labeling method as regularization. This includes the information that neighboring pixels are more likely to have the same label.

3. **Labels.** For the labeling problem, the set of labels must be known beforehand and these prototypical elements should properly represent the data.

The first item must be considered separately for each dataset. We do not want to pay much attention to this point. Therefore, we choose an intuitive dissimilarity function for each feature space. In the case of RGB values, for example, this is the Euclidean distance. The main focus of this work will be on the last two points. Depending on the labeling method, regularization may depend on the dissimilarity function and thus on the feature space. We will choose a labeling method where feature space and label assignment are decoupled so that the regularization can be considered independent of the dataset. As regularization we will examine label statistics in the form of filter statistics. Labels, on the other hand, depend on the dissimilarity function and, based on this function, they should reflect the range of the data well. In addition, they should also allow a labeling in good accordance with the regularization, i.e. the labels should also be chosen depending on the regularization and thus on the labeling method itself. Since it is difficult or sometimes impossible to determine such labels in advance, we will follow the

approach of learning labels during the labeling process whenever no labels are given. For this we will follow established clustering methods.

As our underlying labeling method we will use the assignment flow approach introduced in [ÅPSS17]. This has the following reasons. The approach follows a strict separation between the feature space containing the data and the label assignments. As mentioned above, this allows the regularization to be considered independently of the feature space. Moreover, the approach allows a spatial regularization that is unbiased by pixel position. Furthermore, the strict separation of feature space and label assignments enables any metric space to be used as a feature space. Here we are particularly interested in Riemannian manifolds that appear in practice. Examples include the space of positive definite matrices [TPM08] or the one-dimensional sphere [RPF97]. Another interesting aspect of the assignment flow approach is that it assumes data to lie on vertices of a general (undirected) graph. This does not necessarily have to be a grid graph, as is common for image domains, but can also be, for example, a triangulated manifold or an image presegmented as superpixels [Ach+12]. Thus the application is not limited to 2D/3D images.

1.2 Related Work

Image Labeling

A large number of labeling methods model the labeling problem as a variational problem of form

$$\min_{\ell\colon I\to L} E(\ell), \qquad \text{with} \qquad E(\ell) = E_{\text{data}}(\ell) + E_{\text{reg}}(\ell), \tag{1.1}$$

where I is the image domain, L is the finite set of labels, E_{data} is a data fidelity term and E_{reg} is a regularization term. The image domain can be continuous or discrete. In [LS11], for example, a continuous domain and a convex relaxation of problems of the form (1.1) are considered. In the field of computer vision discrete models prevail. Prominent examples can be found in the context of inference tasks of probabilistic graphical models like Markov random fields [WJ08], where minimization (1.1) corresponds to maximum a posteriori inference with respect to the probability distribution $p(\ell) = \frac{1}{Z}\exp(-E(\ell))$. In that case, the vertices of a graph correspond to the image domain, i.e. the pixel locations. The regularization term measures the compatibility of labels along edges or cliques of higher order. This interaction between the pixels makes the labeling problem a hard combinatorial problem. Thus, in practical applications relaxations are used as approximations. One noteworthy example is the local polytope relaxation, which is a particular linear programming relaxation of the labeling problem [Wer07]. On the other hand, min-cut/max-flow algorithms such as the α-expansion algorithm [BVZ01, BK04] are widely used because they provide sufficiently accurate solutions to many large-scale problems in an efficient way. For further methods we refer to [Kap+15], which provides a comprehensive overview and evaluation.

The assignment flow approach [ÅPSS17], on which we base our work, differs from the above methods in two ways. First, there is no energy that is (approximately) minimized by the solution of the assignment flow. Although the assignment flow was introduced as an approximation to the gradient flow of a functional, the functional is not defined for the limits of the assignment flow. Second, the data fidelity term and the regularization term do not enter

additively into the labeling process as in (1.1), but are both described by operators that appear in the vector field of the assignment flow by concatenation.

Replicator Dynamics

The vector field of the assignment flow is a coupled system of replicator equations. Replicator equations are ordinary differential equations that were introduced in mathematical biology to model natural selection [TJ78] and since then have appeared in various literature on evolutionary game theory, e.g. [Fri91, HS03]. From the perspective of differential geometry, replicator equations are related to a particular Riemannian geometry on the probability simplex, known in information geometry as Fisher-Rao geometry [Ama16], and can be understood as Riemannian gradient flows with respect to this geometry [MS18]. With respect to convergence statements for replicator equations, we mention [LA83] and [BS99], which consider replicator equations with symmetrical linear fitness and show convergence to a stationary point of a quadratic function on the probability simplex.

Filter Statistics

Early seminal publications on the generative aspects of filter statistics in image processing include [ZM97] and [PS00]. In the former, filters appear as parameters of a Gibbs-Boltzmann distribution. A Gibbs reaction-diffusion for this distribution induces filter statistics in an image over time. In the second publication, textures are synthesized by projecting an image onto a manifold formed by several hundred filter constraints. A more recent work on filter statistics is [BGGS17], where filters are learned as parameters of TV-like regularizers for image denoising problems.

Clustering

A naive approach to determining labels for image labeling is to ignore the spatial context and group the feature vectors in the feature space using a clustering method and subsequently determine a representative for each cluster, e.g. by averaging. One of the best known methods is k-means clustering, which can be found in various literature. We merely refer to the survey [Teb07]. k-means tries to minimize the mean square distance to the cluster centers by averaging within the clusters and thus automatically provides representatives for each cluster. Another centroid-based method is k-center clustering [Har11], which is applicable in any metric space. It attempts to minimize the maximum distance to the cluster centers. In addition, there are density-based clustering approaches such as mean-shift [CM02, FH75], which determine cluster centroids with respect to an empirical density, similar to k-means. By generalizing means to Riemannian (Karcher) means, mean-shift can be extended to data on a Riemannian manifold [SM09].

A clustering method widely used in statistics is the EM algorithm [MP00], which is modeled on a mixed distribution and thus belongs to the distribution-based clustering methods. In addition to the class probabilities, the parameters of the class-conditional distributions are also determined. The authors in [BMDG05] investigated the case where the class-conditional distributions belong to an exponential family and are represented by a Bregman divergence, whereby the parameters of the distributions are single points (representatives) in the feature space.

Examples of methods that provide only a segmentation of the data and no representatives are connectivity-based methods that decide the grouping based solely on pairwise distances between the data. These include spectral clustering [SM00] and hierarchical clustering [Mül11, Bat88].

1.3 Contribution and Overview

In this thesis we deal with the assignment flow approach. We investigate its properties mathematically and we also present two extensions, one towards regularization with label statistics and one towards unsupervised label learning. Our main contributions are:

- **Analysis of the asymptotic behavior of the assignment flow.**
  Since the introduction of the assignment flow, its convergence has hardly been studied. The convergence towards integral assignment probabilities for data in general position has merely been postulated and empirically verified. In Chapter 3, we carry out a purely mathematical investigation of the convergence using the recently introduced S-parameterization [SS20]. We prove the postulated statement under weak assumptions on the parameters. In addition, we analyze the stability of the limits depending on the data. For limits that fulfill a certain stability condition, we derive a criterion with which convergence towards these limits can be determined at an early stage and which thus provides a convergence guarantee. We also show that the assignment flow converges towards these stable limits with an exponential convergence rate. We extend the stability analysis to the discrete scheme [Sav+17] and verify the validity of the convergence guarantee for this discrete case as well. This contribution to the convergence analysis is mostly published in the preprint [ZZS20].

- **A new approach for enforcing label statistics by global convex constraints.**
  We consider convex constraints on the filter statistics of the assignment probabilities. These constraints induce label statistics on the resulting assignments. We incorporate these constraints into the assignment flow approach by modifying its vector field using logarithmic barrier functions. Suitable filters for the constraints can be learned by positive and negative examples. We formulate the filter learning as a generalized eigendecomposition. This contribution is published in the conference paper [ZRS18] and it is chronologically the first contribution in this thesis.

- **A novel approach for unsupervised label learning in the context of manifold-valued images.**
  We consider the case that no labels for the assignment flow are given in advance and must be determined from data as well. In order to enable the labels to comply with the regularization, we follow the approach of adapting labels in time during the labeling process. To get reasonable initial labels in an efficient way, they are determined by a greedy approximation of k-center clustering, which provides a performance guarantee. For the label adaptation we follow established clustering methods by adapting the labels to the assignment probabilities. We present two different flows in the feature space for the adaptation, each of which is coupled with the assignment flow. We interpolate the two resulting approaches using an additional parameter, so that we obtain a one-parameter family of unsupervised assignment flows. The approach is applicable for data on a

Riemannian manifold. We concretize the approach for a number of manifolds occurring in applications, such as the set of positive definite matrices. The first version of the approach is published in the conference paper [Zer+18]. The elaborated one-parameter family of unsupervised assignment flows can be found in the journal paper [ZZPS20a].

This thesis is organized as follows. We provide mathematical background for subsequent work in **Chapter 2**. This includes an overview of differential geometry, which is used in several parts of the thesis, as well as a presentation of the assignment flow approach [ÅPSS17] on which the main part of the thesis is based. In addition, we also provide definitions and statements that are needed for the stability analysis of equilibria of a dynamic system. These statements are needed for the study of the asymptotic behavior of the assignment flow.

In **Chapter 3** we investigate the asymptotic behavior of the assignment flow. For this purpose, we first analyze in Section 3.2 the connection between the limits of the S-flow and the limits of the assignment flow. Then we prove in Section 3.3 the convergence of the S-flow towards an equilibrium assuming symmetry of the parameters. We also demonstrate with an example that the convergence statement need not hold without symmetry. In Section 3.4 we investigate the stability of the equilibria of the S-flow. We show under weak assumptions on the parameters that only integral equilibria can be stable and conclude that the S-flow almost surely converges towards integral equilibria. For exponentially stable equilibria we estimate the basin of attraction in Section 3.4.2. In Section 3.5 we investigate the discrete scheme [Sav+17] and carry out a stability analysis for its equilibria. At the end of the chapter we discuss a generalization of the S-flow and mention how the stability statements apply to the generalization.

In **Chapter 4**, we discuss an extension of the assignment flow approach, with which label statistics can be induced. The label statistics are presented in Section 4.2.1 as convex constraints on linear filter statistics. Afterwards, the learning of filters is discussed in Section 4.2.2. The extension of the assignment flow by the label statistics is done in Section 4.2.3. At the end, the approach will be evaluated on the basis of academic experiments in Section 4.3 and an outlook is given in Section 4.4.

Chapter 5 contains our extension of the assignment flow for unsupervised label learning. First, established clustering methods are summarized in Section 5.2 and extended in Section 5.3 for the case of data lying on a Riemannian manifold. The variants on manifolds are then coupled with the assignment flow in Section 5.4 and interpolated with one parameter. The resulting approach is compared to another extension of the assignment flow for unsupervised clustering, namely the self-assignment flow [ZZPS20b], in Section 5.5. We work out our approach for different manifolds in Section 5.6 and evaluate the resulting approaches with numerous experiments in Section 5.7.

In **Chapter 6** at the end of this thesis we provide a concluding summary of our results and give an outlook for potential research directions.

1.4 Notation

Sets

The set of natural numbers, integers, real numbers and complex numbers are notated as usual with $\mathbb{N}$, $\mathbb{Z}$, $\mathbb{R}$ and $\mathbb{C}$. Moreover, $\mathbb{N}_0$ is the set of natural numbers including 0, i.e. $\mathbb{N}_0 = \mathbb{N} \cup \{0\}$.

$\mathbb{R}_{\geq 0}$ and $\mathbb{R}_{>0}$ denote the set of nonnegative real numbers and positive real numbers. Open intervals are denoted with parentheses, e.g. $(0, 1)$, and closed intervals with square brackets, e.g. $[0, 1]$. The latter should not be confused with the finite set $[n] \coloneqq \{1, \dots, n\}$. In this thesis arg max and arg min are sets and not elements, i.e.

$$\arg\max_{j \in J} \; a_j = \left\{ j^* \in J \colon a_{j^*} = \max_{j \in J} a_j \right\}. \tag{1.2}$$

The set $\mathrm{C}^k(U, V)$ consists of all k-times continuously differentiable functions $f \colon U \to V$. Calligraphic capital letters usually denote manifolds or metric spaces. In particular, examples are the open probability simplex $\mathcal{S}$ (2.35), the assignment manifold $\mathcal{W}$ (2.85) and the manifold of positive definite $d \times d$ matrices $\mathcal{P}_d$. An exception are the vertex neighborhoods $\mathcal{N}_i$ (2.71).

Matrix and Vector Notations

Vectors $v \in \mathbb{R}^n$ are understood as column vectors. The angle brackets $\langle \cdot, \cdot \rangle$ denote the Euclidean inner product, i.e. $\langle u, v \rangle = u^\top v$ for vectors $u, v \in \mathbb{R}^n$ and $\langle U, V \rangle = \operatorname{tr}(U^\top V)$ for matrices $U, V \in \mathbb{R}^{m \times n}$. The j-th standard basis vector in $\mathbb{R}^n$ is represented by $e_j = (0, \dots, 1, \dots, 0)^\top$. The dimension of e_j should be clear from the context. $\mathbb{1}_n \in \mathbb{R}^n$ is the constant vector of ones and $\mathbb{1}_A \in \mathbb{R}^n$ is the indicator vector for a subset $A \subseteq [n]$, i.e. $(\mathbb{1}_A)_j = 1$ if $j \in A$ and 0 otherwise. Multiplication pq and division $\frac{p}{q}$ for vectors $p, q \in \mathbb{R}^n$ are to be understood entrywise as well as the elementary functions $\sqrt{p}$, $e^q = \exp(q)$ and $\log(q)$. The entrywise product of matrices is indicated by $\odot$. In some circumstances, this operator is also used for vectors, if it serves the readability. The $\otimes$ operator denotes the Kronecker product of two matrices or vectors. $\operatorname{Diag}(v) \in \mathbb{R}^{n \times n}$ is the diagonal matrix with the diagonal $v \in \mathbb{R}^n$. For a matrix $W \in \mathbb{R}^{m \times n}$, $W_i \in \mathbb{R}^n$ denotes the i-th row (as column vector) and $W^j \in \mathbb{R}^m$ denotes the j-th column.

Maps and Operators

The maps $\exp_p$ and $\exp_W$ refer to the maps (2.90) and (2.100a) and must not be confused with the entrywise exponentiation exp (without subscript). The exponential mapping on a Riemannian manifold is given as Exp_p and the matrix exponential is given as expm. For a function f, $\mathrm{d}f$ will represent either the differential (2.4) or the Gâteaux (directional) derivative. For the differential we use the notation $\mathrm{d}f_p(v)$, while for the Gâteaux derivative we use the notation $\mathrm{d}f(x)[v]$. Since both represent the same quantity, this should not lead to confusion. More generally we enclose the arguments of a linear operator with square brackets, e.g. $L[V]$. We do this especially with the replicator operator $R_W \colon \mathbb{R}^{m \times n} \to \mathbb{R}^{m \times n}$ (2.100b) in order to distinguish it from its matrix representation, which we also denote as $R_W \in \mathbb{R}^{mn \times mn}$. For a function $f \colon \mathcal{M} \to \mathbb{R}$ on a Riemannian manifold, $\operatorname{grad} f$ denotes the Riemannian gradient. The Euclidean gradient (with respect to the coordinates of the ambient Euclidean space) is denoted by ∇f instead.

2 Preliminaries

In this chapter we present the preliminary material for the subsequent work. First, we summarize basic definitions and statements from differential geometry, which are needed for the discussion of unsupervised label learning on Riemannian manifolds in Chapter 5. Furthermore, we discuss additional topics related to Riemannian geometry like information geometry which prepare the introduction of the assignment flow presented afterwards. The latter constitutes the main part of the preliminary work, since all following chapters are based on it. Finally, we give an overview of definitions and statements from stability theory of dynamical systems. These statements are used to investigate the assignment flow in Chapter 3.

2.1 Differential Geometry

We summarize relevant topics related to differential geometry in this section. Besides basic definitions, this particularly includes a discussion of flows and divergence functions on manifolds as well as the Riemannian mean. We also present a brief look into information geometry and we discuss a class of Riemannian manifolds motivated by constrained optimization problems.

2.1.1 Basic Definitions

We give a brief overview of differential geometry to specify some terminology and notation. In this overview, we assume the reader to be familiar with manifolds. For a detailed description of the topic we refer to the textbooks [Lee13, Lee06, Lee18].

Vector and Covector Fields on Smooth Manifolds

Let $\mathcal{M}$ be a smooth n-dimensional manifold. We denote the set of all smooth functions $f\colon \mathcal{M} \to \mathbb{R}$ by $\mathrm{C}^\infty(\mathcal{M}, \mathbb{R})$. The *tangent space* of $\mathcal{M}$ at $p \in \mathcal{M}$ is the n-dimensional vector space $T_p\mathcal{M}$ consisting of all derivations of $\mathrm{C}^\infty(\mathcal{M}, \mathbb{R})$ at p. That means, $vf \in \mathbb{R}$ is a directional derivative of $f \in \mathrm{C}^\infty(\mathcal{M}, \mathbb{R})$ at p for any $v \in T_p\mathcal{M}$. Using local coordinates $x^1, \dots, x^n$, any tangent vector $v \in T_p\mathcal{M}$ can be written as

$$v = \sum_{i \in [n]} v^i \frac{\partial}{\partial x^i}\Big|_p \qquad \text{with } v^1, \dots, v^n \in \mathbb{R}. \tag{2.1}$$

Whenever $\mathcal{M}$ is embedded into an $\mathbb{R}^N$, we may also consider $T_p\mathcal{M}$ as the set of all vectors $v \in \mathbb{R}^N$ tangent to $\mathcal{M}$ at p by identifying directions with corresponding directional derivatives. The *tangent bundle* of $\mathcal{M}$ is defined as disjoint union of all tangent spaces, i.e.

$$T\mathcal{M} \coloneqq \bigcup_{p \in \mathcal{M}} \{p\} \times T_p\mathcal{M}, \tag{2.2}$$

and it forms a smooth $2n$-dimensional manifold. A *vector field* X on $\mathcal{M}$ is a function $\mathcal{M} \to T\mathcal{M}$ such that $X_p \in T_p\mathcal{M}$ for all $p \in \mathcal{M}$, where subscripts indicate the evaluation of the vector field and $\{p\} \times T_p\mathcal{M}$ is identified with $T_p\mathcal{M}$. In local coordinates, vector fields take the form

$$X = \sum_{i \in [n]} X^i \frac{\partial}{\partial x^i} \qquad \text{with } X^1, \dots, X^n \in \mathrm{C}^\infty(\mathcal{M}, \mathbb{R}). \tag{2.3}$$

We denote the set of all smooth vector fields on $\mathcal{M}$ by $\mathfrak{X}(\mathcal{M})$. With regard to the cotangent space $T_p^*\mathcal{M}$ consisting of all linear forms $T_p\mathcal{M} \to \mathbb{R}$, the cotangent bundle $T^*\mathcal{M}$ and the set $\mathfrak{X}^*(\mathcal{M})$ of all smooth covector fields are defined correspondingly. For a smooth map $F\colon \mathcal{M} \to \mathcal{N}$ between manifolds, the *differential* of F at p is the linear map $\mathrm{d}F_p\colon T_p\mathcal{M} \to T_{F(p)}\mathcal{N}$ given by

$$\mathrm{d}F_p(v)(f) = v(f \circ F) \qquad \text{for all } v \in T_p\mathcal{M} \text{ and } f \in \mathrm{C}^\infty(\mathcal{N}, \mathbb{R}). \tag{2.4}$$

For functions $f \in \mathrm{C}^\infty(\mathcal{M}, \mathbb{R})$, the differential induces a smooth covector field $\mathrm{d}f \in \mathfrak{X}^*(\mathcal{M})$ with $\mathrm{d}f(X) = Xf$ for vector fields $X \in \mathfrak{X}(\mathcal{M})$.

Riemannian Metric, Distance and Gradient

A *Riemannian metric* g on $\mathcal{M}$ is a family of inner products g_p for each tangent space $T_p\mathcal{M}$ such that g_p depends smoothly on p, i.e., the function $g(X, Y)$ given by $p \mapsto g_p(X_p, Y_p)$ is smooth for any smooth vector fields $X, Y \in \mathfrak{X}(\mathcal{M})$. The pair $(\mathcal{M}, g)$ is called *Riemannian manifold*. By means of the Riemannian metric some concepts from Euclidean geometry can be applied to manifolds. It enables to define the length of a tangent vector by

$$\|v\|_g \coloneqq \sqrt{g_p(v, v)} \qquad \forall v \in T_p\mathcal{M} \tag{2.5}$$

as well as the length of a (piecewise) smooth curve $\gamma\colon [a, b] \to \mathcal{M}$ by

$$L_g(\gamma) \coloneqq \int_a^b \|\dot\gamma(t)\|_g \,\mathrm{d}t, \tag{2.6}$$

where $\dot\gamma(t) = \mathrm{d}\gamma_t\left(\frac{\mathrm{d}}{\mathrm{d}t}\right) \in T_{\gamma(t)}\mathcal{M}$ denotes the directional derivative along $\gamma(t)$. Furthermore, the *Riemannian distance* on $\mathcal{M}$ is defined by

$$d_g(p, q) \coloneqq \inf\left\{L_g(\gamma)\colon \text{piecewise smooth curve } \gamma\colon [0, 1] \to \mathcal{M} \text{ with } \gamma(0) = p \text{ and } \gamma(1) = q\right\}. \tag{2.7}$$

Finally, the *Riemannian gradient* of a function $f \in \mathrm{C}^\infty(\mathcal{M}, \mathbb{R})$ is the vector field

$$\operatorname{grad} f \in \mathfrak{X}(\mathcal{M}) \tag{2.8a}$$

defined by

$$g(\operatorname{grad} f, X) = df(X) = Xf \qquad \forall X \in \mathfrak{X}(\mathcal{M}). \tag{2.8b}$$

Using the linear tangent-cotangent isomorphism[1]

$$\widehat{g}\colon \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}^*(\mathcal{M}), \qquad \widehat{g}(X)(Y) = g(X, Y), \quad \forall X, Y \in \mathfrak{X}(\mathcal{M}) \tag{2.9}$$

that associates with a vector field $X \in \mathfrak{X}(\mathcal{M})$ the covector field $\widehat{g}(X) = g(X, \cdot)$, the Riemannian gradient (2.8b) reads

$$\operatorname{grad} f = \widehat{g}^{-1}(df). \tag{2.10}$$

[1] The maps $\widehat{g}$ and $\widehat{g}^{-1}$ are sometimes denoted with $\flat$ and $\sharp$ in the literature ('musical isomorphism'). We stick to the notation from [Lee13] here.

Geodesics and Exponential Map

An *affine connection* ∇ on $\mathcal{M}$ is a map[2]

$$\nabla\colon \mathfrak{X}(\mathcal{M}) \times \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}(\mathcal{M}), \qquad (X, Y) \mapsto \nabla_X Y, \tag{2.11}$$

which is $C^\infty(\mathcal{M})$-linear in X, $\mathbb{R}$-linear in Y and satisfies the product rule

$$\nabla_X(fY) = f\nabla_X Y + (Xf)Y \quad \text{for } f \in \mathrm{C}^\infty(\mathcal{M}, \mathbb{R}). \tag{2.12}$$

The affine connection describes directional derivatives of a vector field Y and it allows to define the *acceleration* $\nabla_{\dot\gamma}\dot\gamma$ of a curve γ. A *geodesic* with respect to ∇ is a curve γ with vanishing acceleration, i.e. $\nabla_{\dot\gamma}\dot\gamma \equiv 0$. Geodesics are the generalization of straight lines from Euclidean spaces. For any $p \in \mathcal{M}$ and $v \in T_p\mathcal{M}$ there exists a (locally) unique geodesic γ_v through $p = \gamma_v(0)$ with velocity $v = \dot\gamma_v(0)$. The *exponential map* at p

$$\mathrm{Exp}_p\colon V_p \to \mathcal{M}, \qquad v \mapsto \mathrm{Exp}_p(v) = \gamma_v(1) \tag{2.13}$$

with domain given by

$$V_p = \left\{v \in T_p\mathcal{M}\colon \gamma_v \text{ is defined on } [0,1]\right\} \tag{2.14}$$

is a diffeomorphism from some open neighborhood of $0 \in T_p\mathcal{M}$ into its image. Its inverse $\mathrm{Log}_p = \mathrm{Exp}_p^{-1}$ is called *logarithmic map*. On any Riemannian manifold $(\mathcal{M}, g)$ there exists a unique connection ∇^g called *Levi-Civita connection* which is both compatible with g, i.e.

$$Xg(Y,Z) = g(\nabla^g_X Y, Z) + g(X, \nabla^g_X Z) \qquad \text{for all } X, Y, Z \in \mathfrak{X}(\mathcal{M}), \tag{2.15}$$

and torsion-free, i.e.

$$\nabla^g_X Y - \nabla^g_Y X \equiv [X, Y]. \tag{2.16}$$

Here, the *Lie bracket* $[X, Y] \in \mathfrak{X}(\mathcal{M})$ denotes the vector field satisfying

$$[X, Y]f = X(Yf) - Y(Xf) \qquad \text{for } f \in \mathrm{C}^\infty(\mathcal{M}, \mathbb{R}). \tag{2.17}$$

Geodesics with respect to the Levi-Civita connection are locally length minimizing, i.e., the minimizing curve in (2.7) is a geodesic. Whenever no affine connection is specified, geodesics and exponential maps are considered with respect to the Levi-Civita connection.

2.1.2 Flows

We give a short definition of flows on smooth manifolds $\mathcal{M}$. For a more detailed treatment of flows we refer to [Lee13, Chapter 9].

Definition 2.1.1 (flow)**.** A *flow* on $\mathcal{M}$ is a continuous map $\phi\colon \mathcal{D} \to \mathcal{M}$, where $\mathcal{D} \subseteq \mathbb{R} \times \mathcal{M}$ is an open subset with $\{0\} \times \mathcal{M} \subset \mathcal{D}$, that satisfies the following group laws: for all $p \in \mathcal{M}$,

$$\phi(0, p) = p, \tag{2.18}$$

[2] We use the symbol ∇ to denote the Euclidean gradient in the subsequent sections and chapters. Since affine connections are used only in this preliminary section in order to define geodesics, no notation conflict can occur.

and for all $s, t \in \mathbb{R}$ such that $(s, p), \big(t, \phi(s,p)\big), (s+t, p) \in \mathcal{D}$, it holds

$$\phi\big(t, \phi(s,p)\big) = \phi(s+t, p). \tag{2.19}$$

If $\mathcal{D} = \mathbb{R} \times \mathcal{M}$, then ϕ is called a *global flow*.

A global flow is also termed (continuous-time) *dynamical system* (cf. [BS02]). In this work, we are interested in (global) flows generated by smooth vector fields in the following sense.

Definition 2.1.2. Let $X \in \mathfrak{X}(\mathcal{M})$ be a smooth vector field on $\mathcal{M}$.

(a) An *integral curve* of X is a differentiable curve $\gamma\colon \mathcal{I} \to \mathcal{M}$ on an open interval $\mathcal{I} \subseteq \mathbb{R}$ fulfilling

$$\dot{\gamma}(t) = X_{\gamma(t)} \qquad \text{for all } t \in \mathcal{I}. \tag{2.20}$$

(b) A smooth flow $\phi\colon \mathcal{D} \to \mathcal{M}$ is said to be generated by X if $\phi^{(p)} \coloneqq \phi(\cdot, p)$ is an integral curve of X for all $p \in \mathcal{M}$.

Simply put, a flow generated by a vector field is a family of integral curves of that vector field. An integral curve is called *maximal* if it cannot be extended to an integral curve on any larger open interval. The following lemma states that a maximal integral curve exists for all $t \in \mathbb{R}$ if it stays within a compact subset. This criterion can be used to verify that a flow generated by a given vector field is global.

Lemma 2.1.3 (Escape Lemma [Lee13, Lemma 9.19])**.** *Let $\gamma\colon \mathcal{I} \to \mathcal{M}$ be a maximal integral curve of a vector field $X \in \mathfrak{X}(\mathcal{M})$ on a smooth manifold $\mathcal{M}$. Suppose $\mathcal{I}$ has a finite upper bound, i.e. $b = \sup \mathcal{I} < \infty$. Then for any $t_0 \in \mathcal{I}$, $\gamma\big([t_0, b)\big)$ is not contained in any compact subset of $\mathcal{M}$.*

In the subsequent sections and chapters we will use the terms (global) flow and dynamical system interchangeably for the actual flow as well as for the ODE (2.20) generating this flow.

2.1.3 Statistical Models from the Viewpoint of Information Geometry

We give a brief overview of a subset of information geometry which deals with manifolds of probability distributions. The considered class of Riemannian manifolds are statistical models equipped with the Fisher information metric. For our overview we follow the textbook [CU14] while also using some aspects from [AJLS17, Ama16].

The Framework

Let $(\mathcal{X}, \Sigma, \mu)$ be a measure space, that means, $\mathcal{X}$ is a set, Σ is a σ-algebra of subsets of $\mathcal{X}$, and $\mu\colon \Sigma \to \mathbb{R} \cup \{\pm\infty\}$ is a measure on $\mathcal{X}$. The set $\mathcal{X}$ is called *sample space*. In case of discrete distributions, we consider the counting measure μ on a finite or countable set $\mathcal{X}$, and in case of continuous distributions, μ is the Lebesgue measure on a set $\mathcal{X} \subseteq \mathbb{R}^d$. In the following we regard parameterized surfaces in the space of probability densities on $\mathcal{X}$ with respect to μ, i.e. in the space

$$\mathcal{P}(\mathcal{X}, \mu) = \Big\{ f \in \mathrm{L}^1(\mathcal{X}, \mu)\colon f \geq 0,\ \int_{\mathcal{X}} f \,\mathrm{d}\mu = 1 \Big\}. \tag{2.21}$$

Definition 2.1.4 (statistical model)**.** Let Ξ be an open subset of $\mathbb{R}^n$ or, more generally, a smooth manifold of dimension n. We call the set Ξ the *parameter space*. Further, let

$$\Xi \to \mathcal{P}(\mathcal{X}, \mu), \qquad \xi \mapsto p_\xi \tag{2.22}$$

be an injective C^1-map such that[3]

$$\frac{\partial p_\xi}{\partial \xi^1}, \dots, \frac{\partial p_\xi}{\partial \xi^n} \tag{2.23}$$

are linearly independent functions. The parameterized family

$$\mathcal{M} = \left\{ p_\xi \colon \xi \in \Xi \right\} \subset \mathcal{P}(\mathcal{X}, \mu) \tag{2.24}$$

is then called a *(parametric) statistical model*[4] of dimension n.

The inverse map of (2.22), i.e. $p_\xi \mapsto \xi$, can be considered as local coordinates turning $\mathcal{M}$ into a manifold. In the following we consider only probability density functions with full support, i.e., we regard statistical models which are subsets of

$$\mathcal{P}_+(\mathcal{X}, \mu) = \left\{ f \in \mathcal{P}(\mathcal{X}, \mu) \colon f > 0 \right\} \subset \mathcal{P}(\mathcal{X}, \mu). \tag{2.25}$$

To define a Riemannian metric on $\mathcal{M}$, we introduce a few notations first. We denote the expectation of a function $f \in L^1(\mathcal{X}, p_\xi \mu)$ with respect to the probability measure $p_\xi \mu$ by

$$\mathbb{E}_\xi[f] \coloneqq \int_{\mathcal{X}} f(x) p_\xi(x) \, \mathrm{d}\mu(x). \tag{2.26}$$

Further we use the *log-likelihood function* $l(\xi) \coloneqq \log p_\xi$ as well as the shorthand $\partial_i \coloneqq \frac{\partial}{\partial \xi^i}$. The functions $\partial_i l(\xi)$, $i \in [n]$ are called *score functions* and they fulfill $\mathbb{E}_\xi[\partial_i l(\xi)] = 0$ assuming $\partial_i l(\xi) \in L^1(\mathcal{X}, p_\xi \mu)$. In the following definition we assume $\partial_i l(\xi) \in L^2(\mathcal{X}, p_\xi \mu)$ for all $i \in [n]$ and all $\xi \in \Xi$.

Definition 2.1.5. The *Fisher information metric* on a statistical model $\mathcal{M} = \{p_\xi \colon \xi \in \Xi\}$ is defined by

$$g_\xi(\partial_i, \partial_j) = \mathbb{E}_\xi\big[\partial_i l(\xi) \partial_j l(\xi)\big], \qquad \forall i, j \in [n], \qquad \forall \xi \in \Xi. \tag{2.27}$$

The Fisher information metric, also called *Fisher-Rao metric*, defines a Riemannian structure on $\mathcal{M}$, which is distinct for the following reasons. Up to scaling, this Riemannian metric is unique in that it is invariant under sufficient statistics – cf. [AN00, Theorem 2.6]. In particular, it is invariant under reparametrizations of the sample space which are trivial examples of sufficient statistics. Additionally, the metric (2.27) transforms covariantly under reparametrizations of the parameter space [CU14, Theorem 1.6.5], so that the geometry (e.g. Riemannian distance) induced by the Fisher-Rao metric is independent of the parameterization of the statistical model $\mathcal{M}$.

Using the Fisher-Rao metric g, a one-dimensional family of affine connections on $\mathcal{M}$ can be defined as follows.

[3] Since $\xi \mapsto p_\xi \in L^1(\mathcal{X}, \mu)$ maps into a Banach space, differentiability means Fréchet differentiability and the (partial) derivatives are taken as Gâteaux-derivatives [AJLS17, Definition C.1].

[4] A statistical model may also refer to the corresponding set of probability measures on $\mathcal{X}$, i.e. the set $\{p_\xi \mu \colon \xi \in \Xi\}$.

Definition 2.1.6. The *α-connection* $\nabla^{(\alpha)}$, $\alpha \in [-1,1]$ on a statistical model $\mathcal{M} = \{p_\xi \colon \xi \in \Xi\}$ is defined by

$$g\big(\nabla^{(\alpha)}_{\partial_i}\partial_j, \partial_k\big) = \mathbb{E}_\xi\Big[\big(\partial_i\partial_j l + \tfrac{1-\alpha}{2}(\partial_i l)(\partial_j l)\big)(\partial_k l)\Big] \qquad \text{for all } i,j,k \in [n]. \tag{2.28}$$

By the linearity of affine connections the above equation describes the α-connection completely.

This family of connections includes the Levi-Civita connection ∇^g of $(\mathcal{M}, g)$ given by $\alpha = 0$. The two marginal connections are the mixture connection $\nabla^{(m)}$ given by $\alpha = -1$ and the exponential connection $\nabla^{(e)}$ given by $\alpha = 1$.

Some Notable Examples

We start our examples with a class of statistical models which includes many common families of probability distributions like the (multivariate) normal distribution, exponential distribution, Bernoulli distribution and Poisson distribution.

Definition 2.1.7. An *exponential family* is a parametric statistical model given by probability density functions of the form

$$p_\theta(x) = \exp\big(\gamma(x) + \langle\theta, f(x)\rangle - \psi(\theta)\big), \tag{2.29}$$

where $\theta = (\theta^1, \dots, \theta^n)$ are called *natural parameters* with domain

$$\Theta = \left\{\theta \in \mathbb{R}^n \colon \int_{\mathcal{X}} \exp\big(\gamma(x) + \langle\theta, f(x)\rangle\big)\,\mathrm{d}\mu(x) < \infty\right\}, \tag{2.30}$$

and $\gamma(x), f_1(x), \dots, f_n(x)$ are functions on $\mathcal{X}$ such that the functions $1, f_1, \dots, f_n$ are linearly independent. The function

$$\psi(\theta) = \log\Big(\int_{\mathcal{X}} \exp\big(\gamma(x) + \langle\theta, f(x)\rangle\big)\,\mathrm{d}\mu(x)\Big) \tag{2.31}$$

is called *log-partition function*.

We note that the log-partition function ψ is strictly convex on the convex set Θ. Calculating the Fisher-Rao metric (2.27) for an exponential family gives

$$g_\theta(\partial_i, \partial_j) = \mathrm{Cov}_\theta(f_i, f_j) = \partial_i\partial_j\psi(\theta) \qquad \forall i,j \in [n], \tag{2.32}$$

where Cov_θ denotes the covariance with respect to the probability measure $p_\theta\mu$. Furthermore, applying (2.28) with $\alpha = 1$ results in $g_\theta\big(\nabla^{(e)}_{\partial_i}\partial_j, \partial_k\big) = 0$. The latter equation is referred to as $\nabla^{(e)}$-flatness, i.e., a statistical model given by an exponential family is $\nabla^{(e)}$-flat.

Next we discuss categorial distributions which we will use later. These distributions form an exponential family, but we may also choose a different parameterization than (2.29).

Example 2.1.8. We consider discrete probability distributions on the finite sample space $\mathcal{X} = \{x^1, \dots, x^n\}$. In particular, we use the counting measure on $\mathcal{X}$, i.e. $\mu(A) = |A|$ for $A \subseteq \mathcal{X}$. We focus on the parameterization

$$p_\xi(x^i) = \begin{cases} \xi^i, & i \in [n-1], \\ 1 - \sum_{j\in[n-1]} \xi^j, & i = n, \end{cases} \tag{2.33}$$

with parameter space

$$\Xi = \Big\{(\xi^1, \dots, \xi^{n-1}) \in \mathbb{R}^{n-1}_{>0} \colon \sum_{i \in [n-1]} \xi^i < 1\Big\}. \tag{2.34}$$

The statistical model $\mathcal{M} = \{p_\xi \colon \xi \in \Xi\}$ can be identified with the open probability simplex

$$\mathcal{S} \coloneqq \big\{p \in \mathbb{R}^n \colon p > 0,\ \langle p, \mathbb{1}_n \rangle = 1\big\}. \tag{2.35}$$

The Fisher-Rao metric (2.27) now takes the form

$$g_\xi(\partial_i, \partial_j) = \begin{cases} \frac{1}{p_\xi(x^i)} + \frac{1}{p_\xi(x^n)}, & i = j, \\ \frac{1}{p_\xi(x^n)}, & i \neq j \end{cases} \qquad \text{for } i, j \in [n-1]. \tag{2.36}$$

2.1.4 Hessian Riemannian Geometry

Now we discuss a particular class of Riemannian manifolds which are motivated by a constrained minimization problem and presented in [ABB04]. These manifolds are convex subsets of $\mathbb{R}^n$ equipped with a Riemannian metric induced by the Hessian matrix of a strictly convex function. For these manifolds, the tangent spaces, the Riemannian metric and the Riemannian gradient can be expressed in the coordinates of the ambient space $\mathbb{R}^n$, which simplifies their use in practice.

The Framework

Let $C \subseteq \mathbb{R}^n$ be an open and convex subset. Further, let $A \in \mathbb{R}^{m \times n}$ be a full-rank matrix with $m \leq n$ and let $b \in \mathbb{R}^m$. The convex set

$$\mathcal{M} = \{x \in C \colon Ax = b\} \tag{2.37}$$

forms a smooth manifold whose tangent space at any $x \in \mathcal{M}$ is given by

$$T_x\mathcal{M} = \ker A = \{v \in \mathbb{R}^n \colon Av = 0\}. \tag{2.38}$$

With the goal of solving the constrained minimization problem

$$\min_{x \in \overline{\mathcal{M}}} f(x) \qquad \text{with} \quad f \in \mathrm{C}^1(\mathbb{R}^n, \mathbb{R}) \tag{2.39}$$

using a interior point flow, a Riemannian structure on $\mathcal{M}$ is defined so that the trajectories generated by the Riemannian gradient descent vector field

$$\dot{x}(t) = -\operatorname{grad} f|_{\mathcal{M}}(x(t)), \qquad x(0) \in \mathcal{M}, \tag{2.40}$$

remain in the relative interior $\mathcal{M}$ of the feasible set $\overline{\mathcal{M}}$. To this end, we consider the Hessian of a strictly convex function of a particular class defined below. For this purpose we denote the domain of a function $\varphi \colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ by $\operatorname{dom}\varphi = \{x \in \mathbb{R}^n \colon \varphi(x) < \infty\}$ and its interior by $\operatorname{int}(\operatorname{dom}\varphi)$.

Definition 2.1.9 (Legendre type)**.** Let $\varphi \colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $\varphi \not\equiv \infty$ be lower semicontinuous and convex. φ is *of Legendre type* if

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

(ii) $\|\nabla\varphi(x^j)\| \to \infty$ as $j \to \infty$ for every sequence $(x^j) \subset \operatorname{int}(\operatorname{dom}\varphi)$ converging to a boundary point of $\operatorname{dom}\varphi$.

Definition 2.1.10 (Legendre metric)**.** Let $\varphi\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$ be a function such that

(i) φ is of Legendre type with $\operatorname{int}(\operatorname{dom}\varphi) = C$,

(ii) $\varphi|_C \in \mathrm{C}^2(C, \mathbb{R})$ and the Hessian matrix $\nabla^2\varphi(x)$ is positive definite for all $x \in C$,

(iii) the mapping $C \ni x \mapsto \nabla^2\varphi(x)$ is locally Lipschitz continuous.

Then the Riemannian metric $g_x(u, v) = \langle u, G(x)v\rangle$ on C given by the Hessian $G(x) = \nabla^2\varphi(x)$ is called the *Legendre metric* induced by φ.

The Legendre metric defines a Riemannian metric on $\mathcal{M} \subset C$ by restriction. The Riemannian gradient of a function $f \in \mathrm{C}^1(C, \mathbb{R})$ with respect to this metric takes the form

$$\operatorname{grad} f|_{\mathcal{M}}(x) = \Pi_x G(x)^{-1}\nabla f(x) \in T_x\mathcal{M}, \tag{2.41}$$

where $\Pi_x \in \mathbb{R}^{n\times n}$ denotes the g_x-orthogonal projection onto $T_x\mathcal{M} = \ker A$ given by

$$\Pi_x = I_n - G(x)^{-1}A^\top\left(AG(x)^{-1}A^\top\right)^{-1}A \tag{2.42}$$

and ∇f denotes the Euclidean gradient of f.

Besides the Riemannian structure on $\mathcal{M}$, the Legendre type function φ also induces a coordinate transform from $\mathcal{M}$ to a subset of $\ker A$. For some choices of φ, this subset turns out to be the whole $\ker A$, which enables to transform the constrained flow (2.40) to an unconstrained flow in the vector space $\ker A$.

Definition 2.1.11 (Legendre transform)**.** Let φ fulfill the assumption of Definition 2.1.10. The *Legendre transform coordinates mapping* on $\mathcal{M}$ associated with φ is defined by

$$\Psi_\varphi\colon \mathcal{M} \to \mathcal{M}^* := \Psi_\varphi(\mathcal{M}) \subseteq \ker A, \quad x \mapsto \Psi_\varphi(x) = \Pi\nabla\varphi(x), \tag{2.43}$$

where $\Pi = I_n - A^\top(AA^\top)^{-1}A \in \mathbb{R}^{n\times n}$ denotes the orthogonal projection onto $\ker A$.

Theorem 2.1.12 ([ABB04, Theorem 5.1])**.** *Under the above definition and assumptions, $\mathcal{M}^*$ is a convex and (relatively) open subset of* $\ker A$*, and Ψ_φ is a C^1-diffeomorphism from $\mathcal{M}$ to $\mathcal{M}^*$.*

Some Notable Examples

For later reference we now discuss a couple of examples of this framework. First we sketch a systematic method presented in [ABB04] to construct explicit Legendre metrics. Assume that the convex set is given of the form

$$C = \left\{x \in \mathbb{R}^n\colon f_i(x) > 0 \quad \forall i \in I\right\}, \tag{2.44}$$

where I is a finite set and $f_i \in \mathrm{C}^3(\mathbb{R}^n, \mathbb{R})$ is a concave function for all $i \in I$. Further assume that these functions fulfill the following nondegeneracy condition:

$$\operatorname{span}\left\{\nabla f_i(x)\colon i \in I\right\} = \mathbb{R}^n \qquad \text{for all } x \in C. \tag{2.45}$$

Then a function φ fulfilling the assumptions of Definition 2.1.10 is given by

$$\varphi(x) = \sum_{i \in I} \theta(f_i(x)), \tag{2.46}$$

where $\theta \colon \mathbb{R} \to \mathbb{R} \cup \{\infty\}$ is a Legendre type function satisfying

(i) either $\operatorname{dom} \theta = \mathbb{R}_{>0}$ or $\operatorname{dom} \theta = \mathbb{R}_{\geq 0}$,

(ii) $\theta \in \mathrm{C}^3(\mathbb{R}_{>0}, \mathbb{R})$,

(iii) $\lim_{s \to 0^+} \theta'(s) = -\infty$,

(iv) $\theta'' > 0$ on $\mathbb{R}_{>0}$,

(v) either θ is nonincreasing or all f_i, $i \in I$ are affine functions.

The function θ is called *Legendre kernel* of φ. Examples of such kernels are the Boltzmann-Shannon entropy

$$\theta_0(s) = s \log(s) - s \qquad \text{with} \qquad \operatorname{dom} \theta_0 = \mathbb{R}_{\geq 0} \tag{2.47a}$$

as well as the log-barrier

$$\theta_1(s) = -\log(s) \qquad \text{with} \qquad \operatorname{dom} \theta_1 = \mathbb{R}_{>0}. \tag{2.47b}$$

We apply this construction to the open probability simplex

$$\mathcal{S} = \left\{ p \in \mathbb{R}^n_{>0} \colon \langle \mathbb{1}_n, p \rangle = 1 \right\}, \tag{2.48}$$

whose tangent space is given by

$$T_p \mathcal{S} = \left\{ v \in \mathbb{R}^n \colon \langle \mathbb{1}_n, v \rangle = 0 \right\} =: T_0. \tag{2.49}$$

Using the entropy θ_0 as kernel, we get the Legendre type function

$$\varphi(p) = \sum_{j \in [n]} p_j \log p_j - p_j \tag{2.50}$$

with $\operatorname{dom} \varphi = \mathbb{R}^n_{\geq 0}$. The induced Legendre metric

$$g_p(u, v) = \left\langle u, \operatorname{Diag}(p)^{-1} v \right\rangle \qquad \forall u, v \in T_0 \tag{2.51}$$

is the Fisher-Rao metric (2.36) represented in the coordinates of the surrounding $\mathbb{R}^n$. The equation (2.41) for the Riemannian gradient on $\mathcal{S}$ now takes the form

$$\operatorname{grad} f(p) = R_p \nabla f(p) \qquad \text{with} \qquad R_p \coloneqq \operatorname{Diag}(p) - p p^\top. \tag{2.52}$$

The Legendre transform coordinate mapping (2.43)

$$\Psi_\varphi \colon \mathcal{S} \to T_0, \qquad \Psi_\varphi(p) = \Pi_0 \log(p) \qquad \text{with} \qquad \Pi_0 \coloneqq I_n - \tfrac{1}{n} \mathbb{1}_n \mathbb{1}_n^\top \tag{2.53}$$

is a diffeomorphism between the manifold $\mathcal{S}$ and the vector space T_0. Its inverse is

$$\Psi_\varphi^{-1}\colon T_0 \to \mathcal{S}, \qquad \Psi_\varphi^{-1}(v) = \frac{e^v}{\langle e^v, \mathbb{1}_n\rangle}, \tag{2.54}$$

where exponentiation is applied componentwise.

Finally, we mention the set of symmetric positive definite matrices,

$$\mathcal{P}_d \coloneqq \left\{X \in \mathbb{R}^{d\times d}\colon X^\top = X,\ X \text{ is positive definite}\right\}, \tag{2.55}$$

which may be interpreted to be of the form (2.37). We regard it as an open subset of the vector space of symmetric matrices

$$\mathrm{Sym}_d \coloneqq \left\{X \in \mathbb{R}^{d\times d}\colon X^\top = X\right\}. \tag{2.56}$$

The tangent space at $X \in \mathcal{P}_d$ is therefore $T_X\mathcal{P}_d = \mathrm{Sym}_d$. The function

$$\varphi(X) = \begin{cases} -\log\det(X), & \text{if } X \in \mathcal{P}_d \\ \infty, & \text{else} \end{cases} \tag{2.57}$$

is a Legendre type function on the vector space Sym_d and the induced Legendre metric on $\mathcal{P}_d$,

$$g_X(U,V) = \operatorname{tr}\left(X^{-1}UX^{-1}V\right), \tag{2.58}$$

is the so-called *affine invariant Riemannian metric* [PFA06]. This metric may also be referred to as Fisher information metric, since it is the metric (2.27) corresponding to zero-mean multivariate Gaussian distributions [Ama16, Section 4.5]. The Riemannian gradient of a function $f\colon \mathcal{P}_d \to \mathbb{R}$ with respect to this metric takes the form

$$\operatorname{grad} f(X) = X\nabla f(X)X \in \mathrm{Sym}_d, \qquad X \in \mathcal{P}_d, \tag{2.59}$$

where $\nabla f(X) \in \mathrm{Sym}_d$ denotes the Euclidean gradient at X, and the matrices X and $\nabla f(X)$ are multiplied as usual.

2.1.5 Divergence Functions

In this subsection we present divergence functions which we will use as surrogate functions for (squared) distances. We follow the presentation of [ABB04] and [Ama16]. For a more detailed description of Bergman divergences we refer to [CZ97, BB97].

Bregman divergences are distance-like functions of the form

$$D_\varphi\colon \operatorname{dom}\varphi \times \operatorname{int}(\operatorname{dom}\varphi) \to \mathbb{R}_{\geq 0}, \qquad D_\varphi(x,y) \coloneqq \varphi(x) - \varphi(y) - \langle \varphi(y), x-y\rangle, \tag{2.60}$$

induced by convex functions $\varphi\colon \mathbb{R}^d \to \mathbb{R}\cup\{\infty\}$ of Legendre type – cf. Definition 2.1.9. These divergences D_φ satisfy

$$D_\varphi \geq 0 \quad \text{and} \quad D_\varphi(x,y) = 0 \ \Leftrightarrow\ x = y, \tag{2.61}$$

which shows that D_φ behaves like a distance, but symmetry $D_\varphi(x,y) = D_\varphi(y,x)$ is not required and generally does not hold. If the function φ is a Bregman function as defined below, then the divergence D_φ is also useful for convergence proofs.

Definition 2.1.13 (Bregman function [ABB04, Definition 4.1]). Let $\varphi\colon \mathbb{R}^n \to \mathbb{R} \cup \{\infty\}$, $\varphi \not\equiv \infty$ be lower semicontinuous and convex. φ is called *Bregman function with zone C* if

(i) $\operatorname{dom}\varphi = \overline{C}$,

(ii) φ is continuous and strictly convex on $\overline{C}$,

(iii) φ is continuously differentiable on C,

(iv) $\{y \in C\colon D_\varphi(a, y) \leq \gamma\}$ is bounded for any $a \in \overline{C}$ and any $\gamma \in \mathbb{R}$,

(v) $D_\varphi(y, y^j) \to 0$ as $j \to \infty$ for all $y \in \overline{C}$ and for every sequence $(y^j) \subset C$ with $y^j \to y$.

Lemma 2.1.14 ([ABB04, Lemma 4.3]). *If φ is a Bregman function with zone C then for all $y \in \overline{C}$ and all sequences $(y^j) \subset C$ with $D_\varphi(y, y^j) \to 0$ we have $y^j \to y$.*

Now we discuss divergence functions on manifolds. Let $(\mathcal{M}, g)$ be a d-dimensional Riemannian manifold. A function $D\colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$ is a proper divergence function defined on $\mathcal{M}$ if, for any chart $U \subset \mathcal{M}$ with local coordinates $x\colon U \to \mathbb{R}^d$ and $p, q \in U$, the function

$$\widetilde{D}(x, y) = \widetilde{D}\big(x(p), x(q)\big) = D(p, q) \tag{2.62a}$$

satisfies (2.61) and recovers the positive definite metric tensor by

$$D(p, q) \approx \frac{1}{2} \sum_{i,j \in [d]} g_{ij}(p) z^i z^j, \tag{2.62b}$$

for $z = x(q) - x(p)$ and small $\|z\|$. We mention some examples for later reference. Firstly, the function $D(p, q) = \frac{1}{2} d_g(p, q)^2$ defines a *canonical* divergence function on a Riemannian manifold $(\mathcal{M}, g)$ in terms of the squared Riemannian distance d_g^2. A more general definition of a canonical divergence is used in information geometry, where the Levi-Civita connection is replaced by another affine connection in order to define a divergence function through affine geodesics and corresponding squared distances. We refer to [AN00, Section 3.4], [AC10] and [AJLS17, Section 4.4] for background and further details.

Moreover, with regard to the framework of Section 2.1.4, the Bregman divergence D_φ is a proper divergence function on the Riemannian manifold $(\mathcal{M}, g)$ whose Riemannian metric g is the Legendre metric induced by the same Legendre type function φ – cf. Definition 2.1.10. As specific examples we continue the two examples from Section 2.1.4. The Fisher-Rao metric (2.51) on the open probability simplex $\mathcal{S}$ is induced by the entropy (2.50). The corresponding Bregman divergence is the so-called *Kullback–Leibler divergence*

$$D_{\mathrm{KL}}(p, q) = \sum_{j \in [n]} p_j \log\left(\tfrac{p_j}{q_j}\right) + p_j - q_j = \sum_{j \in [n]} p_j \log\left(\tfrac{p_j}{q_j}\right), \qquad p \in \overline{\mathcal{S}}, \quad q \in \mathcal{S}. \tag{2.63}$$

Since the entropy (2.50) is a Bregman function according to Definition 2.1.13, the KL divergence satisfies the convergence property stated in Lemma 2.1.14. We will use this property for the convergence proof of the assignment flow in Section 3.3. The second example is the affine invariant metric (2.58) on the manifold $\mathcal{P}_d$ of positive definite matrices. The Riemannian structure is induced by the log-determinant (2.57) which also induces the log-det divergence

$$D_\varphi(X, Y) = -\log\det(X) + \log\det(Y) + \operatorname{tr}\left(Y^{-1}(X - Y)\right) \tag{2.64a}$$

$$= -\log\det\left(Y^{-1}X\right) + \operatorname{tr}\left(Y^{-1}X\right) - d. \tag{2.64b}$$

A symmetrization of this divergence results in the *Stein divergence* [Sra13]

$$D_{\mathrm{S}}(X,Y) = \tfrac{1}{2}D_\varphi\big(X, \tfrac{X+Y}{2}\big) + \tfrac{1}{2}D_\varphi\big(Y, \tfrac{X+Y}{2}\big) \tag{2.65a}$$
$$= \log\det\big(\tfrac{X+Y}{2}\big) - \tfrac{1}{2}\log\det(XY), \tag{2.65b}$$

which is also a proper divergence function on $\mathcal{P}_d$. We will study this and other concrete divergence functions in connection with unsupervised label learning in Section 5.6, where in some cases alternative divergences are required that serve as surrogate functions for the (squared) Riemannian distance and are easier to evaluate computationally.

2.1.6 Riemannian Mean

We shortly mention the Riemannian mean which generalizes the arithmetic mean from Euclidean spaces to Riemannian manifolds.

Let $(\mathcal{M}, g)$ be a Riemannian manifold. The *weighted Riemannian mean* [Jos17, Def. 6.9.1] of a collection of points $p^1, \dots, p^n \in \mathcal{M}$ with respect to weights $w = (w_1, \dots, w_n) \in \overline{\mathcal{S}}$ is a point $q \in \mathcal{M}$ satisfying

$$J_w(q) = \inf_{p\in\mathcal{M}} J_w(p), \qquad J_w(p) = \frac{1}{2}\sum_{i\in[n]} w_i d_g^2(p^i, p), \tag{2.66}$$

where $d_g(q,p)$ denotes the Riemannian distance (2.7). We have [Jos17, Lemma 6.9.4]

$$\operatorname{grad} J_w(p) = -\sum_{i\in[n]} w_i \operatorname{Exp}_p^{-1}(p^i) \in T_p\mathcal{M}, \tag{2.67}$$

where Exp_p denotes the Riemannian exponential map (2.13). Hence the optimality condition for q reads

$$\sum_{i\in[n]} w_i \operatorname{Exp}_q^{-1}(p^i) = 0. \tag{2.68}$$

This equation is typically solved by the *mean shift* (fixed point) *iteration*

$$q^{(t+1)} = \operatorname{Exp}_{q^{(t)}}\Big(\sum_{i\in[n]} w_i \operatorname{Exp}_{q^{(t)}}^{-1}(p^i)\Big), \quad t \in \mathbb{N}_0 \tag{2.69}$$

with a suitable initialization $q^{(0)}$.

Eventually we note that the Riemannian mean can be generalized to a Riemannian mean with respect to a divergence function D by replacing $\frac{1}{2}d_g^2$ in (2.66) by the divergence D. The Riemannian gradient and the optimality condition $\operatorname{grad} J_w(q) = 0$ change accordingly.

2.2 Assignment Flow

In this section, we review the assignment flow introduced in [ÅPSS17] as a geometric approach to the image labeling problem. This flow evolves on the manifold of row-stochastic matrices whose elements assign prior data to observed measurements. The strict separation of assignment manifold and feature space allows the data to lie in any metric space, while geometric

averaging on the assignment manifold results in an unbiasedly spatially regularized labeling. We refer to [Sch20] for an overview of recent work on the assignment flow.

We start the review with a discussion of input data and assumptions of the framework. After presenting the assignment manifold and the lifting map onto it, we define the similarity map which steers the assignment flow. Eventually, we discuss the numerical integration of the assignment flow.

Prerequisite

Let $(\mathcal{M}, d_{\mathcal{M}})$ be a metric space and let data be given by observations

$$\{z^i\}_{i\in I} \subset \mathcal{M}. \tag{2.70}$$

The set I is assumed to form the vertex set of an undirected graph $G = (I, E)$ with $E \subset I \times I$ which describes the spatial relation of the data (2.70). Most of the time we consider manifold-valued images so that G will be the grid graph of the image domain and $\mathcal{M}$ will be a Riemannian manifold representing the feature space of the image. In that setting, the set I corresponds to pixel locations and $z^i \in \mathcal{M}$ is the datum given at pixel $i \in I$. The graph G defines neighborhoods

$$\mathcal{N}_i = \big\{k \in I \colon (i,k) \in E\big\} \cup \{i\} \tag{2.71}$$

with which we associate weights $\{\omega_{ik} \colon k \in \mathcal{N}_i\}$ satisfying

$$\omega_{ik} > 0 \quad \forall k \in \mathcal{N}_i, \qquad \sum_{k\in\mathcal{N}_i} \omega_{ik} = 1 \qquad \forall i \in I. \tag{2.72}$$

For example, these can simply be chosen to be the uniform weights

$$\omega_{ik} = \frac{1}{|\mathcal{N}_i|} \quad \forall k \in \mathcal{N}_i \quad \forall i \in I. \tag{2.73}$$

Besides the observations (2.70), we assume predefined prototypes (class representatives)

$$M = \{m^j\}_{j\in J} \subset \mathcal{M} \tag{2.74}$$

to be given, which we also call *labels*. We will discuss in Chapter 5 how these prototypes can be learned and adapted within the framework. Last but not least we consider a dissimilarity function $D_{\mathcal{M}} \colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}_{\geq 0}$. For instance, this can be the distance function $d_{\mathcal{M}}$. In case of a Riemannian manifold $\mathcal{M}$, we may also choose a divergence function on $\mathcal{M}$ as discussed in Section 2.1.5.

We consider the *data labeling* problem, i.e., the task of assigning each node $i \in I$ a single label $m^{j(i)} \in M$. This labeling should meet two partly conflicting conditions. On the one hand it should be faithful to the data in the sense that $D_{\mathcal{M}}(z^i, m^{j(i)})$ is as small as possible. On the other hand it should be spatially coherent, meaning that nodes in the same neighborhood $\mathcal{N}_i$ should get the same label if possible. By setting

$$W_{ij} = \begin{cases} 1, & \text{if label } m^j \text{ is assigned to node } i \in I, \\ 0, & \text{otherwise,} \end{cases} \tag{2.75}$$

an labeling can be represented as an element of

$$\mathcal{W}^* := \big\{W \in \{0,1\}^{m\times n} \colon W\mathbb{1}_n = \mathbb{1}_m\big\}, \tag{2.76}$$

where m and n denote the number of vertices of the underlying graph G and the number of labels indexed by J, respectively,

$$m = |I|, \qquad n = |J|. \tag{2.77}$$

The approach [ÅPSS17] considers the probabilistic view of the labeling problem, i.e., $W_{ij} = \Pr(m^j|z^i)$ represents the probability of assigning label m^j to node $i \in I$. Each row W_i of the matrix W is then an element of the probability simplex

$$\Delta_n := \big\{p \in \mathbb{R}^n_{\geq 0} \colon \langle p, \mathbb{1}_n\rangle = 1\big\}. \tag{2.78}$$

The assignment manifold discussed next is based on this view.

Assignment Manifold

The relative interior of Δ_n is the open probability simplex

$$\mathcal{S} = \operatorname{rint}(\Delta_n) = \big\{p \in \mathbb{R}^n_{>0} \colon \langle p, \mathbb{1}_n\rangle = 1\big\}. \tag{2.79}$$

As discussed in Section 2.1.4, this is a smooth Riemannian manifold with tangent spaces

$$T_p\mathcal{S} = T_0 = \big\{v \in \mathbb{R}^n \colon \langle v, \mathbb{1}_n\rangle = 0\big\}, \qquad p \in \mathcal{S}, \tag{2.80}$$

orthogonal projection $\Pi_0 \colon \mathbb{R}^n \to T_0$ given by the matrix

$$\Pi_0 = I_n - \tfrac{1}{n}\mathbb{1}_n\mathbb{1}_n^\top \in \mathbb{R}^{n\times n}, \tag{2.81}$$

and Riemannian structure given by the Fisher-Rao metric (2.51)

$$g_p(u,v) = \big\langle u, \operatorname{Diag}(p)^{-1}v\big\rangle = \sum_{j\in[n]} \frac{u_j v_j}{p_j} \qquad \text{for } u, v \in T_0. \tag{2.82}$$

The barycenter of $\mathcal{S}$ is the uniform distribution

$$C_{\mathcal{S}} := \tfrac{1}{n}\mathbb{1}_n. \tag{2.83}$$

The *assignment manifold*[5] is defined as the product space

$$\mathcal{W} := \underbrace{\mathcal{S} \times \cdots \times \mathcal{S}}_{m \text{ times}} \tag{2.84}$$

which we identify with the space

$$\mathcal{W} = \big\{W \in \mathbb{R}^{m\times n}_{>0} \colon W\mathbb{1}_n = \mathbb{1}_m\big\}. \tag{2.85}$$

[5] The assignment manifold is sometimes called multinomial manifold in the literature – cf. [LL05].

Thus, any element $W \in \mathcal{W}$ is a row-stochastic matrix $W \in \mathbb{R}^{m\times n}$ with rows $W_i \in \mathcal{S}$ representing the assignment probabilities for node $i \in I$. We call the vectors $W_i \in \mathcal{S}$ *assignment vectors*. The tangent space

$$\mathcal{T}_0 := \underbrace{T_0 \times \cdots \times T_0}_{m \text{ times}} \tag{2.86}$$

can be identified with the space

$$\mathcal{T}_0 = \big\{V \in \mathbb{R}^{m\times n} \colon V\mathbb{1}_n = 0\big\}. \tag{2.87}$$

Thus, $V_i \in T_0$ for all rows of $V \in \mathcal{T}_0 \subset \mathbb{R}^{m\times n}$. The orthogonal projection $\Pi_0 \colon \mathbb{R}^{m\times n} \to \mathcal{T}_0$ is given by row-wise application of the projection (2.81). The induced Riemannian metric on $\mathcal{W}$ is given by

$$g_W(U,V) = \sum_{i\in I} g_{W_i}(U_i, V_i) \qquad \text{for } U, V \in \mathcal{T}_0 \tag{2.88}$$

which we also name Fisher-Rao metric. The uninformative distribution

$$C_{\mathcal{W}} := \mathbb{1}_m C_{\mathcal{S}}^\top = \tfrac{1}{n}\mathbb{1}_m\mathbb{1}_n^\top \tag{2.89}$$

forms the barycenter of the assignment manifold $\mathcal{W}$.

Lifting Map

The authors of [ÅPSS17] proposed to approximate the Riemannian exponential map on $\mathcal{S}$ by the *lifting map*[6]

$$\exp_p \colon T_0 \to \mathcal{S}, \qquad (p,v) \mapsto \exp_p(v) = \frac{pe^v}{\langle p, e^v\rangle}, \tag{2.90}$$

where exponentiation and multiplication are applied entrywise. Unlike the Riemannian exponential map which is induced by the Levi-Civita connection, the lifting map $\exp_p$ is defined on the entire tangent space T_0, and it is more convenient for numerical computations. Before we continue with the properties of the lifting map, we mention two more maps related to this map. The *replicator operator* is a linear mapping $R_p \colon \mathbb{R}^n \to T_0$ given by the matrix

$$R_p = \operatorname{Diag}(p) - pp^\top \in \mathbb{R}^{n\times n}, \qquad p \in \mathcal{S}. \tag{2.91}$$

It is connected to replicator equations from evolutionary game dynamics [HS03] which take the form

$$\dot p = R_p[F(p)] \qquad \text{with a vector field } F \colon \mathcal{S} \to \mathbb{R}^n. \tag{2.92}$$

This map also appears in the Riemannian gradient (2.52) and it satisfies

$$R_p = R_p\Pi_0 = \Pi_0 R_p. \tag{2.93}$$

The second map is the exponential map corresponding to the α-connection with $\alpha = 1$ from information geometry – cf. Definition 2.1.6. The map takes the form [AJLS17, Section 2.4]

$$\operatorname{Exp}_p \colon T_0 \to \mathcal{S}, \qquad \operatorname{Exp}_p(v) = \frac{pe^{\frac{v}{p}}}{\left\langle p, e^{\frac{v}{p}}\right\rangle}, \qquad p \in \mathcal{S}. \tag{2.94}$$

[6] The lifting map $\exp_p$ is denoted with lower case letters in order to distinguish it from the (Riemannian) exponential map Exp_p (2.13).

This exponential map is defined on the entire tangent space T_0 and it has the inverse

$$\operatorname{Exp}_p^{-1}\colon \mathcal{S} \to T_0, \qquad \operatorname{Exp}_p^{-1}(q) = R_p\big[\log \tfrac{q}{p}\big]. \tag{2.95}$$

With these two maps the lifting map (2.90) can also be written as

$$\exp_p = \operatorname{Exp}_p \circ R_p. \tag{2.96}$$

We note that the lifting map extends to a map $\exp_p\colon \mathbb{R}^n = T_0 \oplus \mathbb{R}\mathbb{1}_n \to \mathcal{S}$ which does not depend on the constant component of the argument, due to (2.96) and (2.93).

The following lemma summarizes some properties of the lifting map (2.90) which we will use in the subsequent subsection as well as in Chapter 3.

Lemma 2.2.1 ([SS20, Lemma 3.1])**.** *The following properties hold for* $\exp_p\colon \mathbb{R}^n \to \mathcal{S}$*.*

(a) The restriction of $\exp_p$ *to* T_0 *is a diffeomorphism between* T_0 *and* $\mathcal{S}$*. Its inverse* $\exp_p^{-1}\colon \mathcal{S} \to T_0$ *is given by*

$$\exp_p^{-1}(q) = \Pi_0 \log \tfrac{q}{p}. \tag{2.97}$$

(b) The differential of $\exp_p$ *at* $v \in \mathbb{R}^n$ *is given by*

$$\mathrm{d}\exp_p(v)[u] = R_{\exp_p(v)}[u] \qquad \text{for all } u \in \mathbb{R}^n. \tag{2.98}$$

(c) If $\mathbb{R}^n$ *is viewed as an Abelian group, then* $\exp\colon \mathbb{R}^n \times \mathcal{S} \to \mathcal{S}$ *given by* $(v, p) \mapsto \exp_p(v)$ *defines a Lie-group action, i.e.*

$$\exp_p(u+v) = \exp_{\exp_p(v)}(u) \quad \text{and} \quad \exp_p(0) = p \qquad \text{for all } v, u \in \mathbb{R}^n \text{ and } p \in \mathcal{S}. \tag{2.99}$$

We extend the lifting map (2.90) and the replicator operator (2.91) to corresponding maps for the assignment manifold $\mathcal{W}$ by applying these maps row-wise, i.e.

$$\exp_W\colon \mathbb{R}^{m\times n} \to \mathcal{W}, \qquad \exp_W(V) \coloneqq \big(\exp_{W_1}(V_1), \ldots, \exp_{W_m}(V_m)\big), \tag{2.100a}$$

$$R_W\colon \mathbb{R}^{m\times n} \to \mathcal{T}_0, \qquad R_W[V] \coloneqq \big(R_{W_1}[V_1], \ldots, R_{W_m}[V_m]\big), \tag{2.100b}$$

for $W \in \mathcal{W}$. The statements of Lemma 2.2.1 also apply accordingly to these maps.

Likelihood Map and Similarity Map

We use the lifting map (2.100a) to define the components of the assignment flow. Starting from the *distance matrix*

$$D = \Big(D_{\mathcal{M}}(z^i, m^j)\Big)_{i\in I, j\in J} \in \mathbb{R}^{m\times n}_{\geq 0} \tag{2.101}$$

storing the dissimilarities between observations $\{z^i\}_{i\in I}$ (2.70) and prototypes $\{m^j\}_{j\in J}$ (2.74) with respect to a dissimilarity function $D_{\mathcal{M}}$, we define the *likelihood map* $L\colon \mathcal{W} \to \mathcal{W}$ by lifting the distance matrix onto the assignment manifold, i.e.

$$L(W) \coloneqq \exp_W\big(-\tfrac{1}{\rho}D\big) \in \mathcal{W}. \tag{2.102}$$

The scaling parameter $\rho > 0$ is used for normalizing the apriori unknown scale of the components of D. The rows $D_i \in \mathbb{R}^n_{\geq 0}$ of the distance matrix D are called *distance vectors*. Similarly, *likelihood vectors* are the rows $L_i(W) \in \mathcal{S}$ of the likelihood matrix $L(W) \in \mathcal{W}$.

Next we define the *geometric mean* of assignment vectors

$$G_i^\omega(W) := \operatorname{Exp}_{W_i}\left(\sum_{k\in\mathcal{N}_i} \omega_{ik} \operatorname{Exp}_{W_i}^{-1}(W_k)\right) \tag{2.103a}$$

$$= \exp_{C_\mathcal{S}}\left(\sum_{k\in\mathcal{N}_i} \omega_{ik} \exp_{C_\mathcal{S}}^{-1}(W_k)\right) \in \mathcal{S}, \qquad i \in I, \tag{2.103b}$$

with respect to the neighborhoods (2.71) and weights (2.72). This map is an approximation of the weighted Riemannian mean of the assignment vectors $\{W_k : k \in \mathcal{N}_i\} \subset \mathcal{S}$ with respect to the weights $(\omega_{ik})_{k\in\mathcal{N}_i} \in \Delta_{|\mathcal{N}_i|}$ – cf. Section 2.1.6. It is a closed-form solution of the equation (2.68), where the Riemannian exponential map is replaced by the exponential map (2.94). The map (2.103) defines a smoothing operator $G^\omega : \mathcal{W} \to \mathcal{W}$. Applying this map to likelihood vectors yields the *similarity map*

$$S : \mathcal{W} \to \mathcal{W}, \qquad S(W) = G^\omega(L(W)) \tag{2.104}$$

with *similarity vectors* $S_i(W) = G_i^\omega(L(W)) \in \mathcal{S}$, $i \in I$. With regard to the data labeling problem, the likelihood map (2.102) can be interpreted as a data fidelity term, while the smoothing map (2.103) serves as regularization.

Assignment Flow

The *assignment flow* is induced by the system of nonlinear ODEs

$$\dot{W} = R_W[S(W)], \qquad W(0) = C_\mathcal{W}, \tag{2.105}$$

which is a coupled system of replicator equations

$$\dot{W}_i = R_{W_i}[S_i(W)] = W_i\big(S_i(W) - \langle W_i, S_i(W)\rangle \mathbb{1}_n\big), \qquad W_i(0) = C_\mathcal{S} \qquad \forall i \in I. \tag{2.106}$$

A labeling, i.e. an integral assignment, is obtained from the flow (2.105) by a trivial rounding of $W(T) \in \mathcal{W}$ for a sufficiently large $T > 0$. It was conjectured [ÅPSS17, Conjecture 1] that, for data in 'general position' as they are typically observed in real scenarios, the assignment flow $W(t)$ converges to an integral assignment $W^* \in \mathcal{W}^*$ as $t \to \infty$. We will confirm this conjecture in Chapter 3 under suitable assumptions on the weights $\{\omega_{ik} : k \in \mathcal{N}_i\}$. For this, we will use a reparameterization of the assignment flow (2.105) which was recently proposed in [SS20]. By collecting the weights (2.72) into a matrix $\Omega = (\omega_{ik})_{i,k\in I} \in \mathbb{R}^{m\times m}$, where we set $\omega_{ik} = 0$ for $k \notin \mathcal{N}_i$, the assignment flow can be parameterized as follows.

Proposition 2.2.2 (S-parameterization [SS20, Proposition 3.6])**.** *The assignment flow* (2.105) *is equivalent to the system*

$$\dot{S} = R_S[\Omega S], \qquad S(0) = \exp_{C_\mathcal{W}}\big(-\tfrac{1}{\rho}\Omega D\big), \tag{2.107a}$$

$$\dot{W} = R_W[S], \qquad W(0) = C_\mathcal{W}. \tag{2.107b}$$

Remark 2.2.3. The assignment flow was motivated in [ÅPSS17] as the Riemannian gradient ascent flow of the potential

$$J(W) = \langle W, S(W) \rangle. \tag{2.108}$$

The corresponding Riemannian gradient is given by $\operatorname{grad} J(W) = R_W[\nabla J(W)]$ – cf. equation (2.52). The approximation $\nabla J(W) \approx S(W)$ of the Euclidean gradient $\nabla J(W)$ resulted in the assignment flow (2.105). We note that it was shown in [SS20, Theorem 3.5] that there exists no potential $J \in \mathrm{C}^1(\mathcal{W}, \mathbb{R})$ such that $\operatorname{grad} J(W) = R_W[S(W)]$. In contrast, the S-flow (2.107a) possesses a potential under a symmetry assumption on Ω.

Numerical Integration

We adopt the idea from [Sav+17, ZSPS20] to parameterize the assignment flow by $W(t) = \exp_{C_\mathcal{W}}(V(t))$, where $V(t)$ is a flow on the tangent space $\mathcal{T}_0$ solving

$$\dot{V} = \Pi_0[S(W)], \quad W = \exp_{C_\mathcal{W}}(V), \quad V(0) = 0. \tag{2.109}$$

This reparameterization can be interpreted as Legendre transform coordinates – cf. equation (2.54). Its validity can be easily verified using Lemma 2.2.1(b):

$$\dot{W} = \frac{\mathrm{d}}{\mathrm{d}t} \exp_{C_\mathcal{W}}(V) \overset{(2.98)}{=} R_W[\dot{V}] = R_W \circ \Pi_0[S(W)] \overset{(2.93)}{=} R_W[S(W)], \tag{2.110a}$$

$$W(0) = \exp_{C_\mathcal{W}}(V(0)) = \exp_{C_\mathcal{W}}(0) \overset{(2.99)}{=} C_\mathcal{W}. \tag{2.110b}$$

Since the flow (2.109) is a flow in the vector space $\mathcal{T}_0$, any classical Runge–Kutta scheme can be applied for numerical integration of (2.109). Applying the explicit Euler scheme with step size $h > 0$ and using $W = \exp_{C_\mathcal{W}}(V)$ as well as Lemma 2.2.1(c) results in the following numerical scheme for the assignment flow [Sav+17]:

$$W^{(t+1)} = \exp_{W^{(t)}}\big(hS(W^{(t)})\big), \qquad W^{(0)} = C_\mathcal{W}, \qquad t \in \mathbb{N}_0. \tag{2.111}$$

For an overview of different numerical integration schemes for the assignment flow we refer to [ZSPS20].

2.3 Dynamical Systems and Stability

We give an overview of the theory of dynamical systems with a focus on stability of their equilibria. The statements from this section will be used to investigate the assignment flow (2.107) and its discretization (2.111) in Chapter 3.

Informally, a dynamical system on a set $\mathcal{M}$ is a mapping $\phi\colon T \times \mathcal{M} \to \mathcal{M}$ fulfilling the group law

$$\phi(0, x) = x, \qquad \phi(s + t, x) = \phi\big(s, \phi(t, x)\big) \qquad \text{for all } x \in \mathcal{M} \text{ and } s, t \in T. \tag{2.112}$$

Here, the time set T is either $\mathbb{R}$, $\mathbb{R}_{\geq 0}$, $\mathbb{Z}$ or $\mathbb{N}_0$. A point $x^* \in \mathcal{M}$ is called *equilibrium* of the dynamical system if $\phi(t, x^*) = x^*$ for all $t \in T$.

In case of $T = \mathbb{R}$ or $T = \mathbb{R}_{\geq 0}$, the system is called continuous-time dynamical system or flow – cf. Section 2.1.2. For $T = \mathbb{Z}$ or $T = \mathbb{N}_0$ we have a discrete-time dynamical system.

We regard continuous-time and discrete-time dynamical systems separately. Firstly, we discuss continuous-time dynamical systems described by autonomous differential equations, i.e., $\phi(t, x_0) = x(t)$ is the trajectory of the system

$$\dot{x}(t) = F(x(t)), \qquad x(0) = x_0. \tag{2.113}$$

An equilibrium x^* is then a zero of the vector field F, i.e. $F(x^*) = 0$.

Secondly, we consider discrete-time dynamical systems generated by maps $F\colon \mathcal{M} \to \mathcal{M}$, i.e.

$$\phi(t, x) = (\underbrace{F \circ \cdots \circ F}_{t \text{ times}})(x), \qquad t \in \mathbb{N}_0. \tag{2.114}$$

In this case the equilibria coincide with the fixed points of the map F.

2.3.1 Continuous-Time Dynamical Systems from Differential Equations

We summarize some definitions and important statements which can be found in various textbooks on dynamical systems, e.g. [Tes12, SC16, Per01, GH02]. This theory concerns dynamical systems generated by vector fields $F \in \mathrm{C}^1(U, \mathbb{R}^n)$ on an open subset $U \subseteq \mathbb{R}^n$. We start with the definition of stability of an equilibrium.

Definition 2.3.1 (stability)**.** Let x^* be an equilibrium point of the system $\dot{x}(t) = F(x(t))$, i.e. $F(x^*) = 0$.

(a) x^* is called *stable*, if for every open neighborhood $U_0 \ni x^*$ there exists a smaller neighborhood U_1 such that if $x(t_0) \in U_1$ for some $t_0 \in \mathbb{R}$, then $x(t) \in U_0$ for all $t \geq t_0$. If x^* is not stable, it is called *unstable*.

(b) x^* is called *attracting*, if there exists an open neighborhood $U^* \ni x^*$, such that $x(t) \to x^*$ as $t \to \infty$ if $x(t_0) \in U^*$ for some $t_0 \in \mathbb{R}$.

(c) x^* is called *asymptotically stable*, if it is stable and attracting.

(d) x^* is called *exponentially stable*, if there exists constants $\alpha, \beta, \delta > 0$ such that

$$\|x(t) - x^*\| \leq \alpha e^{-\beta(t-t_0)} \quad \text{if } \|x(t_0) - x^*\| < \delta \text{ for some } t_0 \in \mathbb{R}. \tag{2.115}$$

We note that exponential stability implies asymptotic stability. The stability of an equilibrium point x^* is particularly relevant for the following reason. The *region of attraction of* x^* is the set of all initial points x_0 for which the trajectory of $\dot{x}(t) = F(x(t))$ starting at x_0 converges to x^* as $t \to \infty$. If x^* is attracting then its region of attraction is an open neighborhood of x^* [BS02, Lemma 2.6.11]. This means that the convergence towards x^* is stable under small perturbations of the initial point x_0. In contrast, a small perturbation of x_0 could prevent convergence towards an unstable equilibrium as discussed in the context of invariant manifolds below.

An important tool for stability analysis is the linearization of the system $\dot{x}(t) = F(x(t))$ at an equilibrium point. The next theorem therefore handles the linear case and summarizes statements from [Per01, Section 1.4, Section 1.9]. It also provides the first example of invariant manifolds. In this context, a set $\mathcal{M}$ is invariant under a flow if any trajectory of the flow starting in $\mathcal{M}$ also remains in $\mathcal{M}$ for all time points.

Theorem 2.3.2. *Let $A \in \mathbb{R}^{n\times n}$. We consider the linear system*

$$\dot{x}(t) = Ax(t), \quad x(t_0) = x_0. \tag{2.116}$$

(a) The system has the unique solution $x(t) = e^{A(t-t_0)}x_0$ with $e^A = \exp(A)$ denoting the matrix exponential.

(b) Let $\{v_j\} \subset \mathbb{C}^n$ be the generalized eigenvectors of A with corresponding eigenvalues $\{\lambda_j\} \subset \mathbb{C}$, i.e. $v_j \in \ker(\lambda_j I_n - A)^k$ for some $k \in \mathbb{N}$. Then the subspaces

$$E_\mathrm{s} \coloneqq \operatorname{span}\big\{\operatorname{Re}(v_j), \operatorname{Im}(v_j) \colon \operatorname{Re}(\lambda_j) < 0\big\}, \tag{2.117a}$$
$$E_\mathrm{u} \coloneqq \operatorname{span}\big\{\operatorname{Re}(v_j), \operatorname{Im}(v_j) \colon \operatorname{Re}(\lambda_j) > 0\big\}, \tag{2.117b}$$
$$E_\mathrm{c} \coloneqq \operatorname{span}\big\{\operatorname{Re}(v_j), \operatorname{Im}(v_j) \colon \operatorname{Re}(\lambda_j) = 0\big\} \tag{2.117c}$$

are invariant under the flow $\dot{x} = Ax$ and we have

$$\mathbb{R}^n = E_\mathrm{s} \oplus E_\mathrm{u} \oplus E_\mathrm{c}. \tag{2.118}$$

(c) We have $\lim_{t\to\infty} e^{A(t-t_0)}x_0 = 0$ if and only if $x_0 \in E_\mathrm{s}$. In this case, there exist $\alpha, \beta > 0$, such that

$$\left\|e^{A(t-t_0)}x_0\right\| \le \alpha e^{-\beta(t-t_0)}\|x_0\|. \tag{2.119}$$

(d) If $x_0 \in E_\mathrm{u}$, then there exist $\alpha, \beta > 0$ with

$$\left\|e^{A(t-t_0)}x_0\right\| \ge \alpha e^{\beta(t-t_0)}\|x_0\|. \tag{2.120}$$

Especially, $\lim_{t\to\infty} \left\|e^{A(t-t_0)}x_0\right\| = \infty$ if $x_0 \neq 0$.

Definition 2.3.3. The invariant subspaces (2.117) are called the *stable*, *unstable* and *center subspace* of the linear system $\dot{x} = Ax$.

For a hyperbolic equilibrium point as defined below, the following theorem provides a link between a (nonlinear) system and its linearization at that equilibrium. Informally, the theorem states that a system $\dot{x} = F(x)$ behaves locally like its linearization $\dot{y} = \frac{\partial F}{\partial x}(x^*)y$ near a hyperbolic equilibrium x^*.

Definition 2.3.4 (hyperbolic)**.** An equilibrium point x^* of the system $\dot{x}(t) = F(x(t))$ is called *hyperbolic* if all eigenvalues of the Jacobian matrix $\frac{\partial F}{\partial x}(x^*)$ have nonvanishing real part.

Theorem 2.3.5 (Hartman-Grobman Theorem)**.** *Let $F \in \mathrm{C}^1(U, \mathbb{R}^n)$ with $U \subseteq \mathbb{R}^n$ being an open subset. Suppose x^* is a hyperbolic equilibrium for $\dot{x}(t) = F(x(t))$. Let $A = \frac{\partial F}{\partial x}(x^*)$. Then there exist an open neighborhood $\tilde{U} \ni x^*$, an open neighborhood $V \ni 0$, and a homeomorphism $\Psi \colon \tilde{U} \to V$ with $\Psi(x^*) = 0$ such that*

$$\Psi(x(t)) = e^{A(t-t_0)}\Psi(x(t_0)) \tag{2.121}$$

for all initial values $x(t_0) \in \tilde{U}$ of $\dot{x} = F(x)$ and all times t such that $x(t)$ is defined and belongs to $\tilde{U}$. Furthermore, the maps Ψ and Ψ^{-1} are α-Hölder continuous with some $\alpha = \alpha(A) \in (0, 1)$.

The proof of the Hartman-Grobman theorem can be found for example in [Per01, Section 2.8]. For the Hölder continuity statement we refer to [BR09]. A consequence of the α-Hölder continuity is that an exponential convergence rate of the linearized system implies an exponential convergence rate of the original system, i.e., if $\|\Psi(x(t))\| \leq \gamma e^{-\beta t}$ with $\gamma, \beta > 0$ then

$$\|x(t) - x^*\| = \|\Psi^{-1}(\Psi(x(t))) - \Psi^{-1}(0)\| \leq C\|\Psi(x(t))\|^{\alpha} \leq C\gamma^{\alpha} e^{-\alpha\beta t}. \tag{2.122}$$

The following theorem gives further insight based on the linearization of a system. It states the existence of invariant manifolds that help in the analysis of stability and convergence.

Theorem 2.3.6 (The Center Manifold Theorem)**.** *Let $F \in \mathrm{C}^r(U, \mathbb{R}^n)$ where $U \subseteq \mathbb{R}^n$ is an open subset and $r \geq 1$. Suppose that $F(x^*) = 0$ and the Jacobian matrix $\frac{\partial F}{\partial x}(x^*)$ has n_s eigenvalues with negative real part, n_u eigenvalues with positive real part, and $n_\mathrm{c} = n - n_\mathrm{s} - n_\mathrm{u}$ eigenvalues with zero real part. Further, let E_c, E_s and E_u be the center, stable and unstable subspaces of the linearized system $\dot{y}(t) = \frac{\partial F}{\partial x}(x^*)y(t)$. Then*

(i) there exists an n_c-dimensional center manifold $\mathcal{M}_\mathrm{c}(x^*)$ of class C^{r-1} tangent to the center subspace E_c at x^*,

(ii) there exists an n_s-dimensional stable manifold $\mathcal{M}_\mathrm{s}(x^*)$ of class C^r tangent to the stable subspace E_s at x^*,

(iii) there exists an n_u-dimensional unstable manifold $\mathcal{M}_\mathrm{u}(x^*)$ of class C^r tangent to the unstable subspace E_u at x^*,

such that $\mathcal{M}_\mathrm{c}(x^*)$, $\mathcal{M}_\mathrm{s}(x^*)$ and $\mathcal{M}_\mathrm{u}(x^*)$ are invariant under the flow $\dot{x}(t) = F(x(t))$. While the stable and unstable manifold are unique, the center manifold need not be.

We refer to [GH02, Theorem 3.2.1] or [Per01, Section 2.7] for this theorem. We note that if the flow starts at a point on the stable manifold $\mathcal{M}_\mathrm{s}(x^*)$ then it converges to x^* with an exponential rate as $t \to \infty$. There are two other invariant manifolds which are useful for convergence analysis. Unfortunately, they are not listed in the textbooks mentioned above. These manifolds were introduced in [Kel66, Kel67]. We present them in the following theorem.

Theorem 2.3.7. *In addition to the statements of Theorem 2.3.6*

(i) there exists a $(n_\mathrm{c} + n_\mathrm{s})$-dimensional center-stable manifold $\mathcal{M}_\mathrm{cs}(x^*)$ of class C^{r-1} tangent to the subspace $E_\mathrm{c} \oplus E_\mathrm{s}$ at x^*,

(ii) there exists a $(n_\mathrm{c} + n_\mathrm{u})$-dimensional center-unstable manifold $\mathcal{M}_\mathrm{cu}(x^*)$ of class C^{r-1} tangent to the subspace $E_\mathrm{c} \oplus E_\mathrm{u}$ at x^*,

such that the manifolds $\mathcal{M}_\mathrm{cs}(x^*)$ and $\mathcal{M}_\mathrm{cu}(x^*)$ are invariant under the flow $\dot{x}(t) = F(x(t))$. As it is the case for the center manifold, the center-stable manifold and the center-unstable manifold are not unique in general.

We note that in case of non-isolated equilibria there exists a center-stable manifold $\mathcal{M}_\mathrm{cs}(K)$, where K is a compact subset of a manifold of equilibria on which the dimensions n_s, n_c and n_u are constant [Fen79, Theorem 9.1]. This manifold is a center-stable manifold for all $x^* \in K$. The same applies to the center-unstable manifold $\mathcal{M}_\mathrm{cu}(K)$ and the center manifold $\mathcal{M}_\mathrm{c}(K)$.

The manifolds $\mathcal{M}_c(K)$, $\mathcal{M}_{cs}(K)$ and $\mathcal{M}_{cu}(K)$ might not be unique in general but they possess the following weak uniqueness property. If an integral curve $x(t)$ of the vector field F stays within a sufficiently small neighborhood of K for $t \geq 0$, then this integral curve lies in any center-stable manifold $\mathcal{M}_{cs}(K)$. Especially, every $\mathcal{M}_{cs}(K)$ (locally) contains all trajectories converging to some $x^* \in K$ as $t \to \infty$. The same is true of the center-unstable manifold $\mathcal{M}_{cu}(K)$ with integral curves staying in a neighborhood of K for $t \leq 0$. Similarly, any center manifold $\mathcal{M}_c(K)$ contains all integral curves remaining in a small neighborhood of K for all $t \in \mathbb{R}$. This includes cycles near an equilibrium x^*.

The following example demonstrates the non-uniqueness of the center manifold. The equilibrium $x^* = 0$ of the vector field

$$\dot{x}_1 = x_1^2, \qquad \dot{x}_2 = x_2 \tag{2.123}$$

has a unique one-dimensional unstable manifold and a (non-unique) one-dimensional center manifold which are illustrated in Figure 2.1. Since the center manifold and the center-stable

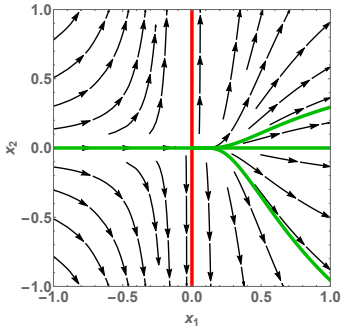


Figure 2.1: (Non-)Uniqueness of the center manifold. The plot shows the phase portrait of the vector field (2.123) as well as the invariant manifolds for the equilibrium $x^* = 0$. The unstable manifold is shown as red curve, while any green curve represents a center manifold. Note that all center manifolds coincide in the left half plane.

manifold coincide in this case, the corresponding uniqueness property applies here, i.e., the center manifold contains all trajectories converging to $x^* = 0$. This can be seen in the left half plane in Figure 2.1 where all center manifolds coincide.

As an application of the center-stable manifold, we discuss the case when an equilibrium point x^* possesses a nontrivial unstable subspace, i.e. $n_u \geq 1$. In this case the center-stable manifold $\mathcal{M}_{cs}(x^*)$ is low-dimensional, i.e., a small perturbation of a point in $\mathcal{M}_{cs}(x^*)$ might throw it out of the manifold. Since $\mathcal{M}_{cs}(x^*)$ contains all trajectories converging to x^*, this means a small perturbation of the initial point x_0 of a trajectory can prevent the convergence towards the (unstable) equilibrium x^*.

As direct consequence of the Center Manifold Theorem we get the following statement.

Theorem 2.3.8. *Let x^* be an equilibrium point of the system $\dot{x}(t) = F(x(t))$ with $F \in C^1(U, \mathbb{R}^n)$.*

(a) If all eigenvalues of the Jacobian matrix $\frac{\partial F}{\partial x}(x^*)$ have negative real part, then x^* is exponentially stable.

(b) If the Jacobian matrix $\frac{\partial F}{\partial x}(x^*)$ has an eigenvalue with positive real part, then x^* is unstable.

We note that the stability criteria presented in the above theorem concern flows $\dot{x}(t) = F(x(t))$ on an open subset of $\mathbb{R}^n$. Since we want to use these statements for flows on $\overline{\mathcal{W}}$, we need a few more arguments. In [SC16, Section 6.8.4], a direct proof of Theorem 2.3.8(b) is sketched. This proof gives more insight which becomes handy for our analysis. Therefore, we summarize the core arguments of this direct proof in the following proposition.

Proposition 2.3.9. *Let x^* be an equilibrium point of $\dot{x}(t) = F(x(t))$ with $F \in \mathrm{C}^1(U, \mathbb{R}^n)$. Then*

(a) There exist a sufficiently small $\varepsilon_1 > 0$ and a (real) similarity transform

$$V^{-1} \frac{\partial F}{\partial x}(x^*) V = \begin{pmatrix} A_{sc} & 0 \\ 0 & A_u \end{pmatrix} = A, \tag{2.124}$$

such that

(i) $\mathrm{Re}(\lambda) \leq 0$ for all eigenvalues λ of A_{sc},

(ii) $\mathrm{Re}(\lambda) > 0$ for all eigenvalues λ of A_u,

(iii) $\langle y_{sc}, A_{sc} y_{sc} \rangle \leq \frac{\varepsilon_1}{4} \|y_{sc}\|^2$,

(iv) $\langle y_u, A_u y_u \rangle \geq \varepsilon_1 \|y_u\|^2$.

(b) Suppose $\frac{\partial F}{\partial x}(x^*)$ has at least one eigenvalue λ with $\mathrm{Re}(\lambda) > 0$. Considering an affine coordinate transform $y = V^{-1}(x - x^*)$ with $V \in \mathrm{GL}_n(\mathbb{R})$ from (a), the resulting flow

$$\dot{y} = G(y) = V^{-1} F(Vy + x^*), \tag{2.125}$$

which has the equilibrium $y^* = 0$ with $\frac{\partial G}{\partial y}(0) = V^{-1} \frac{\partial F}{\partial x}(x^*) V = A$, has the following property. There exist $\eta > 0$, $\delta > 0$ and $\varepsilon > 0$, such that if the flow starts at some point in the open truncated cone

$$U_{\eta,\delta} = \left\{ y = \left(\begin{smallmatrix} y_{sc} \\ y_u \end{smallmatrix}\right) \in \mathbb{R}^n \colon \|y_{sc}\|^2 < \eta \|y_u\|^2,\ \|y\| < \delta \right\} \subset B_\delta(0) = \left\{ y \in \mathbb{R}^n \colon \|y\| < \delta \right\}, \tag{2.126}$$

then the solution will not cross the conical portion of $\partial U_{\eta,\delta}$, i.e.

$$\left\{ y \in \mathbb{R}^n \colon \|y_{sc}\|^2 = \eta \|y_u\|^2,\ \|y\| < \delta \right\}, \tag{2.127}$$

and it fulfills $\|y(t)\| \geq \|y(0)\| e^{\varepsilon t}$ as long as $y(t) \in U_{\eta,\delta}$, i.e., $y(t)$ leaves the ball $B_\delta(0)$ at some time point. Especially, the equilibrium $y^* = 0$ is unstable. This property is accordingly transferred to the equilibrium x^* of $\dot{x}(t) = F(x(t))$ using $x(t) = Vy(t) + x^*$.

The proof of this proposition can be found in the appendix. We note that if $\frac{\partial F}{\partial x}(x^*)$ is diagonalizable with real eigenvalues then the similarity transform in Proposition 2.3.9(a) is just the diagonalization. In general, if $v \in \mathbb{C}^n$ is an eigenvector of $\frac{\partial F}{\partial x}(x^*)$ corresponding to an eigenvalue $\lambda \in \mathbb{C}$ with $\mathrm{Re}(\lambda) > 0$ and $V \in \mathrm{GL}_n(\mathbb{R})$ is given by Proposition 2.3.9(a), then $V^{-1} \mathrm{Re}(v) = \left(\begin{smallmatrix} 0 \\ y_u \end{smallmatrix}\right)$ and $V^{-1} \mathrm{Im}(v) = \left(\begin{smallmatrix} 0 \\ \tilde{y}_u \end{smallmatrix}\right)$.

Last but not least we mention a theorem about the asymptotic behavior of a flow. It states, under certain conditions, the convergence towards an invariant set which is not necessarily a singleton.

Theorem 2.3.10 (LaSalle Invariance Principle)**.** *Let $F\colon U \to \mathbb{R}^n$ be a vector field on an open subset $U \subseteq \mathbb{R}^n$ defining the system*

$$\dot{x}(t) = F(x(t)), \quad x(0) = x_0 \in U. \tag{2.128}$$

Let $U_c \subset U$ be a compact positively invariant set with respect to (2.128) *and let $V \in \mathrm{C}^1(U, \mathbb{R})$ be a Lyapunov function on U_c, i.e., $\dot{V}(x) \coloneqq \langle \nabla V(x), F(x) \rangle \geq 0$ on U_c. Further, let $\mathcal{I}$ be the largest invariant set contained in*

$$\left\{ x \in U_c \colon \dot{V}(x) = 0 \right\}. \tag{2.129}$$

If $x_0 \in U_c$, then $x(t) \to \mathcal{I}$ as $t \to \infty$.

References for this theorem are [HC08, Theorem 3.3] or [Wig03, Theorem 8.3.1]. We only need the statement that the limit points of (2.128) are contained in (2.129) if $x_0 \in U_c$. We will apply this theorem to the S-flow (2.107a) with the compact invariant set $\overline{\mathcal{W}}$ in Chapter 3.

2.3.2 Discrete-Time Dynamical Systems from Maps

We present definitions and statements regarding the stability of equilibria of discrete-time dynamical systems $x^{(t+1)} = F(x^{(t)})$, where $F\colon U \to U$ is a C^1 map on an open subset $U \subseteq \mathbb{R}^n$. As reference we consider the textbooks [Gal07, Ela05].

We have the same notions of stability as in the continuous-time case with corresponding adaptation of the definition.

Definition 2.3.11 (cf. [Ela05, Definition 4.2])**.** Let $x^* \in \mathbb{R}^n$ be an equilibrium point of the system $x^{(t+1)} = F(x^{(t)})$, i.e. $F(x^*) = x^*$.

(a) x^* is called *stable*, if for every neighborhood $U_0 \ni x^*$ there exists a smaller neighborhood U_1 such that if $x^{(t_0)} \in U_1$ for some $t_0 \in \mathbb{Z}$, then $x^{(t)} \in U_0$ for all $t \geq t_0$. If x^* is not stable, it is called *unstable*.

(b) x^* is called *attracting*, if there exists a neighborhood $U^* \ni x^*$, such that $\lim_{t\to\infty} x^{(t)} = x^*$ if $x^{(t_0)} \in U^*$ for some $t_0 \in \mathbb{Z}$.

(c) x^* is called *asymptotically stable*, if it is stable and attracting.

(d) x^* is called *exponentially stable*, if there exists constants $\alpha, \delta > 0$ and $\gamma \in (0, 1)$ such that

$$\|x^{(t)} - x^*\| \leq \alpha \gamma^{t-t_0} \quad \text{if } \|x^{(t_0)} - x^*\| < \delta \text{ for some } t_0 \in \mathbb{Z}. \tag{2.130}$$

As in the continuous-time case, exponential stability implies asymptotic stability. Analogously to Theorem 2.3.8, the stability of an equilibrium point x^* of the system $x^{(t+1)} = F(x^{(t)})$ can be determined via the spectrum of the Jacobian $\frac{\partial F}{\partial x}(x^*)$. However, the difference is that the eigenvalue conditions $\mathrm{Re}(\lambda) < 0$ and $\mathrm{Re}(\lambda) > 0$ are replaced by $|\lambda| < 1$ and $|\lambda| > 1$, respectively.

Theorem 2.3.12. *Let x^* be an equilibrium point of the system $x^{(t+1)} = F(x^{(t)})$.*

(a) If all eigenvalues λ of the Jacobian matrix $\frac{\partial F}{\partial x}(x^*)$ have absolute value less than 1, i.e. $|\lambda| < 1$, then x^* is exponentially stable.

(b) If the Jacobian matrix $\frac{\partial F}{\partial x}(x^*)$ has an eigenvalue λ with $|\lambda| > 1$, then x^* is unstable.

The statements of the theorem can be found in [Ela05, Corollary 4.34, Theorem 4.38]. Again, the theorem concerns dynamical systems on an open subset of $\mathbb{R}^n$. Since we will regard dynamical systems on $\overline{\mathcal{W}}$, we worked out a discrete-time version of Proposition 2.3.9. The proof of the following proposition is in the appendix.

Proposition 2.3.13. *Let x^* be an equilibrium point of $x^{(t+1)} = F(x^{(t)})$. Then*

(a) There exist a sufficiently small $\varepsilon_1 > 0$ and a (real) similarity transform

$$V^{-1}\frac{\partial F}{\partial x}(x^*)V = \begin{pmatrix} A_{sc} & 0 \\ 0 & A_u \end{pmatrix} = A, \tag{2.131}$$

such that

(i) $|\lambda| \leq 1$ for all eigenvalues λ of A_{sc},

(ii) $|\lambda| > 1$ for all eigenvalues λ of A_u,

(iii) $\|A_{sc}y_{sc}\|^2 \leq (1 + \frac{\varepsilon_1}{4})^2\|y_{sc}\|^2$,

(iv) $\|A_u y_u\|^2 \geq (1 + \varepsilon_1)^2\|y_u\|^2$.

(b) Suppose $\frac{\partial F}{\partial x}(x^*)$ has at least one eigenvalue λ with $|\lambda| > 1$. Considering an affine coordinate transform $y = V^{-1}(x - x^*)$ with $V \in \mathrm{GL}_n(\mathbb{R})$ from (a), the resulting map

$$G(y) = V^{-1}\big(F(Vy + x^*) - x^*\big), \tag{2.132}$$

which has the fixed point $y^* = 0$ with $\frac{\partial G}{\partial y}(0) = V^{-1}\frac{\partial F}{\partial x}(x^*)V = A$, has the following property. There exist $\eta > 0$ and $\delta > 0$, such that if $y^{(0)}$ is a point in the open truncated cone $U_{\eta,\delta} = C_\eta \cap B_\delta(0)$ with

$$C_\eta = \left\{y = \left(\begin{smallmatrix} y_{sc} \\ y_u \end{smallmatrix}\right) \in \mathbb{R}^n \colon \|y_{sc}\|^2 < \eta\|y_u\|^2\right\}, \tag{2.133a}$$

$$B_\delta(0) = \big\{y \in \mathbb{R}^n \colon \|y\| < \delta\big\}, \tag{2.133b}$$

then we have

(i) $\big\|G(y^{(0)})\big\|^2 \geq (1 + \varepsilon_1)\big\|y^{(0)}\big\|^2$,

(ii) $G(y^{(0)}) \in C_\eta$.

Hence, the sequence $y^{(t+1)} = G(y^{(t)})$ eventually leaves the ball $B_\delta(0)$ at some time point. Especially, the equilibrium $y^* = 0$ is unstable. This property is accordingly transferred to the equilibrium point x^* of $F(x)$ using $x^{(t)} = Vy^{(t)} + x^*$.

We note that for some other theorems presented for continuous-time dynamical systems there exist corresponding versions for discrete-time dynamics. Since we will not use them, we only mention them. In the discrete-time case, the stable subspace E_s of $\frac{\partial F}{\partial x}(x^*)$ corresponds to the eigenvalues λ with $|\lambda| < 1$. Similarly, the center subspace E_c and the unstable subspaces E_u correspond to $|\lambda| = 1$ and $|\lambda| > 1$, respectively [Gal07, Definition 4.3]. An equilibrium x^* is hyperbolic if $\frac{\partial F}{\partial x}(x^*)$ has no eigenvalue λ with $|\lambda| = 1$ [Gal07, Definition 4.5]. Accordingly, discrete-time versions of the Center Manifold Theorem [Wig03, Section 3.3] and the Hartman-Grobman Theorem [Ela05, Theorem D.2] exist. However, both theorems assume F to be a diffeomorphism.

3 Asymptotic Behavior of the Assignment Flow

3.1 Introduction and Overview

Motivation

The assignment flow, which we have summarized in Section 2.2, was introduced in [ÅPSS17] as a geometric approach to the image labeling problem. It was conjectured [ÅPSS17, Conjecture 1] that, for data in 'general position' as they are typically observed in real scenarios (e.g. no symmetry due to additive noise), the assignment flow converges to an integral assignment. This has been empirically verified in numerous experiments. Mathematically, however, convergence has hardly been studied. In this chapter we want to remedy this shortcoming and perform a detailed mathematical analysis of the asymptotic behavior of assignment flow. We prove the conjecture under a positivity and symmetry assumption on the parameters $\{\omega_{ik} : k \in \mathcal{N}_i\}$, $i \in I$. To this end we use the S-parameterization (2.107), which was introduced in [SS20]. This parameterization has the advantage that it is defined on an open neighborhood of $\overline{\mathcal{W}}$ containing the assignment manifold $\mathcal{W}$ (2.85). This enables the common theory on continuous-time dynamic systems to be used for analysis – cf. Section 2.3. Besides the continuous-time assignment flow, we also investigate the discrete scheme (2.111) introduced in [Sav+17]. For both cases we derive a criterion to verify convergence towards an exponentially stable equilibrium of the S-flow.

Organization

The detailed organization of this chapter is as follows. First, we recall the S-parameterization in Section 3.2 and establish the relation between the asymptotic behavior of the assignment flow and that of the S-flow, so that the rest of the chapter is mainly concerned with the latter. In Section 3.3 we prove under a symmetry assumption on the parameters $\Omega = (\omega_{ik})_{ik}$ that the S-flow always converges towards an equilibrium. We also show by an example that this convergence statement need not apply without symmetry. Afterwards we examine the stability of equilibria of the S-flow (Section 3.4). Under a positivity assumption on the diagonal of the parameter matrix Ω, we show that nonintegral equilibria are unstable and conclude that the set of initial points for which the S-flow converges towards those equilibria is negligible. Furthermore, in Section 3.4.2 we give an estimation of the basin of attraction for exponentially stable equilibria, which can be used as a convergence guarantee criterion. In Section 3.5 we investigate the discrete scheme (2.111). We give estimates on the approximation error and we show that our stability statements as well as the estimation for the basins of attraction from the continuous case also apply to the discrete scheme. We discuss in Section 3.6 a generalization of the S-flow, where not only the rows of the matrix $S \in \mathbb{R}^{m \times n}$ are averaged, but general linear

maps on $\mathbb{R}^{m\times n}$ are considered. For this generalization, the previous convergence and stability statements are transferred. Finally, we summarize the results in Section 3.7.

Notation

In this chapter we use the same notation as in Section 2.2, i.e., I is the set of data indices and $m = |I|$ is the number of data points. J is the set of label indices and $n = |J|$ is the number of labels. $\mathcal{W} \subset \mathbb{R}^{m\times n}$ is the assignment manifold (2.85) and $\mathcal{T}_0$ is its tangent space. $\mathcal{W}^* = \overline{\mathcal{W}} \cap \{0,1\}^{m\times n}$ denotes the set of integral assignments. $\exp_W$ is the lifting map (2.100a). For a matrix $S \in \mathbb{R}^{m\times n}$, the i-th row is denoted by S_i, i.e., $(\Omega S)_i$ is the i-th row of the matrix product ΩS with a matrix $\Omega \in \mathbb{R}^{m\times m}$.

Furthermore, the k-th standard basis vector is denoted by $e_k = (0,\dots,1,\dots,0)^\top \in \mathbb{R}^n$. The dimension of this vector should be clear from the context, e.g. $e_i \in \mathbb{R}^{|I|}$ for a data index $i \in I$ and $e_j \in \mathbb{R}^{|J|}$ for a label index $j \in J$. The constant vector of ones is denoted by $\mathbb{1}_n = (1,\dots,1)^\top \in \mathbb{R}^n$ and $I_n \in \mathbb{R}^{n\times n}$ denotes the identity matrix. For a subset $A \subset [n]$ its indicator vector is denoted by $\mathbb{1}_A \in \{0,1\}^n$. The support of a vector $v \in \mathbb{R}^n$ is denoted by $\operatorname{supp} v = \{i \in [n] : v_i \neq 0\}$, i.e. $\operatorname{supp} W_i = \{j \in J : W_{ij} \neq 0\}$ for an assignment matrix $W \in \overline{\mathcal{W}}$. For two matrices of the same size, $A \odot B$ denotes the Hadamard (entrywise) product. For matrices $A \in \mathbb{R}^{m\times n}$, $B \in \mathbb{R}^{p\times q}$, the matrix $A \otimes B \in \mathbb{R}^{mp\times nq}$ denotes the Kronecker product of these matrices with submatrices $A_{ij}B \in \mathbb{R}^{p\times q}$, $i \in [m]$, $j \in [n]$ (cf. e.g. [Van00]). The latter two operators will be analogously applied to vectors, i.e. $v \otimes w \in \mathbb{R}^{mn}$ for vectors $v \in \mathbb{R}^m$, $w \in \mathbb{R}^n$.

3.2 The S-Parametrization and its Implication on the Assignment Flow

We recall the assignment flow, which was presented in Section 2.2. Given a distance matrix $D \in \mathbb{R}^{m\times n}_{\geq 0}$ and a weighted adjacency matrix $\Omega = (\omega_{ik})_{ik} \in \mathbb{R}^{m\times m}$, the assignment flow is defined by the system

$$\dot{W} = R_W[S(W)], \qquad W(0) = C_{\mathcal{W}} = \tfrac{1}{n}\mathbb{1}_m\mathbb{1}_n^\top, \tag{3.1}$$

where the similarity map $S(W) = G^\omega(L(W))$ is a concatenation of the likelihood map $L : \mathcal{W} \to \mathcal{W}$, $L(W) = \exp_W(-D)$ and a smoothing operator $G^\omega : \mathcal{W} \to \mathcal{W}$ (2.103) depending on the weights ω_{ik}. In [SS20] it was shown that the similarity matrix $S(W) \in \mathcal{W}$ considered as a function in time, i.e. $S(t) := S(W(t))$, is itself a solution of a system of ODEs and the assignment flow (3.1) can thus be reparameterized as follows.

Proposition 3.2.1 (S-parametrization [SS20, Proposition 3.6])**.** *The assignment flow* (3.1) *is equivalent to the system*

$$\dot{S} = R_S[\Omega S], \qquad S(0) = S_0 = \exp_{C_{\mathcal{W}}}(-\Omega D), \tag{3.2a}$$

$$\dot{W} = R_W[S], \qquad W(0) = C_{\mathcal{W}}. \tag{3.2b}$$

More precisely, $W(t)$, $t \geq 0$ solves (3.1) *if and only if it solves* (3.2)*.*

The flow induced by the equation (3.2a) is called the S-flow. Based on the S-parameterization we will analyze the asymptotic behavior of the assignment flow (3.1). The first implication from the S-parameterization is the following representation of the assignment flow.

Proposition 3.2.2. *Let $t \mapsto S(t)$ be the trajectory of the S-flow* (3.2a)*. Then the trajectory of the assignment flow* (3.2b) *is given by*

$$W(t) = \exp_{C_\mathcal{W}}\left(\int_0^t S(\tau)\,\mathrm{d}\tau\right) = \exp_{C_\mathcal{W}}\left(\int_0^t \Pi_0[S(\tau)]\,\mathrm{d}\tau\right), \tag{3.3}$$

where $\Pi_0 \colon \mathbb{R}^{m\times n} \to \mathcal{T}_0$ is the orthogonal projection onto the tangent space.

Proof. We set $I_S(t) = \int_0^t S(\tau)\mathrm{d}\tau$. Then $W(t) = \exp_{C_\mathcal{W}}(I_S(t))$ and Lemma 2.2.1 implies

$$\dot{W}(t) = \mathrm{d}\exp_{C_\mathcal{W}}(I_S(t))\left[\dot{I}_S(t)\right] \overset{(2.98)}{=} R_{\exp_{C_\mathcal{W}}(I_S(t))}\left[\dot{I}_S(t)\right] = R_{W(t)}[S(t)], \tag{3.4a}$$

$$W(0) = \exp_{C_\mathcal{W}}(0) \overset{(2.99)}{=} C_\mathcal{W}, \tag{3.4b}$$

i.e., equation (3.2b) is fulfilled. The second equation of (3.3) follows from (2.96) and (2.93). □

Before we will discuss the implication of the equation (3.3), we show the following proposition, which legitimates to regard $S(t)$ for all time points $t \in \mathbb{R}$ and therefore to inspect the limit $\lim_{t\to\infty} S(t)$.

Proposition 3.2.3 (global existence and uniqueness)**.** *The solutions $W(t)$, $S(t)$ to* (3.2) *are unique and exist for all $t \in \mathbb{R}$.*

Proof. The hyperplanes $\{S\colon \sum_j S_{ij} = 1\}$ for $i \in I$ and $\{S\colon S_{ij} = 0\}$ for $i \in I$, $j \in J$ are invariant under the flow (3.2a), due to the replicator operator R_S. Hence, $S(t)$ stays in $\overline{\mathcal{W}}$ for $S_0 \in \overline{\mathcal{W}}$. By Lemma 2.1.3 (or [Tes12, Corollary 2.16]), the trajectory $t \mapsto S(t)$ exists for all $t \in \mathbb{R}$. Equation (3.3) then implies the existence of $W(t)$ for all $t \in \mathbb{R}$. The uniqueness of the solutions follow by the local Lipschitz continuity of the right-hand side of (3.2a) and (3.2b), respectively. □

Remark 3.2.4. While the domain of the assigment flow (3.1) is limited to $\mathcal{W}$, the S-flow (3.2a) can be considered on $\overline{\mathcal{W}}$. Furthermore, the domain of the S-flow can be extended to an open set U with $\overline{\mathcal{W}} \subset U \subseteq \mathbb{R}^{m\times n}$. In the latter case, although the existence for all $t \in \mathbb{R}$ is no longer guaranteed, this simplifies the stability analysis of equilibria $S^* \in \overline{\mathcal{W}}$, since common theory of dynamical systems can be applied – cf. Section 2.3. This will be used in Section 3.4.

Now, we come to the main statement of this section, which motivates the detailed analysis of the S-flow in the subsequent sections. The following statement shows how the convergence of the assignment flow can be deduced by the convergence of the S-flow.

Proposition 3.2.5. *Let*

$$\mathcal{V}_j = \left\{p \in \Delta_n\colon p_j > p_l,\ \forall l \in [n] \setminus \{j\}\right\}, \quad j \in [n] \tag{3.5}$$

denote the Voronoi cells of the vertices of Δ_n in Δ_n and suppose $\lim_{t\to\infty} S_i(t) = S_i^* \in \Delta_n$, for any $i \in I$. Then the following assertions hold.

(a) If $S_i^* \in \mathcal{V}_{j^*(i)}$ for some label index $j^* = j^*(i) \in J$, then there exist constants $\alpha_i, \beta_i > 0$ such that

$$\|W_i(t) - e_{j^*(i)}\|_1 \leq \alpha_i e^{-\beta_i t}, \quad \forall t \geq 0. \tag{3.6}$$

In particular,

$$\lim_{t\to\infty} W_i(t) = e_{j^*(i)}. \tag{3.7}$$

(b) One has

$$\int_0^\infty \|S_i(t) - S_i^*\|_1 \,\mathrm{d}t < \infty \quad \Longrightarrow \quad \lim_{t\to\infty} W_i(t) = W_i^* \quad \textit{with} \quad \operatorname{supp}(W_i^*) = \arg\max_{j\in J} S_{ij}^*. \tag{3.8}$$

Proof.

(a) Let $\beta_i := \frac{1}{2}\min\left\{S_{ij^*(i)}^* - S_{ij}^*\right\}_{j\neq j^*(i)} > 0$. Since

$$\lim_{t\to\infty} S_{ij^*(i)}(t) - S_{ij}(t) = S_{ij^*(i)}^* - S_{ij}^* \geq 2\beta_i > 0 \quad \forall j \in J \setminus \{j^*(i)\}, \tag{3.9}$$

there exists $t_1 \geq 0$ such that

$$S_{ij^*(i)}(t) - S_{ij}(t) > \beta_i \quad \forall t \geq t_1, \quad \forall j \in J \setminus \{j^*(i)\}. \tag{3.10}$$

We estimate

$$\left\|W_i(t) - e_{j^*(i)}\right\|_1 \tag{3.11a}$$

$$= 1 - W_{ij^*(i)} + \sum_{j\neq j^*(i)} W_{ij} = 2 - 2W_{ij^*(i)} \overset{(3.3)}{=} 2 - 2\frac{\exp\left(\int_0^t S_{ij^*(i)}(\tau)\,\mathrm{d}\tau\right)}{\sum_{j\in J}\exp\left(\int_0^t S_{ij}(\tau)\,\mathrm{d}\tau\right)} \tag{3.11b}$$

$$= 2\frac{\sum_{j\neq j^*(i)}\exp\left(\int_0^t S_{ij}(\tau)\,\mathrm{d}\tau\right)}{\sum_{j\in J}\exp\left(\int_0^t S_{ij}(\tau)\,\mathrm{d}\tau\right)} \tag{3.11c}$$

$$= 2\frac{\sum_{j\neq j^*(i)}\exp\left(\int_0^t \left(S_{ij}(\tau) - S_{ij^*(i)}(\tau)\right)\mathrm{d}\tau\right)}{1 + \sum_{j\neq j^*(i)}\exp\left(\int_0^t \left(S_{ij}(\tau) - S_{ij^*(i)}(\tau)\right)\mathrm{d}\tau\right)} \tag{3.11d}$$

$$\leq 2\sum_{j\neq j^*(i)}\exp\left(\int_0^t \left(S_{ij}(\tau) - S_{ij^*(i)}(\tau)\right)\mathrm{d}\tau\right) \tag{3.11e}$$

$$= 2\sum_{j\neq j^*(i)}\exp\left(\int_0^{t_1} \left(S_{ij}(\tau) - S_{ij^*(i)}(\tau)\right)\mathrm{d}\tau + \int_{t_1}^t \big(\overbrace{S_{ij}(\tau) - S_{ij^*(i)}(\tau)}^{<-\beta_i}\big)\,\mathrm{d}\tau\right) \tag{3.11f}$$

$$\leq 2\sum_{j\neq j^*(i)}\exp\left(\int_0^{t_1} \left(S_{ij}(\tau) - S_{ij^*(i)}(\tau)\right)\mathrm{d}\tau\right)\cdot e^{-\beta_i(t-t_1)} \tag{3.11g}$$

$$= \underbrace{2e^{\beta_i t_1}\sum_{j\neq j^*(i)}\exp\left(\int_0^{t_1} \left(S_{ij}(\tau) - S_{ij^*(i)}(\tau)\right)\mathrm{d}\tau\right)}_{=:\alpha_i > 0}\cdot e^{-\beta_i t}, \tag{3.11h}$$

which proves (3.6).

(b) Let $J^*(i) \coloneqq \arg\max_{j \in J} S^*_{ij}$. For any $j, l \in J^*(i)$, we have $S^*_{ij} = S^*_{il}$ and therefore

$$\int_0^\infty \big|S_{ij}(t) - S_{il}(t)\big| \,\mathrm{d}t \leq \int_0^\infty \big|S_{ij}(t) - S^*_{ij}\big| \,\mathrm{d}t + \int_0^\infty \big|S_{il}(t) - S^*_{il}\big| \,\mathrm{d}t \tag{3.12a}$$

$$\leq 2 \int_0^\infty \|S_i(t) - S^*_i\|_1 \,\mathrm{d}t < \infty, \tag{3.12b}$$

where the last inequality follows from the hypothesis of (3.8). Thus, the improper integral $\int_0^\infty \big(S_{ij}(t) - S_{il}(t)\big) \,\mathrm{d}t \in \mathbb{R}$ exists. If $j \in J^*(i)$, we obtain

$$W_{ij}(t) \overset{(3.3)}{=} \frac{\exp\Big(\int_0^t S_{ij}(\tau)\,\mathrm{d}\tau\Big)}{\sum_{l \in J} \exp\Big(\int_0^t S_{il}(\tau)\,\mathrm{d}\tau\Big)} \tag{3.13a}$$

$$= \Bigg(1 + \sum_{l \in J \setminus J^*(i)} \exp\Big(\overbrace{\int_0^t \big(S_{il}(\tau) - S_{ij}(\tau)\big)\,\mathrm{d}\tau}^{\to -\infty}\Big) + \sum_{l \in J^*(i) \setminus \{j\}} \exp\Big(\int_0^t \big(S_{il}(\tau) - S_{ij}(\tau)\big)\,\mathrm{d}\tau\Big)\Bigg)^{-1} \tag{3.13b}$$

$$\longrightarrow \Bigg(1 + \sum_{l \in J^*(i) \setminus \{j\}} \exp\Big(\int_0^\infty \big(S_{il}(\tau) - S_{ij}(\tau)\big)\,\mathrm{d}\tau\Big)\Bigg)^{-1} \in (0, 1] \quad \text{for} \quad t \to \infty, \tag{3.13c}$$

whereas for any $j \in J \setminus J^*(i)$

$$W_{ij}(t) \overset{(3.3)}{=} \frac{\exp\Big(\int_0^t S_{ij}(\tau)\,\mathrm{d}\tau\Big)}{\sum_{l \in J} \exp\Big(\int_0^t S_{il}(\tau)\,\mathrm{d}\tau\Big)} \tag{3.14a}$$

$$= \Bigg(\overbrace{\sum_{l \in J \setminus J^*(i)} \exp\Big(\int_0^t \big(S_{il}(\tau) - S_{ij}(\tau)\big)\,\mathrm{d}\tau\Big)}^{\geq 0} + \sum_{l \in J^*(i)} \exp\Big(\underbrace{\int_0^t \big(S_{il}(\tau) - S_{ij}(\tau)\big)\,\mathrm{d}\tau}_{\to \infty}\Big)\Bigg)^{-1} \longrightarrow 0 \quad \text{for} \quad t \to \infty. \tag{3.14b}$$

(3.14c)

□

Proposition 3.2.5(a) states that if any subvector of the S-flow converges to a point in a Voronoi cell (3.5), then the corresponding subvector of $W(t)$ converges exponentially fast to the corresponding integral assignment. If the S-flow itself already converges to an integral assignment, then $W(t)$ converges to the same assignment. We will show under some mild

conditions on the parameters Ω that the S-flow converges almost surely to integral assignments (cf. Corollary 3.4.9 below).

Proposition 3.2.5(b) handles the case when the limit point S_i^* lies at the border of adjacent Voronoi cells, that is the set $\arg\max_{j\in J} S_{ij}^*$ is *not* a singleton. In this case, one can only state that $W_i(t)$ converges to some (possibly nonintegral) point W_i^* without being able to predict precisely this limit based on S_i^* alone. In contrast to (a), we also have to assume that the convergence of the S-flow is fast enough—see the hypothesis of (3.8). This assumption is reasonable, however, because it is satisfied whenever S_i^* is subvector of a *hyperbolic* equilibrium point of the S-flow (cf. Remark 3.4.11 below).

Example 3.2.6. We briefly demonstrate what may happen when the assumption of (3.8) is violated. Suppose $S_i(t)$ and S_i^* are given by

$$S_i(t) = \begin{pmatrix} \frac{1}{2} - \frac{1}{t+1} \\ \frac{1}{2} - \frac{2}{t+1} \\ \frac{3}{t+1} \end{pmatrix} \longrightarrow S_i^* = \begin{pmatrix} \frac{1}{2} \\ \frac{1}{2} \\ 0 \end{pmatrix} \quad \text{for} \quad t \to \infty. \tag{3.15}$$

The first component of $S_i(t)$ converges faster than the second component. Since $\|S_i(t) - S_i^*\|_1 = \frac{6}{t+1}$, the convergence rate assumption of (3.8) does not hold. Calculating $W_i(t)$ due to (3.3) gives

$$W_i(t) = \frac{1}{1 + \frac{1}{t+1} + (t+1)^4 e^{-\frac{1}{2}t}} \begin{pmatrix} 1 \\ \frac{1}{t+1} \\ (t+1)^4 e^{-\frac{1}{2}t} \end{pmatrix} \longrightarrow W_i^* = \begin{pmatrix} 1 \\ 0 \\ 0 \end{pmatrix} \quad \text{for} \quad t \to \infty, \tag{3.16}$$

i.e. $W_i(t)$ still converges, but we have $\operatorname{supp} W_i^* \subsetneq \arg\max_{j\in J} S_{ij}^*$ unlike the statement of (3.8). This example also shows that, in the case of Proposition 3.2.5(b), the limit W_i^* may depend on the trajectory $S_i(t)$, rather than only on the limit point S_i^* as in case (a).

Proposition 3.2.5 makes explicit that the S-flow largely determines the asymptotic behavior of $W(t)$. The next sections, therefore, focus on the S-flow (3.2a) and on its dependency on the parameters Ω. Even though the adjacency matrix Ω will be row-stochastic in the later applications of the assignment flow, we will assume a general matrix $\Omega \in \mathbb{R}^{m\times m}$ in the ongoing analysis of the S-flow and possibly restrict it in the following statements.

3.3 Convergence of the S-Flow towards an Equilibrium

In this section we investigate whether the S-flow converges towards a single point for $t \to \infty$. We show first, under a symmetry assumption on the weights, that the S-flow always converges towards an equilibrium. Afterwards we present some academical examples that demonstrate that convergence towards a point can fail if the symmetry assumption is violated. This also confirms that this assumption is not too strong.

3.3.1 Convergence to an Equilibrium for Symmetric Weights

In this subsection we make the basic assumption that the parameter matrix Ω has the form

$$\Omega = \operatorname{Diag}(w)^{-1}\widehat{\Omega}, \quad \text{with} \quad w \in \mathbb{R}^m_{>0} \quad \text{and} \quad \widehat{\Omega}^\top = \widehat{\Omega} \in \mathbb{R}^{m\times m}. \tag{3.17}$$

Matrices of the form (3.17) include as special cases where Ω itself is symmetric or where symmetric (nonnegative) weights $\widehat{\Omega}$ were normalized, i.e. $w = \widehat{\Omega}\mathbb{1}_m$. An instance of the latter one are nonnegative uniform weights with symmetric neighborhoods, i.e.

$$\omega_{ik} = \tfrac{1}{|\mathcal{N}_i|} \quad \forall k \in \mathcal{N}_i \qquad \text{and} \qquad k \in \mathcal{N}_i \quad \Leftrightarrow \quad i \in \mathcal{N}_k, \tag{3.18}$$

which we will consider in the subsequent chapters.

Under the symmetry assumption (3.17), we will prove the convergence of the S-flow (3.2a) towards an equilibrium. We will adapt the proof from [LA83] where the authors showed the convergence of the replicator dynamics $\dot{p} = R_p(Ap)$ on a (single) simplex Δ_n with a symmetric matrix A.

Remark 3.3.1. In case of (3.17), the S-flow becomes a Riemannian gradient flow with respect to the (weighted) Fisher-Rao metric

$$g^w_W(U,V) = \sum_{i\in I} w_i \Big\langle \tfrac{U_i}{\sqrt{W_i}}, \tfrac{V_i}{\sqrt{W_i}} \Big\rangle, \quad U, V \in \mathcal{T}_0, \tag{3.19}$$

i.e. $\dot{S} = R_S[\Omega S] = \operatorname{grad} f(S)$ with the potential $f(S) = \frac{1}{2}\langle S, \widehat{\Omega} S\rangle$. This property will be the main ingredient of the convergence proof.

Theorem 3.3.2 (convergence to equilibria)**.** *Let Ω be of the form* (3.17)*. Then the S-flow* (3.2a) *converges to an equilibrium point $S^* = S^*(S_0) \in \overline{\mathcal{W}}$, for any initial value $S_0 \in \mathcal{W}$.*

Proof. First we note that, for Ω of the form (3.17), the S-flow reads component-wise

$$\dot{S}_{ij} = \frac{1}{w_i} S_{ij}\Big((\widehat{\Omega}S)_{ij} - \langle S_i, (\widehat{\Omega}S)_i\rangle \Big) \quad \forall j \in J, \quad \forall i \in I, \tag{3.20}$$

and therefore equilibria $S^* \in \overline{\mathcal{W}}$ are characterized by

$$(\widehat{\Omega}S^*)_{ij} = \langle S^*_i, (\widehat{\Omega}S^*)_i\rangle \quad \forall j \in \operatorname{supp} S^*_i \quad \forall i \in I. \tag{3.21}$$

Now, let $\Lambda^+(S_0) \subset \overline{\mathcal{W}}$ be the (positive) limit set of the S-flow starting in S_0, i.e.

$$\Lambda^+(S_0) = \big\{ S^* \in \overline{\mathcal{W}} \colon \exists (t_k)_{k\in\mathbb{N}} \subset \mathbb{R}_{\geq 0},\ t_k \to \infty,\ S(t_k) \to S^* \big\}. \tag{3.22}$$

The set $\Lambda^+(S_0) \neq \emptyset$ is non-empty since $\overline{\mathcal{W}}$ is compact.

Lemma 3.3.3. *Every point $S^* \in \Lambda^+(S_0)$ of the limit set* (3.22) *is an equilibrium of the S-flow.*

Proof. We use Theorem 2.3.10. Note that $\overline{\mathcal{W}}$ is compact and positively invariant with respect to the S-flow. The function

$$f \colon \mathbb{R}^{m\times n} \to \mathbb{R}, \qquad f(S) = \langle S, \widehat{\Omega}S\rangle \tag{3.23}$$

is a Lyapunov function of the S-flow (3.2a) on $\overline{\mathcal{W}}$, since

$$\begin{aligned} \frac{\mathrm{d}}{\mathrm{d}t} f(S(t)) = 2\langle \widehat{\Omega}S, \dot{S}\rangle = 2\sum_{i\in I} \langle (\widehat{\Omega}S)_i, \dot{S}_i\rangle &\overset{\langle \mathbb{1}_n, \dot{S}_i\rangle = 0}{=} 2\sum_{i\in I} \Big\langle (\widehat{\Omega}S)_i - \langle S_i, (\widehat{\Omega}S)_i\rangle \mathbb{1}_n, \dot{S}_i \Big\rangle \\ &\overset{(3.2a),(3.17)}{=} \sum_{i\in I} \frac{2}{w_i} \sum_{j\in J} S_{ij}\Big((\widehat{\Omega}S)_{ij} - \langle S_i, (\widehat{\Omega}S)_i\rangle \Big)^2 \geq 0 \end{aligned} \tag{3.24}$$

with equality only if S satisfies the equilibrium criterion (3.21). By Theorem 2.3.10 this equality holds for every point $S^* \in \Lambda^+(S_0)$, i.e., S^* is an equilibrium. □

Next, we introduce some additional notation. Let $S^* \in \Lambda^+(S_0)$ be an equilibrium. The weighted Kullback-Leibler divergence is defined by

$$D_{\mathrm{KL}}^w(S^*, S) := \begin{cases} -\sum_{i\in I} w_i \sum_{j\in \operatorname{supp} S_i^*} S_{ij}^* \log \frac{S_{ij}}{S_{ij}^*}, & \text{if } \operatorname{supp} S^* \subseteq \operatorname{supp} S, \\ \infty, & \text{else.} \end{cases} \tag{3.25a}$$

$$= \sum_{i\in I} w_i D_{\mathrm{KL}}(S_i^*, S_i) \tag{3.25b}$$

with weights $w \in \mathbb{R}^m_{>0}$ from (3.17) and the supports

$$\operatorname{supp} S = \{(i,j) \in I \times J \colon S_{ij} \neq 0\}, \tag{3.26a}$$

$$\operatorname{supp} S_i = \{j \in J \colon S_{ij} \neq 0\}. \tag{3.26b}$$

Analogously to [LA83], we regard the index sets

$$J_0(i) = \left\{j \in J \colon (\widehat{\Omega} S^*)_{ij} = \langle S_i^*, (\widehat{\Omega} S^*)_i \rangle\right\}, \tag{3.27a}$$

$$J_-(i) = \left\{j \in J \colon (\widehat{\Omega} S^*)_{ij} > \langle S_i^*, (\widehat{\Omega} S^*)_i \rangle\right\}, \tag{3.27b}$$

$$J_+(i) = \left\{j \in J \colon (\widehat{\Omega} S^*)_{ij} < \langle S_i^*, (\widehat{\Omega} S^*)_i \rangle\right\}, \tag{3.27c}$$

and define (continuous) functions $Q \colon \overline{\mathcal{W}} \to \mathbb{R}_{\geq 0}$ and $V \colon \overline{\mathcal{W}} \to \mathbb{R}_{\geq 0} \cup \{\infty\}$ by

$$Q(S) = \sum_{i\in I} w_i \sum_{j \in J_+(i)} S_{ij}, \tag{3.28a}$$

$$V(S) = D_{\mathrm{KL}}^w(S^*, S) + 2Q(S). \tag{3.28b}$$

The equilibrium criterion (3.21) implies

$$\operatorname{supp} S_i^* \subseteq J_0(i) \quad \text{and} \quad J_-(i), J_+(i) \subseteq J \setminus \operatorname{supp} S_i^* \qquad \forall i \in I, \tag{3.29}$$

i.e. $V(S^*) = Q(S^*) = 0$. Using the Lyapunov function (3.23), we have the following.

Lemma 3.3.4 (cf. [LA83, Proposition 2]). *There exists $\varepsilon > 0$ such that, if $\|S(t) - S^*\| < \varepsilon$ and $f(S(t)) < f(S^*)$ with f given in* (3.23)*, then $\frac{\mathrm{d}}{\mathrm{d}t} V(S(t)) < 0$.*

For reasons of readability, we postpone the proof of the lemma and continue the proof of the theorem. Let $S^* \in \Lambda^+(S_0)$ be the equilibrium from above and let $(t_k)_{k\in\mathbb{N}}$ be a corresponding sequence due to (3.22). We show that $D_{\mathrm{KL}}^w(S^*, S(t)) \to 0$ for $t \to \infty$, which is equivalent to the assertion $S(t) \to S^*$ to be shown – cf. Section 2.1.5.

Choose $\varepsilon > 0$ according to the Lemma 3.3.4. There exists[1] $\varepsilon_1 > 0$ such that the (relatively) open set $U = \{S \in \overline{\mathcal{W}} \colon V(S) < \varepsilon_1\}$ is contained in $\{S \in \overline{\mathcal{W}} \colon \|S - S^*\| < \varepsilon\}$. The function $t \mapsto f(S(t))$ is strictly increasing unless the orbit $\{S(t) \colon t \geq 0\}$ consists of an equilibrium. Hence, $f(S(t)) < f(S^*)$ for all $t \geq 0$. Since $S(t_k) \to S^*$, we get $S(t_{k_0}) \in U$ for some $k_0 \in \mathbb{N}$. Since then $t \mapsto V(S(t))$ is decreasing, i.e. $V(S(t)) < V(S(t_{k_0})) < \varepsilon_1$ for all $t > t_{k_0}$. Because $V(S(t))$ is decreasing and $V(S(t_k)) \to V(S^*) = 0$, we get

$$0 \leq D_{\mathrm{KL}}^w(S^*, S(t)) \leq V(S(t)) \to 0 \quad \text{for} \quad t \to \infty, \tag{3.30}$$

which implies $S(t) \to S^*$ for $t \to \infty$. To complete the proof, it remains to prove Lemma 3.3.4.

[1] This follows by the fact that $S \to S^*$ if $D_{\mathrm{KL}}^w(S^*, S) \to 0$.

Proof of Lemma 3.3.4. Since $S(t) \in \mathcal{W}$ for all $t \geq 0$, we have $D_{\mathrm{KL}}^{w}(S^*, S(t)) < \infty$. Hence, we can compute the time derivate

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} D_{\mathrm{KL}}^{w}(S^*, S(t)) &= -\sum_{i\in I} w_i \sum_{j\in \operatorname{supp} S_i^*} S_{ij}^* \frac{\dot{S}_{ij}}{S_{ij}} && \text{(3.31a)}\\
&= -\sum_{i\in I} \sum_{j\in \operatorname{supp} S_i^*} S_{ij}^* \big((\widehat{\Omega} S)_{ij} - \langle S_i, (\widehat{\Omega} S)_i \rangle \big) && \text{(3.31b)}\\
&\overset{\sum_j S_{ij}^*=1}{=} \langle S, \widehat{\Omega} S \rangle - \langle S^*, \widehat{\Omega} S \rangle \overset{(3.17)}{=} \langle S, \widehat{\Omega} S \rangle - \langle S, \widehat{\Omega} S^* \rangle && \text{(3.31c)}\\
&= \langle S, \widehat{\Omega} S \rangle - \langle S^*, \widehat{\Omega} S^* \rangle + \langle S^*, \widehat{\Omega} S^* \rangle - \langle S, \widehat{\Omega} S^* \rangle && \text{(3.31d)}\\
&\overset{\sum_j S_{ij}=1}{=} \langle S, \widehat{\Omega} S \rangle - \langle S^*, \widehat{\Omega} S^* \rangle + \sum_{i\in I} \sum_{j\in J} S_{ij} \Big(\langle S_i^*, (\widehat{\Omega} S^*)_i \rangle - (\widehat{\Omega} S^*)_{ij} \Big) && \text{(3.31e)}\\
&\overset{(3.27)}{=} \underbrace{\langle S, \widehat{\Omega} S \rangle - \langle S^*, \widehat{\Omega} S^* \rangle}_{f(S)-f(S^*)<0} + \sum_{i\in I} \sum_{j\in J_-(i)} S_{ij} \Big(\underbrace{\langle S_i^*, (\widehat{\Omega} S^*)_i \rangle - (\widehat{\Omega} S^*)_{ij}}_{<0} \Big) \\
&\quad + \sum_{i\in I} \sum_{j\in J_+(i)} S_{ij} \Big(\underbrace{\langle S_i^*, (\widehat{\Omega} S^*)_i \rangle - (\widehat{\Omega} S^*)_{ij}}_{>0} \Big). && \text{(3.31f)}
\end{aligned}$$

We now focus on $Q(S)$ (3.28a) that is added to the KL-divergence to define $V(S)$ in (3.28b). We have for each $j \in J_+(i)$

$$\langle S_i, (\widehat{\Omega} S)_i \rangle - (\widehat{\Omega} S)_{ij} \longrightarrow \langle S_i^*, (\widehat{\Omega} S^*)_i \rangle - (\widehat{\Omega} S^*)_{ij} > 0 \quad \text{as} \quad S \to S^*. \tag{3.32}$$

Since the limit is positive, there exists $\varepsilon > 0$ such that $\|S - S^*\| < \varepsilon$ implies

$$\langle S_i, (\widehat{\Omega} S)_i \rangle - (\widehat{\Omega} S)_{ij} \geq \frac{3}{4} \Big(\langle S_i^*, (\widehat{\Omega} S^*)_i \rangle - (\widehat{\Omega} S^*)_{ij} \Big) \quad \forall j \in J_+(i), \quad \forall i \in I. \tag{3.33}$$

Consequently,

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} Q(S(t)) &\overset{\substack{(3.2a)\\(3.28a)}}{=} \sum_{i\in I} \sum_{j\in J_+(i)} S_{ij} \Big((\widehat{\Omega} S)_{ij} - \langle S_i, (\widehat{\Omega} S)_i \rangle \Big) && \text{(3.34a)}\\
&\leq -\frac{3}{4} \sum_{i\in I} \sum_{j\in J_+(i)} S_{ij} \Big((\widehat{\Omega} S^*)_{ij} - \langle S_i^*, (\widehat{\Omega} S^*)_i \rangle \Big). && \text{(3.34b)}
\end{aligned}$$

Substituting (3.31) and (3.34) into (3.28b), we finally obtain

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t} V(S(t)) &= \frac{\mathrm{d}}{\mathrm{d}t} D_{\mathrm{KL}}^{w}(S^*, S(t)) + 2 \cdot \frac{\mathrm{d}}{\mathrm{d}t} Q(S(t)) && \text{(3.35a)}\\
&\leq \langle S, \widehat{\Omega} S \rangle - \langle S^*, \widehat{\Omega} S^* \rangle + \sum_{i\in I} \sum_{j\in J_-(i)} S_{ij} \Big(\langle S_i^*, (\widehat{\Omega} S^*)_i \rangle - (\widehat{\Omega} S^*)_{ij} \Big) \\
&\qquad - \frac{1}{2} \sum_{i\in I} \sum_{j\in J_+(i)} S_{ij} \Big(\langle S_i^*, (\widehat{\Omega} S^*)_i \rangle - (\widehat{\Omega} S^*)_{ij} \Big) && \text{(3.35b)}\\
&< 0. && \text{(3.35c)}
\end{aligned}$$

□

3.3.2 Academic Examples for Non-Convergence towards Equilibria

In this subsection, we construct a family of S-flows (3.2a) in terms of a class of nonnegative parameter matrices Ω, that may violate assumption (3.17) which underlies Theorem 3.3.2. Accordingly, for a small problem size $n = 3$, we explicitly specify flows that exhibit one of the following behaviors:

1. $t \mapsto S(t)$ converges towards a point $S^* \in \overline{\mathcal{W}}$ as $t \to \infty$;
2. $t \mapsto S(t)$ is periodic with some period $t_1 > 0$;
3. $t \mapsto S(t)$ neither converges to a point nor is periodic.

These cases are discussed below as Example 3.3.9 and illustrated by Figure 3.1. They demonstrate that assumption (3.17) is not too strong, because violation may easily imply that the flow fails to converge to an equilibrium.

Let $\mathcal{D}$ denote the set of *doubly stochastic, circulant* matrices. We consider the case $m = |I| = |J| = n$ and therefore have $\mathcal{D} \subset \overline{\mathcal{W}}$. Let

$$P \in \{0,1\}^{n \times n}, \qquad P_{ij} = \begin{cases} 1, & \text{if } i - j \equiv 1 \pmod{n}, \\ 0, & \text{else} \end{cases} \tag{3.36}$$

denote the permutation matrix that represents the n-cycle $(1, \dots, n)$. Then $\mathcal{D}$ is the convex hull of the matrices $\{P, P^2, \dots, P^n\}$ with $P^n = I_n$, and any element $M \in \mathcal{D}$ admits the representation

$$M = \sum_{k \in [n]} \mu_k P^k \quad \text{with} \quad \mu \in \Delta_n. \tag{3.37}$$

Since the matrices $P, P^2, \dots, P^n \in \mathbb{R}^{n \times n}$ are linearly independent, the vector $\mu \in \Delta_n$ is uniquely determined. We will call μ the *representative* of $M \in \mathcal{D}$. The following lemma characterizes two matrix products on $\mathcal{D}$ in terms of the corresponding matrix representatives.

Lemma 3.3.5. *Let $\mu^{(1)}, \mu^{(2)} \in \Delta_n$ be the representatives of any two matrices $M^{(1)}, M^{(2)} \in \mathcal{D}$. Then the element-wise Hadamard product and the ordinary matrix product, respectively, are given by*

$$M^{(1)} \odot M^{(2)} = \sum_{k \in [n]} \eta_k P^k \quad \textit{with} \quad \eta = \mu^{(1)} \odot \mu^{(2)} \in \mathbb{R}^n_{\geq 0}, \tag{3.38a}$$

$$M^{(1)} M^{(2)} = \sum_{k \in [n]} \mu_k P^k \quad \textit{with} \quad \mu = M^{(1)} \mu^{(2)} \in \Delta_n. \tag{3.38b}$$

Proof. We note that the k-th power of P is given by

$$(P^k)_{ij} = \begin{cases} 1, & \text{if } i - j \equiv k \pmod{n}, \\ 0, & \text{else.} \end{cases} \tag{3.39}$$

This implies $P^k \odot P^l = \delta_{kl} P^k$ for $k, l \in [n]$, with δ_{kl} denoting the Kronecker delta, and

$$M^{(1)} \odot M^{(2)} = \left(\sum_{k \in [n]} \mu_k^{(1)} P^k \right) \odot \left(\sum_{l \in [n]} \mu_l^{(2)} P^l \right) = \sum_{k,l \in [n]} \mu_k^{(1)} \mu_l^{(2)} P^k \odot P^l = \sum_{k \in [n]} \mu_k^{(1)} \mu_k^{(2)} P^k. \tag{3.40}$$

As for (3.38b), we compute

$$M^{(1)}M^{(2)} = \sum_{k,j\in[n]} \mu_k^{(1)}\mu_j^{(2)}P^{k+j} = \sum_{i\in[n]} \sum_{k+j\equiv i \,(\mathrm{mod}\, n)} \mu_k^{(1)}\mu_j^{(2)}P^{i} \tag{3.41a}$$

$$\overset{(3.39)}{=} \sum_{i\in[n]}\sum_{k\in[n]} \mu_k^{(1)}(P^k\mu^{(2)})_i\, P^i = \sum_{i\in[n]} (M^{(1)}\mu^{(2)})_i\, P^i. \tag{3.41b}$$

□

The following proposition shows that, in case of $\Omega \in \mathcal{D}$, the S-flow on $\mathcal{D}$ can be expressed by the evolution of the corresponding representative.

Proposition 3.3.6. *Let $\Omega \in \mathcal{D}$ and suppose the S-flow* (3.2a) *is initialized at $S(0) \in \mathcal{D}$. Then the solution $S(t) \in \mathcal{D}$ evolves on $\mathcal{D}$ for all $t \in \mathbb{R}$. In addition, the corresponding representative $p(t) \in \Delta_n$ of $S(t) = \sum_{k\in[n]} p(t)P^k$ satisfies the replicator equation*

$$\dot{p} = R_p(\Omega p). \tag{3.42}$$

Proof. Let $S = \sum_{k\in[n]} p_k P^k \in \mathcal{D}$ with $p \in \Delta_n$. Lemma 3.3.5 implies

$$S \odot (\Omega S) = \sum_{k\in[n]} p_k(\Omega p)_k\, P^k. \tag{3.43}$$

Therefore, for any $i \in [n]$,

$$\langle S_i, (\Omega S)_i\rangle = \langle \mathbb{1}_n, S_i \odot (\Omega S)_i\rangle = \left\langle \mathbb{1}_n, \big(S\odot(\Omega S)\big)_i \right\rangle = \sum_{k\in[n]} p_k(\Omega p)_k \underbrace{\langle \mathbb{1}_n, (P^k)_i\rangle}_{=1} = \langle p, \Omega p\rangle. \tag{3.44}$$

Since this equation holds for any $i \in [n]$, the right-hand side of the S-flow (3.2a) can be rewritten as

$$R_S[\Omega S] = S \odot (\Omega S) - \langle p, \Omega p\rangle S \overset{(3.43)}{=} \sum_{k\in[n]} \Big(p_k(\Omega p)_k\, P^k - \langle p, \Omega p\rangle p_k P^k\Big) \tag{3.45a}$$

$$= \sum_{k\in[n]} v_k P^k \qquad \text{with} \qquad v = p \odot (\Omega p) - \langle p, \Omega p\rangle p = R_p(\Omega p). \tag{3.45b}$$

For $p \in \Delta_n$, the vector v is tangent to Δ_n, i.e., $\dot{S} = \sum_k v_k P^k$ is tangent to $\mathcal{D}$. Hence, $S(t)$ stays in $\mathcal{D}$ for all $t \in \mathbb{R}$. By (3.45), $\dot{S} = \sum_{k\in[n]} \dot{p}_k P^k = R_S[\Omega S]$ is determined by $\dot{p} = v = R_p(\Omega p)$, whose solution $p(t)$ evolves on Δ_n. □

The following proposition introduces a restriction of parameter matrices $\Omega \in \mathcal{D}$ that ensures, for any such Ω, that the product $\prod_{j\in[n]} p_j$ changes monotonously depending on the flow (3.42).

Proposition 3.3.7. *Let $\Omega = \sum_{k\in[n]} \mu_k P^k \in \mathcal{D}$ be parametrized by*

$$\mu = \alpha e_n + \frac{\beta}{n}\mathbb{1}_n + \sum_{k<\left\lfloor \frac{n}{2}\right\rfloor} \gamma_k(e_k - e_{n-k}) \in \Delta_n, \qquad \alpha, \beta, \gamma_1, \ldots, \gamma_{\lfloor \frac{n}{2}\rfloor - 1} \in \mathbb{R}. \tag{3.46}$$

Suppose $p(t) \in \mathcal{S} = \mathrm{rint}(\Delta_n)$ solves (3.42). *Then*

$$\frac{\mathrm{d}}{\mathrm{d}t} \prod_{j\in[n]} p_j(t) \begin{cases} < 0, & \text{if } \alpha > 0 \\ = 0, & \text{if } \alpha = 0 \\ > 0, & \text{if } \alpha < 0 \end{cases}, \quad \text{for} \quad p(t) \neq \tfrac{1}{n}\mathbb{1}_n. \tag{3.47}$$

Proof. Set $\pi_p \coloneqq \prod_{j\in[n]} p_j$. By virtue of (3.42) and $\langle \mathbb{1}_n, \Omega p\rangle = \langle \Omega^\top \mathbb{1}_n, p\rangle = 1$ (Ω is doubly stochastic and $p \in \Delta_n$), we have

$$\frac{\mathrm{d}}{\mathrm{d}t}\pi_p = \pi_p \sum_{j\in[n]} \big((\Omega p)_j - \langle p, \Omega p\rangle\big) = \pi_p\big(1 - n\langle p, \Omega p\rangle\big). \tag{3.48}$$

Hence, since $\pi_p > 0$ for $p \in \mathcal{S}$, $\frac{\mathrm{d}}{\mathrm{d}t}\pi_p$ has the same sign as $\frac{1}{n} - \langle p, \Omega p\rangle$. Regarding the term

$$\langle p, \Omega p\rangle = \sum_{k\in[n]} \mu_k \langle p, P^k p\rangle, \tag{3.49}$$

we have the following three cases:

(α) for all $k < n$, the inequality $\langle p, P^k p\rangle \le \langle p, p\rangle = \langle p, P^n p\rangle$ holds, with equality if and only if $p = \frac{1}{n}\mathbb{1}_n$;

(β) $\sum_{k\in[n]} \langle p, P^k p\rangle = \langle p, \mathbb{1}_{n\times n}\, p\rangle = 1$;

(γ) for all $k \in [n]$, $\langle p, P^k p\rangle = \langle p, P^{n-k} p\rangle$, since $P^{-1} = P^\top$.

Inserting (3.46) into (3.49) and applying (α), (β), (γ) gives

$$\langle p, \Omega p\rangle = \alpha\langle p, p\rangle + \beta\frac{1}{n} \qquad \text{and} \qquad \langle p, p\rangle > \frac{1}{n}\sum_{k\in[n]} \langle p, P^k p\rangle = \frac{1}{n} \quad \text{for} \quad p \neq \tfrac{1}{n}\mathbb{1}_n. \tag{3.50}$$

Since $\langle \mu, \mathbb{1}_n\rangle = \alpha + \beta = 1$, we further obtain

$$\langle p, \Omega p\rangle \begin{cases} > \frac{1}{n}, & \text{if } \alpha > 0 \\ = \frac{1}{n}, & \text{if } \alpha = 0 \\ < \frac{1}{n}, & \text{if } \alpha < 0 \end{cases}, \quad \text{for all} \quad p \in \Delta_n \setminus \{\tfrac{1}{n}\mathbb{1}_n\}. \tag{3.51}$$

Combining (3.51) and (3.48) yields (3.47). □

Remark 3.3.8. Based on Proposition 3.3.7, we observe: If $\alpha > 0$, then $p(t)$ moves towards the (relative) boundary of the simplex Δ_n, for any $p(0) \neq \frac{1}{n}\mathbb{1}_n$. If $\alpha < 0$, then $p(t)$ converges towards the barycenter $\frac{1}{n}\mathbb{1}_n$. For $\alpha = 0$, the product $\prod_{j\in[n]} p_j(t)$ is constant over time.

The scalars γ_k in (3.46) steer the skew-symmetric part of Ω. Consequently, if $\gamma_k = 0$ for all k, then Ω is symmetric and the S-flow converges to a single point by Theorem 3.3.2. Depending on the skew-symmetric part, the S-flow may not converge to a point, as Example 3.3.9 below will demonstrate for few explicit instances and $n = 3$. Note that, in this case $n = 3$, (3.46) describes a parametrization rather than a restriction of $\Omega \in \mathcal{D}$.

Example 3.3.9. Let $n = 3$. The matrix $\Omega \in \mathcal{D}$ takes the form

$$\mu = \alpha e_3 + \tfrac{\beta}{3}\mathbb{1}_3 + \gamma(e_1 - e_2), \tag{3.52}$$

$$\Omega = \begin{pmatrix} \mu_3 & \mu_2 & \mu_1 \\ \mu_1 & \mu_3 & \mu_2 \\ \mu_2 & \mu_1 & \mu_3 \end{pmatrix} = \alpha \begin{pmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} + \frac{\beta}{3}\begin{pmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 1 & 1 \end{pmatrix} + \gamma \begin{pmatrix} 0 & -1 & 1 \\ 1 & 0 & -1 \\ -1 & 1 & 0 \end{pmatrix} \tag{3.53}$$

with the constraint $\mu \in \Delta_3$, i.e.

$$\alpha + \beta = 1, \quad \alpha + \frac{\beta}{3} \geq 0, \quad \frac{\beta}{3} \geq |\gamma|. \tag{3.54}$$

We examine the behavior of the flow (3.42), depending on the parameters α and γ. Note, that the flow does not depend on the parameter β that merely ensures Ω to be row-stochastic.

Case $\alpha < 0$. As already discussed (Remark 3.3.8), $p(t)$ converges to the barycenter in this case. Depending on γ, this may happen with ($\gamma \neq 0$) or without ($\gamma = 0$) a spiral as depicted by Figure 3.1 (a) and (b).

Case $\alpha = 0$. We distinguish the two cases $\gamma = 0$ and $\gamma \neq 0$. If $\gamma = 0$, then we have $\Omega = \frac{1}{3}\mathbb{1}_{3\times 3}$ and therefore $\dot{p} = R_p(\Omega p) \equiv 0$, i.e., each point $p^* \in \Delta_3$ is an equilibrium. In contrast, if $\gamma \neq 0$, then we have the (standard) rock-paper-scissors dynamics [SG16, Chapter 10]:

$$\dot{p} = \gamma \begin{pmatrix} p_1(p_3 - p_2) \\ p_2(p_1 - p_3) \\ p_3(p_2 - p_1) \end{pmatrix} \neq 0, \quad \text{for } p \in \Delta_3 \setminus \{e_1, e_2, e_3, \tfrac{1}{3}\mathbb{1}_3\}. \tag{3.55}$$

Starting at a point $p_0 \in \operatorname{rint}\Delta_3 \setminus \{\frac{1}{3}\mathbb{1}_3\}$, the curve $t \mapsto p(t)$ moves along the closed curve $\left\{p \in \Delta_3 \colon \prod_j p_j = \prod_j p_{0,j}\right\}$, i.e., the curve $t \mapsto p(t)$ is periodic; see Figure 3.1 (c).

Case $\alpha > 0$. We distinguish again the two cases $\gamma = 0$ and $\gamma \neq 0$. If $\gamma = 0$, then the flow reduces to $\dot{p} = \alpha R_p p$ whose solution converges to

$$\lim_{t\to\infty} p(t) = \tfrac{1}{|J^*|} \sum_{j \in J^*} e_j \in \Delta_3, \quad \text{with} \quad J^* = \operatorname*{arg\,max}_{j \in [3]} p_j(0). \tag{3.56}$$

As for the remaining case $\alpha > 0$ and $\gamma \neq 0$, we distinguish $\alpha > |\gamma|$ and $\alpha \leq |\gamma|$ as illustrated by Figure 3.1 (e), (f) and (g). If $\alpha \leq |\gamma|$, then we have a generalized rock-paper-scissors game [SG16, Chapter 10]. The curve $t \mapsto p(t)$ spirals towards the boundary of the simplex Δ_3 and does not converge to a single point. In contrast, if $\alpha > |\gamma|$, then the flow converges to a point on the boundary. In fact, the vertices of the simplex are attractors.

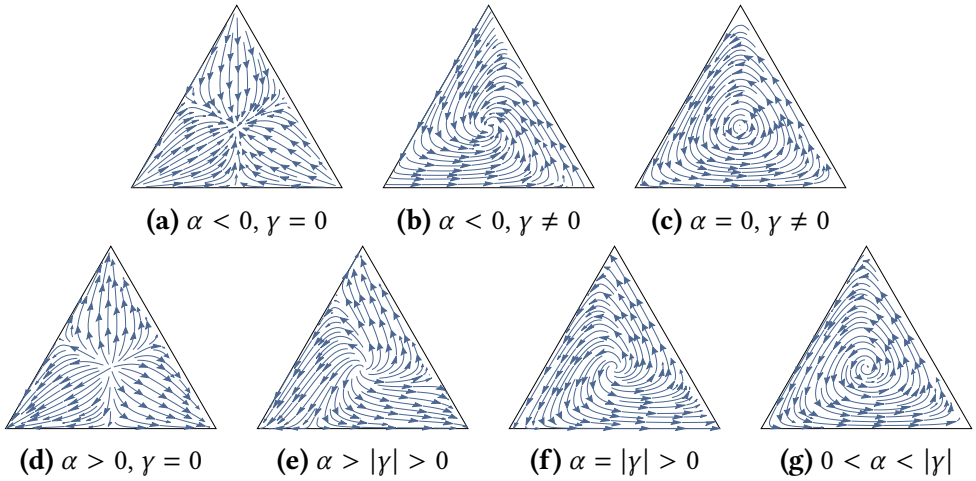
Example 3.3.9 is devoted to the S-flow (3.2a) that parametrizes the assignment flow (3.2b), as specified by Proposition 3.2.2. The following examples illustrate how the assignment flow may behave if the S-flow does not converge to an equilibrium point.

Example 3.3.10. This example continues Example 3.3.9. Accordingly, we consider the case $n = 3$ and assume $\Omega \in \mathcal{D}$. Let the distance matrix D be given by

$$D = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}. \tag{3.57}$$

Then, if $\Omega \in \mathcal{D}$, the initial value $S(0) = \exp_{C_W}(-\Omega D)$ of the S-flow (3.2a) lies in $\mathcal{D}$ as well. Hence, the above observations of Example 3.3.9 for the S-flow hold. The resulting assignment flow $t \mapsto W(t)$ then also evolves in $\mathcal{D}$ which can be verified using (3.3). As for the averaging parameters Ω, we consider the following three matrices in $\mathcal{D}$:

$$\Omega_{\text{center}} = \begin{pmatrix} 0 & 0 & 1 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \end{pmatrix}, \quad \Omega_{\text{cycle}} = \frac{1}{3}\begin{pmatrix} 1 & 0 & 2 \\ 2 & 1 & 0 \\ 0 & 2 & 1 \end{pmatrix}, \quad \Omega_{\text{spiral}} = \frac{1}{5}\begin{pmatrix} 2 & 0 & 3 \\ 3 & 2 & 0 \\ 0 & 3 & 2 \end{pmatrix}. \tag{3.58}$$



(a) $\alpha < 0, \gamma = 0$ **(b)** $\alpha < 0, \gamma \neq 0$ **(c)** $\alpha = 0, \gamma \neq 0$

(d) $\alpha > 0, \gamma = 0$ **(e)** $\alpha > |\gamma| > 0$ **(f)** $\alpha = |\gamma| > 0$ **(g)** $0 < \alpha < |\gamma|$

Figure 3.1: Phase portraits for the flows $\dot{p} = R_p(\Omega p)$ of Example 3.3.9. Ω is parameterized as specified by (3.53). Parameter α controls whether the flow evolves towards the barycenter ($\alpha < 0$) as in **(a)** and **(b)**, or towards the boundary of the simplex ($\alpha > 0$) as in **(d)**-**(g)**. Parameter γ controls the rotational component of the flow. In **(c)**, the flow neither evolves towards the barycenter nor towards the boundary, and the rotational component of the flow causes periodic orbits. If $\alpha > 0$, then the convergence of the flow depends on the size of γ. If $0 \leq |\gamma| < \alpha$ as in **(d)** and **(e)**, then the flow converges to a point on the boundary. If $|\gamma| \geq \alpha$ as in **(f)** and **(g)**, then the flow spirals towards the boundary without converging to a single point.

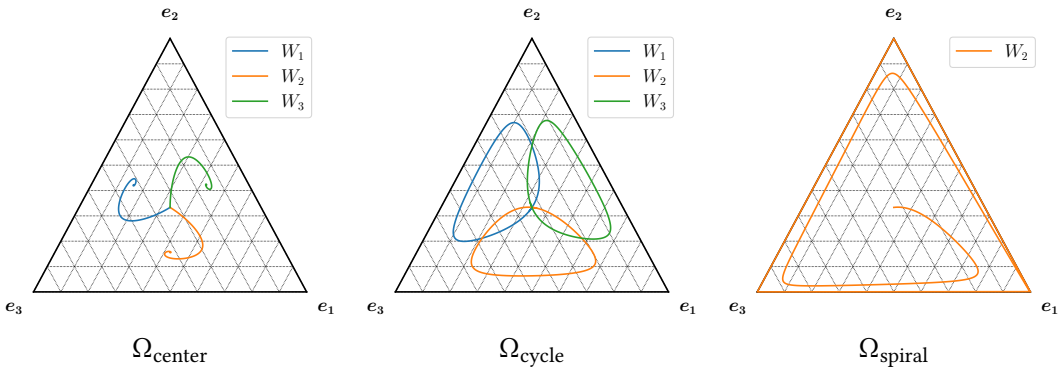


Figure 3.2: Trajectories of the assignment flows in example 3.3.10. The input data is given in (3.57) and (3.58). The flow for the matrix Ω_{center} converges to a point in the interior of the assignment manifold. This limit point differs from the barycenter. The trajectory for the averaging matrix Ω_{cycle} is a closed curve. The trajectory for Ω_{spiral} is spiraling towards the boundary of the simplex. For the sake of clarity, the trajectory of only one data point is plotted for Ω_{spiral}. The trajectories for the other data points can be obtained from that one by permuting the label indices.

Figure 3.2 displays the trajectories of the assignment flow for these averaging matrices. The symmetry of these plots results from $W(t) \in \mathcal{D}$.

Matrix Ω_{center} corresponds to the parameters $(\alpha, \beta, \gamma) = (-\frac{1}{2}, \frac{3}{2}, \frac{1}{2})$ of (3.53), for which the S-flow converges to the barycenter. As a consequence, $W(t)$ converges to a point in $\mathcal{W} \setminus \{C_{\mathcal{W}}\}$.

Matrix Ω_{cycle} corresponds to the parameters $(\alpha, \beta, \gamma) = (0, 1, \frac{1}{3})$, for which the S-flow has periodic orbits. Since these orbits are symmetric around the barycenter, i.e. $\int_0^{t_1} \big(S(t) - C_{\mathcal{W}}\big)\mathrm{d}t = 0$ with t_1 being the period of the trajectory, the trajectory $t \mapsto W(t)$ is also periodic as a consequence of equation (3.3).

Finally, matrix Ω_{spiral} corresponds to the parameters $(\alpha, \beta, \gamma) = (0.1, 0.9, 0.3)$, for which the S-flow spirals towards the boundary of the simplex. It is not clear apriori if $t \mapsto W(t)$ does not converge to a single point either. The trajectory of $W(t)$ shown by Figure 3.2 suggests that the assignment flow also spirals towards the boundary of the simplex without converging to a single point.

Remark 3.3.11. Examples 3.3.9 and 3.3.10 considered the special case $m = |I| = |J| = n = 3$. We observed in further experiments similar behaviors also in the case $n < m$. For example, it can be verified, for $n = 2$ and $\Omega = \Omega_{\text{cycle}}$ from Example 3.3.10, that the S-flow possesses a (unstable) limit cycle, i.e. a periodic orbit.

The above examples also demonstrate that several symmetries in the input data are required, e.g. $\Omega \in \mathcal{D}$ and $S_0 \in \mathcal{D}$, in order to obtain nonconvergent orbits. However, small perturbations like numerical errors or the omnipresent noise in *real* data will break these symmetries. Therefore, it is very unlikely to observe such behavior of the S-flow and the assignment flow, respectively, in practice.

3.4 Stability Analysis for Equilibria of the S-Flow

In this section we will analyze the stability of equilibria of the S-flow. For this, we will use statements provided in Section 2.3.1. In particular, we will utilize the linearization of a system, i.e., we consider the Jacobian matrix of the vector field of the S-flow and examine its eigenvalues. We will show under a positivity assumption on the diagonal of the parameter matrix Ω that nonintegral equilibria are unstable and infer from this that the set of initial values of the S-flow, for which the flow converge to a nonintegral equilibrium, is negligible. At the end of this section, we will take a closer look at attracting equilibria and derive estimates for their basin of attraction. As in the previous section, we assume a general matrix $\Omega \in \mathbb{R}^{m \times m}$ and, if necessary, restrict it in the following statements.

3.4.1 Stability Analysis by Linearization

In this subsection we will use the shorthand F for the vector field of the S-flow, i.e., the S-flow reads

$$\dot{S} = F(S) = R_S[\Omega S], \qquad S(0) = S_0 \in \mathcal{W}. \tag{3.59}$$

We will analyze the stability of equilibria by regarding the spectrum of the Jacobian matrix $\frac{\partial F}{\partial S}(S^*)$ at an equilibrium $S^* \in \overline{\mathcal{W}}$, i.e. $F(S^*) = 0$. First we have to characterize the equilibria and calculate the corresponding Jacobian matrix.

Proposition 3.4.1 (equilibria). *Let $\Omega \in \mathbb{R}^{n\times n}$ be an arbitrary matrix.*

(a) A point $S^* \in \overline{\mathcal{W}}$ is an equilibrium of the S-flow (3.59) *if and only if*

$$(\Omega S^*)_{ij} = \langle S_i^*, (\Omega S^*)_i \rangle, \qquad \forall j \in \operatorname{supp} S_i^*, \qquad \forall i \in I, \tag{3.60}$$

i.e., the subvectors $(\Omega S^*)_i$ are constant on $\operatorname{supp} S_i^*$, for each $i \in I$.

(b) Every point $S^* \in \mathcal{W}^*$ is an equilibrium of the S-flow (3.59).

(c) Let $J_+ \subseteq J$ be a non-empty subset of indices, and let $\mathbb{1}_{J_+} \in \mathbb{R}^{|J|}$ be the corresponding indicator vector with components $(\mathbb{1}_{J_+})_j = 1$ if $j \in J_+$ and $(\mathbb{1}_{J_+})_j = 0$ otherwise. Then $S^* = \frac{1}{|J_+|}\mathbb{1}_m \mathbb{1}_{J_+}^\top$ is an equilibrium. In particular, the barycenter $C_{\mathcal{W}} = \frac{1}{m}\mathbb{1}_m\mathbb{1}_n^\top$ corresponding to $J_+ = J$ is an equilibrium.

(d) The set $\mathcal{E} = \{S^* \in \overline{\mathcal{W}} \colon F(S^*) = 0\}$ of all equilibria of the S-flow (3.59) *is a compact subset.*

(e) For an index set $\Sigma \subseteq I \times J$, let

$$\mathcal{E}_\Sigma = \left\{S^* \in \mathcal{E} \colon \operatorname{supp} S^* = \Sigma\right\} \subset \mathcal{E} \tag{3.61}$$

denote the set of equilibria with given support Σ. This set is the relative interior of a convex polytope. In particular, the set of all equilibria is a finite union of convex polytopes, i.e.

$$\mathcal{E} = \bigcup_{\Sigma \subseteq I\times J} \mathcal{E}_\Sigma = \bigcup_{\Sigma \subseteq I\times J} \overline{\mathcal{E}_\Sigma}. \tag{3.62}$$

Proof.

(a) Each equation of the system (3.59) has the form

$$\dot S_{ij} = S_{ij}\big((\Omega S)_{ij} - \langle S_i, (\Omega S)_i\rangle\big), \quad i \in I,\ j \in J. \tag{3.63}$$

$\dot S_{ij} = 0$ implies that either $S_{ij} = S_{ij}^* = 0$, i.e. $j \notin \operatorname{supp}(S_i^*)$, or the term in the round brackets is zero, which is (3.60).

(b) The replicator matrix (2.91) satisfies $R_{e_j} = 0,\ \forall j \in J$. This implies $R_{S^*} = 0$ and in turn $R_{S^*}[\Omega S^*] = 0$.

(c) Since $\Omega S^* = \frac{1}{|J_+|}(\Omega \mathbb{1}_m)\mathbb{1}_{J_+}^\top$, the subvectors $(\Omega S^*)_i,\ i \in I$ are constant on $J_+ = \operatorname{supp} S_i^*$, which implies by (a) that S^* is an equilibrium point.

(d) As the zero set of the continuous vector field F, the set $\mathcal{E}$ is a closed subset of the compact set $\overline{\mathcal{W}}$ and therefore itself compact.

(e) Using the equilibrium criterion (3.60), equilibria $S^* \in \overline{\mathcal{W}}$ are described by the following set of linear equality and inequality constraints:

$$\left.\begin{aligned} (\Omega S^*)_{ij} - (\Omega S^*)_{il} &= 0 && \forall j,l \in \operatorname{supp} S_i^* \\ \langle S_i^*, \mathbb{1}_n\rangle &= 1 && \\ S_{ij}^* &= 0 && \forall j \in J \setminus \operatorname{supp} S_i^* \\ S_{ij}^* &> 0 && \forall j \in \operatorname{supp} S_i^* \end{aligned}\right\} \quad \forall i \in I. \tag{3.64}$$

If $\operatorname{supp} S^*$ is prescribed, then (3.64) defines the relative interior of a convex polytope. □

Remark 3.4.2. The set of equilibria characterized by Proposition 3.4.1 (b) and (c) may not exhaust the set of all equilibrium points for a general parameter matrix Ω. For example, $\mathcal{E}_\Sigma$ in (e) does not have to be a singleton. However, we will show below that, under certain mild conditions, any such additional equilibrium point must be unstable.

Lemma 3.4.3 (Jacobian matrix)**.** *Let $F(S)$ denote the vector field defining the S-flow* (3.59)*. Then, after stacking S row-wise, the Jacobian matrix of F is given by*

$$\frac{\partial F}{\partial S} = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_m \end{pmatrix} + \begin{pmatrix} R_{S_1} & & \\ & \ddots & \\ & & R_{S_m} \end{pmatrix} \cdot \Omega \otimes I_n \tag{3.65}$$

with block matrices $B_i = \operatorname{Diag}\big((\Omega S)_i\big) - \langle S_i, (\Omega S)_i\rangle I_n - S_i(\Omega S)_i^\top$ and R_{S_i} given by (2.91)*.*

Proof. The subvectors of F have the form

$$F_i(S) = R_{S_i}(\Omega S)_i = \big(\operatorname{Diag}(S_i) - S_i S_i^\top\big)(\Omega S)_i, \qquad i \in I. \tag{3.66}$$

Hence, the directional derivative of F_i along $V \in \mathbb{R}^{m\times n}$ is given by

$$\begin{aligned} \mathrm{d}F_i(S)[V] &= \tfrac{\mathrm{d}}{\mathrm{d}t} F_i(S+tV)|_{t=0} && \text{(3.67a)}\\ &= \big(\operatorname{Diag}(V_i) - V_i S_i^\top - S_i V_i^\top\big)(\Omega S)_i + R_{S_i}(\Omega V)_i && \text{(3.67b)}\\ &= \big(\operatorname{Diag}\big((\Omega S)_i\big) - \langle S_i, (\Omega S)_i\rangle I_n - S_i(\Omega S)_i^\top\big)V_i + R_{S_i}(\Omega V)_i && \text{(3.67c)}\\ &= B_i V_i + R_{S_i}(\Omega V)_i. && \text{(3.67d)} \end{aligned}$$

We have $\mathrm{d}F(S)[V] = \frac{\partial F}{\partial S}\operatorname{vec}(V)$ with $\operatorname{vec}(V) \in \mathbb{R}^{mn}$ denoting the vector that results from stacking the row vectors (subvectors) of V. Comparing both sides of this equation, with the block matrices of the left-hand side given by (3.67), implies (3.65). □

Proposition 3.4.4 (eigenvalues of the Jacobian matrix)**.** *Let $S^* \in \overline{\mathcal{W}}$ be an equilibrium of the S-flow* (3.59)*, i.e. $F(S^*) = R_{S^*}[\Omega S^*] = 0$. Then regarding the spectrum $\sigma\big(\frac{\partial F}{\partial S}(S^*)\big)$, the following assertions hold.*

(a) A subset of the spectrum is given by

$$\sigma\big(\tfrac{\partial F}{\partial S}(S^*)\big) \supseteq \bigcup_{i\in I}\big\{-\langle S_i^*, (\Omega S^*)_i\rangle\big\} \cup \big\{(\Omega S^*)_{ij} - \langle S_i^*, (\Omega S^*)_i\rangle\big\}_{j\in J\setminus\operatorname{supp}(S_i^*)}. \tag{3.68}$$

This relation becomes an equation if S^* is integral, i.e. $S^* \in \mathcal{W}^*$.

(b) If $S^* = \frac{1}{|J_+|}\mathbb{1}_m\mathbb{1}_{J_+}^\top$ with $J_+ \subseteq J$ and $|J_+| \geq 2$, then

$$\sigma\big(\tfrac{\partial F}{\partial S}(S^*)\big) = \bigcup_{i\in I}\Big\{-\tfrac{(\Omega\mathbb{1}_n)_i}{|J_+|}\Big\} \cup \bigcup_{\lambda\in\sigma(\Omega)}\big\{\tfrac{\lambda}{|J_+|}\big\}. \tag{3.69}$$

(c) Assume the parameter matrix Ω with nonnegative diagonal elements ω_{ii}, $i \in I$. If $S_i^* \notin \{0,1\}^n$ and $\omega_{ii} > 0$ hold for some $i \in I$, then the Jacobian matrix has at least one eigenvalue with positive real part.

Proof.

(a) Since $\sigma\big(\frac{\partial F}{\partial S}(S^*)^\top\big) = \sigma\big(\frac{\partial F}{\partial S}(S^*)\big)$, we may alternatively regard the transpose of the Jacobian

$$\frac{\partial F}{\partial S}(S^*)^\top = \begin{pmatrix} B_1^\top & & \\ & \ddots & \\ & & B_m^\top \end{pmatrix} + \Omega^\top \otimes I_n \cdot \begin{pmatrix} R_{S_1^*} & & \\ & \ddots & \\ & & R_{S_m^*} \end{pmatrix} \tag{3.70}$$

with $B_i^\top = \operatorname{Diag}\big((\Omega S^*)_i\big) - \langle S_i^*, (\Omega S^*)_i\rangle I_n - (\Omega S^*)_i S_i^{*\top}$. We have for each $i \in I$,

$$B_i^\top \mathbb{1}_n = -\langle S_i^*, (\Omega S^*)_i\rangle \mathbb{1}_n, \qquad R_{S_i^*}\mathbb{1}_n = 0, \tag{3.71a}$$

$$B_i^\top e_j = \big((\Omega S^*)_{ij} - \langle S_i^*, (\Omega S^*)_i\rangle\big) e_j, \quad R_{S_i^*} e_j = 0, \quad \forall j \in J \setminus \operatorname{supp} S_i^*. \tag{3.71b}$$

Hence, the transposed Jacobian possesses the following eigenpairs:

$$\frac{\partial F}{\partial S}(S^*)^\top \cdot e_i \otimes \mathbb{1}_n = -\langle S_i^*, (\Omega S^*)_i\rangle \cdot e_i \otimes \mathbb{1}_n, \quad \forall i \in I, \tag{3.72a}$$

$$\frac{\partial F}{\partial S}(S^*)^\top \cdot e_i \otimes e_j = \big((\Omega S^*)_{ij} - \langle S_i^*, (\Omega S^*)_i\rangle\big) \cdot e_i \otimes e_j, \quad \forall j \in J \setminus \operatorname{supp} S_i^*, \quad \forall i \in I. \tag{3.72b}$$

If $S^* \in \mathcal{W}^*$, then $|\operatorname{supp} S_i^*| = 1$ for each $i \in I$ and therefore (3.72) specifies all mn eigenpairs and the entire spectrum, which proves (3.68). In this case, the eigenvectors of $\frac{\partial F}{\partial S}(S^*)$ can also be stated explicitly: Since $R_{S^*} = 0$, we have

$$\frac{\partial F}{\partial S}(S^*) = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_m \end{pmatrix}. \tag{3.73}$$

Each block B_i fulfills

$$B_i S_i^* = -\langle S_i^*, (\Omega S^*)_i\rangle S_i^*, \tag{3.74a}$$

$$B_i(e_j - S_i^*) = \big((\Omega S^*)_{ij} - \langle S_i^*, (\Omega S^*)_i\rangle\big)(e_j - S_i^*) \quad \forall j \in J \setminus \operatorname{supp} S_i^*. \tag{3.74b}$$

Hence, the corresponding eigenvectors of $\frac{\partial F}{\partial S}(S^*)$ are

$$e_i \otimes S_i^*, \quad e_i \otimes (e_j - S_i^*), \quad \forall j \in J \setminus \operatorname{supp} S_i^*, \quad \forall i \in I. \tag{3.75}$$

(b) Since $\Omega S^* = \frac{1}{|J_+|}(\Omega \mathbb{1}_m)\mathbb{1}_{J_+}^\top$ for $S^* = \frac{1}{|J_+|}\mathbb{1}_m \mathbb{1}_{J_+}^\top$, we have

$$B_i = (\Omega \mathbb{1}_m)_i \cdot \Big(\tfrac{1}{|J_+|}\operatorname{Diag}(\mathbb{1}_{J_+}) - \tfrac{1}{|J_+|} I_n - \tfrac{1}{|J_+|^2}\mathbb{1}_{J_+}\mathbb{1}_{J_+}^\top\Big), \tag{3.76}$$

$$R_{S_i^*} = \tfrac{1}{|J_+|}\operatorname{Diag}(\mathbb{1}_{J_+}) - \tfrac{1}{|J_+|^2}\mathbb{1}_{J_+}\mathbb{1}_{J_+}^\top \tag{3.77}$$

for all $i \in I$, i.e., the Jacobian matrix simplifies to

$$\frac{\partial F}{\partial S}(S^*) = \operatorname{Diag}(\Omega \mathbb{1}_m) \otimes B_0 + \Omega \otimes R_{S_1}$$
$$\text{with} \quad B_0 = \tfrac{1}{|J_+|}\operatorname{Diag}(\mathbb{1}_{J_+}) - \tfrac{1}{|J_+|} I_n - \tfrac{1}{|J_+|^2}\mathbb{1}_{J_+}\mathbb{1}_{J_+}^\top. \tag{3.78}$$

Let $\{(\lambda_i, w_i)\}_{i\in\tilde{I}} \subset \mathbb{C}\times\mathbb{C}^m$ be the set of all eigenpairs of Ω indexed by $\tilde{I}$, and let $\{v_1,\dots,v_{|J_+|-1}\}$ be a basis of $\{v\in\mathbb{R}^n\colon \langle v, \mathbb{1}_{J_+}\rangle = 0,\ \operatorname{supp}(v)\subseteq J_+\}$. Note that $|\tilde{I}| < m$ if and only if Ω is not diagonalizable. A short calculation shows

$$B_0 e_j = -\tfrac{1}{|J_+|}e_j, \qquad R_{S_1} e_j = 0, \qquad \forall j\in J\setminus J_+, \tag{3.79a}$$

$$B_0 \mathbb{1}_{J_+} = -\tfrac{1}{|J_+|}\mathbb{1}_{J_+}, \quad R_{S_1}\mathbb{1}_{J_+} = 0, \tag{3.79b}$$

$$B_0 v_j = 0, \qquad R_{S_1} v_j = \tfrac{1}{|J_+|}v_j, \quad \forall j\in\{1,\dots,|J_+|-1\}. \tag{3.79c}$$

Hence, the Jacobian has the following $mn-(m-|\tilde{I}|)(|J_+|-1)$ eigenpairs:

$$\Big(-\tfrac{(\Omega\mathbb{1}_m)_i}{|J_+|}, e_i\otimes e_j\Big), \quad \forall j\in J\setminus J_+, \quad \forall i\in I, \tag{3.80a}$$

$$\Big(-\tfrac{(\Omega\mathbb{1}_m)_i}{|J_+|}, e_i\otimes \mathbb{1}_{J_+}\Big), \quad \forall i\in I, \tag{3.80b}$$

$$\Big(\tfrac{\lambda_i}{|J_+|}, w_i\otimes v_j\Big), \quad \forall j\in\{1,\dots,|J_+|-1\}, \quad \forall i\in\tilde{I}. \tag{3.80c}$$

If $|\tilde{I}| = m$, we thus have a complete set of mn eigenpairs. If $|\tilde{I}| < m$, we may consider a diagonalizable perturbation $\widetilde{\Omega}$ of Ω. By the same argument, we get a complete set of eigenpairs for the perturbed Jacobian matrix. Consequently, we obtain (3.69) by continuity of the spectrum.

(c) We show that the real and imaginary parts of the corresponding eigenvector lie in the linear subspace

$$\mathcal{T}_+ = \mathcal{T}_+(S^*) = \big\{V\in\mathbb{R}^{mn}\colon \langle V_i, \mathbb{1}_n\rangle = 0,\ \operatorname{supp} V_i\subseteq \operatorname{supp} S_i^*,\ \forall i\in I\big\} \tag{3.81}$$

To this end, we show the two inclusions

$$\operatorname{im} R_{S^*} \subseteq \mathcal{T}_+ \subseteq \ker B, \tag{3.82}$$

where R_{S^*} and B denote the block diagonal matrices

$$B = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_m \end{pmatrix}, \quad R_{S^*} = \begin{pmatrix} R_{S_1^*} & & \\ & \ddots & \\ & & R_{S_m^*} \end{pmatrix}. \tag{3.83}$$

As for the first inclusion, we use the orthogonal projection onto $\mathcal{T}_+$ given by

$$\Pi_{\mathcal{T}_+} = \begin{pmatrix} \Pi_{\mathcal{T}_+,1} & & \\ & \ddots & \\ & & \Pi_{\mathcal{T}_+,m} \end{pmatrix} \quad \text{with} \quad \Pi_{\mathcal{T}_+,i} = \operatorname{Diag}(\mathbb{1}_{J_i}) - \tfrac{1}{|J_i|}\mathbb{1}_{J_i}\mathbb{1}_{J_i}^\top, \quad J_i = \operatorname{supp} S_i^* \quad \forall i\in I. \tag{3.84}$$

One can verify that $\Pi_{\mathcal{T}_+} R_{S^*} = R_{S^*}$ which implies $\operatorname{im} R_{S^*}\subseteq \operatorname{im}\Pi_{\mathcal{T}_+} = \mathcal{T}_+$, i.e. the first inclusion of (3.82).

As for the second inclusion, we have to take into account that S^* is an equilibrium point, i.e. by (3.60)

$$(\Omega S^*)_{ij} = \langle S_i^*, (\Omega S^*)_i\rangle \quad \forall j\in\operatorname{supp} S_i^* \quad \forall i\in I. \tag{3.85}$$

Since B is a block diagonal matrix, it suffices to examine each block

$$B_i = \operatorname{Diag}\big((\Omega S^*)_i\big) - \langle S_i^*, (\Omega S^*)_i\rangle I_n - S_i^* (\Omega S^*)_i^\top \tag{3.86}$$

separately. Since

$$B_i e_j = (\Omega S^*)_{ij} e_j - \langle S_i^*, (\Omega S^*)_i\rangle e_j - (\Omega S^*)_{ij} S_i^* = -\langle S_i^*, (\Omega S^*)_i\rangle S_i^*, \quad \forall j \in \operatorname{supp} S_i^* \tag{3.87}$$

is independent of $j \in \operatorname{supp} S_i^*$, we get $B_i v = 0$ for any $v \in \mathbb{R}^n$ with $\langle v, \mathbb{1}_n\rangle = 0$ and $\operatorname{supp} v \subseteq \operatorname{supp} S_i^*$. This verifies the second inclusion of (3.82).

As a consequence of the two inclusions (3.82), any eigenvector V of $R_{S^*}(\Omega \otimes I_n)$ corresponding to a nonvanishing eigenvalue $\lambda \neq 0$ has a real and imaginary part lying in $\operatorname{im} R_{S^*} \subseteq \mathcal{T}_+ \subseteq \ker B$. Therefore, (λ, V) is also an eigenpair of $\frac{\partial F}{\partial S}(S^*) = B + R_{S^*}(\Omega \otimes I_n)$. It remains to show that

$$R_{S^*}(\Omega \otimes I_n) = \begin{pmatrix} \omega_{11} R_{S_1^*} & \cdots & \omega_{1m} R_{S_1^*} \\ \vdots & & \vdots \\ \omega_{m1} R_{S_m^*} & \cdots & \omega_{mm} R_{S_m^*} \end{pmatrix} \tag{3.88}$$

has at least one eigenvalue with positive real part. Since the trace

$$\operatorname{tr}\big(R_{S^*}(\Omega \otimes I_n)\big) = \sum_{i\in I} \omega_{ii} \operatorname{tr}\big(R_{S_i^*}\big) = \sum_{i \in I} \underbrace{\omega_{ii} \sum_{j \in J} (S_{ij}^* - S_{ij}^{*\,2})}_{\begin{cases} \geq 0, & \forall i \in I \\ > 0, & \text{for some } i \in I \end{cases}} > 0 \tag{3.89}$$

is positive by assumption, the existence of such an eigenvalue is guaranteed.

□

Next, we apply Theorem 2.3.8 and Proposition 3.4.4 in order to classify the equilibria of the S-flow.

Corollary 3.4.5 (stabilty of equilibria)**.** *Let Ω be a nonnegative matrix with positive diagonal entries. Then, regarding the equilibria $S^* \in \overline{\mathcal{W}}$ of the S-flow* (3.59)*, the following assertions hold.*

(a) $S^* \in \mathcal{W}^*$ is exponentially stable if, for all $i \in I$,

$$(\Omega S^*)_{ij} < (\Omega S^*)_{ij^*(i)} \quad \text{for all } j \in J \setminus \{j^*(i)\} \quad \text{with} \quad \{j^*(i)\} = \operatorname*{arg\,max}_{j \in J} S_{ij}^*. \tag{3.90}$$

(b) $S^* \in \mathcal{W}^*$ is unstable if, for some $i \in I$,

$$(\Omega S^*)_{ij} > (\Omega S^*)_{ij^*(i)} \quad \text{for some } j \in J \setminus \{j^*(i)\} \quad \text{with} \quad \{j^*(i)\} = \operatorname*{arg\,max}_{j \in J} S_{ij}^*. \tag{3.91}$$

(c) All equilibria $S^* \notin \mathcal{W}^*$ are unstable.

Proof.

(a) We apply Theorem 2.3.8(a) that provides a condition for stability of the S-flow, regarded as flow on an open subset of $\mathbb{R}^{m\times n}$. Since the stability also holds on subsets, this shows stability of the S-flow on $\overline{\mathcal{W}}$.

By Proposition 3.4.4(a), the spectrum of $\frac{\partial F}{\partial S}(S^*)$, for $S^* \in \mathcal{W}^*$, is given by the right-hand side of (3.68) and, since Ω is nonnegative, is clearly negative if condition (3.90) holds.

(b) We take eigenvectors into account and invoke Proposition 2.3.9(b). By (3.75), the eigenvectors are

$$e_i e_{j^*(i)}^\top \in \mathbb{R}^{m\times n}, \quad e_i(e_j - e_{j^*(i)})^\top \in \mathcal{T}_0, \quad \forall j \in J \setminus \{j^*(i)\}, \quad \forall i \in I, \tag{3.92}$$

and if the eigenvalue $\lambda = (\Omega S^*)_{ij} - (\Omega S^*)_{ij^*(i)}$ is positive, then the corresponding eigenvector $V = e_i(e_j - e_{j^*(i)})^\top \in \mathcal{T}_0$ is tangent to $\overline{\mathcal{W}}$ at S^*. By Proposition 2.3.9(b), there exists an open truncated cone $C \subset \mathbb{R}^{m\times n}$ such that $\delta \cdot V \in C$, for sufficiently small $\delta > 0$, and the S-flow (3.59) is repelled from S^* within $S^* + C$. The (relatively) open subset $(S^* + C) \cap \overline{\mathcal{W}} \subset \overline{\mathcal{W}}$ is non-empty, since it contains $S^* + \delta \cdot V$. This shows the instability of S^*.

(c) By the assumption on Ω, there is an eigenvalue with positive real part due to Proposition 3.4.4(c), and the real and imaginary part of the corresponding eigenvector lie in

$$\mathcal{T}_+ = \left\{V \in \mathcal{T}_0 \colon \operatorname{supp} V \subseteq \operatorname{supp} S^*\right\} \subseteq \mathcal{T}_0. \tag{3.93}$$

So the argument of (b) applies here as well using the real part of the eigenvector.

□

Remark 3.4.6 (selection of stable equilibria)**.** For S^* to be exponentially stable, Corollary 3.4.5(a) requires that every averaged subvector $(\Omega S^*)_i$ has the same component as maximal component, as does the corresponding subvector S_i^*. This means that the Ω-weighted average of the vectors S_j^* within the neighborhood $j \in \mathcal{N}_i$ lies in the Voronoi-cell $\mathcal{V}_{j^*(i)}$ (3.5) corresponding to S_i^*. Thus, Corollary 3.4.5 provides a mathematical and intuitively plausible definition of 'spatially coherent' segmentations of given data, that can be determined by means of the assignment flow.

We now turn our attention to the consequences of the instability of equilibria.

Proposition 3.4.7. *Let Ω be nonnegative with positive diagonal entries, and let $S^* \in \overline{\mathcal{W}}$ be an equilibrium point of the S-flow* (3.59) *which satisfies one of the instability criteria of Corollary 3.4.5 (b) or (c). Then the set of starting points $S_0 \in \mathcal{W}$ for which the S-flow converges to S^* has measure zero in $\mathcal{W}$.*

Proof. By Theorem 2.3.7, there exists a center-stable manifold $\mathcal{M}_{\mathrm{cs}}(S^*)$ which is invariant under the S-flow and tangent to $E_{\mathrm{c}} \oplus E_{\mathrm{s}}$ at S^*. Here, E_{c} and E_{s} denote the center and stable subspace of $\frac{\partial F}{\partial S}(S^*)$, respectively. Any trajectory of the S-flow converging to S^* lies in $\mathcal{M}_{\mathrm{cs}}(S^*)$. Therefore, it suffices to show that the dimension of the manifold $\mathcal{M}_{\mathrm{cs}}(S^*) \cap \mathcal{W}$ is smaller than

the dimension of $\mathcal{W}$. Note that $\mathcal{M}_{\mathrm{cs}}(S^*) \cap \mathcal{W}$ is a manifold since both $\mathcal{M}_{\mathrm{cs}}(S^*)$ and $\mathcal{W}$ are invariant under the S-flow. We have

$$\dim\big(\mathcal{M}_{\mathrm{cs}}(S^*) \cap \mathcal{W}\big) = \dim\big((E_{\mathrm{c}} \oplus E_{\mathrm{s}}) \cap \mathcal{T}_0\big) = \dim(\mathcal{T}_0) - \dim(E_{\mathrm{u}} \cap \mathcal{T}_0) \tag{3.94a}$$

$$= \dim(\mathcal{W}) - \dim(E_{\mathrm{u}} \cap \mathcal{T}_0), \tag{3.94b}$$

where E_{u} denotes the unstable subspace of $\frac{\partial F}{\partial S}(S^*)$. Since $\frac{\partial F}{\partial S}(S^*)$ has an eigenvalue with positive real part and a corresponding eigenvector lying in $\mathcal{T}_0$ (cf. proof of Corollary 3.4.5), we have $\dim(E_{\mathrm{u}} \cap \mathcal{T}_0) \geq 1$ and therefore $\dim\big(\mathcal{M}_{\mathrm{cs}}(S^*) \cap \mathcal{W}\big) \leq \dim(\mathcal{W}) - 1$. □

Theorem 3.4.8. *Let Ω be a nonnegative matrix with positive diagonal entries. Then the set of starting points $S_0 \in \mathcal{W}$ for which the S-flow* (3.59) *converges to a nonintegral equilibrium $S^* \in \overline{\mathcal{W}}$, has measure zero in $\mathcal{W}$.*

Proof. Let $\mathcal{E}$ be the set of all equilibria of the S-flow in $\overline{\mathcal{W}}$, which is a compact subset of $\overline{\mathcal{W}}$. If $\mathcal{E}$ contains only isolated points, i.e., $\mathcal{E}$ is finite, then the statement follows from Proposition 3.4.7. In order to take also into account nonfinite sets $\mathcal{E}$ of equilibria, we apply the more general [Fen79, Theorem 9.1]. Some additional notation is introduced first.

Let $\mathcal{E}_\Sigma$ be the set of equilibria with given support $\Sigma \subseteq I \times J$ as in Proposition 3.4.1(e). The set $\mathcal{E}_\Sigma$ is the relative interior of a convex polytope and therefore a manifold of equilibria. Further, we define for $n_{\mathrm{s}}, n_{\mathrm{c}}, n_{\mathrm{u}} \in \mathbb{N}_0$ with $n_{\mathrm{s}} + n_{\mathrm{c}} + n_{\mathrm{u}} = mn$ the set

$$\mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})} = \big\{S^* \in \mathcal{E} \colon \dim E_{\mathrm{s}}(S^*) = n_{\mathrm{s}},\ \dim E_{\mathrm{c}}(S^*) = n_{\mathrm{c}},\ \dim E_{\mathrm{u}}(S^*) = n_{\mathrm{u}}\big\}, \tag{3.95}$$

where $E_{\mathrm{c}}(S^*)$, $E_{\mathrm{s}}(S^*)$ and $E_{\mathrm{u}}(S^*)$ denote the center, stable and unstable subspace of $\frac{\partial F}{\partial S}(S^*)$. This set can be written as countable union of compact sets which can be seen as follows. The map

$$\mathcal{E} \to \big\{x \in \mathbb{R}^{mn} \colon x_1 \leq x_2 \leq \cdots \leq x_{mn}\big\}, \qquad S^* \mapsto \mathrm{Re}\left(\lambda\big(\tfrac{\partial F}{\partial S}(S^*)\big)\right), \tag{3.96}$$

where $\lambda(\cdot)$ denotes the vector of eigenvalues, is a continuous map[2] on a compact set and therefore proper, i.e., preimages of compact sets under the map (3.96) are compact. It is clear that the set $U_{\mathrm{s}} \times U_{\mathrm{c}} \times U_{\mathrm{u}}$ with

$$U_{\mathrm{s}} = \big\{x \in \mathbb{R}^{n_{\mathrm{s}}} \colon x_1 \leq \cdots \leq x_{n_s} < 0\big\}, \tag{3.97a}$$

$$U_{\mathrm{c}} = \big\{x \in \mathbb{R}^{n_{\mathrm{c}}} \colon x = 0\big\}, \tag{3.97b}$$

$$U_{\mathrm{u}} = \big\{x \in \mathbb{R}^{n_{\mathrm{u}}} \colon 0 < x_1 \leq \cdots \leq x_{n_u}\big\} \tag{3.97c}$$

can be written as countable union of compact sets. The preimage of this set under the map (3.96) is $\mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})}$.

To complete the proof, we now argue similar to the proof of Proposition 3.4.7: the existence of nontrivial unstable subspaces for nonintegral equilibria implies that the center-stable manifold has a smaller dimension.

Let Σ be the support of any nonintegral equilibrium and let $\mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})}$ be such that the intersection $\mathcal{E}_\Sigma \cap \mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})} \neq \emptyset$ is non-empty. As seen in the proof of Corollary 3.4.5(c), we

[2] This follows by the topological continuity of the spectrum – cf. e.g. [HJ12, Appendix D].

have $E_{\mathrm{u}}(S^*) \cap \mathcal{T}_0 \neq \{0\}$ for any $S^* \in \mathcal{E}_\Sigma$, i.e. $n_{\mathrm{u}} \geq 1$. Since both $\mathcal{E}_\Sigma$ and $\mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})}$ can be written as countable union of compact sets, this is also the case for their intersection, i.e., we have

$$\mathcal{E}_\Sigma \cap \mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})} = \bigcup_{l \in \mathbb{N}} K_l \tag{3.98}$$

with $K_l \subseteq \mathcal{E}_\Sigma$ compact. For any $l \in \mathbb{N}$, there exists a center-stable manifold $\mathcal{M}_{\mathrm{cs}}(K_l)$ containing K_l, which is invariant under the S-flow and tangent to $E_{\mathrm{c}}(S^*) \oplus E_{\mathrm{s}}(S^*)$ at any $S^* \in K_l$ [Fen79, Theorem 9.1]. Any trajectory of the S-flow converging to a point $S^* \in K_l$ lies in $\mathcal{M}_{\mathrm{cs}}(K_l)$. Hence, analogous to the proof of Proposition 3.4.7, we have

$$\dim\big(\mathcal{M}_{\mathrm{cs}}(K_l) \cap \mathcal{W}\big) = \dim(\mathcal{W}) - \dim(E_{\mathrm{u}}(S^*) \cap \mathcal{T}_0) \leq \dim(\mathcal{W}) - 1, \tag{3.99}$$

with any $S^* \in K_l$, i.e., $\mathcal{M}_{\mathrm{cs}}(K_l) \cap \mathcal{W}$ has measure zero in $\mathcal{W}$. All trajectories converging to an equilibrium $S^* \in \mathcal{E}_\Sigma \cap \mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})}$ are contained in the countable union $\bigcup_{l\in\mathbb{N}} \mathcal{M}_{\mathrm{cs}}(K_l) \cap \mathcal{W}$, which has measure zero as well. Since there are only finitely many such sets $\mathcal{E}_\Sigma \cap \mathcal{E}_{(n_{\mathrm{s}},n_{\mathrm{c}},n_{\mathrm{u}})}$, this completes the proof. □

In view of Theorem 3.4.8, the following corollary that additionally takes into account assumption (3.17) of Theorem 3.3.2, is obvious.

Corollary 3.4.9 (convergence to integral assignments)**.** *Let Ω be a nonnegative matrix with positive diagonal entries which also fulfills the symmetry assumption* (3.17)*. Then the set of starting points $S_0 \in \mathcal{W}$, for which the S-flow* (3.59) *does not converge to an integral assignment $S^* \in \mathcal{W}^*$, has measure zero. If Ω is additionally invertible, then the set of distance matrices $D \in \mathbb{R}^{m\times n}$ for which the assignment flow* (3.1) *does not converge to an integral assignment has measure zero as well.*

We briefly discuss the assumption of Theorem 3.4.8. We consider a small dynamical system that violates the basic assumption, that all diagonal entries of the parameter matrix Ω of the S-flow (3.59) are positive. As a consequence, an entire line of nonintegral equilibrium points S^* is locally attracting the flow.

Example 3.4.10. Let $m = |I| = 3$ and $n = |J| = 2$, and let the parameters of the S-flow (3.59) be given by the row-stochastic matrix

$$\Omega = (\omega_{ik})_{i,k\in[3]} = \frac{1}{4}\begin{pmatrix} 0 & 2 & 2 \\ 1 & 2 & 1 \\ 1 & 1 & 2 \end{pmatrix}, \quad i \in I. \tag{3.100}$$

One easily checks that any point S^* on the line

$$\mathcal{L} = \left\{ \begin{pmatrix} p & 1-p \\ 1 & 0 \\ 0 & 1 \end{pmatrix} : p \in [0,1] \right\} \subset \overline{\mathcal{W}} \tag{3.101}$$

is an equilibrium of the S-flow satisfying $F(S^*) = 0$. In particular, this includes nonintegral points with $p \in (0,1)$. The eigenvalues of the Jacobian matrix are given by[3]

$$\sigma\big(\tfrac{\partial F}{\partial S}(S^*)\big) = \big\{0, -\tfrac{1}{2}, -\tfrac{p+2}{4}, -\tfrac{p}{2}, -\tfrac{1-p}{2}, -\tfrac{3-p}{4}\big\} \subset \mathbb{R}_{\leq 0} \tag{3.102}$$

[3] The eigenvalue 0 results from non-isolated equilibria, while the remaining five eigenvalues are given by Proposition 3.4.4(a).

and are nonpositive. The phase portrait depicted by Figure 3.3 illustrates that $\mathcal{L}$ locally attracts the flow.

This small example demonstrates that violation of the basic assumption of Theorem 3.4.8—here, specifically, ω_{11} of (3.100) is *not* positive—leads to S-flows for which Theorem 3.4.8 does not apply: there is an (relatively) open set of starting points $S_0 \in \mathcal{W}$ for which the S-flow converges to nonintegral equilibria $S^* \in \mathcal{L}$.

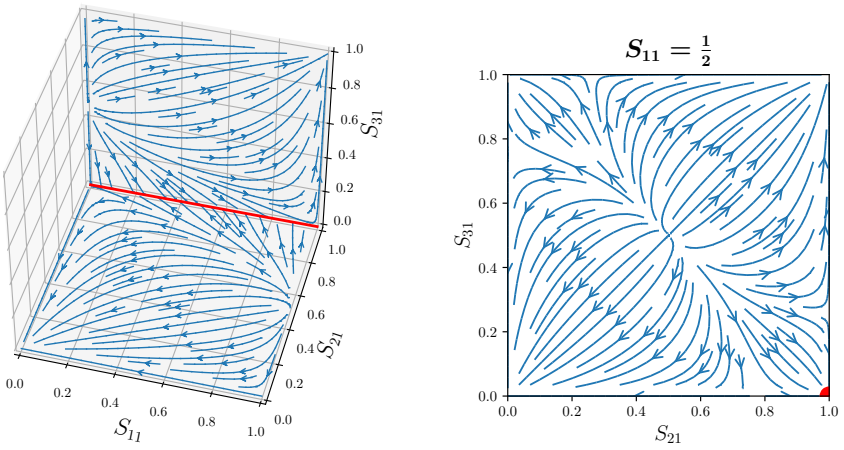


Figure 3.3: Phase portrait for the flow of Example 3.4.10. We graphically depict the S-flow with Ω given by (3.100), by its first column. This describes the flow completely, since $n = |J| = 2$. The left panel shows the phase portrait of the flow within the planes $\{S_{21} = 1\}$ and $\{S_{31} = 0\}$. The plane $\{S_{11} = \frac{1}{2}\}$ is depicted by the right panel. The line $\mathcal{L}$ of equilibria given by (3.101) is marked red and located in the lower right vertex in the right plot. The phase portrait illustrates that this line attracts the flow within a small neighborhood.

We end this subsection with a remark on hyperbolic equilibria.

Remark 3.4.11 (consequences of hyperbolicity)**.** If $S^* \in \overline{\mathcal{W}}$ is a *hyperbolic* equilibrium point, i.e., the Jacobian matrix has no eigenvalue with vanishing real part, then the S-flow locally behaves as its linearization near S^* by the Hartman-Grobman theorem (Theorem 2.3.5). Since a linear flow can only converge with an exponential convergence rate, this is also the case for the S-flow (3.59). More precisely, if the S-flow converges to a hyperbolic equilibrium $S^* \in \overline{\mathcal{W}}$ then there exist $\alpha, \beta > 0$ such that $\|S(t) - S^*\| \leq \alpha e^{-\beta t}$ irrespective of whether S^* is stable or not. A direct consequence is $\int_0^\infty \|S_i(t) - S_i^*\|_1 \mathrm{d}t < \infty$ for all $i \in I$, i.e., assumption of Proposition 3.2.5(b) automatically holds if S^* is hyperbolic.

3.4.2 Estimation of Basins of Attraction

We have shown in Corollary 3.4.5 that if a point $S^* \in \mathcal{W}^*$ satisfies the stability criterion (3.90) then there exists an open neighborhood of S^* such that the S-flow starting in this neighborhood eventually converge to S^* with an exponential convergence rate. In this subsection we will quantify this statement.

Proposition 3.4.12 (attracting region). *Let Ω be a nonnegative matrix with positive diagonal entries, and let $S^* \in \mathcal{W}^*$ satisfy* (3.90). *Furthermore, set*

$$A(S^*) \coloneqq \bigcap_{i\in I} \bigcap_{j\neq j^*(i)} \left\{S \in \mathbb{R}^{m\times n} \colon (\Omega S)_{ij} < (\Omega S)_{ij^*(i)}\right\} \quad \text{with} \quad \{j^*(i)\} = \arg\max_{j\in J} S^*_{ij}, \tag{3.103}$$

which is an open convex polytope containing S^*. Finally, let $\varepsilon > 0$ be small enough such that

$$B_\varepsilon(S^*) \coloneqq \left\{S \in \overline{\mathcal{W}} \colon \max_{i\in I} \|S_i - S_i^*\|_1 < \varepsilon\right\} \subset \left(A(S^*) \cap \overline{\mathcal{W}}\right). \tag{3.104}$$

Then, regarding the S-flow (3.59), *the following holds: If $S(t_0) \in B_\varepsilon(S^*)$ for some point in time t_0, then $S(t) \in B_\varepsilon(S^*)$ for all $t \geq t_0$ and $\lim_{t\to\infty} S(t) = S^*$. Moreover, we have*

$$\|S_i(t) - S_i^*\|_1 \leq \|S_i(t_0) - S_i^*\|_1 \cdot e^{-\beta_i (t-t_0)}, \quad \forall i \in I, \tag{3.105a}$$

where

$$\beta_i = \min_{S\in\overline{B_\delta(S^*)}\cap\overline{\mathcal{W}}} S_{ij^*(i)} \cdot \min_{j\neq j^*(i)} \left((\Omega S)_{ij^*(i)} - (\Omega S)_{ij}\right) > 0 \tag{3.105b}$$

and $\delta > 0$ is chosen small enough such that $S(t_0) \in \overline{B_\delta(S^*)} \subset B_\varepsilon(S^*)$.

Proof. For each $i \in I$, we have with $S_i^* = e_{j^*(i)}$

$$\begin{aligned}
\frac{\mathrm{d}}{\mathrm{d}t}\|S_i - S_i^*\|_1 &\overset{\sum_j S_{ij}=1}{=} \frac{\mathrm{d}}{\mathrm{d}t}\Big(1 - S_{ij^*(i)} + \sum_{j\neq j^*(i)} S_{ij}\Big) && (3.106a)\\
&= \frac{\mathrm{d}}{\mathrm{d}t}(2 - 2S_{ij^*(i)}) && (3.106b)\\
&\overset{(3.2a)}{=} -2S_{ij^*(i)}\left((\Omega S)_{ij^*(i)} - \langle S_i, (\Omega S)_i\rangle\right) && (3.106c)\\
&\leq -2S_{ij^*(i)}\Big((\Omega S)_{ij^*(i)} - S_{ij^*(i)}(\Omega S)_{ij^*(i)} - \max_{j\neq j^*(i)} (\Omega S)_{ij} \sum_{j\neq j^*(i)} S_{ij}\Big) && (3.106d)\\
&= -2S_{ij^*(i)}(1 - S_{ij^*(i)})\Big((\Omega S)_{ij^*(i)} - \max_{j\neq j^*(i)} (\Omega S)_{ij}\Big) && (3.106e)\\
&\overset{(3.106b)}{=} -S_{ij^*(i)}\|S_i - S_i^*\|_1 \cdot \min_{j\neq j^*(i)} \left((\Omega S)_{ij^*(i)} - (\Omega S)_{ij}\right). && (3.106f)
\end{aligned}$$

Choosing $\delta > 0$ such that $S(t_0) \in \overline{B_\delta(S^*)} \subset B_\varepsilon(S^*) \subset A(S^*)$, it follows that β_i given by (3.105b) is positive. Consequently

$$\frac{\mathrm{d}}{\mathrm{d}t}\|S_i - S_i^*\|_1 \leq -\beta_i \|S_i - S_i^*\|_1 \tag{3.107}$$

and by Gronwall's Lemma (3.105a) holds. Hence, $\max_{i\in I} \|S_i - S_i^*\|_1$ monotonically decreases as long as $S(t) \in \overline{B_\delta(S^*)}$. This guarantees that $S(t)$ stays in $\overline{B_\delta(S^*)} \subset B_\varepsilon(S^*)$ and converges towards S^*. □

Note that if $S(t)$ is close to S^*, then the convergence rate (3.105) of $S(t)$ is approximately governed by

$$\beta_i \approx \min_{j\neq j^*(i)} \left((\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij}\right). \tag{3.108}$$

Proposition 3.4.12 provides a criterion for terminating the numerical integration of the S-flow and subsequent 'save' rounding to an integral solution. For this purpose, the following proposition provides an estimate of ε defining (3.104).

Proposition 3.4.13. *Let the same assumptions hold for $\Omega = (\omega_{ik}) \in \mathbb{R}^{m\times m}$ and $S^* \in \mathcal{W}^*$ as in Proposition 3.4.12. A value $\varepsilon > 0$ that is sufficient small for the inclusion* (3.104) *to hold, is given by*

$$\varepsilon_{\text{est}} = \min_{i\in I} \min_{j\neq j^*(i)} 2 \cdot \frac{(\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij}}{(\Omega \mathbb{1}_m)_i + (\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij}} > 0. \tag{3.109}$$

Proof. Let $S \in \overline{\mathcal{W}}$ be a point such that

$$\max_{i\in I} \|S_i - S_i^*\|_1 < \varepsilon = \varepsilon_{\text{est}}. \tag{3.110}$$

We have to show that $S \in A(S^*)$, with $A(S^*)$ given by (3.103).

Since $\|S_i - S_i^*\|_1 = 2 - 2S_{ij^*(i)}$, we have

$$S_{ij^*(i)} > 1 - \frac{\varepsilon}{2}, \qquad S_{ij} \leq \sum_{l\neq j^*(i)} S_{il} = 1 - S_{ij^*(i)} < \frac{\varepsilon}{2}, \quad \forall j \neq j^*(i). \tag{3.111}$$

Hence, for any $i \in I$ and any $j \neq j^*(i)$, we get with $j^*(k)$, $k \in I$ similarly defined as $j^*(i)$ in (3.90),

$$\begin{aligned}
&(\Omega S)_{ij^*(i)} - (\Omega S)_{ij} \\
&\quad = \sum_{k\in I} \omega_{ik} S_{kj^*(i)} - \sum_{k\in I} \omega_{ik} S_{kj} && \text{(3.112a)} \\
&\quad = \sum_{\substack{k\in I \\ j^*(k)=j^*(i)}} \omega_{ik} \overbrace{S_{kj^*(i)}}^{>1-\frac{\varepsilon}{2}} + \sum_{\substack{k\in I \\ j^*(k)\neq j^*(i)}} \omega_{ik} \overbrace{S_{kj^*(i)}}^{\geq 0} - \sum_{\substack{k\in I \\ j^*(k)=j}} \omega_{ik} \overbrace{S_{kj}}^{\leq 1} - \sum_{\substack{k\in I \\ j^*(k)\neq j}} \omega_{ik} \overbrace{S_{kj}}^{<\frac{\varepsilon}{2}}, && \text{(3.112b)}
\end{aligned}$$

and by dropping the second nonnegative summand,

$$> \left(1 - \frac{\varepsilon}{2}\right) \sum_{\substack{k\in I \\ j^*(k)=j^*(i)}} \omega_{ik} - \sum_{\substack{k\in I \\ j^*(k)=j}} \omega_{ik} - \frac{\varepsilon}{2} \sum_{\substack{k\in I \\ j^*(k)\neq j}} \omega_{ik} \tag{3.112c}$$

and using that the rows of S^* are unit vectors,

$$\begin{aligned}
&= \left(1 - \frac{\varepsilon}{2}\right)(\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij} - \frac{\varepsilon}{2}\left((\Omega \mathbb{1}_m)_i - (\Omega S^*)_{ij}\right) && \text{(3.112d)} \\
&= (\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij} - \frac{\varepsilon}{2}\left((\Omega \mathbb{1}_m)_i + (\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij}\right) && \text{(3.112e)} \\
&\overset{(3.109)}{\geq} 0. && \text{(3.112f)}
\end{aligned}$$

This verifies $S \in A(S^*)$. □

Figure 3.4 illustrates the sets $A(S^*)$ and $B_\varepsilon(S^*)$ defined by (3.103) and (3.104), for some examples in the simple case of two data points and two labels. The beige and green regions in the left panel illustrate that the condition $S(t_0) \in A(S^*)$ neither guarantees that the S-flow converges to S^* nor to stay in $A(S^*)$. This demonstrates the need for the sets $B_\varepsilon(S^*)$, shown as shaded squares in Figure 3.4. Note that $B_\varepsilon(S^*) \neq \emptyset$ only if $S^* \in A(S^*) \neq \emptyset$, i.e. if the stability condition (3.90) is fulfilled.

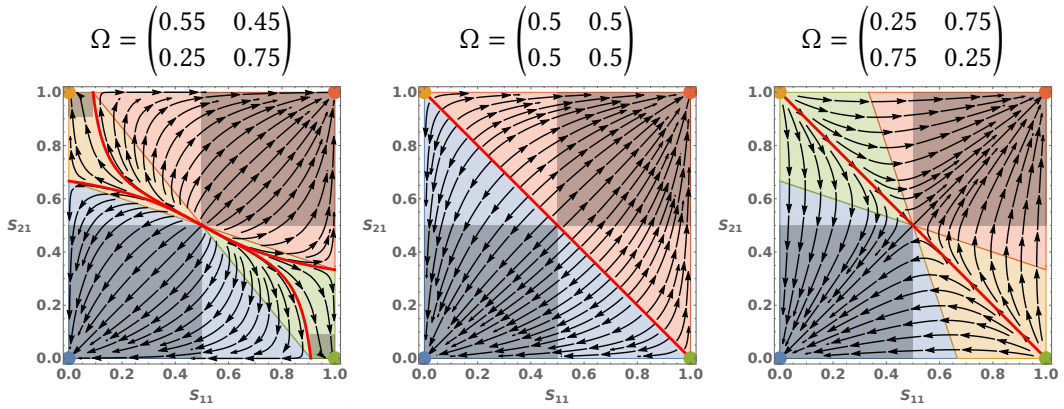


Figure 3.4: Illustration of the approximation of the basins of attraction for the case $m = n = 2$. The plots show the phase portrait of the S-flow (3.59) for three different row-stochastic matrices Ω. The four points $S^* \in \mathcal{W}^*$ are marked with $\{$●,●,●,●$\}$, the corresponding sets $A(S^*)$ (3.103) are shown as colored regions, and the balls $B_\varepsilon(S^*)$ (3.104) around the equilibria for which convergence to the equilibria is guaranteed are shown as shaded squares, with $\varepsilon = \varepsilon_{\text{est}}(S^*, \Omega)$ from (3.109). Finally, the boundary between the basins of attraction is marked with a thick red curve. In the center and right panel, only the constant labelings $S^* \in \{\left(\begin{smallmatrix} 0 & 1 \\ 0 & 1 \end{smallmatrix}\right), \left(\begin{smallmatrix} 1 & 0 \\ 1 & 0 \end{smallmatrix}\right)\}$ fulfill the stability criterion (3.90), i.e. $S^* \in A(S^*)$. As for the other two points $S^* \in \mathcal{W}^*$, we have either $A(S^*) = \emptyset$ (center panel) or $S^* \notin A(S^*) \neq \emptyset$ (right panel).

If *uniform* weights Ω are used for averaging, then the estimate (3.109) can be cast into a simple form that no longer depends on S^*.

Corollary 3.4.14. *Let $\Omega = (\omega_{ik}) \in \mathbb{R}^{m \times m}$ be given by uniform weights $\omega_{ik} = \frac{1}{|\mathcal{N}_i|}$, $k \in \mathcal{N}_i$, $i \in I$. Then the value $\varepsilon > 0$ that achieves the inclusion* (3.104) *can be chosen as*

$$\varepsilon_{\text{unif}} = \frac{2}{1 + \max_{i \in I} |\mathcal{N}_i|} > 0. \tag{3.113}$$

Proof. Let $j^*(i)$ be defined as in (3.90). We have

$$(\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij} = \frac{|\{k \in \mathcal{N}_i \colon j^*(k) = j^*(i)\}| - |\{k \in \mathcal{N}_i \colon j^*(k) = j\}|}{|\mathcal{N}_i|}. \tag{3.114}$$

Since (3.114) is positive by assumption and its numerator is integral, we get

$$(\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij} \geq \frac{1}{|\mathcal{N}_i|}. \tag{3.115}$$

Monotonicity of the function $x \mapsto \frac{x}{1+x}$ implies

$$2 \cdot \frac{(\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij}}{1 + (\Omega S^*)_{ij^*(i)} - (\Omega S^*)_{ij}} \geq 2 \cdot \frac{\frac{1}{|\mathcal{N}_i|}}{1 + \frac{1}{|\mathcal{N}_i|}} = \frac{2}{1 + |\mathcal{N}_i|} \tag{3.116}$$

and hence $\varepsilon_{\text{unif}} \leq \varepsilon_{\text{est}}$, with ε_{est} given by (3.109). The assertion, therefore, follows from Proposition 3.4.13. □

We will demonstrate the statements of Proposition 3.4.12 and Proposition 3.4.13 in Example 3.5.14 in the next section, after we have verified the validity of Proposition 3.4.12 also for the discrete scheme discussed next.

3.5 Analysis of the Geometric Euler Discretization

In this section we examine the discretization of the assignment flow. We confine ourselves to the simplest geometric scheme worked out by [Sav+17] for numerically integrating the assignment flow (3.1) which reads

$$W^{(t+1)} = \exp_{W^{(t)}}\big(hS(W^{(t)})\big), \quad t \in \mathbb{N}_0, \qquad W^{(0)} = C_{\mathcal{W}} \tag{3.117}$$

where $h > 0$ denotes a fixed step size and the iteration step $t \in \mathbb{N}_0$ represents the point of time th. The similarity map $S(W)$ (2.104) can be expressed by [SS20, Lemma 3.2]

$$S(W) = \exp_{C_{\mathcal{W}}}\big(\Omega \exp^{-1}_{C_{\mathcal{W}}}(W) - \Omega D\big), \tag{3.118}$$

where $\exp^{-1}_{C_{\mathcal{W}}} : \mathcal{W} \to \mathcal{T}_0$ is the inverse map of $\exp_{C_{\mathcal{W}}}$ restricted to $\mathcal{T}_0$ – cf. Lemma 2.2.1(a). Applying the scheme [Sav+17] to the S-flow (3.59) that has the same structure as (3.1), yields the iteration

$$S^{(t+1)} = \exp_{S^{(t)}}\big(h\Omega S^{(t)}\big), \quad t \in \mathbb{N}_0, \qquad S^{(0)} = S(C_{\mathcal{W}}) = \exp_{C_{\mathcal{W}}}(-\Omega D). \tag{3.119}$$

Analogous to Section 3.2, we establish a relation between the discrete scheme of the assignment flow (3.117) and the discrete scheme of the S-flow (3.119). Furthermore, we show that the continuous flows can be approximated arbitrarily accurately on a bounded time interval by their discretization. Finally, we will verify the stability statements in Corollary 3.4.5 and Proposition 3.4.12 also for the discrete scheme of the S-flow.

3.5.1 Relation Between the Discrete Schemes of the Two Flows

This subsection is the discrete counterpart of Section 3.2, i.e., we establish the corresponding statements for the discrete schemes (3.117) and (3.119).

Lemma 3.5.1. *The sequences $(W^{(t)})_{t\in\mathbb{N}_0}$ and $(S^{(t)})_{t\in\mathbb{N}_0}$ generated by* (3.117) *and* (3.119) *satisfy the relation*

$$S^{(t)} = S(W^{(t)}) \qquad \forall t \in \mathbb{N}_0, \tag{3.120}$$

where $S(W)$ denotes the similarity map (3.118).

Proof. We use mathematical induction. The case $t = 0$, i.e. $S^{(0)} = S(C_{\mathcal{W}}) = S(W^{(0)})$, is clear.

For $t > 0$, we have

$$S(W^{(t+1)}) \overset{(3.117)}{=} S\Big(\exp_{W^{(t)}}\big(hS(W^{(t)})\big)\Big) \tag{3.121a}$$
$$\overset{(2.99)}{=} S\Big(\exp_{C_{\mathcal{W}}}\Big(hS(W^{(t)}) + \exp_{C_{\mathcal{W}}}^{-1}(W^{(t)})\Big)\Big) \tag{3.121b}$$
$$\overset{(3.118)}{=} \exp_{C_{\mathcal{W}}}\Big(h\Omega S(W^{(t)}) + \Omega \exp_{C_{\mathcal{W}}}^{-1}(W^{(t)}) - \Omega D\Big) \tag{3.121c}$$
$$\overset{(2.99)}{=} \exp_{\exp_{C_{\mathcal{W}}}\left(\Omega \exp_{C_{\mathcal{W}}}^{-1}(W^{(t)}) - \Omega D\right)}\big(h\Omega S(W^{(t)})\big) \tag{3.121d}$$
$$\overset{(3.118)}{=} \exp_{S(W^{(t)})}\big(h\Omega S(W^{(t)})\big) \tag{3.121e}$$
$$\overset{(3.120)}{=} \exp_{S^{(t)}}\big(h\Omega S^{(t)}\big) \tag{3.121f}$$
$$\overset{(3.119)}{=} S^{(t+1)}. \tag{3.121g}$$

□

A corollary of this lemma is the following discrete version of the S-parameterization (3.2).

Corollary 3.5.2. *The scheme* (3.117) *is equivalent to the system*

$$S^{(t+1)} = \exp_{S^{(t)}}\big(h\Omega S^{(t)}\big), \quad t \in \mathbb{N}_0, \qquad S^{(0)} = \exp_{C_{\mathcal{W}}}(-\Omega D), \tag{3.122a}$$
$$W^{(t+1)} = \exp_{W^{(t)}}\big(hS^{(t)}\big), \quad t \in \mathbb{N}_0, \qquad W^{(0)} = C_{\mathcal{W}}. \tag{3.122b}$$

Remark 3.5.3. The system (3.122) is numerically more stable than (3.117), since the calculation of $S(W)$ (3.118) involves the logarithm, which causes numerical issues for assignments near the boundary of $\mathcal{W}$.

With the representation (3.122), we obtain the discrete counterparts of Proposition 3.2.2 and Proposition 3.2.5.

Proposition 3.5.4. *Let* $(S^{(t)})_{t\in\mathbb{N}_0}$ *be the sequence generated by* (3.119)*. Then the sequence* (3.117) *is given by*

$$W^{(t)} = \exp_{C_{\mathcal{W}}}\left(h\sum_{k=0}^{t-1} S^{(k)}\right) = \exp_{C_{\mathcal{W}}}\left(h\sum_{k=0}^{t-1} \Pi_0\big[S^{(k)}\big]\right). \tag{3.123}$$

Proof. We use mathematical induction. We have $W^{(0)} = \exp_{C_{\mathcal{W}}}(0) = C_{\mathcal{W}}$ and

$$\exp_{C_{\mathcal{W}}}\left(h\sum_{k=0}^{t} S^{(k)}\right) \overset{(2.99)}{=} \exp_{\exp_{C_{\mathcal{W}}}\left(h\sum_{k=0}^{t-1} S^{(k)}\right)}\big(hS^{(t)}\big) \tag{3.124a}$$
$$\overset{(3.123)}{=} \exp_{W^{(t)}}\big(hS^{(t)}\big) \tag{3.124b}$$
$$\overset{(3.122b)}{=} W^{(t+1)}. \tag{3.124c}$$

□

Proposition 3.5.5. *Let* $\mathcal{V}_i$ (3.5) *be the Voronoi cells of the vertices of* Δ_n *in* Δ_n *and suppose* $\lim_{t\to\infty} S_i^{(t)} = S_i^* \in \Delta_n$*, for any* $i \in I$*. Then the following assertions hold.*

(a) If $S_i^* \in \mathcal{V}_{j^*(i)}$ for some label index $j^* = j^*(i) \in J$, then there exist constants $\alpha_i, \beta_i > 0$ such that

$$\left\| W_i^{(t)} - e_{j^*(i)} \right\|_1 \leq \alpha_i e^{-\beta_i h t} \quad \forall t \in \mathbb{N}. \tag{3.125}$$

In particular,

$$\lim_{t\to\infty} W_i^{(t)} = e_{j^*(i)} \in \Delta_n. \tag{3.126}$$

(b) One has

$$\sum_{t\in\mathbb{N}} \left\| S_i^{(t)} - S_i^* \right\|_1 < \infty \quad \Longrightarrow \quad \lim_{t\to\infty} W_i^{(t)} = W_i^* \quad \text{with} \quad \operatorname{supp}(W_i^*) = \arg\max_{j\in J} S_{ij}^*. \tag{3.127}$$

The proof is analogous to the proof of Proposition 3.2.5, where the integral $\int_0^t$ is replaced by the sum $h\sum_{k=0}^{t-1}$. Therefore we skip the proof.

3.5.2 Approximation Error

Here, we establish estimates on the approximation error of the discrete schemes (3.117) and (3.119).

Proposition 3.5.6. *Let $L > 0$ be the Lipschitz constant of the mapping F (3.59) defining the S-flow. Then there exists a constant $C > 0$ such that the solution $S(t)$ to the S-flow (3.59) and the sequence $(S^{(t)})_{t\in\mathbb{N}_0}$ generated by (3.119) satisfy the relation*

$$\left\| S(th) - S^{(t)} \right\| \leq C_S(th)h \quad \forall t \in \mathbb{N}_0 \qquad \text{with} \quad C_S(th) = \frac{C}{2L}(1+Lh)(e^{Lth} - 1). \tag{3.128}$$

Proof. For any $t \in \mathbb{N}_0$, we set

$$Y^{(t)}(\tau) = F_\tau(S^{(t)}) \overset{(3.119)}{=} \exp_{S^{(t)}}(\tau\Omega S^{(t)}) \tag{3.129a}$$

and thus have

$$Y^{(t)}(h) = S^{(t+1)}, \qquad Y^{(t)}(0) = S^{(t)}, \qquad Y^{(0)}(0) = S_0. \tag{3.129b}$$

Formula (2.98) implies

$$\dot{Y}^{(t)}(\tau) = \frac{\mathrm{d}}{\mathrm{d}\tau} \exp_{S^{(t)}}(\tau\Omega S^{(t)}) = R_{Y^{(t)}(\tau)}\left[\Omega S^{(t)}\right] = G(Y^{(t)}), \tag{3.130}$$

where we defined the shorthand $G(Y^{(t)})$.

Now, with $S(t)$ solving the S-flow (3.59), we estimate with any $T \geq th$,

$$
\begin{aligned}
&S(T) - Y^{(t)}(T-th) && \text{(3.131a)}\\
&\quad = \quad S(th) - Y^{(t)}(0) + \int_0^{T-th} \frac{\mathrm{d}}{\mathrm{d}\tau}\big(S(th+\tau) - Y^{(t)}(\tau)\big)\mathrm{d}\tau && \text{(3.131b)}\\
&\quad \overset{(3.130),(3.59)}{=} S(th) - Y^{(t)}(0) + \int_0^{T-th} \Big(F\big(S(th+\tau)\big) - G\big(Y^{(t)}(\tau)\big)\Big)\mathrm{d}\tau && \text{(3.131c)}\\
&\quad = \quad S(th) - Y^{(t)}(0) + \int_0^{T-th} \Big(F\big(S(th+\tau)\big) - F\big(Y^{(t)}(\tau)\big)\Big)\mathrm{d}\tau \\
&\qquad + \int_0^{T-th} \Big(F\big(Y^{(t)}(\tau)\big) - G\big(Y^{(t)}(\tau)\big)\Big)\mathrm{d}\tau && \text{(3.131d)}\\
&\quad = \quad S(th) - Y^{(t)}(0) + \int_0^{T-th} \Big(F\big(S(th+\tau)\big) - F\big(Y^{(t)}(\tau)\big)\Big)\mathrm{d}\tau \\
&\qquad + \int_0^{T-th} \int_0^{\tau} \frac{\mathrm{d}}{\mathrm{d}\tau}\Big(F\big(Y^{(t)}(\tau)\big) - G\big(Y^{(t)}(\tau)\big)\Big)\Big|_{\tau=\lambda}\mathrm{d}\lambda\mathrm{d}\tau && \text{(3.131e)}\\
&\quad = \quad S(th) - Y^{(t)}(0) + \int_0^{T-th} \Big(F\big(S(th+\tau)\big) - F\big(Y^{(t)}(\tau)\big)\Big)\mathrm{d}\tau \\
&\qquad + \int_0^{T-th} \int_0^{\tau} \Big(dF\big(Y^{(t)}(\lambda)\big)\big[G\big(Y^{(t)}(\lambda)\big)\big] - dG\big(Y^{(t)}(\lambda)\big)\big[G\big(Y^{(t)}(\lambda)\big)\big]\Big)\mathrm{d}\lambda\mathrm{d}\tau. && \text{(3.131f)}
\end{aligned}
$$

We regard the integrand of the last integral. Since the mappings F given by (3.59) and G given by (3.130) are both C^1 and $\overline{\mathcal{W}}$ is compact, there exists a constant C, independent of $S^{(t)} \in \overline{\mathcal{W}}$ in (3.130), such that

$$
\big\|\mathrm{d}F(Y)[G(Y)] - \mathrm{d}G(Y)[G(Y)]\big\| \leq C \quad \forall Y \in \overline{\mathcal{W}}. \tag{3.132}
$$

Hence,

$$
\begin{aligned}
&\big\|S(T) - Y^{(t)}(T-th)\big\| \\
&\quad \leq \big\|S(th) - Y^{(t)}(0)\big\| + \int_0^{T-th} \big\|F\big(S(th+\tau)\big) - F\big(Y^{(t)}(\tau)\big)\big\|\mathrm{d}\tau + C\int_0^{T-th}\int_0^{\tau}\mathrm{d}\lambda\mathrm{d}\tau && \text{(3.133a)}\\
&\quad \leq \big\|S(th) - Y^{(t)}(0)\big\| + L\int_0^{T-th} \|S(th+\tau) - Y^{(t)}(\tau)\|\mathrm{d}\tau + \frac{C}{2}(T-th)^2. && \text{(3.133b)}
\end{aligned}
$$

Applying Gronwall's inequality [Tes12, Lemma 2.7] yields

$$
\big\|S(T) - Y^{(t)}(T-th)\big\| \leq \Big(\big\|S(th) - Y^{(t)}(0)\big\| + \frac{C}{2}(T-th)^2\Big)e^{L(T-th)} \tag{3.134}
$$

and setting $T = (t+1)h$

$$
\begin{aligned}
\big\|S\big((t+1)h\big) - S^{(t+1)}\big\| &\overset{(3.129b)}{=} \big\|S\big((t+1)h\big) - Y^{(t)}(h)\big\| \\
&\leq \Big(\big\|S(th) - Y^{(t)}(0)\big\| + \frac{Ch^2}{2}\Big)e^{Lh} \overset{(3.129b)}{\leq} \Big(\big\|S(th) - S^{(t)}\big\| + \frac{Ch^2}{2}\Big)e^{Lh}. \qquad \text{(3.135)}
\end{aligned}
$$

Thus, mathematical induction over $t \in \mathbb{N}$ results in

$$\left\|S(th) - S^{(t)}\right\| \leq \frac{Ch^2}{2} \sum_{k \in [t]} e^{Lkh} = \frac{Ch^2}{2} e^{Lh} \frac{e^{Lth} - 1}{e^{Lh} - 1} \qquad \forall t \in \mathbb{N}. \tag{3.136}$$

Eventually, using $e^{Lh} \geq 1 + Lh$ gives

$$\frac{e^{Lh}}{e^{Lh} - 1} = 1 + \frac{1}{e^{Lh} - 1} \leq 1 + \frac{1}{Lh} = \frac{1 + Lh}{Lh} \tag{3.137}$$

and consequently the desired inequality. □

Proposition 3.5.7. *Let $C_S(th)$ be the constant from Proposition 3.5.6. Further, let $\tilde{L} > 0$ be the Lipschitz constant of the lifting map $\exp_{C_W}$ and let $\tilde{C} = \max_{S \in \overline{\mathcal{W}}} \|F(S)\|$ for the mapping F* (3.59) *defining the S-flow. Then the solution $W(t)$ to the assignment flow* (3.1) *and the sequence $(W^{(t)})_{t \in \mathbb{N}_0}$ generated by* (3.117) *satisfy the relation*

$$\left\|W(th) - W^{(t)}\right\| \leq C_W(th)h \quad \forall t \in \mathbb{N}_0 \qquad \text{with} \quad C_W(th) = \tilde{L} \cdot th \cdot \left(C_S(th) + \tfrac{\tilde{C}}{2}\right). \tag{3.138}$$

Proof. By Proposition 3.2.2 and Proposition 3.5.4, we have

$$W(th) = \exp_{C_W}\left(\int_0^{th} S(\tau)\,\mathrm{d}\tau\right), \qquad W^{(t)} = \exp_{C_W}\left(h \sum_{k=0}^{t-1} S^{(k)}\right). \tag{3.139}$$

Hence, the Lipschitz continuity of $\exp_{C_W}$ implies

$$\left\|W(th) - W^{(t)}\right\| \leq \tilde{L}\left\|\int_0^{th} S(\tau)\,\mathrm{d}\tau - h\sum_{k=0}^{t-1} S^{(k)}\right\| \tag{3.140a}$$

$$= \tilde{L}\left\|\sum_{k=0}^{t-1}\int_{kh}^{(k+1)h} S(\tau) - S^{(k)}\,\mathrm{d}\tau\right\| \tag{3.140b}$$

$$\leq \tilde{L}\sum_{k=0}^{t-1}\int_{kh}^{(k+1)h} \left\|S(\tau) - S^{(k)}\right\|\mathrm{d}\tau. \tag{3.140c}$$

Each summand of the last term can be bounded by

$$\begin{aligned}
\int_{kh}^{(k+1)h} \left\|S(\tau) - S^{(k)}\right\|\mathrm{d}\tau &= \int_{kh}^{(k+1)h} \left\|S(kh) - S^{(k)} + \int_{kh}^{\tau} \overbrace{\dot{S}(\lambda)}^{=F(S(\lambda))}\,\mathrm{d}\lambda\right\|\mathrm{d}\tau && \text{(3.141a)}\\
&\leq \int_{kh}^{(k+1)h}\left(\left\|S(kh) - S^{(k)}\right\| + \int_{kh}^{\tau}\left\|F(S(\lambda))\right\|\mathrm{d}\lambda\right)\mathrm{d}\tau && \text{(3.141b)}\\
&\leq h\left\|S(kh) - S^{(k)}\right\| + \int_{kh}^{(k+1)h}(\tau - kh)\tilde{C}\,\mathrm{d}\tau && \text{(3.141c)}\\
&\overset{(3.128)}{\leq} C_S(th)h^2 + \frac{\tilde{C}h^2}{2}. && \text{(3.141d)}
\end{aligned}$$

Combining (3.140) and (3.141) results in (3.138). □

Remark 3.5.8. Proposition 3.5.6 and Proposition 3.5.7 imply that the discretizations (3.119) and (3.117) can approximate the continuous flows $S(t)$ and $W(t)$ on a bounded time interval $t \in [0, T]$ as accurately as required by sufficiently decreasing the step size $h > 0$.

3.5.3 Stability Analysis for Equilibria

In this subsection we will use the shorthand $F_h(S) = \exp_S(h\Omega S)$ for the mapping defining the discrete scheme (3.119), i.e., the discrete scheme of the S-flow reads

$$S^{(t+1)} = F_h\big(S^{(t)}\big), \quad F_h(S) = \exp_S(h\Omega S), \quad t \in \mathbb{N}_0, \qquad S^{(0)} = \exp_{C_{\mathcal{W}}}(-\Omega D). \tag{3.142}$$

We will consider this scheme as discrete-time dynamical systems (cf. Section 2.3.2) and analyze the stability of its equilibria $S^* \in \overline{\mathcal{W}}$, which are the fixed points of the map F_h, i.e. $F_h(S^*) = S^*$. Analogous to Section 3.4, we will examine the eigenvalues of the Jacobian matrix $\frac{\partial F_h}{\partial S}(S^*)$ for this analysis. We will adapt certain statements from Section 3.4 to the discrete scheme (3.142).

Proposition 3.5.9 (equilibria)**.** *Let $\Omega \in \mathbb{R}^{m\times m}$ be an arbitrary matrix. A point $S^* \in \overline{\mathcal{W}}$ is an equilibrium of* (3.142) *if and only if*

$$(\Omega S^*)_{ij} = \langle S_i^*, (\Omega S^*)_i\rangle \quad \forall j \in \operatorname{supp} S_i^* \quad \forall i \in I, \tag{3.143}$$

i.e., the equilibria of the discrete scheme are precisely the equilibria of the S-flow – cf. Proposition 3.4.1(a).

Proof. We have

$$\begin{aligned} F_h(S^*) = S^* \quad &\Leftrightarrow \quad \frac{S_{ij}^* e^{h(\Omega S^*)_{ij}}}{\langle S_i^*, e^{h(\Omega S^*)_i}\rangle} = S_{ij}^* \qquad \forall j \in J \quad \forall i \in I \\ &\Leftrightarrow \quad e^{h(\Omega S^*)_{ij}} = \left\langle S_i^*, e^{h(\Omega S^*)_i}\right\rangle \qquad \forall j \in \operatorname{supp} S_i^* \quad \forall i \in I. \end{aligned} \tag{3.144}$$

The last equation states that the vector $e^{h(\Omega S^*)_i}$ is constant on $\operatorname{supp} S_i^*$. This is equivalent to the vector $(\Omega S^*)_i$ being constant on $\operatorname{supp} S_i^*$, i.e. to equation (3.143). □

Lemma 3.5.10 (Jacobian matrix)**.** *Let F_h denote the mapping in* (3.142)*. Then, after stacking S row-wise, the Jacobian matrix of F_h is given by*

$$\frac{\partial F_h}{\partial S} = \begin{pmatrix} \tilde{B}_1 & & \\ & \ddots & \\ & & \tilde{B}_m \end{pmatrix} + h \begin{pmatrix} R_{F_{h,1}(S)} & & \\ & \ddots & \\ & & R_{F_{h,m}(S)} \end{pmatrix} \cdot \Omega \otimes I_n \tag{3.145}$$

with block matrices $R_p = \operatorname{Diag}(p) - pp^\top$ and

$$\tilde{B}_i = \operatorname{Diag}(Z_i) - F_{h,i}(S) Z_i^\top, \qquad Z_i = \frac{e^{h(\Omega S)_i}}{\langle S_i, e^{h(\Omega S)_i}\rangle}. \tag{3.146}$$

Proof. The subvectors of F_h have the form

$$F_{h,i}(S) = \frac{S_i \odot e^{h(\Omega S)_i}}{\langle S_i, e^{h(\Omega S)_i}\rangle}, \tag{3.147}$$

Hence, the directional derivative of $F_{h,i}$ along $V \in \mathbb{R}^{m\times n}$ is given by

$$\begin{aligned}
\mathrm{d}F_{h,i}(S)[V] &= \tfrac{\mathrm{d}}{\mathrm{d}t} F_{h,i}(S+tV)|_{t=0} && \text{(3.148a)}\\
&= \frac{h(\Omega V)_i \odot S_i \odot e^{h(\Omega S)_i} + V_i \odot e^{h(\Omega S)_i}}{\langle S_i, e^{h(\Omega S)_i}\rangle} \\
&\qquad - \frac{S_i \odot e^{h(\Omega S)_i}}{\big(\langle S_i, e^{h(\Omega S)_i}\rangle\big)^2}\Big\langle h(\Omega V)_i \odot S_i + V_i, e^{h(\Omega S)_i}\Big\rangle && \text{(3.148b)}\\
&= h(\Omega V)_i \odot F_{h,i}(S) + V_i \odot Z_i \\
&\qquad - F_{h,i}(S)\Big\langle h(\Omega V)_i, F_{h,i}(S)\Big\rangle - F_{h,i}(S)\langle V_i, Z_i\rangle && \text{(3.148c)}\\
&= \underbrace{\Big(\mathrm{Diag}(Z_i) - F_{h,i}(S)Z_i^\top\Big)}_{=\tilde{B}_i} V_i + hR_{F_{h,i}(S)}(\Omega V)_i. && \text{(3.148d)}
\end{aligned}$$

We have $\mathrm{d}F_h(S)[V] = \frac{\partial F_h}{\partial S}\,\mathrm{vec}(V)$ with $\mathrm{vec}(V) \in \mathbb{R}^{mn}$ denoting the vector that results from stacking the row vectors (subvectors) of V. Comparing both sides of this equation, with the block matrices of the left-hand side given by (3.148), implies (3.145). □

Proposition 3.5.11 (eigenvalues of Jacobian matrix)**.** *Let $S^* \in \overline{\mathcal{W}}$ be an equilibrium of* (3.142)*, i.e. $F_h(S^*) = S^*$. Then regarding the spectrum $\sigma\big(\frac{\partial F_h}{\partial S}(S^*)\big)$, the following assertions hold.*

(a) If $S^* \in \mathcal{W}^*$, then

$$\sigma\big(\tfrac{\partial F_h}{\partial S}(S^*)\big) = \{0\} \cup \bigcup_{i\in I}\Big\{e^{h(\Omega S^*)_{ij} - h(\Omega S^*)_{ij^*(i)}}\Big\}_{j\in J\setminus\{j^*(i)\}} \tag{3.149a}$$

and the corresponding eigenvectors are given by

$$e_i S_i^{*\top} \in \mathbb{R}^{m\times n}, \quad e_i(e_j - S_i^*)^\top \in \mathcal{T}_0 \qquad \forall j \in J \setminus \operatorname{supp} S_i^* \quad \forall i \in I. \tag{3.149b}$$

(b) If $S^* = \frac{1}{|J_+|}\mathbb{1}_m \mathbb{1}_{J_+}^\top$ with $J_+ \subseteq J$ and $|J_+| \geq 2$, then

$$\sigma\big(\tfrac{\partial F_h}{\partial S}(S^*)\big) = \{0\} \cup \bigcup_{i\in I}\Big\{e^{-\frac{h}{|J_+|}(\Omega\mathbb{1}_m)_i}\Big\} \cup \bigcup_{\lambda\in\sigma(\Omega)}\Big\{1 + \tfrac{h\lambda}{|J_+|}\Big\}. \tag{3.150}$$

(c) Assume the parameter matrix Ω with nonnegative diagonal elements ω_{ii}, $i \in I$. If $S_i^* \notin \{0,1\}^n$ and $\omega_{ii} > 0$ hold for some $i \in I$, then the Jacobian matrix has at least one eigenvalue λ with $|\lambda| > 1$. The real and imaginary part of the corresponding eigenvector lies in

$$\mathcal{T}_+ = \big\{V \in \mathcal{T}_0 \colon \operatorname{supp} V \subseteq \operatorname{supp} S^*\big\}. \tag{3.151}$$

Proof. We start with some properties of the block matrices $\tilde{B}_i$ (3.146). Since $Z_i \odot S_i = F_{h,i}(S)$, we get

$$\tilde{B}_i S_i = 0, \tag{3.152a}$$

$$\tilde{B}_i e_j = Z_{ij}\big(e_j - F_{h,i}(S)\big) = \frac{e^{h(\Omega S)_{ij}}}{\langle S_i, e^{h(\Omega S)_i}\rangle}\big(e_j - F_{h,i}(S)\big) \qquad \forall j \in J. \tag{3.152b}$$

Now, let $S = S^*$ be an equilibrium, i.e. $F_h(S^*) = S^*$. Equations (3.152) imply

$$\tilde{B}_i(e_j - S_i^*) = \frac{e^{h(\Omega S^*)_{ij}}}{\langle S_i^*, e^{h(\Omega S^*)_i}\rangle}(e_j - S_i^*) \qquad \forall j \in J. \tag{3.153}$$

By the equilibrium criterion (3.143), we have $Z_{ij} = 1$ for $j \in \operatorname{supp} S_i^*$. Hence, the equation (3.152b) also implies

$$\tilde{B}_i v = v \qquad \forall v \in \mathcal{T}_{+,i} = \left\{v \in \mathbb{R}^n \colon \langle v, \mathbb{1}_n\rangle = 0,\ \operatorname{supp} v \subseteq \operatorname{supp} S_i^*\right\}. \tag{3.154}$$

We will use these equations to prove each particular statement.

(a) Since $R_{S^*} = 0$, we have

$$\tfrac{\partial F_h}{\partial S}(S^*) = \tilde{B} = \begin{pmatrix} \tilde{B}_1 & & \\ & \ddots & \\ & & \tilde{B}_m \end{pmatrix}. \tag{3.155}$$

By equations (3.152a) and (3.153), the mn eigenpairs are given by

$$\tfrac{\partial F_h}{\partial S}(S^*) \cdot e_i \otimes S_i^* = 0, \tag{3.156a}$$

$$\begin{aligned}\tfrac{\partial F_h}{\partial S}(S^*) \cdot e_i \otimes (e_j - S_i^*) &= \frac{e^{h(\Omega S^*)_{ij}}}{\langle S_i^*, e^{h(\Omega S^*)_i}\rangle} e_i \otimes (e_j - S_i^*) \\ &= e^{h(\Omega S^*)_{ij} - h(\Omega S^*)_{ij^*(i)}} e_i \otimes (e_j - S_i^*) \qquad \forall j \in J \setminus \{j^*(i)\}.\end{aligned} \tag{3.156b}$$

(b) Let $\{(\lambda_i, w_i)\}_{i\in\tilde{I}} \subset \mathbb{C} \times \mathbb{C}^m$ be the set of all eigenpairs of Ω. We have $|\tilde{I}| < m$ if and only if Ω is not diagonalizable. Further, let $\{v_1, \dots, v_{|J_+|-1}\}$ be a basis of the vector space $\left\{v \in \mathbb{R}^n \colon \langle v, \mathbb{1}_{J_+}\rangle = 0,\ \operatorname{supp}(v) \subseteq J_+\right\}$. We have $R_{S_i^*} v_j = \frac{1}{|J_+|} v_j$. Using the equations (3.152a), (3.153) and (3.154), we get $mn - (m - |\tilde{I}|)(|J_+| - 1)$ eigenpairs:

$$\tfrac{\partial F_h}{\partial S}(S^*) \cdot e_i \otimes S_i^* = 0 \qquad \forall i \in I, \tag{3.157a}$$

$$\tfrac{\partial F_h}{\partial S}(S^*) \cdot e_i \otimes (e_j - S_i^*) = e^{-\frac{h}{|J_+|}(\Omega \mathbb{1}_m)_i} e_i \otimes (e_j - S_i^*) \qquad \forall j \in J \setminus J_+, \quad \forall i \in I, \tag{3.157b}$$

$$\tfrac{\partial F_h}{\partial S}(S^*) \cdot w_i \otimes v_j = \left(1 + \tfrac{\lambda h}{|J_+|}\right) w_i \otimes v_j, \qquad \forall i \in \tilde{I}, \quad \forall j \in \{1, \dots, |J_+| - 1\}. \tag{3.157c}$$

If $|\tilde{I}| = m$, we have a full set of mn eigenpairs. If $|\tilde{I}| < m$, the stated spectrum follows by a perturbation of the matrix Ω and the continuity of the spectrum as argued in the proof of Proposition 3.4.4(b).

(c) Analogous to the proof of Proposition 3.4.4(c), we have the inclusions

$$\operatorname{im} R_{S^*} \subseteq \mathcal{T}_+ \subseteq \ker\left(\tilde{B} - I_{mn}\right), \tag{3.158}$$

where the second inclusion follows by (3.154). In that proof it was also shown that the matrix $R_{S^*}(\Omega \otimes I_n)$ possesses an eigenvalue λ_0 with $\operatorname{Re}(\lambda_0) > 0$ and the real and imaginary part of the corresponding eigenvector lying in $\mathcal{T}_+$. The inclusions (3.158) imply, that this eigenvector is an eigenvector of $\tfrac{\partial F_h}{\partial S}(S^*) = \tilde{B} + hR_{S^*}(\Omega \otimes I_n)$ to the eigenvalue $\lambda = 1 + h\lambda_0$. Eventually, we have $|\lambda| = |1 + h\lambda_0| \geq \operatorname{Re}(1 + h\lambda_0) > 1$.

□

Corollary 3.5.12 (stability of equilibria). *Let Ω be a nonnegative matrix with positive diagonal entries. Then, regarding the equilibria $S^* \in \overline{\mathcal{W}}$ of (3.142), the following assertions hold.*

(a) $S^* \in \mathcal{W}^*$ is exponentially stable if, for all $i \in I$,

$$(\Omega S^*)_{ij} < (\Omega S^*)_{ij^*(i)} \quad \text{for all } j \in J \setminus \{j^*(i)\} \quad \text{with} \quad \{j^*(i)\} = \arg\max_{j \in J} S^*_{ij}. \tag{3.159}$$

(b) $S^* \in \mathcal{W}^*$ is unstable if, for some $i \in I$,

$$(\Omega S^*)_{ij} > (\Omega S^*)_{ij^*(i)} \quad \text{for some } j \in J \setminus \{j^*(i)\} \quad \text{with} \quad \{j^*(i)\} = \arg\max_{j \in J} S^*_{ij}. \tag{3.160}$$

(c) All equilibria $S^* \notin \mathcal{W}^*$ are unstable.

Note that the same stability statements hold for the discrete scheme (3.142) as for the continuous S-flow (3.59). The proof of this corollary proceeds analogously to the proof of Corollary 3.4.5 with Theorem 2.3.8, Proposition 3.4.4 and Proposition 2.3.9 replaced by the discrete counterparts Theorem 2.3.12, Proposition 3.5.11 and Proposition 2.3.13.

Proposition 3.5.13. *Let Ω, $S^* \in \mathcal{W}^*$, $A(S^*)$ and $B_\varepsilon(S^*)$ be as in Proposition 3.4.12. Then, for the sequence $(S^{(t)})_{t \in \mathbb{N}_0}$ generated by (3.119), the following holds. If $S^{(t_0)} \in B_\varepsilon(S^*)$ for some time point $t_0 \in \mathbb{N}$, then $S^{(t)} \in B_\varepsilon(S^*)$ for all $t \geq t_0$ and $\lim_{t\to\infty} S^{(t)} = S^*$. Moreover, we have*

$$\left\| S_i^{(t)} - S_i^* \right\|_1 \leq \left\| S_i^{(t_0)} - S_i^* \right\|_1 \cdot \gamma_i^{t-t_0} \tag{3.161}$$

with $\gamma_i \in (0,1)$, for each $i \in I$.

Proof. Let

$$\beta_i = \beta_i(S) \coloneqq \min \left\{ (\Omega S)_{ij^*(i)} - (\Omega S)_{ij} \right\}_{j \neq j^*(i)}. \tag{3.162}$$

For $S \in A(S^*)$, we have $\beta_i(S) > 0$ and with $S_i^* = e_{ij^*(i)}$, $F_{h,i}(S) \in \Delta_n$,

$$\begin{aligned}
\left\| F_{h,i}(S) - S_i^* \right\|_1 &= 2 - 2F_{h,ij^*(i)}(S) && \text{(3.163a)} \\
&= 2 - 2\frac{S_{ij^*(i)}}{S_{ij^*(i)} + \sum_{j \neq j^*(i)} S_{ij} e^{h(\Omega S)_{ij} - h(\Omega S)_{ij^*(i)}}} && \text{(3.163b)} \\
&\leq 2 - 2\frac{S_{ij^*(i)}}{S_{ij^*(i)} + (1 - S_{ij^*(i)}) e^{-h\beta_i}} && \text{(3.163c)} \\
&= \|S_i - S_i^*\|_1 \underbrace{\frac{e^{-h\beta_i}}{S_{ij^*(i)} + (1 - S_{ij^*(i)}) e^{-h\beta_i}}}_{<1}. && \text{(3.163d)}
\end{aligned}$$

Choosing $\delta > 0$ with $S^{(t_0)} \in \overline{B_\delta(S^*)} \subset B_\varepsilon(S^*)$, we set

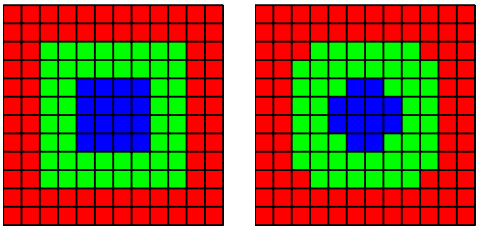
$$\gamma_i = \max_{S \in \overline{B_\delta(S^*)}} \frac{e^{-h\beta_i(S)}}{S_{ij^*(i)} + (1 - S_{ij^*(i)}) e^{-h\beta_i(S)}} \in (0,1). \tag{3.164}$$

and thus get $\|F_{h,i}(S) - S_i^*\|_1 \leq \gamma_i \|S_i - S_i^*\|_1$ for $S \in \overline{B_\delta(S^*)}$, which implies $F_h(\overline{B_\delta(S^*)}) \subseteq \overline{B_\delta(S^*)}$ and the exponential convergence rate (3.161) of $S^{(t)}$. □

We conclude this subsection with an illustrative application of Corollary 3.5.12 and Proposition 3.5.13. We design and construct a small academical example that, despite its simplicity, illustrates the following important points:

- the attracting region $B_\varepsilon(S^*)$ due to Corollary 3.4.14, here for the special case of uniform averaging parameters Ω (and likewise more generally for nonuniform Ω (Proposition 3.4.13)), that enables to terminate the numerical scheme and rounding to the *correct* labeling;
- the influence of Ω on the spatial shape of patterns created through data labeling, which provides the basis for pixel-accurate 'semantic' image labeling;
- undesired asymptotic behavior of the *numerically* integrated assignment flow—cf. Remark 3.5.15 below—cannot occur when using proper geometric numerical integration, like the scheme (3.117) or any scheme devised by [ZSPS20].

Example 3.5.14. We consider a 12×12 RGB image $u \colon I \to [0,1]^3$ shown by Figure 3.5.



input image S^*

Figure 3.5: Illustration of input and output of Example 3.5.14. The input image consisting of three colors, which was used for computing the distance matrix D, is shown on the left. This distance matrix was used to initialize the S-flow, whose limit is illustrated by the image on the right. This is a minimal example that demonstrates how stability conditions (3.90) constrain spatial shape.

The three unit vectors e_j, $j \in J = [3]$ define the labels that are marked by the colors red, green and blue. For spatial regularization we used 3×3 neighborhoods $\mathcal{N}_i$, $i \in I$ with uniform weights $\omega_{ik} = \frac{1}{|\mathcal{N}_i|}$, $k \in \mathcal{N}_i$, with shrunken neighborhoods if they intersect the boundary of the underlying quadratic domain. The distance matrix D that initializes the S-flow by $S_0 = \exp_{C_\mathcal{W}}(-\Omega D)$, was set to $D_{ij} = 10 \cdot \|u_i - e_j\|$, $i \in I$, $j \in J$.

Adopting the termination criterion from [ÅPSS17], we numerically integrated the S-flow using the scheme (3.119), until iteration T when the average entropy dropped below 10^{-3}, i.e.

$$-\frac{1}{m \log n} \sum_{i \in I, j \in J} S_{ij}^{(T)} \log S_{ij}^{(T)} < 10^{-3}. \tag{3.165}$$

The resulting assignment $S^{(T)}$ was rounded to the integral assignment $S^* \in \mathcal{W}^*$ depicted by the right panel of Figure 3.5. We observe the following.

(i) The resulting labeling S^* differs from the input image although *exact* (integral) input data are used.

This conforms to Corollary 3.5.12(b), which enables to recognize the input data as *unstable*. As a consequence, the green and blue labels at the corners of the corresponding quadrilateral shapes in the input data are replaced by the flow. The resulting labeling S^* is stable, as one easily verifies using Corollary 3.5.12(a).

(ii) Using the estimate (3.113) which is the special case of (3.109) in the case of uniform weights, we computed

$$\varepsilon_{\text{est}} = \varepsilon_{\text{unif}} = 0.2. \tag{3.166}$$

Since the distance between S^* and the assignment $S^{(T)}$ obtained after terminating numerical integration due to (3.165), satisfied

$$\max_{i \in I} \left\| S_i^{(T)} - S_i^* \right\|_1 \approx 0.00196 < \varepsilon_{\text{est}}, \tag{3.167}$$

we had the *guarantee* due to Proposition 3.5.13 that $S^{(t)}$ converges for $t > T$ to S^*, i.e. that no label indicated by $S^{(T)}$ can change anymore. In addition, Proposition 3.5.5 guarantees that the discrete schema (3.117) of the assignment flow converges to S^* as well. With regard to Proposition 3.5.6, the estimate (3.167) implies for sufficiently small step size $h > 0$ that the continuous S-flow $S(hT)$ also lies in the attracting region $B_\varepsilon(S^*)$. Proposition 3.4.12 then states the convergence of the S-flow to S^*. Eventually, the continuous assignment flow (3.1) converge to S^* by Proposition 3.2.5.

Remark 3.5.15 (numerical integration and asymptotic behavior)**.** The authors of [ÅPSS17] adopted a numerical scheme from [LA83] which, when adapted and applied to (3.1), was shown in [BFPS17] to always converge to a constant solution as $t \to \infty$, i.e., a *single* label is assigned to every pixel, which clearly is an unfavorable property. This strange asymptotic behavior resulted from the fact that the adaption of the discrete scheme of [LA83] implicitly uses *different* step sizes for updating the flow S_i at different locations $i \in I$.

Our results in this chapter show that the continuous-time assignment flow does not exhibit this asymptotic behavior, under appropriate assumptions on the parameter matrix Ω. In addition, point (ii) above and Proposition 3.5.13 show that using a proper geometric scheme from [ZSPS20] turns condition (3.165) into a sound criterion for terminating the numerical scheme, followed by save rounding to an integral labeling.

3.6 Generalization of the S-Flow

In this section we consider a general class of flows of the form

$$\dot{S} = R_S\big[\mathbf{V}[S]\big], \quad S(0) = S_0, \tag{3.168}$$

where $\mathbf{V}\colon \mathbb{R}^{m \times n} \to \mathbb{R}^{m \times n}$ is a linear map, i.e., we have $\operatorname{vec}(\mathbf{V}[S]) = M_{\mathbf{V}} \operatorname{vec}(S)$ with a matrix $M_{\mathbf{V}} \in \mathbb{R}^{mn \times mn}$. We will write this map component-wise in the form

$$\mathbf{V}_{ij}[S] = \sum_{k \in I} \sum_{l \in J} v_{ij,kl} S_{kl}, \qquad v_{ij,kl} \in \mathbb{R}. \tag{3.169}$$

The S-flow (3.59) is therefore a special case with

$$v_{ij,kl} = \begin{cases} \omega_{ik}, & \text{if } j = l, \\ 0, & \text{else.} \end{cases} \tag{3.170}$$

While the S-flow only considers interactions between pixels, (3.168) also allows interactions between labels.

Since this class of flows might be of interest in future application, we transfer certain statements from Section 3.3 and Section 3.4 and present them in brief below. Since the proofs do not change significantly and therefore do not provide any additional insight, we omit them.

Proposition 3.6.1 (equilibria). *A point $S^* \in \overline{\mathcal{W}}$ is an equilibrium point of* (3.168) *if and only if*

$$\mathbf{V}_{ij}[S^*] = \left\langle S_i^*, \mathbf{V}_i[S^*] \right\rangle \qquad \forall j \in \operatorname{supp} S_i^*, \quad \forall i \in I. \tag{3.171}$$

Theorem 3.6.2 (convergence to equilibria). *Assume $\mathbf{V}$ fulfills the symmetry assumption*

$$v_{ij,kl} = \frac{\widehat{v}_{ij,kl}}{w_i} \quad \text{with} \quad \widehat{v}_{ij,kl} = \widehat{v}_{kl,ij}, \quad w_i > 0 \qquad \forall i,k \in I, \quad \forall j,l \in J. \tag{3.172}$$

Then the flow (3.168) *converges to an equilibrium $S^* = S^*(S_0) \in \overline{\mathcal{W}}$ for any initial value $S_0 \in \mathcal{W}$.*

Lemma 3.6.3 (Jacobian matrix). *Let $F(S) = R_S[\mathbf{V}[S]]$ denote the vector field defining the flow* (3.168). *Then, after stacking S row-wise, the Jacobian matrix of F is given by*

$$\frac{\partial F}{\partial S} = \begin{pmatrix} B_1 & & \\ & \ddots & \\ & & B_m \end{pmatrix} + \begin{pmatrix} R_{S_1} & & \\ & \ddots & \\ & & R_{S_m} \end{pmatrix} \cdot M_{\mathbf{V}} \tag{3.173}$$

with block matrices $B_i = \operatorname{Diag}\left(\mathbf{V}_i[S]\right) - \langle S_i, \mathbf{V}_i[S]\rangle I_n - S_i \mathbf{V}_i[S]^\top$ and $R_{S_i} = \operatorname{Diag}(S_i) - S_i S_i^\top$.

Proposition 3.6.4 (eigenvalues of Jacobian matrix). *Let $S^* \in \overline{\mathcal{W}}$ be an equilibrium point of* (3.168). *Then regarding the spectrum $\sigma\left(\frac{\partial F}{\partial S}(S^*)\right)$, the following assertions hold.*

(a) If $S^* \in \mathcal{W}^*$, then

$$\sigma\left(\tfrac{\partial F}{\partial S}(S^*)\right) = \bigcup_{i \in I} \left\{ -\langle S_i^*, \mathbf{V}_i[S^*]\rangle \right\} \cup \left\{ \mathbf{V}_{ij}[S^*] - \langle S_i^*, \mathbf{V}_i[S^*]\rangle \right\}_{j \in J \setminus \operatorname{supp} S_i^*} \tag{3.174}$$

and the corresponding eigenvectors are given by

$$e_i S_i^{*\top} \in \mathbb{R}^{m \times n}, \quad e_i (e_j - S_i^*)^\top \in \mathcal{T}_0 \qquad \forall j \in J \setminus \operatorname{supp} S_i^* \quad \forall i \in I. \tag{3.175}$$

(b) All nonvanishing eigenvalues of $R_{S^*} M_{\mathbf{V}}$ are also eigenvalues of $\frac{\partial F}{\partial S}(S^*) = B + R_{S^*} M_{\mathbf{V}}$ with the same corresponding eigenvectors. Further, we have

$$\operatorname{tr}\left(R_{S^*} M_{\mathbf{V}}\right) = \sum_{i \in I} \sum_{j \in J} S_{ij}^* \cdot \Big(v_{ij,ij} - \sum_{l \in J} v_{ij,il} S_{il}^* \Big). \tag{3.176}$$

If the operator $\mathbf{V}$ fulfills

$$v_{ij,ij} > v_{ij,il} \qquad \forall l \in J \setminus \{j\} \qquad \forall j \in J \qquad \forall i \in I \tag{3.177}$$

and the equilibrium S^* is nonintegral, i.e. $S^* \notin \mathcal{W}^*$, then the trace (3.176) *is positive and thus $\frac{\partial F}{\partial S}(S^*)$ has an eigenvalue with positive real part. The real and imaginary part of the corresponding eigenvector lies in*

$$\mathcal{T}_+ = \left\{ V \in \mathcal{T}_0 \colon \operatorname{supp} V \subseteq \operatorname{supp} S^* \right\}. \tag{3.178}$$

We note that in the case of the S-flow (3.170) the condition (3.177) means the positivity of the diagonal of the parameter matrix $\Omega = (\omega_{ik})$.

Corollary 3.6.5 (stability of equilibria)**.** *Let* $\mathbf{V}$ *fulfill* $v_{ij,kl} \geq 0$ *and* (3.177)*. Then regarding the equilibria* $S^* \in \overline{\mathcal{W}}$ *of the flow* (3.168)*, the following assertions hold.*

(a) $S^* \in \{0,1\}^{m \times n}$ *is exponentially stable if, for all* $i \in I$*,*

$$\mathbf{V}_{ij}[S^*] < \mathbf{V}_{ij^*(i)}[S^*] \quad \text{for all } j \in J \setminus \{j^*(i)\} \quad \text{with} \quad \{j^*(i)\} = \arg\max_{j \in J} S^*_{ij}. \tag{3.179}$$

(b) $S^* \in \{0,1\}^{m \times n}$ *is unstable if, for some* $i \in I$*,*

$$\mathbf{V}_{ij}[S^*] > \mathbf{V}_{ij^*(i)}[S^*] \quad \text{for some } j \in J \setminus \{j^*(i)\} \quad \text{with} \quad \{j^*(i)\} = \arg\max_{j \in J} S^*_{ij}. \tag{3.180}$$

(c) All equilibria $S^* \notin \mathcal{W}^*$ *are unstable.*

Theorem 3.6.6. *Let* $\mathbf{V}$ *fulfill* $v_{ij,kl} \geq 0$ *and* (3.177)*. Then the set of starting points* $S_0 \in \mathcal{W}$ *for which the flow* (3.168) *converges to a nonintegral equilibrium* $S^* \in \overline{\mathcal{W}}$*, has measure zero in* $\mathcal{W}$*.*

3.7 Summary

In this chapter, we analyzed the convergence of the assignment flow based on the S-parametrization [SS20]. We showed that the limits of the assignment flow are obtained by rounding the limits of the S-flow to integral assignments, whenever rounding is uniquely possible (Proposition 3.2.5). In particular, integral limits of the S-flow are also limits of the assignment flow. We focused then on the convergence of the S-flow. We proved under a symmetry assumption on the parameters Ω that the S-flow always converges towards an equilibrium (Theorem 3.3.2). This symmetry assumption is fulfilled for uniform weights (3.18) with according neighborhoods, which we will use in the following chapters. We also demonstrated in Example 3.3.9 that convergence can fail without this symmetry assumption. Later in this chapter we examined the stability of equilibria of the S-flow in order to clarify to which equilibrium the flow may converge. We showed under positivity assumption on the diagonal of the parameter matrix Ω that nonintegral equilibria are unstable (Corollary 3.4.5) and concluded that the convergence towards them is only valid for a null set of initial values of the S-flow (Theorem 3.4.8). This positivity assumption is fulfilled for uniform weights (3.18) as well. For equilibria that meet the stability criterion (3.90) and are thus exponentially stable, we derived an estimate for their basin of attraction (Proposition 3.4.12, Proposition 3.4.13) that can be used in practical application as a termination criterion and provide a guarantee for rounding. After the investigation of the continuous flow, we dealt with the geometric Euler discretization [Sav+17]. We saw that on a bounded time interval the discrete scheme approximates the continuous flow accurately for sufficiently small step size $h > 0$ (Proposition 3.5.6, Proposition 3.5.7). Furthermore, we verified the stability statements on equilibria of the continuous S-flow also for the discrete scheme (Corollary 3.5.12). In particular, the estimates for the basin of attraction of the exponentially stable equilibria also apply and thus provide convergence guarantees for the discrete scheme as well (Proposition 3.5.13). This convergence guarantee was used in Example 3.5.14 to

demonstrate that the assignment flow converges to a nonconstant labeling. At the end of the chapter we discussed a generalization of the continuous S-flow, where the fitness interacts not only between pixels but also between labels. For this flow we listed how certain statements previously stated for the S-flow are transferred. These include conditions for guaranteed convergence to equilibria (Theorem 3.6.2) as well as the instability of nonintegral equilibria (Corollary 3.6.5, Theorem 3.6.6).

4 Label Statistics by Global Filtering Constraints

4.1 Introduction and Overview

Motivation

The *discriminative* power of *filter statistics* for object detection and classification is well known [Low04, MY09] and has been widely explored in the literature. The *generative* power of filter statistics for representing image structure, on the other hand, has been less explored during the recent years. Early seminal work on generative aspects of filter statistics includes [ZM97] and [PS00] and many references in these papers. In the former case, heavy-tailed empirical filter statistics are imposed on the variational problem of learning the parameters of a Gibbs-Boltzmann distribution. In the latter case, several hundred filter constraints form nonlinear submanifolds (level sets) onto which a given image has to be projected. While both works impressively demonstrate the generative power of filter statistics, exploiting these statistics as prior knowledge for inference and reproducibility of results has remained a challenge from the viewpoint of algorithm design and numerical optimization.

In this chapter, we focus on a mathematically sound and numerically tractable approach to impose filter statistics on *labeled* image structure, i.e., we consider filter constraints in *label space* rather than in image space. The simplest constraint, for example, imposes lower and upper bounds on the area occupied by some label, without specifying the corresponding locations. More general constraints arise from replacing the 'identity filter' by linear filters learned offline through a simple generalized eigenvalue technique and imposing similar linear statistical constraints. To incorporate such constraints into an algorithm for image labeling in a numerically *tractable and reproducible* way, we adopt the assignment flow approach presented in Section 2.2 and take *filtered label statistics* into account by using standard log-barrier functions [NN94, BV04].

Organization

Section 4.2 details our contribution. First we mention the considered global filter constraints (Section 4.2.1) and then discuss how we learn filters for label statistics (Section 4.2.2) and take corresponding empirical constraints into account during inference for image labeling (Section 4.2.3). We do not focus on any specific application in this chapter. Rather, the proof-of-concept experiments discussed in Section 4.3 are supposed to demonstrate how statistics gathered by linear filters of small support can *enhance image labeling*, represent *primitive shape information* and support *spatial pattern formation*, by extending the assignment flow approach through corresponding *convex* constraints.

Notation

We use the notation from Section 2.2, i.e., n is the number of labels and $J = [n] = \{1, \dots, n\}$ is the set of label indices. The set of pixel indices is $I = [m]$, where $m = m_1 m_2$ is the number of pixels of an $m_1 \times m_2$ image. For a matrix $W \in \mathbb{R}^{m \times n}$, $W_i \in \mathbb{R}^n$ is the i-th row and $W^j \in \mathbb{R}^m$ is the j-th column, with the latter representing a $m_1 \times m_2$ grayscale image. This also applies to $p \times p$ filters $f \in \mathbb{R}^{p^2 \times n}$, i.e., $f^j \in \mathbb{R}^{p^2}$ represents a $p \times p$ filter. The convolution of the $p \times p$ filter f^j with the $m_1 \times m_2$ image W^j is denoted with $f^j * W^j \in \mathbb{R}^m$, whereby the filter result is again regarded as a vector.

4.2 Label Assignment with Global Constraints

In this section, we discuss how filter constraints can be incorporated into the assignment flow approach from Section 2.2 which evolves on the assignment manifold $\mathcal{W} \subset \mathbb{R}^{m \times n}$. First we present the considered global filter constraints on assignment matrices $W \in \overline{\mathcal{W}}$ and then explain the adaption of the vector field of the assignment flow taking these constraints into account.

4.2.1 Filter Constraints

In order to incorporate some prior knowledge about labelings, we consider linear $p \times p$ filters $f \in \mathcal{F} \subset \mathbb{R}^{p^2 \times n}$ operating on assignment matrices $W \in \overline{\mathcal{W}}$. For each label $j \in J$, we have a $p \times p$ filter $h^j \in \mathbb{R}^{p^2}$ in the usual sense, and the filter operation is given by

$$f * W := \sum_{j \in J} f^j * W^j, \tag{4.1}$$

with the common convolution of the 'label images' W^j, $j \in J$ with a $p \times p$ filter on the right-hand side. The space of filters $\mathcal{F}$ will be specified in Section 4.2.2. To avoid complications at and close to the boundary of the image region, we only take into account filter results $(f * W)_i$ at interior pixels i where the $p \times p$ filter support (centered at i) does not overlap with the boundary. At the remaining pixels the filter result is set to 0.

The filter result $(f * W)_i$ at a pixel $i \in I$ depends on the assignment within a $p \times p$ neighborhood of i and hence reflects the local spatial relation of the labels. Our objective is to control label assignments by constraining the filter results for a set

$$\left\{ f^{(k)} \in \mathcal{F} : k \in K \right\} \tag{4.2}$$

of $|K|$ filters in order to take into account statistical prior information about the local geometry of labelings. Motivated by [BGGS17], where the ℓ^1-norm of the filter results of a grayscale image was considered in connection with non-smooth sparse regularization, we consider here the ℓ^2-norm of filter results which conforms to our *smooth* label assignment approach of Section 2.2.

Specifically, we consider *global convex constraints* of the form

$$c_{\mathrm{low}} \leq \tfrac{1}{m} W^\top \mathbb{1}_m \leq c_{\mathrm{up}}, \tag{4.3a}$$

$$\| f^{(k)} * W \|_2 \leq d^{(k)}, \qquad k \in K, \tag{4.3b}$$

where the parameter vectors $c_{\text{low}}, c_{\text{up}} \in \mathbb{R}^n$ impose lower and upper cardinality bounds for the assignment of each label $j \in J = [n]$ to the range of pixels $I = [m]$, whereas the parameters $d^{(k)}$ of (4.3b) constrain the output energy of each filter $f^{(k)}$, $k \in K$. To ensure that the region of feasible assignments W has a non-empty interior, we require

$$c_{\text{low}} < c_{\text{up}}, \qquad \langle c_{\text{low}}, \mathbb{1}_n \rangle < 1 < \langle c_{\text{up}}, \mathbb{1}_n \rangle \qquad \text{and} \qquad d^{(k)} > 0,\ k \in K. \tag{4.4}$$

As alternative to the ℓ^2-norm defining (4.3b), we also used a smooth approximation of the ℓ^1-norm denoted by

$$\|x\|_{1,\varepsilon} = \sum_{i \in I} |x_i|_\varepsilon, \qquad |x_i|_\varepsilon = \sqrt{x_i^2 + \varepsilon^2} - \varepsilon. \tag{4.5}$$

4.2.2 Space of Filters and Filter Learning

We discuss how to choose filters for given classes of labelings and corresponding example data. First of all, we restrict the space of all possible $p \times p$ filters in order to eliminate some redundant degrees of freedom. To this end, we consider the decomposition of the space of all filters

$$\mathbb{R}^{p^2 \times n} = \mathcal{F}_0 \oplus \mathcal{F}_1 \oplus \mathcal{F}_2 \tag{4.6}$$

into subspaces given by[1]

$$\mathcal{F}_0 = \big\{ f \in \mathbb{R}^{p^2 \times n} \colon f^j = f^l,\ \operatorname{mean}(f^j) = 0,\ \forall j, l \in J \big\}, \tag{4.7a}$$

$$\mathcal{F}_1 = \big\{ f \in \mathbb{R}^{p^2 \times n} \colon \langle f_i, \mathbb{1}_n \rangle = 0,\ \forall i \in [p^2],\ \operatorname{mean}(f^j) = 0,\ \forall j \in J \big\}, \tag{4.7b}$$

$$\mathcal{F}_2 = \big\{ f \in \mathbb{R}^{p^2 \times n} \colon f_i = f_k,\ \forall i, k \in [p^2] \big\}. \tag{4.7c}$$

These spaces are orthogonal to each other with respect to the Euclidean inner product. The space $\mathcal{F}_2$ consists of all filters f that are constant for each label, i.e. $f^j = c_j \cdot \mathbb{1}_{p^2}$ for each $j \in J$ with $c \in \mathbb{R}^n$. The space $\mathcal{F}_0$ consists of all zero-mean filters, which do not distinguish between labels. For any filter $f \in \mathcal{F}_0$, we have $f * W = f^1 * \sum_j W^j = f^1 * \mathbb{1}_m = 0$ for all $W \in \overline{\mathcal{W}}$, i.e., the subspace $\mathcal{F}_0$ does not represent any useful information for our purpose.

Thus, we can choose either $\mathcal{F} = \mathcal{F}_1$ or $\mathcal{F} = \mathcal{F}_1 \oplus \mathcal{F}_2$ as the actual space of filters. Our choice is

$$\mathcal{F} = \mathcal{F}_1, \qquad \dim \mathcal{F} = \dim \mathcal{F}_1 = (p^2 - 1)(n - 1) \tag{4.8}$$

for two reasons. Firstly, we use this framework for segmentation, where larger homogenous regions occur (e.g., background). Filters which return a small ℓ^2-norm for such labelings have (approximately) a zero-mean and therefore belong to $\mathcal{F}_1$. Secondly, we will use an inner point approach which requires a feasible initialization. In case of zero-mean filters, we can simply use constant assignments as initial assignment. As a result, we do not need an additional initialization process on which the final result might depend. This conforms to the philosophy to start the assignment flow without any bias at the barycenter $C_{\mathcal{W}} \in \mathcal{W}$.

For learning the filters, we assume that sets $\mathcal{I}^+, \mathcal{I}^- \subset \overline{\mathcal{W}}$ for favorable and unfavorable label assignments are given. We are looking for filters $f \in \mathcal{F}$ such that $\|f * W\|_2$ is smaller for

[1] Notation: Filters f are matrix-valued (image vectors × labels) with rows f_i and columns f^j. Superscripts in brackets $f^{(k)}$ index members of a collection of filters.

$W \in \mathcal{I}^+$ than for $W \in \mathcal{I}^-$. For simplicity, we choose

$$\frac{\operatorname{mean}_{W \in \mathcal{I}^+} \|f * W\|_2^2}{\operatorname{mean}_{W \in \mathcal{I}^-} \|f * W\|_2^2} < 1 \tag{4.9}$$

as criterion for filters $f \in \mathcal{F}$, which leads to a generalized eigenvalue problem. Specifically, let $\left\{e_i^{\mathcal{F}} : i = 1, \dots, \dim \mathcal{F}\right\}$ be an orthonormal basis of $\mathcal{F}$ and consider the map

$$M \colon \overline{\mathcal{W}} \to \mathbb{R}^{\dim \mathcal{F} \times \dim \mathcal{F}}, \quad M(W)_{ij} = \left\langle e_i^{\mathcal{F}} * W, e_j^{\mathcal{F}} * W \right\rangle. \tag{4.10}$$

Then we have

$$\frac{\operatorname{mean}_{W \in \mathcal{I}^+} \|f * W\|_2^2}{\operatorname{mean}_{W \in \mathcal{I}^-} \|f * W\|_2^2} = \frac{x^\top A^+ x}{x^\top A^- x} \tag{4.11}$$

with $A^\pm = \operatorname{mean}_{W \in \mathcal{I}^\pm} M(W)$ and $f = \sum_i x_i \, e_i^{\mathcal{F}}$. As a consequence, a set of linearly independent filters satisfying the criterion (4.9) is given by the generalized eigenvectors of the matrix pencil (A^+, A^-) corresponding to eigenvalues less than 1. The filters corresponding to eigenvalues greater than 1 might also focus on useful features as can be seen, for example, in Figure 4.7 (d), where the last 16 filters correspond to eigenvalues greater than 1. These filters can be used additionally, since they further restrict the assignment and therefore may prevent some assignments, which were not taken into account by $\mathcal{I}^-$.

Having determined a set of filters as generalized eigenvectors, we normalize them in a post-processing step so as to meet the condition $f * W \in [-1, 1]^m$ for all $W \in \overline{\mathcal{W}}$, i.e.

$$\|f\|_{\mathcal{F}} = 1, \qquad \|f\|_{\mathcal{F}} \coloneqq \max \left\{ - \sum_{i \in [p^2]} \min_{j \in J} f_{ij}, \sum_{i \in [p^2]} \max_{j \in J} f_{ij} \right\}. \tag{4.12}$$

4.2.3 Adaption of the Assignment Flow

In order to take into account the constraints (4.3), we follow the approach in [AL10] by using log-barrier functions [NN94, BV04]. Given the parameters in (4.4) and filters $f^{(k)}$, $k \in K$, these functions read

$$B_{\text{low}}(W) = - \left\langle \mathbb{1}_n, \log \left(\tfrac{1}{m} W^\top \mathbb{1}_m - c_{\text{low}} \right) \right\rangle, \tag{4.13a}$$

$$B_{\text{up}}(W) = - \left\langle \mathbb{1}_n, \log \left(c_{\text{up}} - \tfrac{1}{m} W^\top \mathbb{1}_m \right) \right\rangle, \tag{4.13b}$$

$$B_{\text{filter}}(W) = - \sum_{k \in K} \log \left((d^{(k)})^2 - \|f^{(k)} * W\|_2^2 \right). \tag{4.13c}$$

Summing up these functions yields the overall barrier function

$$B(W) = B_{\text{low}}(W) + B_{\text{up}}(W) + B_{\text{filter}}(W) \tag{4.14}$$

for the constraints (4.3). We complement the similarity map (2.104)

$$S_\tau(W) = S(W) + \tau \nabla B(W), \qquad \tau > 0 \tag{4.15}$$

which steers the assignment flow (2.105). This adaptation is due to the observation that the assignment flow was originally introduced as a Riemannian gradient flow of $J(W) = \langle W, S(W) \rangle$ with $\nabla J(W) \approx S(W)$ and that (4.15) corresponds to the gradient of $J_\tau(W) = J(W) + \tau B(W)$.

The barrier parameter $\tau > 0$ goes to 0 during the evolution of the flow, so that the modified assignment flow takes the form

$$\dot{W}(t) = R_{W(t)}\big[S_{\tau(t)}(W(t))\big] = R_{W(t)}\big[S(W) + \tau(t)\nabla B(W)\big], \tag{4.16}$$

where $\tau : \mathbb{R}_{\geq 0} \to \mathbb{R}_{>0}$ is a monotonously decreasing function with $\lim_{t\to\infty} \tau(t) = 0$. This flow is initialized at

$$W_i(0) = c_{\text{low}} + \frac{1 - \langle c_{\text{low}}, \mathbb{1}_n\rangle}{\langle c_{\text{up}}, \mathbb{1}_n\rangle - \langle c_{\text{low}}, \mathbb{1}_n\rangle}(c_{\text{up}} - c_{\text{low}}), \qquad \forall i \in I. \tag{4.17}$$

Since we use zero-mean filters due to (4.8) and (4.7), the initialization (4.17) is strictly feasible for (4.3).

It remains to specify the gradients of the barrier functions that are required to evaluate the vector field in (4.16). The gradient $\nabla B(W) \in \mathbb{R}^{m\times n}$ of the barrier function (4.14) is given by

$$\nabla B_{\text{low}}(W)_i = -\frac{1}{m}\frac{1}{\frac{1}{m}W^\top \mathbb{1}_m - c_{\text{low}}}, \quad \nabla B_{\text{up}}(W)_i = \frac{1}{m}\frac{1}{c_{\text{up}} - \frac{1}{m}W^\top \mathbb{1}_m} \tag{4.18}$$

for each pixel $i \in I$, and by

$$\nabla B_{\text{filter}}(W)^j = 2\sum_{k\in K} \frac{f^{(k),j} \star (f^{(k)} * W)}{(d^{(k)})^2 - \|f^{(k)} * W\|_2^2} \tag{4.19}$$

for any label $j \in J$ with $f^{(k),j} \in \mathbb{R}^{p^2}$ being the j-th layer (column) of the k-th filter. Here, $\star$ denotes the cross-correlation operation, i.e. convolution with the mirrored filter. This convolution is performed on the whole image with zero-padding.

If the approximated ℓ^1-norm (4.5) is used instead of the ℓ^2-norm, the barrier function takes the form

$$B_{\text{filter}}(W) = -\sum_{k\in K} \log\left(d^{(k)} - \|f^{(k)} * W\|_{1,\varepsilon}\right) \tag{4.20}$$

with the Euclidean gradient given by

$$\nabla B_{\text{filter}}(W)^j = \sum_{k\in K} \frac{f^{(k),j} \star \nabla\|f^{(k)} * W\|_{1,\varepsilon}}{d^{(k)} - \|f^{(k)} * W\|_{1,\varepsilon}}, \qquad \nabla\|x\|_{1,\varepsilon} = \frac{x}{\sqrt{x\cdot x + \varepsilon^2 \mathbb{1}_m}}, \tag{4.21}$$

where the operations of the latter right-hand side apply componentwise.

4.3 Experiments

In this section, we investigate the influence of the filter constraints on the labeling result. We test the new approach on several academic labeling scenarios and compare the results to those obtained without using these constraints.

Setup

We represent assignments by choosing for each label $j \in J$ some color $c^{(j)} \in [0,1]^3$ in the RGB color space. Then an assignment $W \in \overline{\mathcal{W}} \subset \mathbb{R}^{m \times n}$ is represented by the color image $u \in \mathbb{R}^{m \times 3}$ given by $u_i = \sum_{j \in J} W_{ij} c^{(j)} \in [0,1]^3$ for each pixel $i \in I$.

We consider *three different data sets* in order to check the effect of constraints on **(a)** primitive shape information, **(b)** spatial relations (inclusion of regions), and **(c)** the separation of fore- and background each defined by *several* labels.

The first data set **(a)** contains binary rectangles and ellipses. Filters of size 3×3 were trained for rectangles against ellipses, i.e., all assignments in a set $\mathcal{I}^+$ represent rectangles, while assignments in the complement set $\mathcal{I}^-$ represent ellipses. The second data set **(b)** comprises three labels (white, orange, black) forming white ellipses overlapped by orange ellipses on black background (see Figure 4.2 for illustration). All ellipses have varying radii, orientation and position. Filters of size 5×5 were used to separate the positive class $\mathcal{I}^+$ defined by inclusions of regions, whereas these topological relations are violated in the negative class $\mathcal{I}^-$. The third data set **(c)** consists of Voronoi diagrams, with each polygon labeled by either one of three foreground labels (red, green, blue) and likewise for the background (black, gray, white). Both foreground and background are connected and the foreground is located in the center of the image domain. The negative class $\mathcal{I}^-$ is defined by randomly labeled Voronoi diagrams.

Implementation Details

For numerical integration, we used the geometric explicit Euler discretization (2.111) with according modifications. An adaptive step size $h > 0$ determined by backtracking line search was used for the discretization in order to keep the assignment matrices $W^{(t)}$ feasible. As maximal step size we chose $h = 1$, which produced satisfying results. As is common for interior point methods, we used an exponential function $\tau(t) = \tau_0 e^{-\alpha t}$ with $\alpha > 0$ as the barrier parameter function. In our experiments, $\alpha = 0.1$ and $\tau_0 = 100$ turned out to be an appropriate choice. We terminated the iteration either after 5000 steps or when both the average entropy $-\frac{1}{m \log n} \sum_{i \in I} \sum_{j \in J} W_{ij}^{(t)} \log W_{ij}^{(t)}$ dropped below the threshold 10^{-3} and τ dropped below 10^{-10}.

Results

(a) For the binary data set (rectangles/ellipses), the filter space $\mathcal{F}$ has dimension 8. We used the four filters corresponding to the eigenvalues less than 1. Inspecting these filters reveals discrete versions of the partial derivatives ∂_{xy}, ∂_{xxy}, ∂_{xyy} and ∂_{xxyy}. The upper bounds for the filter constraints were set to $d = 2\|h * W\|$, where W is an assignment representing two rectangles. Figure 4.1 illustrates that using the filter constraints enables to remove noise, to regularize the rectangle, and to rectify the ellipse by *imposing local shape constraints*.

(b) For the second data set, we used all 48 filters obtained by the generalized eigenvalue problem. The first 24 filters corresponding to eigenvalues less than 1 contribute to separating the orange region from the background. The remaining 24 filters regularize the boundary of the white region (see Figure 4.3). The upper bounds for the filter constraints were set to $d = \max_{W \in \mathcal{I}^+} \|f * W\|$. The bounds c_{low} and c_{up} were set in a similar way. We used the distance matrix $D_{ij} = \frac{1}{n} \|\tilde{W}_i - e_j\|_2$, where the matrix $\tilde{W} \in \mathbb{R}^{m \times n}$ was obtained by adding white noise ($\sigma^2 = 4$) to the ground truth assignment. Figure 4.4 demonstrates that the constraints notably improve the results, and that in addition to the filter constraints, cardinality constraints are

essential to preserve thin structure. In order to demonstrate the potential of the constraints for *spatial pattern formation*, we repeated the experiments with *pure noise* as input data. Figure 4.5 demonstrates the strong regularizing effect of the constraints.

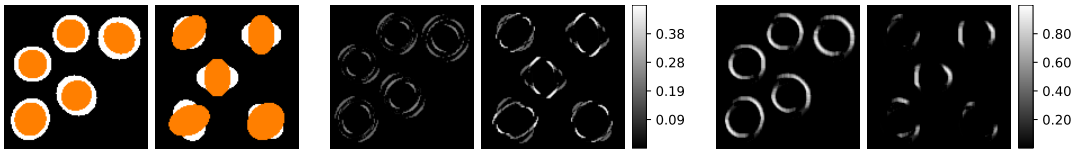
(c) For the Voronoi data set, the filters determined by the eigenvalue problem can be subdivided into three groups: The first 8 filters (eigenvalues ≈ 0.12) contribute to separating the three foreground labels from the three background labels. The next 16 filters (eigenvalues ≈ 0.52) regularize the foreground. The last 16 filters (eigenvalues $\approx 1.46 > 1$) regularize the background (Figure 4.7). The distance matrix and the parameters for the constraints were set as described above for case (b). The results shown by Figure 4.8 (b), (f), (g) and (h) demonstrate the effect of the three groups of filters. Repeating the experiments with pure noise as input data illustrates how *spatial patterns* are induced by the constraints (Figure 4.9).

(a) (b) (c) (d) (e) (f) (g)

Figure 4.1: Representing and enforcing rectangular structure. (a) shows the original gray-scale image and **(b)** shows a noisy version of it, which was used as input data. **(c)** and **(d)** show the labeling results *without imposing constraints* obtained through Riemannian averaging over neighborhoods of sizes 3×3 and 7×7 respectively. **(e)** and **(f)** show the results of the new approach (without cardinality constraints) with neighborhood size 3×3 and four filters of size 3×3, which were trained for rectangles against ellipses. These 4 filters prefer horizontal and vertical edges. For **(e)**, the ℓ^2-norm was used for filter constraints. For **(f)** and **(g)**, $\|\cdot\|_{1,\varepsilon}$ with $\varepsilon = 0.1$ and $\varepsilon = 0.01$ was used.

(a) *Positive* labeling examples in $\mathcal{I}^+$. (b) *Negative* labeling examples in $\mathcal{I}^-$.

Figure 4.2: Illustration of the training sets $\mathcal{I}^+, \mathcal{I}^- \subset \overline{\mathcal{W}}$. Positive examples **(a)** are defined by topological relations: orange ellipses are completely contained in the white ones. Negative examples **(b)** are labelings where this topological relation is violated.



(a) (b) (c)

Figure 4.3: Illustration of the generalized eigenvalue filters. (a) shows one label assignment in $\mathcal{I}^+$ and one assignment in $\mathcal{I}^-$, used to illustrate the filter outputs in (b), (c). The computed filters of size 5×5 can be subdivided into two groups. The first 24 filters $f^{(1)}, \ldots, f^{(24)}$ respond to the boundary of the orange and black regions, and they have a large response at the border between the orange and black regions. This is illustrated by **(b)** which shows the absolute value of $f^{(1)} * W$. The last 24 filters $f^{(25)}, \ldots, f^{(48)}$ mainly respond to the boundary of the white regions. **(c)** shows the absolute value of $f^{(48)} * W$.

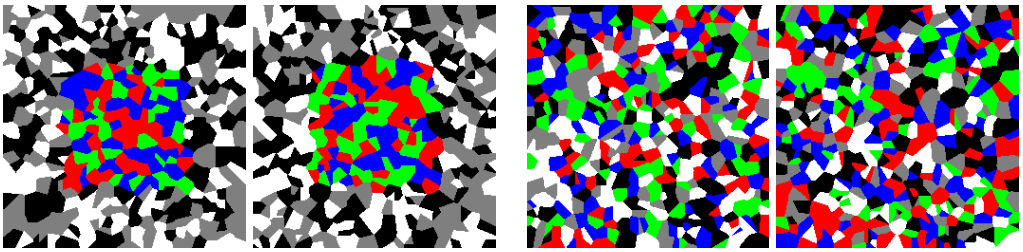
(a) 100% (b) 92.5% (c) 91.0% (d) 89.0% (e) 91.1% (f) 92.4%

(g) 93.7% (h) 93.8% (i) 92.9% (j) 93.2% (k) 92.2% (l) 91.9% (m) 92.1%

Figure 4.4: Experimental results obtained with and without constraints. (a) shows the ground truth assignment. A noisy version of this assignment was used as input data. The percentages of correctly labeled pixels are shown below the images. The results obtained *without constraints* are shown in **(b)**-**(d)** for neighborhood sizes 3×3, 5×5 and 7×7 respectively. For the results **(e)**-**(m)** of the new approach, 3×3 neighborhoods were used for spatial regularization. For **(e)**, only cardinality constraints were used. For **(f)**, only filter constraints were used (48 filters of size 5×5). For **(g)**-**(m)**, *both cardinality constraints and filter constraints* were used. **(g)** and **(h)** were obtained with 24 filters using the ℓ^2-norm and the approximated ℓ^1-norm $\|\cdot\|_{1,\varepsilon}$ with $\varepsilon = 0.01$ respectively. For **(i)** and **(j)**, 48 filters were used. For **(k)**-**(m)**, the distance matrix was rescaled by a factor of 0.01, and 48 filters were used with ℓ^2-norm as well as $\|\cdot\|_{1,\varepsilon}$ with $\varepsilon = 0.1$ and $\varepsilon = 0.01$ respectively.

(a) (b) (c)

Figure 4.5: Spatial pattern formation induced by pure noise and convex label constraints. All experiments were performed using a 3×3 neighborhood for spatial regularization, and using both cardinality constraints and filter constraints based on 48 filters and $\|\cdot\|_{1,\varepsilon}$ with $\varepsilon = 0.01$. Panels **(a)**-**(c)** show random spatial labeling patterns on the right induced by the random noise images on the left. These results demonstrate how filter constraints favor local shape and topological spatial structure on image labelings within our geometric approach to label assignments.



(a) *Positive* labeling examples in $\mathcal{I}^+$. (b) *Negative* labeling examples in $\mathcal{I}^-$.

Figure 4.6: Illustration of the training sets $\mathcal{I}^+, \mathcal{I}^- \subset \overline{\mathcal{W}}$. Both the foreground region and the background region of these Voronoi tilings are defined by three labels: red, green, blue and black, gray, white, respectively. Positive examples in $\mathcal{I}^+$ are defined by approximately square-shaped foreground regions that are simply connected and centered in the middle of the image domain. Negative examples $\mathcal{I}^-$ contain polygons that are randomly labeled and distributed over the image domain.

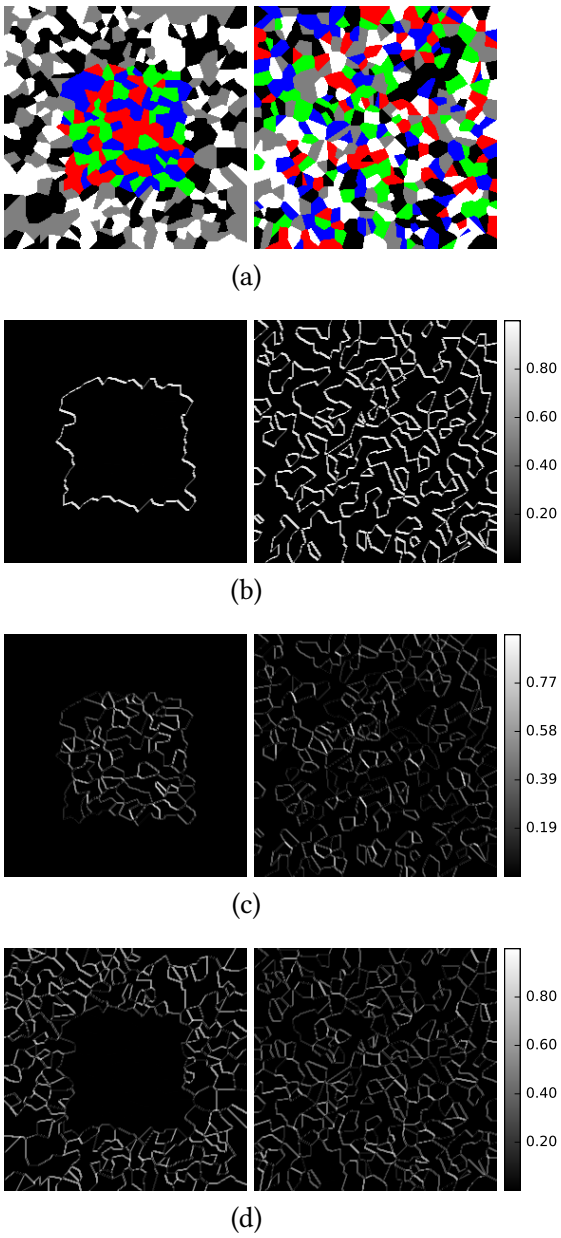
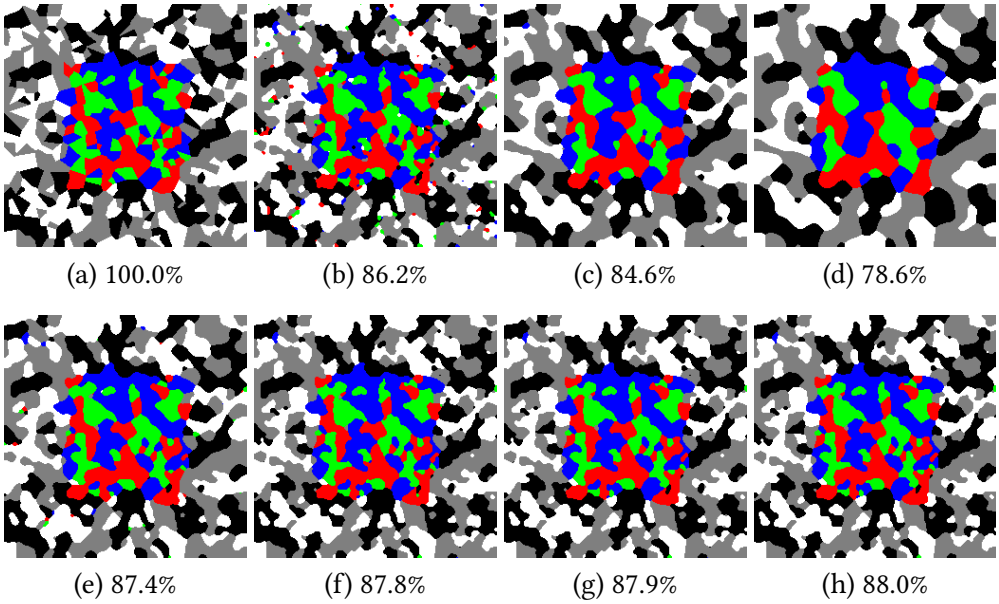


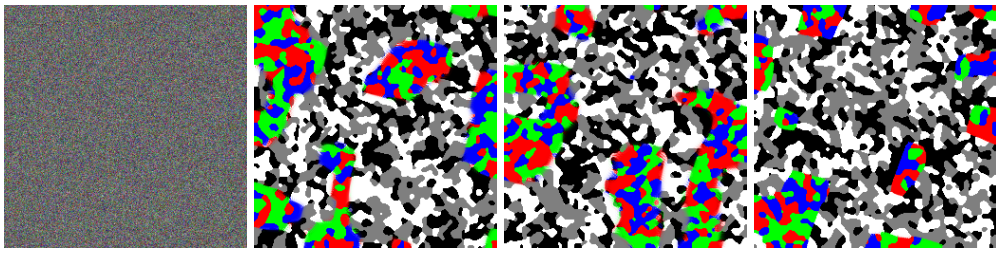
Figure 4.7: Illustration of the generalized eigenvalue filters. (a) shows one assignment in $\mathcal{I}^+$ and $\mathcal{I}^-$, respectively, used to illustrate the filter outputs. The computed filters of size 3×3 can be subdivided into three groups. The first eight filters $f^{(1)}, \ldots, f^{(8)}$ regularize the boundary *between* foreground (red, green, blue) and background (black, gray, white). **(b)** shows the absolute value of the filter result $f^{(4)} * W$ as example. Filters $f^{(9)}, \ldots, f^{(24)}$ regularize the boundaries *within the foreground* as illustrated by **(c)**, which shows the absolute value of $f^{(10)} * W$. Eventually, filters $f^{(25)}, \ldots, f^{(40)}$ regularize the boundaries *within the background* as illustrated by **(d)**.



(a) 100.0% (b) 86.2% (c) 84.6% (d) 78.6%

(e) 87.4% (f) 87.8% (g) 87.9% (h) 88.0%

Figure 4.8: Experimental results obtained with and without constraints. **(a)** shows the ground truth assignment. A noisy version of this assignment was used as input data. The percentages of correctly labeled pixels compared to the ground truth **(a)** are shown below the images. Panels **(b)**-**(d)** show the results obtained *without* constraints using neighborhood sizes 3×3, 5×5 and 7×7 respectively. **(e)** is the result for neighborhood size 5×5, but with a rescaled (factor 100) distance matrix. **(f)**-**(h)** show the results *with* constraints using 8 filters, 24 filters and 40 filters, respectively. These results were computed using 3×3 neighborhoods for spatial regularization and the ℓ^2-norm for the filter constraints.



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

Figure 4.9: Spatial pattern formation induced by pure noise and convex label constraints. **(a)** shows a random assignment W using the same color coding as for the Voronoi polygons. Each panel **(b)**-**(d)** shows the result of labeling a different random input image of type (a). All experiments were done using 3×3 neighborhoods, without cardinality constraints, and with filter constraints based on 40 filters and $\|\cdot\|_{1,\varepsilon}$ with $\varepsilon = 0.1$. The results demonstrate how the filter constraints enforce both the scale and the spatial structure of fore- and background regions that are randomly located due to the pure noise data.

4.4 Conclusion

We extended the assignment flow approach from Section 2.2 in order to incorporate global convex constraints on the labeling result using linear filters in the label space. This extension was mathematically formulated so as to preserve smoothness of the overall approach. We showed how filters can be determined by solving a generalized eigenvalue problem in order to represent statistical prior knowledge about local shape and spatial relation. Experimental results demonstrate the potential of the approach for imposing these constraints onto labelings of noisy image data.

An interesting research question is how filters of small support can be used to represent and enforce the structure of labelings at multiple spatial scales.

Another exciting research direction is how to induce label statistics via weights, which parameterize the similarity map $S(W)$ (2.104), instead of hard constraints and barrier functions as presented in this chapter. To this end, these weights must be generalized for interactions between labels, as in Section 3.6. Whether these weights can be obtained from learned filters $f \in \mathcal{F}$ or how they can be learned using training sets as in Section 4.2.2 must then be examined accordingly.

5 Unsupervised Label Learning on Manifolds

5.1 Introduction and Overview

Motivation

Geometric methods based on manifold models of data and Riemannian geometry are nowadays widely employed in image processing and computer vision [TS16]. Covariance descriptors, in particular, play a prominent role [TS16, CS16]. They are typically applied to the detection and classification of entire images (e.g. faces, texture) or videos (e.g. action recognition), or as descriptors of local image structure. An important task in this context is to compute a codebook of covariance descriptors that can be used for solving a task at hand like, e.g., image classification by nearest-neighbor search [CSBP13], or image labeling [Kap+15] using the codebook descriptors as labels.

The classical approach for the unsupervised learning of feature prototypes ('labels') is the mean-shift iteration [FH75, CM02], which iteratively seeks modes of the feature density distribution through the averaging of features within local neighborhoods. This has been generalized to *manifold-valued* features by [SM09], by replacing ordinary mean-shifts with Riemannian means [Kar77]. Clustering methods such as the mean-shift iteration work *entirely in feature space* and ignore the *spatial structure of label assignments* to data, which is unfavorable in connection with image labeling. Figure 5.1 illustrates why the spatial structure of label assignments should also drive the evolution of labels in feature space for *unsupervised* label learning, if the resulting labels are subsequently used for *supervised* image labeling for which spatial regularization is typically enforced as well. The common way to take into account the spatial structure of label assignments is to augment the feature space by spatial coordinates, e.g., by turning a color feature (r, g, b) into the feature vector (x, y, r, g, b). However, this *merge* of feature space and spatial domain has a conceptual drawback. The *same* color vector $(\overline{r}, \overline{g}, \overline{b})$ observed at two *different* locations $(x_1, y_1, \overline{r}, \overline{g}, \overline{b})$, $(x_2, y_2, \overline{r}, \overline{g}, \overline{b})$ defines two *different* feature vectors, and hence these two feature vectors may be assigned to different prototypes during clustering despite containing the same color information. Furthermore, clustering spatial coordinates into centroids by mean-shifts (together with the features) differs from *unbiased spatial* regularization as performed by variational approaches, graphical models or the assignment flow presented in Section 2.2, where regularization does *not* depend on the location of centroids and the corresponding shape of local density modes.

We show in this chapter how the assignment flow approach can be combined with basic clustering approaches after extending the latter to feature manifolds, to perform *unsupervised label learning* from manifold-valued feature data through spatially regularized label assignment. Our approach is *consistent and natural* in that the very same approach for supervised image labeling is also used for the unsupervised learning of proper labels for this task. This approach

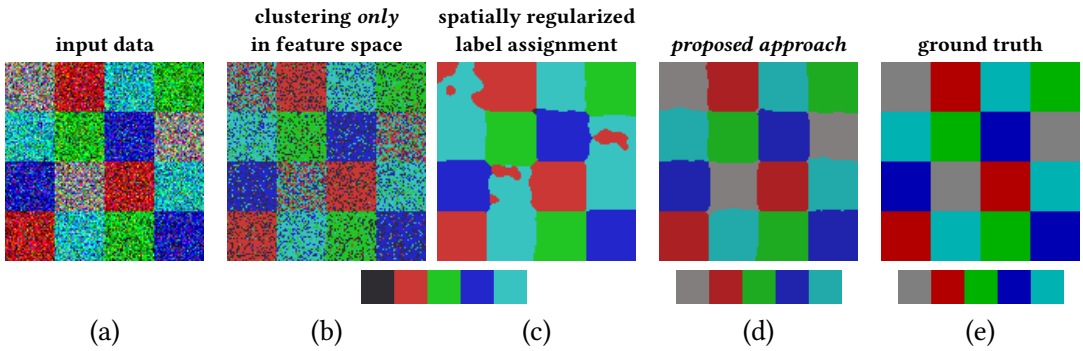


Figure 5.1: Importance of spatially regularized assignments for label *learning*. **(a)** Input data: a synthetic image corrupted by Gaussian noise. **(b)** + **(c)** The classical two-step approach of clustering in feature space first (panel b) followed by *supervised* label assignment (panel c) performs poorly, despite spatial regularization. **(d)** By *coupling* label evolution and spatially regularized assignment, both the label set and the labeled image can be drastically improved. **(e)** Ground truth labeling and label set. Label sets resulting from (b), (d) and (e) are depicted below the respective image labeling results.

has the following properties:

(i) The approach incorporates and performs *unsupervised learning of manifold-valued features*, henceforth called *labels*. The approach applies to any feature manifold equipped with a Riemannian structure for which Riemannian means are well-defined and computationally feasible. Experiments using $\mathbb{S}^1$-valued data (2D orientations), $\mathrm{SO}(3)$-valued data (orthogonal frames) and features on the positive definite matrix manifold (covariance descriptors) illustrate our approach.

(ii) The evolution of labels (unsupervised learning) is driven by *spatially regularized assignments* which are *not* biased toward any spatial centroids. This is accomplished by applying the assignment flow approach.

(iii) The smooth settings of both (i), (ii) enable to define a *smooth coupled flow*

$$(\dot{M}, \dot{W}) = \mathbf{F}(M, W) \tag{5.1}$$

where $\dot{M}$ denotes the evolution of labels and $\dot{W}$ the evolution of spatially regularized label assignments that *interact* through a coupling vector field $\mathbf{F}$. This interaction keeps both domains (i) and (ii) *separate* and hence enables to apply flexibly our approach to various feature manifolds, using the *same* regularized assignment mechanism.

This chapter is based on the joint work with Matthias Zisler, Freddie Åström, Stefania Petra and Christoph Schnörr. A preliminary version of the approach, called 'coupled flow A' in this chapter (see Figure 5.2), was published in the conference paper [Zer+18]. In this chapter, we present a one-parameter family of *unsupervised assignment flows* that smoothly interpolate the 'coupled flow A' and the more general 'coupled flow B', including these two flows as special cases. This approach is published in the journal paper [ZZPS20a].

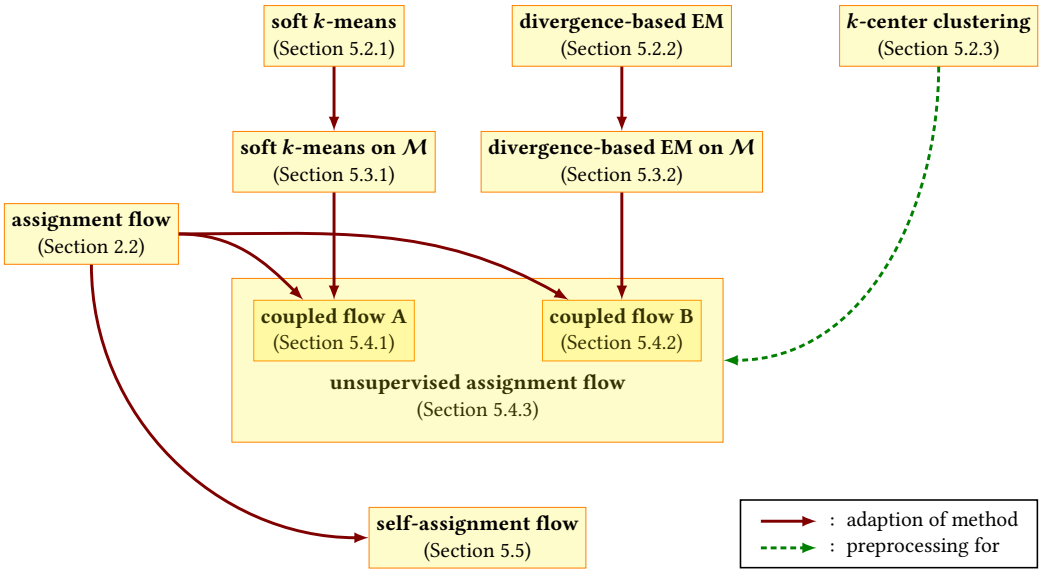


Figure 5.2: Organization of this chapter. Three basic clustering algorithms are summarized in Section 5.2. Two of them, *soft-k-means clustering* and *divergence-based EM-iteration*, are generalized to feature data taking values in a Riemannian manifold $\mathcal{M}$ and coupled with the (spatially regularized) *assignment flow*. A smooth interpolation of the resulting *coupled flow A* and *coupled flow B* finally defines the *unsupervised assignment flow*. For comparison, we discuss the *self-assignment flow* (Section 5.5) as an alternative extension of the assignment flow to the unsupervised scenario.

Organization

We summarize three basic clustering concepts in Section 5.2. While greedy-based k-center clustering (Section 5.2.3) will only serve as a preprocessing step, soft-k-means clustering (Section 5.2.1) and divergence-based EM-iteration (Section 5.2.2), which perform classical label evolution by mean-shift iteration, will be building blocks for subsequent methods (see Figure 5.2). First these two methods are adjusted to manifold-valued data (Section 5.3.1 and Section 5.3.2), and then each of them is modified and coupled with the assignment flow (cf. Section 2.2) that induces a sparsifying effect through spatial regularization. Coupling with soft-k-means leads to the 'coupled flow A' (Section 5.4.1), while coupling with EM-iteration leads to the 'coupled flow B' (Section 5.4.2). Afterwards we provide a more general natural definition of a one-parameter family of *unsupervised assignment flows* (Section 5.4.3) that smoothly interpolates both coupled flows and includes them as special cases. Numerical integration of the unsupervised assignment flow is discussed in Section 5.4.4. In Section 5.5, we present the *self-assignment flow* [ZZPS20b] as an alternative extension of the assignment flow to the unsupervised scenario and compare it to the unsupervised assignment flow from a theoretical point of view. The unsupervised assignment flow is then elaborated for particular feature manifolds (Section 5.6) that will be used as case studies in numerical experiments in Section 5.7.

Notation

Firstly, we use the notation from Section 2.1, i.e., $(\mathcal{M}, g)$ *generally* denotes some Riemannian manifold $\mathcal{M}$ with metric g. Exp_p is the exponential map (2.13) with respect to the Levi-Civita connection on $\mathcal{M}$. The map $\widehat{g}\colon \mathfrak{X}(\mathcal{M}) \to \mathfrak{X}^*(\mathcal{M})$ is the tangent-cotangent isomorphism (2.9) mapping vector fields to covector fields and $\operatorname{grad} f = \widehat{g}^{-1}(\mathrm{d}f)$ denotes the Riemannian gradient of a function $f\colon \mathcal{M} \to \mathbb{R}$.

Secondly, we use the notation from Section 2.2, i.e., the sets I, J denote

$$\begin{aligned} I &\colon \text{ set of data indices,} \\ J &\colon \text{ set of label indices,} \end{aligned} \tag{5.2}$$

with cardinalities $|I|$ and $|J|$. The $(d-1)$-dimensional probability simplex is denoted by

$$\Delta_d = \left\{ p \in \mathbb{R}^d \colon p_j \geq 0,\ j \in [d],\ \langle \mathbb{1}_d, p \rangle = 1 \right\} \subset \mathbb{R}^d. \tag{5.3}$$

$\mathcal{S} = \operatorname{rint}(\Delta_{|J|}) \subset \mathbb{R}^{|J|}$ is the open probability simplex (2.79) and $\mathcal{W} \subset \mathbb{R}^{|I| \times |J|}$ denotes the assignment manifold (2.85). The map $\exp_W$ denotes the lifting map given by (2.90) and (2.100a). $L(W)$ is the likelihood map (2.102) and $S(W)$ is the similarity map (2.104).

Furthermore, it will be convenient to denote the exponential function with vectors as argument in two alternative ways,

$$\exp(x) = e^x \coloneqq (e^{x_1}, \ldots, e^{x_d})^\top. \tag{5.4}$$

This exponential function (*without* subscript) should not be confused with the exponential map Exp_p and the lifting map $\exp_W$ mentioned above.

5.2 Basic Clustering

We briefly summarize in this section the basic iterative schemes

- soft-k-means clustering in Euclidean spaces (Section 5.2.1),
- clustering using mixture distributions, divergence functions and the EM-algorithm (Section 5.2.2), and
- greedy-based clustering in metric spaces (Section 5.2.3).

The first two approaches will be generalized to *manifold-valued* data (features) in Section 5.3 and coupled with the assignment flow for spatial regularization in Section 5.4.

Greedy-based k-center clustering applies to any metric space, in particular to manifolds with the Riemannian distance or a suitable divergence as surrogate distance function. The method has linear complexity and comes with a performance guarantee. Hence this method is suited for fast data selection in a preprocessing step, to obtain an overcomplete codebook (set of prototypes) as initialization for manifold-valued clustering, which subsequently optimizes and sparsifies this codebook in a computationally more sophisticated way.

5.2.1 Euclidean Soft-k-Means Clustering

The content of this subsection can be found in numerous papers and textbooks. We merely refer to the survey [Teb07] and to the bibliography therein.

Given data $z^1, \dots, z^{|I|} \in \mathbb{R}^d$, we consider the task of determining prototypes

$$M = \{m^1, \dots, m^{|J|}\} \subset \mathbb{R}^d \tag{5.5}$$

by minimizing the k-means criterion[1]

$$E(M) = \sum_{i \in I} \min_{j \in J} \|z^i - m^j\|^2 = \sum_{i \in I} \operatorname{vecmin}(D_i(M)), \tag{5.6}$$

where

$$D_i(M) = \big(D_{i1}(M), \dots, D_{i|J|}(M)\big) = \big(\|z^i - m^1\|^2, \dots, \|z^i - m^{|J|}\|^2\big) \tag{5.7}$$

and

$$\operatorname{vecmin}(v) = \min_{j \in [n]} v_j, \qquad v \in \mathbb{R}^n, \quad n \in \mathbb{N}. \tag{5.8}$$

Soft-k-means is based on the *smoothed* objective

$$E_\varepsilon(M) = -\varepsilon \sum_{i \in I} \log \Bigg(\sum_{j \in J} \exp\Big(- \frac{\|z^i - m^j\|^2}{\varepsilon} \Big) \Bigg), \quad \varepsilon > 0 \tag{5.9}$$

which results from approximating the inner minimization problem of evaluating $E(M)$ using the log-exponential function [RW09, p. 27] with smoothing parameter ε. Similar to the basic k-means algorithm, *soft-k-means* clustering solves the stationarity conditions

$$\nabla_{m^j} E_\varepsilon(M) = 0, \quad j \in J \tag{5.10}$$

by fixed point iteration in terms of iteratively computing the **soft-assignments**

$$p^i_{\varepsilon,j}(M) = \frac{\exp\big(-D_{ij}(M)/\varepsilon\big)}{\sum_{k \in J} \exp\big(-D_{ik}(M)/\varepsilon\big)}, \quad q^j_{\varepsilon,i}(M) = \frac{p^i_{\varepsilon,j}(M)}{\sum_{k \in I} p^k_{\varepsilon,j}(M)}, \quad i \in I,\ j \in J \tag{5.11a}$$

with the so-called **mean shifts**

$$m^j = \sum_{i \in I} q^j_{\varepsilon,i}(M) z^i, \quad j \in J. \tag{5.11b}$$

The distributions

$$p^i_\varepsilon(M) \in \Delta_{|J|}, \quad i \in I \tag{5.12}$$

given by (5.11a) represent the soft-assignments $p^i_{\varepsilon,j}(M)$ of each data point z^i, $i \in I$ to each prototype m^j, $j \in J$, whereas the distributions

$$q^j_\varepsilon(M) \in \Delta_{|I|}, \quad j \in J \tag{5.13}$$

determine the convex combinations of data points that determine each prototype m^j by the mean shift (5.11b). Iterating the two steps (5.11) evolves the prototypes M until they reach a local minimum of the objective (5.9).

[1] The symbol 'k' is commonly used in the literature. We prefer in this thesis, however, the more specific symbol J as index set for prototypes and use k (like i, j etc.) as free index.

5.2.2 Divergence Functions and EM-Iteration

An alternative and widely applied approach to clustering utilizes class-conditional distributions $p(x;\theta_j),\ j \in J$ and a corresponding mixture distribution

$$p(x;\Gamma) = \sum_{j\in J} \pi_j p(x;\theta_j) \tag{5.14}$$

as data model, with parameters

$$\Gamma = (\theta,\pi), \qquad \theta = (\theta_1,\dots,\theta_{|J|}), \qquad \pi = (\pi_1,\dots,\pi_{|J|})^\top \in \mathcal{S}, \tag{5.15}$$

with the relative interior of the probability simplex

$$\mathcal{S} = \operatorname{rint}(\Delta_{|J|}) = \big\{p \in \mathbb{R}^{|J|} \colon p_j > 0,\ j \in J,\ \langle \mathbb{1}, p\rangle = 1\big\}. \tag{5.16}$$

Clustering amounts to estimate the parameters Γ. Since the log-likelihood function corresponding to (5.14) is usually involved, maximizing a lower bound through the EM-iteration (EM: expectation-maximization) is the method of choice,

$$p(j|z^i;\Gamma^{(t)}) = \frac{\pi_j^{(t)} p(z^i;\theta_j^{(t)})}{\sum_{l\in J} \pi_l^{(t)} p(z^i;\theta_l^{(t)})}, \qquad j \in J \qquad \textbf{(E-step, soft-assignment)} \tag{5.17a}$$

$$\begin{aligned} \pi_j^{(t+1)} &= \frac{1}{|I|}\sum_{i\in I} p(j|z^i;\Gamma^{(t)}), \\ \theta_j^{(t+1)} &= \arg\max_{\theta_j} \sum_{i\in I} p(j|z^i;\Gamma^{(t)}) \log p(z^i;\theta_j), \end{aligned} \qquad j \in J. \qquad \textbf{(M-step)} \tag{5.17b}$$

for some initialization $\Gamma^{(0)}$. We refer to [MP00] for background and further details.

Banerjee et al. [BMDG05] studied the case where the class-conditional distributions $p(x;\theta_j)$ of (5.14) belong to an exponential family of distributions [Bar78] and, in particular, their representation in terms of a Bregman divergence function D_φ – cf. Section 2.1.5. Then the resulting data model (5.14) reads

$$p(x;\Gamma) = \sum_{j\in J} \pi_j \exp\big(-D_\varphi(f(x),\eta_j)\big) b_\varphi(x), \tag{5.18}$$

where f denotes a sufficient statistic regarded as a feature vector, the factor b_φ accounts for normalization and $\eta_j = \nabla\psi(\theta_j)$ is determined by θ_j through conjugation of the convex

log-partition function $\psi(\theta_j) = \log \int_{\mathcal{X}} p(x;\theta_j)\mathrm{d}x$. The corresponding EM-updates read

$$p(j|z^i;\Gamma^{(t)}) = \frac{\pi_j^{(t)} \exp\big(-D_\varphi(f(z^i),\eta_j^{(t)})\big)}{\sum_{l\in J}\pi_l^{(t)}\exp\big(-D_\varphi(f(z^i),\eta_l^{(t)})\big)}, \qquad j\in J \qquad \textbf{(E-step, soft-assignment)} \tag{5.19a}$$

$$\begin{aligned}\pi_j^{(t+1)} &= \frac{1}{|I|}\sum_{i\in I} p(j|z^i;\Gamma^{(t)}) \\ \eta_j^{(t+1)} &= \arg\min_{\eta_j} \sum_{i\in I} v_{j|i}(\Gamma^{(t)}) D_\varphi\big(f(z^i),\eta_j^{(t)}\big)\end{aligned}, \qquad j\in J. \qquad \textbf{(M-step)} \tag{5.19b}$$

$$v_{j|i}(\Gamma^{(t)}) = \frac{p(j|z^i;\Gamma^{(t)})}{\sum_{k\in I} p(j|z^k;\Gamma^{(t)})}. \tag{5.19c}$$

Moreover, since the Bregman divergence D_φ is induced by a convex function φ of Legendre type, the parameters $\eta_j,\ j\in j$ can be updated by the **mean-shifts**

$$\eta_j^{(t+1)} = \sum_{i\in I} v_{j|i}(\Gamma^{(t)}) f(z^i), \quad j\in J. \tag{5.20}$$

We exploit the above connection to divergence functions in Sections 5.3.2 and 5.4.2.

5.2.3 Greedy-Based k-Center Clustering in Metric Spaces

We adopt a simple greedy algorithm from [Har11] as a *preprocessing step for data reduction*, due to the following properties:

- It works in any *metric space* (X, d_X),
- it has *linear* complexity $O(|J||I|)$ with respect to the problem size $|I|$ which can be large,
- it comes along with a *performance guarantee*.

The task of k-center clustering is as follows. Given data points

$$X_I = \{z^1,\dots,z^{|I|}\} \subset X, \tag{5.21}$$

the objective is to determine a subset

$$M = \{m^1,\dots,m^{|J|}\} \subset X_I \tag{5.22}$$

that solves the combinatorially hard optimization problem

$$E_\infty^* = \min_{M\subset X_I, |M|=|J|} E_\infty(M), \qquad E_\infty(M) = \max_{z\in X_I} d_X(z,M), \tag{5.23}$$

where $d_X(z,M) = \min_{m\in M} d_X(z,m)$. This problem is approximated by a greedy iteration as follows. Starting with a randomly chosen point $m^1 \in X_I$, the remaining $|J|-1$ points $m^2,\dots,m^{|J|}$

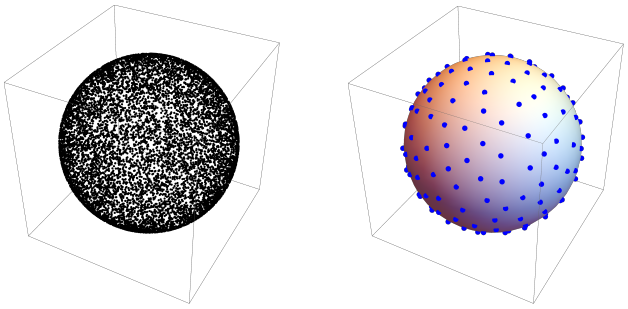


Figure 5.3: Approximation of the k-center clustering objective (5.23). LEFT: 10.000 points on the sphere regarded as manifold equipped with the cosine distance. RIGHT: 200 prototypes determined with linear runtime complexity by metric clustering are *almost uniformly located* in the data set, which qualifies them for unbiased initializations of computationally more involved nonlinear prototype evolutions. This works in any metric space, in particular on feature manifolds using the Riemannian distance or computationally less expensive divergence functions.

are selected by choosing $m^{k+1} \in X_I$ with the largest distance to the points $m^1, \dots, m^k$. This greedy strategy guarantees that the resulting set M is a 2-approximation of the optimum (5.23), i.e. $E_\infty(M) \leq 2E_\infty^*$. We refer to [Har11, Thm. 4.3] for the proof. As a consequence, the subset M of $|J|$ points is almost uniformly distributed in X_I as measured by the metric d_X. Figure 5.3 provides an illustration.

We note that the performance guarantee, i.e. the 2-approximation, follows by the triangle inequality, while the complexity is independent of the properties of a metric space. We will later apply this algorithm in the context of clustering on a Riemannian manifold $\mathcal{M}$, where we have given a smooth (symmetric) divergence function D on $\mathcal{M}$. Most of these divergence functions are squared distances, so that the performance guarantee also holds in this case. One exception might be the rotation-invariant dissimilarity (5.94). Nevertheless, the greedy k-center clustering is applicable but without having a performance guarantee. We will use greedy k-center clustering as preprocessing in order to get an overcomplete set of labels as initial labels for the clustering approaches described in the next section.

5.3 Manifold-Valued Clustering

We generalize the basic iterative clustering schemes of Sections 5.2.1 and 5.2.2 to manifold-valued data.

5.3.1 Manifold-Valued Soft-k-Means Iteration

Let $(\mathcal{M}, g)$ be a smooth Riemannian manifold, and let

$$\left\{z^1, \dots, z^{|I|}\right\} \subset \mathcal{M} \tag{5.24}$$

be given data. We assume a smooth divergence function to be given (cf. Section 2.1.5)

$$D \colon \mathcal{M} \times \mathcal{M} \to \mathbb{R}, \qquad (x, y) \mapsto D(x, y) \tag{5.25}$$

that replaces the Riemannian distance d_g in order to compute Riemannian means more efficiently or even in closed form. Examples are provided in Section 5.6.

We consider the task to determine a set of prototypes

$$M = \{m^1, \dots, m^{|J|}\} \subset \mathcal{M} \tag{5.26}$$

by minimizing an objective function analogous to the soft-k-means objective (5.9),

$$E_\varepsilon(M) := E_\varepsilon(m^1, \dots, m^{|J|}) = -\varepsilon \sum_{i \in I} \log\left(\sum_{j \in J} \exp\left(-\frac{D(z^i, m^j)}{\varepsilon} \right) \right), \quad \varepsilon > 0. \tag{5.27}$$

We next generalize the conditions (5.10). Let $d_j E_\varepsilon(M)$ denote the differential of the function $m^j \mapsto E_\varepsilon(m^1, \dots, m^j, \dots, m^{|J|})$. Then

$$d_j E_\varepsilon(M) = \sum_{i \in I} \underbrace{\frac{\exp\left(-\frac{D(z^i, m^j)}{\varepsilon}\right)}{\sum_{l \in J} \exp\left(-\frac{D(z^i, m^l)}{\varepsilon}\right)}}_{:= p^i_{\varepsilon,j}(M)} d_j D(z^i, m^j) = \sum_{i \in I} p^i_{\varepsilon,j}(M) d_j D(z^i, m^j), \quad j \in J, \tag{5.28}$$

where the **assignment probability vectors** $p^i_\varepsilon(M) \in \Delta_{|J|}$ play the same role as in equations (5.11a) and (5.12). They can be interpreted as *weight functions* depending on the prototypes M: setting temporarily $w_i = p^i_{\varepsilon,j}(M)$, $i \in I$, implies that equation (5.28) has the same structure as the equation on the right of (2.66) after applying the differential on both sides, where we take into account that divergence functions $D(\cdot, \cdot)$ behave like squared distances (Section 2.1.5). Applying the formula (2.10), we obtain the gradients and optimality conditions

$$(\operatorname{grad} E_\varepsilon)_j(M) = \widehat{g}^{-1}\big(d_j E_\varepsilon(M)\big) = \sum_{i \in I} p^i_{\varepsilon,j}(M) \widehat{g}^{-1}\big(d_j D(z^i, m^j)\big) = 0, \quad j \in J. \tag{5.29}$$

Comparing with (5.10) shows that, in the Euclidean case, the mean shift operation (5.11b) is defined by *normalized* weights $q^j_{\varepsilon,i}(M)$ due to (5.11a), conforming to the much more general situation (2.67). While normalization in (5.11a) is a *consequence* of the squared Euclidean distance of the objective (5.9), this may or may not happen in (5.29), depending on the particular manifold $\mathcal{M}$, metric g and divergence function D at hand. Because dividing each optimality condition (5.10) by the corresponding normalization factor in (5.11) does not change the condition, however, and because mean-shift on manifolds is performed with normalized weights, we define

$$p^i_{\varepsilon,j}(M) = \frac{\exp\left(-\frac{D(z^i, m^j)}{\varepsilon}\right)}{\sum_{l \in J} \exp\left(-\frac{D(z^i, m^l)}{\varepsilon}\right)}, \qquad q^j_{\varepsilon,i}(M) = \frac{p^i_{\varepsilon,j}(M)}{\sum_{k \in I} p^k_{\varepsilon,j}(M)}, \quad i \in I,\ j \in J \tag{5.30}$$

and in turn the **mean shift** (fixed point) **iteration**

$$(m^j)^{(t+1)} = \operatorname{Exp}_{(m^j)^{(t)}}\left(-\sum_{i \in I} q^j_{\varepsilon,i}(M^{(t)}) \widehat{g}^{-1}\big(d_j D((m^j)^{(t)}, z^i)\big) \right), \quad j \in J \tag{5.31}$$

analogous to (2.69). The latter equation can be interpreted as a discretization of the system

$$\dot{m}^j(t) = -\alpha \sum_{i \in I} q^j_{\varepsilon,i}(M(t)) \widehat{g}^{-1}\big(d_j D(m^j(t), z^i)\big), \quad j \in J, \qquad \alpha > 0. \tag{5.32}$$

Section 5.6 provides concrete examples for divergence functions on manifolds.

5.3.2 Manifold-Valued EM-Iteration

We consider again the situation (5.24)–(5.26) and adopt the clustering approach of Section 5.2.2. Iteration (5.19) generalizes to

$$p(j|z^i; M^{(t)}) = \frac{\pi_j^{(t)} \exp\big(-D(z^i, (m^j)^{(t)})\big)}{\sum_{l\in J} \pi_l^{(t)} \exp\big(-D(z^i, (m^l)^{(t)})\big)}, \qquad j \in J \qquad \textbf{(E-step, soft-assignment)} \tag{5.33a}$$

$$\begin{aligned} \pi_j^{(t+1)} &= \frac{1}{|I|} \sum_{i\in I} p(j|z^i; M^{(t)}), \\ (m^j)^{(t+1)} &= \operatorname*{arg\,min}_{m^j} \sum_{i\in I} \nu_{j|i}(M^{(t)}) D\big(z^i, (m^j)^{(t)}\big), \end{aligned} \qquad j \in J. \qquad \textbf{(M-step)} \tag{5.33b}$$

$$\nu_{j|i}(M^{(t)}) = \frac{p(j|z^i; M^{(t)})}{\sum_{k\in I} p(j|z^k; M^{(t)})}. \tag{5.33c}$$

Note that we apparently ignore here the connection to class-conditional distributions $p(x;\theta_j)$ of the exponential family that formed the basis for the EM-iteration (5.19). This is not the case, however. Indeed, optimization problem (5.33b) which determines each prototype m^j by minimizing the *expected value* of a squared distance-like function, conforms to the updates (5.19b) and (5.20) of the *expectation parameter* $\eta_j = \nabla\psi(\theta_j) = \mathbb{E}_{\theta_j}[f_i]$, where the expectation is with respect to $p(x;\theta_j)$ and the sufficient statistics $f_i(x)$.

In order to solve problem (5.33b), we proceed analogously to (5.29). Examples with concrete choices of $D(\cdot,\cdot)$ are worked out in Section 5.6.

5.4 Coupling the Assignment Flow and Label Evolution on Feature Manifolds

We show in this section how combining the assignment flow (Section 2.2) and the schemes of Section 5.3 results in *coupled flows* that *simultaneously* perform

- label evolution on a feature manifold, and
- spatially regularized label assignment to given data.

Coupling the assignment flow with the scheme of Section 5.3.1 defines the **coupled flow (CFa)** in Section 5.4.1, whereas coupling the assignment flow with the scheme of Section 5.3.2 defines the **coupled flow (CFb)** in Section 5.4.2. Comparing **(CFa)** and **(CFb)** in Section 5.4.3 shows that both can be interpreted as special instances of a one-parameter family of flows. We will call the latter the **unsupervised assignment flow**.

5.4.1 Spatially Regularized Soft-k-Means on Feature Manifolds

Minimizing the objective function (5.27) induces the assignment probabilities

$$p^i_{\varepsilon,j}(M) = \frac{\exp\big(-\frac{1}{\varepsilon}D(z^i, m^j)\big)}{\sum_{l\in J}\exp\big(-\frac{1}{\varepsilon}D(z^i, m^l)\big)}, \quad i \in I,\ j \in J \tag{5.34}$$

due to (5.28). Regarding the assignment flow, the variables W_{ij} play the same role. The assignment flow (2.105) for W_{ij} reads

$$\dot W_{ij}(t) = W_{ij}(t)\Big(S_{ij}\big(W(t)\big) - \sum_{l\in J} W_{il}(t) S_{il}\big(W(t)\big)\Big), \quad i \in I,\ j \in J, \tag{5.35}$$

where the right-hand side comprises the similarity vectors $S_i(W),\ i \in I$, whose j-th component due to (2.104), (2.103) and (2.102) is given by

$$S_{ij}(W) = \frac{\widehat S_{ij}}{\big\langle \mathbb{1}_{|J|}, \widehat S_i\big\rangle}, \tag{5.36a}$$

$$\widehat S_{ij} = \Big(\prod_{k\in\mathcal{N}_i} L_{kj}(W;M)\Big)^{w_{ik}} = \Big(\prod_{k\in\mathcal{N}_i}\Big(\frac{W_{kj}}{\big\langle W_k, e^{-\frac{1}{\rho}D_k(M)}\big\rangle}\Big)^{w_{ik}}\Big)\exp\Big(\sum_{k\in\mathcal{N}_i} w_{ik}\frac{D(z^k, m^j)}{\rho}\Big), \tag{5.36b}$$

$$D_k(M) = \Big(D\big(z^k, m^1\big), \ldots, D\big(z^k, m^{|J|}\big)\Big). \tag{5.36c}$$

This expression makes explicit how spatial regularization through averaging the given data (in terms of distance vectors) over local neighborhoods, is part of the vector field that drives the assignment flow. As a consequence, label assignments induced by $W(T),\ T \gg 0$, are spatially more coherent.

Hence, we propose to replace in (5.32) the assignment variables $p^i_{\varepsilon,j}(M)$ given by (5.34), where *no* spatial regularization is involved, by the assignment variables W_{ij}. The resulting **coupled flow (CFa)** that simultaneously performs label evolution and label assignment reads

$$\textbf{(CFa)} \quad \begin{cases} \dot m^j(t) = -\alpha\displaystyle\sum_{i\in I} q^j_{\varepsilon,i}(W)\widehat g^{-1}\big(d_j D(z^i, m^j)\big), \quad m^j(0) = m^j_0, \quad \alpha > 0, \quad j \in J, \\ \qquad\qquad q^j_\varepsilon(W) = \dfrac{W^j}{\big\langle \mathbb{1}_{|I|}, W^j\big\rangle}, \quad j \in J \\ \dot W_i(t) = R_{W_i(t)}\big(S_i\big(W(t)\big)\big), \quad W_i(0) = C_{\mathcal{S}}, \quad i \in I, \end{cases} \tag{5.37}$$

with W^j being the j-th row of $W \in \mathcal{W}$ and a user parameter α that enables to adjust the time scale of the label flow induced by $\dot m^j(t),\ j \in J$ relative to the assignment flow induced by $\dot W_i(t),\ i \in I$.

5.4.2 Spatially Regularized EM-Iteration on Feature Manifolds

The scheme of Section 5.3.2 and the update formulas (5.33) suggest an alternative coupling of label evolution and the assignment flow. Equation (2.102) reads

$$L_{ij}(W;M) = \frac{W_{ij}e^{-\frac{1}{\rho}D(z^i,m^j)}}{\sum_{l\in J} W_{il}e^{-\frac{1}{\rho}D(z^i,m^l)}}, \tag{5.38}$$

which agrees with the right-hand side of (5.33a), except for the scaling parameter ρ and the assignment variables W_{ij} in place of the mixture coefficients π_j. Indeed, since there is *no* interaction between different spatial locations $i \in I$ on the right-hand side of (5.38), $L_{ij}(W_i; M)$ can be interpreted as *local* posterior probability of label j given the observation z^i, in agreement with the left-hand side of (5.33a). Likewise, applying the first update equation of (5.33b) to (5.38) yields

$$W_{ij}^{(t+1)} = \frac{1}{|I|} \sum_{i \in I} L_{ij}\big(W^{(t)}; M^{(t)}\big), \quad j \in J \tag{5.39}$$

which does *not* depend on $i \in I$. We therefore take into account spatial regularization by replacing the mixture coefficients $\pi_j,\ j \in J$ by the variables $W_{ij},\ i \in I, j \in J$, that are governed by the assignment flow and hence *do* spatially interact. The resulting **coupled flow (CFb)** reads

$$\textbf{(CFb)} \quad \begin{cases} \dot m^j(t) = -\alpha \displaystyle\sum_{i \in I} v_{j|i}\big(W(t), M(t)\big) \widehat{g}^{-1}\big(d_j D(z^i, m^j(t))\big), \quad m^j(0) = m_0^j,\ \alpha > 0,\ j \in J, \\ \qquad\qquad v_{j|i}(W, M) = \dfrac{L_{ij}(W; M)}{\sum_{k \in I} L_{kj}(W; M)}, \quad L_{ij}(W; M) = \dfrac{W_{ij} e^{-\frac{1}{\rho} D(z^i, m^j)}}{\sum_{l \in J} W_{il} e^{-\frac{1}{\rho} D(z^i, m^l)}}, \\ \dot W_i(t) = R_{W_i(t)}\big(S_i\big(W(t)\big)\big), \quad W_i(0) = C_{\mathcal{S}}, \quad i \in I, \end{cases} \tag{5.40}$$

where $W(t)$ depends on M through (5.36).

5.4.3 Unsupervised Assignment Flow

We examine the relation between the coupled flows (CFa) (5.37) and (CFb) (5.40). Comparing $L_{ij}(W; M)$ given by (5.40) with $q_{\varepsilon,i}^j(W)$ given by (5.37) shows due to $\sum_{l \in J} W_{il} = 1,\ i \in I$ and

$$W_{ij} = \lim_{\rho \to \infty} L_{ij}(W; M) \tag{5.41a}$$

that

$$q_{\varepsilon,i}^j(W) = \lim_{\rho \to \infty} v_{j|i}(W, M). \tag{5.41b}$$

We conclude that **(CFa) is a limit case of (CFb)**. Since the scaling parameter ρ plays a unique role in (2.102), however, we propose to parameterize $L_{ij}(W; M)$ of (5.40) in the *same* way, but with another *independent* parameter $\sigma > 0$ replacing ρ, in order to 'interpolate' smoothly between the coupled flows (CFa) and (CFb) in the sense of (5.41).

As a result, the final form of our approach, called **unsupervised assignment flow (UAF)**, reads

$$\textbf{(UAF)} \quad \begin{cases} \dot m^j(t) = -\alpha \displaystyle\sum_{i \in I} v_{j|i}\big(W(t), M(t)\big) \widehat{g}^{-1}\big(d_j D(z^i, m^j(t))\big), \quad m^j(0) = m_0^j, \quad \alpha > 0,\ j \in J, \\ \qquad\qquad v_{j|i}(W, M) = \dfrac{L_{ij}^\sigma(W; M)}{\sum_{k \in I} L_{kj}^\sigma(W; M)}, \quad L_{ij}^\sigma(W; M) = \dfrac{W_{ij} e^{-\frac{1}{\sigma} D(z^i, m^j)}}{\sum_{l \in J} W_{il} e^{-\frac{1}{\sigma} D(z^i, m^l)}}, \quad \sigma > 0, \\ \dot W_i(t) = R_{W_i(t)}\big(S_i\big(W(t)\big)\big), \quad W_i(0) = C_{\mathcal{S}}, \quad i \in I, \end{cases} \tag{5.42}$$

with user parameters $\alpha > 0$ controlling the relative speed of label vs. assignment evolution, and parameter $\sigma > 0$ as just discussed.

As already mentioned in Section 5.2.3, the greedy k-center clustering provides an overcomplete set of labels in a preprocessing step which is used as initial condition $\{m_0^j\}_{j\in J}$ for the prototype component of the unsupervised assignment flow (5.42).

5.4.4 Geometric Numerical Integration

In this subsection, we detail the iterative scheme that is used in Section 5.7 for numerically integrating the unsupervised assignment flow (5.42). We rewrite these equations more compactly in the form

$$\dot{W}_i(t) = R_{W_i(t)} F_i\big(W(t), M(t)\big), \qquad W_i(0) = C_{\mathcal{S}}, \qquad i \in I, \tag{5.43a}$$

$$\dot{m}^j(t) = G_j\big(W(t), M(t)\big), \qquad m^j(0) = m_0^j, \qquad j \in J, \tag{5.43b}$$

where the dependency of $F_i(W, M) = S_i(W)$ on M is implicitly given via the distance vectors (5.36c) and the dependency of the similarity vectors on these distance vectors – cf. (5.36).

In order to uniformly evaluate our approach for various feature manifolds $\mathcal{M}$, we simply use the Riemannian *explicit* Euler scheme for integrating the prototype evolution flow (5.43b), i.e.,

$$(m^j)^{(t+1)} = \operatorname{Exp}_{(m^j)^{(t)}}\Big(h G_j\big(W^{(t)}, M^{(t)}\big)\Big), \quad j \in J, \tag{5.44}$$

with step size $h > 0$, and Exp_{m^j} is the Riemannian exponential map (2.13) on $\mathcal{M}$. In order to numerically integrate the assignment flow (5.43a), we adapt the geometric *implicit* Euler scheme from [ZSPS20]. It amounts to solving the fixed point equation

$$V^{(t+1)} = h\Pi_0 F\Big(\exp_{W^{(t)}}\big(V^{(t+1)}\big), M^{(t+1)}\Big), \tag{5.45}$$

by an iterative inner loop, where Π_0 denotes the orthogonal projection onto the tangent space $\mathcal{T}_0$, followed by updating

$$W^{(t+1)} = \exp_{W^{(t)}}\big(V^{(t+1)}\big). \tag{5.46}$$

Here, $\exp_W$ denotes the lifting map given by (2.90) and (2.100a).

5.5 Self-Assignment Flows

In this section, we sketch an alternative extension of the (supervised) assignment flow to the unsupervised case. This approach, the so-called **self-assignment flow**, was introduced and rigorously investigated in [ZZPS19b, ZZPS20b]. In this approach, prototypes are no longer involved and instead a pairwise data affinity matrix is used to assign given data to themselves in terms of a self-assignment matrix parameterized by $W \in \overline{\mathcal{W}}$. This self-assignment results in a partitioning of the data, which can optionally be used for computing prototypes.

Self-Assignment Matrices

Instead of regarding the distance matrix $D \in \mathbb{R}_{\geq 0}^{|I|\times|J|}$ between data $\{z^i\}_{i\in I}$ and the (unknown) prototypes $\{m^j\}_{j\in J}$, we assume an affinity matrix

$$K = K^\top \in \mathbb{R}^{|I|\times|I|}, \quad K_{ik} = k_{\mathcal{M}}(z^i, z^k), \quad i, k \in I \tag{5.47}$$

to be given. Its entries measure the similarity of the data points z^i, z^k in terms of a symmetric function $k_{\mathcal{M}}$. A simple example is the negative distance

$$k_{\mathcal{M}}(z^i, z^k) = -d_{\mathcal{M}}(z^i, z^k). \tag{5.48}$$

In case of Euclidean data, i.e. $\mathcal{M} = \mathbb{R}^d$, the Euclidean inner product

$$k_{\mathcal{M}}(z^i, z^k) = \langle z^i, z^k \rangle \tag{5.49}$$

or the Gaussian kernel

$$k_{\mathcal{M}}(z^i, z^k) = \exp\big(-\tfrac{1}{2\sigma^2} d_{\mathcal{M}}(z^i, z^k)^2\big) \tag{5.50}$$

with $d_{\mathcal{M}}(z^i, z^k) = \|z^i - z^k\|$ can also be used. Based on the matrix K, we want to perform a partitioning of the data using latent labels as described next.

Let $W \in \mathcal{W}^* = \overline{\mathcal{W}} \cap \{0,1\}^{|I|\times|J|}$ have full rank, i.e., each (latent) label is present in the labeling. The column vectors W^j, $j \in J$, of W indicate which data points z^i are assigned to the j-th cluster $I_j = \{i \in I \colon W_{ij} = 1\}$. We define the diagonal matrix

$$C(W) \coloneqq \operatorname{Diag}(W^\top \mathbb{1}_{|I|}) = \operatorname{Diag}(n_1, \dots, n_{|J|}) \tag{5.51}$$

with the cardinalities $n_j = |I_j|$ of each cluster as entries. With $A_0(W) \coloneqq WC(W)^{-1}W^\top$ we get

$$\operatorname{tr}\big(KA_0(W)\big) = \operatorname{tr}\big(C(W)^{-1}W^\top KW\big) = \sum_{j\in J} \frac{1}{n_j}(W^\top KW)_{jj} \tag{5.52a}$$

$$= \sum_{j\in J} \frac{1}{n_j} \sum_{i\in I_j} k_{\mathcal{M}}(z^i, z^i) + \sum_{j\in J} \frac{1}{n_j} \sum_{i,k\in I_j \colon i\neq k} k_{\mathcal{M}}(z^i, z^k). \tag{5.52b}$$

The first sum of (5.52b) is constant for common kernel functions like (5.50). The second sum measures the normalized similarity within clusters I_j, $j \in J$. Thus, the problem to partition the data into $|J|$ clusters takes the form

$$\max_W \operatorname{tr}\big(KA_0(W)\big) \qquad \text{subject to} \qquad W \in \mathcal{W}^*, \quad \operatorname{rank}(W) = |J|. \tag{5.53}$$

Before explaining how this partitioning problem can be combined with the assignment flow, we discuss two relaxations of this problem and their interpretation.

Self-Affinity Matrix

For the first relaxation, we simply drop the integrality constraint and the rank constraint, i.e., we regard the *self-affinity matrix*

$$A_0(W) = WC(W)^{-1}W^\top \quad \text{for} \quad W \in \overline{\mathcal{W}}. \tag{5.54}$$

This relaxation can be interpreted as follows. We interpret each entry of the assignment matrix $W \in \overline{\mathcal{W}}$ as posterior probability

$$P(j|i) = W_{ij}, \quad j \in J, \quad i \in I \tag{5.55}$$

of label j, conditioned on the observation of the data point z^i. For the data, we adopt the uniform prior distribution

$$P(i) = \frac{1}{|I|}, \quad i \in I \tag{5.56}$$

according to the unsupervised scenario. By marginalization, we get the label distribution

$$P(j) = \sum_{i \in I} P(j|i)P(i) = \frac{1}{|I|}(W^\top \mathbb{1}_{|I|})_j, \quad j \in J. \tag{5.57}$$

Applying Bayes' rule results in the conditional distribution

$$P(k|j) = \frac{P(j|k)P(k)}{P(j)} = \frac{W_{kj}}{\sum_{i \in I} W_{ij}} = \big(C(W)^{-1}W^\top\big)_{jk}, \quad k \in I, \quad j \in J \tag{5.58}$$

with the roles of data and labels reversed. Eventually, marginalization over the labels results in the probability of the self-assignment $z^i \leftrightarrow z^k$, $i, k \in I$:

$$\sum_{j \in J} P(k|j)P(j|i) = \sum_{j \in J} W_{ij}\big(C(W)^{-1}W^\top\big)_{jk} = \big(WC(W)^{-1}W^\top\big)_{ik} = A_{0,ik}(W). \tag{5.59}$$

So, this expression specifies the probability that two vertices i and k belong to the same cluster. This shows that maximizing the correlation of $A_0(W)$ and the affinity matrix K as in (5.53) amounts to cover the most similar data points by the same clusters.

Self-Influence Matrix

We observe that $C(W) = W^\top W$ for $W \in \mathcal{W}^*$. So as second relaxation we consider the *self-influence matrix*

$$A_1(W) \coloneqq W(W^\top W)^{-1}W^\top \qquad \text{for} \quad W \in \overline{\mathcal{W}} \quad \text{with} \quad \operatorname{rank}(W) = |J|. \tag{5.60}$$

For this relaxation, we have the following interpretation. We assume that the data $\{z^i\}_{i \in I}$ are given in a Euclidean space $\mathcal{M} = \mathbb{R}^d$. We collect the data and the prototypes $\{m^j\}_{j \in J}$ as row vectors into matrices

$$Z = \big(z^1, \dots, z^{|I|}\big)^\top \in \mathbb{R}^{|I| \times d}, \tag{5.61a}$$

$$M = \big(m^1, \dots, m^{|J|}\big)^\top \in \mathbb{R}^{|J| \times d}. \tag{5.61b}$$

Further, we assume the prototypes result from a least-squares fit:

$$M = \operatorname*{arg\,min}_{\tilde{M} \in \mathbb{R}^{|J| \times d}} \frac{1}{2}\big\|Z - W\tilde{M}\big\|^2 = (W^\top W)^{-1}W^\top Z. \tag{5.62}$$

Using these prototypes in turn for predicting data $\hat{Z}$ by assignment yields

$$\hat{Z} = WM = W(W^\top W)^{-1}W^\top Z = \Pi_{\mathrm{im}(W)}Z = A_1(W)Z, \tag{5.63}$$

where $\Pi_{\mathrm{im}(W)}$ denotes the orthogonal projection onto the range of W. Finally, optimizing the assignments $W \in \overline{\mathcal{W}}$ with $\operatorname{rank}(W) = |J|$ in order to obtain the best prediction of the data itself, gives with $A_1(W)^2 = A_1(W)$

$$\operatorname*{arg\,min}_W \frac{1}{2}\big\|A_1(W)Z - Z\big\|^2 = \operatorname*{arg\,max}_W \operatorname{tr}\big(A_1(W)ZZ^\top\big). \tag{5.64}$$

The latter problem is a relaxation of (5.53) with affinity matrix $K = ZZ^\top$ corresponding to the Euclidean inner product (5.49). This relaxation can be interpreted as finding the best $|J|$-dimensional subspace $\mathrm{im}(W)$ spanned by the indicator vectors of the $|J|$ clusters for self-prediction of the given data.

Adaption of the Assignment Flow

The likelihood map in the (supervised) assignment flow takes the form

$$L(W) = \exp_W\big(-\tfrac{1}{\rho}D\big) = \exp_W\big(\tfrac{1}{\rho}\nabla E(W)\big) \qquad \text{with} \qquad E(W) = -\langle D, W\rangle, \tag{5.65}$$

so that the assignment flow maximizes $E(W(t))$ over time (if no averaging is used). In the unsupervised data partitioning problem we want to solve a relaxation of (5.53), i.e., we want to maximize

$$E_s(W) = \mathrm{tr}\big(KA_s(W)\big), \quad W \in \overline{\mathcal{W}}, \quad s \in \{0,1\}. \tag{5.66}$$

So the key idea for combining the partitioning problem with the assignment flow is to adapt the likelihood map by

$$\widetilde{L}(W) = \exp_W\big(\tfrac{1}{\rho}\nabla E_s(W)\big). \tag{5.67}$$

As before, the corresponding similarity map $\widetilde{S}(W) = G^\omega(\widetilde{L}(W))$ is defined by geometric averaging of likelihood vectors $\widetilde{L}_k(W)$ within neighborhoods $k \in \mathcal{N}_i$. Besides replacing the likelihood map, we have to deal with the symmetry of the partitioning problem, e.g., permutations of the label set results in the same clusters. Because of this symmetry, we cannot let the assignment flow start at the barycenter. In order to break this symmetry, we initialize the assignment flow with a small perturbation of the barycenter. Therefore, the resulting self-assignment flow reads

$$\textbf{(SAF)} \qquad \dot{W} = R_W[\widetilde{S}(W)], \qquad W(0) = \exp_{C_{\mathcal{W}}}(-\varepsilon D_0), \quad 0 < \varepsilon \ll 1, \tag{5.68}$$

where the distance matrix $D_0 \in \mathbb{R}^{|I|\times|J|}$ is computed using greedy k-center clustering as explained in Section 5.2.3.

Computing Prototypes

While this approach works without prototypes, we can still compute prototypes using the clusters resulting from (5.68). Let $W^* \in \mathcal{W}^*$ be the assignment resulting from **(SAF)**. Then the j-th prototype m^j can be computed as Riemannian mean of the cluster $I_j = \{i \in I\colon W^*_{ij} = 1\}$, i.e.

$$m^j = \operatorname*{arg\,min}_{m\in\mathcal{M}} \sum_{i\in I_j} d_{\mathcal{M}}(z^i, m)^2, \qquad j \in J. \tag{5.69}$$

In the Euclidean case, this takes the form $M = C(W)^{-1}W^\top Z$, where the matrices Z and M are given in (5.61).

Comparison with the Unsupervised Assignment Flow

We conclude this section by comparing the self-assignment flow **(SAF)** (5.68) with the unsupervised assignment flow **(UAF)** (5.42).

Advantages of (SAF) over (UAF)

The **(SAF)** approach does not involve any prototypes and operates entirely on the assignment manifold $\mathcal{W}$. The only input involving the feature manifold is the affinity matrix K. This reflects the plug-and-play property of the assignment flow. Additionally, the affinity matrix K can be chosen to contain only pairwise distances, so that the **(SAF)** approach can be applied on any metric space $\mathcal{M}$. In comparison, the **(UAF)** approach can be applied only on Riemannian manifolds $\mathcal{M}$ and it comes with the additional effort to implement the computation of the Riemannian gradient of the divergence function, which is needed for the prototype update – cf. equation (5.42).

Advantages of (UAF) over (SAF)

The use of the affinity matrix $K \in \mathbb{R}^{|I|\times|I|}$ as in **(SAF)** instead of the distance matrix $D \in \mathbb{R}^{|I|\times|J|}$ with $|J| \ll |I|$ as in **(UAF)** results in a considerable additional workload, which becomes intractable for already medium sized images. In order to cope with large-scale scenarios, the matrix K then has to be approximated by a low-rank matrix factorization, which only works for positive semi-definite matrices K. We refer to [ZZPS20b] for more details.

In view of manifold valued data, the **(SAF)** approach is inspired by the Euclidean case (cf. the interpretation of the self-influence matrix), while the non-Euclidean case is handled either by using the pairwise distance (5.48) or by embedding the data into a reproducing kernel Hilbert space (RKHS) [HSS08] via a suitable kernel function $k_{\mathcal{M}}$. In particular the latter approach does not properly incorporate the geometry of the feature manifold $\mathcal{M}$. In contrast, the **(UAF)** approach is designed from the Riemannian perspective and therefore reasonably incorporates the geometry of $\mathcal{M}$ into the data partitioning as well as prototype learning.

5.6 Clustering on Particular Feature Manifolds

In Section 5.4, we derived the unsupervised assignment flow (5.42) for a general feature manifold $\mathcal{M}$ together with a geometric numerical integration scheme. In the following subsections, we illustrate the approach by working out details of three concrete feature manifolds. These scenarios will be evaluated numerically in the experiments in Section 5.7.

5.6.1 Rotation Group $\mathrm{SO}(n)$

In this subsection, we study clustering on the Lie group $\mathrm{SO}(n)$ of $n \times n$ rotation matrices. This is a smooth Riemannian manifold whose tangent space at $R \in \mathrm{SO}(n)$ is given by

$$T_R\mathrm{SO}(n) = \{R\Omega \colon \Omega \in \mathfrak{so}(n)\}, \tag{5.70}$$

where $\mathfrak{so}(n) = \{\Omega \in \mathbb{R}^{n\times n} \colon \Omega^\top = -\Omega\}$ denotes the Lie algebra of $\mathrm{SO}(n)$, and with the Riemannian metric given by the Frobenius inner product $g_R(A_1, A_2) = \mathrm{tr}(A_1^\top A_2)$. Based on the matrix exponential expm and matrix logarithm logm [Hig08], the corresponding exponential map and logarithmic map read

$$\mathrm{Exp}_R(R\Omega) = R\,\mathrm{expm}(\Omega), \quad \mathrm{Log}_{R_1}(R_2) = R_1\,\mathrm{logm}(R_1^\top R_2), \tag{5.71}$$

and the Riemannian distance is given by

$$d_{\mathrm{SO}(n)}(R_1, R_2) = \left\| \operatorname{logm}(R_1^\top R_2) \right\|. \tag{5.72}$$

In the specific case $n = 3$, well known formulas in closed form are available [Hig08]. By Rodrigues' formula, the matrix exponential of $A \in \mathfrak{so}(3)$ is given by

$$\operatorname{expm}(A) = I + \operatorname{sinc}(a)A + \tfrac{1}{2}\operatorname{sinc}^2(\tfrac{a}{2})A^2, \qquad a = \sqrt{\tfrac{1}{2}\operatorname{tr}(A^\top A)}, \tag{5.73}$$

with the sinc function

$$\operatorname{sinc}(x) = \begin{cases} \frac{\sin(x)}{x}, & x \neq 0 \\ 1, & x = 0. \end{cases} \tag{5.74}$$

The matrix logarithm of $R \in \mathrm{SO}(3)$ with $\operatorname{tr}(R) = 1 + 2\cos(\theta)$, $|\theta| < \pi$ is given by

$$\operatorname{logm}(R) = \frac{1}{2\operatorname{sinc}(\theta)}(R - R^\top). \tag{5.75}$$

Moreover, the Riemannian distance can be evaluated without computing the matrix logarithm or an eigenvalue decomposition as

$$d_{\mathrm{SO}(3)}(R_1, R_2) = \sqrt{2}\arccos\left(\frac{\operatorname{tr}(R_1^\top R_2) - 1}{2}\right), \quad R_1, R_2 \in \mathrm{SO}(3). \tag{5.76}$$

Regarding the clustering of data $\{R_i\}_{i\in I} \subset \mathrm{SO}(n)$, we use the canonical divergence function $D(R_1, R_2) = \frac{1}{2}d_{\mathrm{SO}(n)}(R_1, R_2)^2$. As a result, the flow of the **(UAF)** (5.42) for the prototypes $S_j \in \mathrm{SO}(n)$ takes the form

$$\begin{aligned} \dot{S}_j(t) &= \alpha \sum_{i\in I} v_{j|i}\big(W(t), S(t)\big) \operatorname{Log}_{S_j(t)}(R_i) \\ &= \alpha \sum_{i\in I} v_{j|i}\big(W(t), S(t)\big) S_j(t) \operatorname{logm}\left(S_j(t)^\top R_i\right), \qquad j \in J. \end{aligned} \tag{5.77}$$

Discretizing this flow due to (5.44) yields the multiplicative update scheme

$$S_j^{(t+1)} = S_j^{(t)} \operatorname{expm}\left(\alpha h \sum_{i\in I} v_{j|i}^{(t)} \operatorname{logm}\left(\left(S_j^{(t)}\right)^\top R_i\right)\right), \quad j \in J. \tag{5.78}$$

5.6.2 Orientations

We consider the task of clustering orientation vector fields in the two-dimensional space. We regard these vector fields as maps from the image domain into the angle space $\mathbb{R}/\pi\mathbb{Z}$, i.e. we identify the line $\{\lambda(\cos\theta, \sin\theta)^\top : \lambda \in \mathbb{R}\} \subset \mathbb{R}^2$ with the angle $\theta \in [0, \pi)$. Let $q\colon \mathbb{R} \to \mathbb{R}/\pi\mathbb{Z}$ be the quotient map $\theta \mapsto \theta \mod \pi$. Rather than operating directly on the quotient manifold $\mathbb{R}/\pi\mathbb{Z}$, we work with representatives of its elements in $\mathbb{R}$. In particular, a flow on the quotient manifold will be given by $q(\vartheta(t))$, where $\vartheta(t)$ is a flow in $\mathbb{R}$. For any two representatives $\theta^1, \theta^2 \in \mathbb{R}$, the induced distance is given by

$$d(\theta^1, \theta^2) = d_{\mathcal{M}}(q(\theta^1), q(\theta^2)) = \min_{\varphi\in\pi\mathbb{Z}} |\theta^1 - \theta^2 + \varphi| \in [0, \tfrac{\pi}{2}], \tag{5.79}$$

and we have $d(\theta^1, \theta^2) = 0$ if and only if $q(\theta^1) = q(\theta^2)$. For the unsupervised assignment flow (5.42), we choose the canonical divergence function $D(x, y) = \frac{1}{2}\big(d_{\mathcal{M}}(x, y)\big)^2$ on $\mathcal{M}$. By (5.79), this corresponds to the dissimilarity function $D(\theta^1, \theta^2) = \frac{1}{2}\big(d(\theta^1, \theta^2)\big)^2$ for representatives $\theta^1, \theta^2 \in \mathbb{R}$. This dissimilarity function is differentiable if $q(\theta^1) \neq q(\theta^2 + \frac{\pi}{2})$, i.e., if the minimizer $\overline{\varphi} \in \arg\min_{\varphi \in \pi\mathbb{Z}} |\theta^1 - \theta^2 + \varphi|$ is unique. In this case, we have $\frac{\partial}{\partial \theta^2} D(\theta^1, \theta^2) = \theta^2 - \theta^1 - \overline{\varphi}$. Now, denoting $\{\theta^i\}_{i \in I} \subset \mathbb{R}$ the representatives of given orientations at pixels $i \in I$ and denoting by $\{\vartheta^j\}_{j \in J} \subset \mathbb{R}$ the representatives of the prototype orientations (labels), the label evolution of (5.42) takes the form

$$\dot{\vartheta}_j(t) = \alpha \cdot \left(\sum_{i \in I} v_{j|i}\big(W(t), \vartheta(t)\big)\big(\theta^i - \varphi_{ij}(t)\big) - \vartheta^j(t) \right) \quad \text{with} \quad \varphi_{ij}(t) \in \operatorname*{arg\,min}_{\varphi \in \pi\mathbb{Z}} \big|\theta^i - \vartheta^j(t) + \varphi\big|. \tag{5.80}$$

Since this flow evolves in $\mathbb{R}^{|J|}$, it can be numerically integrated using classical integration schemes. As mentioned above, the corresponding prototype flow in $\mathcal{M} = \mathbb{R}/_{\pi\mathbb{Z}}$ is then given by $q(\vartheta^j(t))$, $j \in J$.

5.6.3 Positive Definite Matrix Manifold

We consider data given as covariance region descriptors, as introduced in [TPM06]. Details of the corresponding unsupervised assignment flow are worked out in this section. In the subsequent section, we extend the representation to obtain descriptors that are invariant with respect to rotations of the image domain.

Covariance Descriptors

We consider feature maps $f : I \to \mathbb{R}^d$ extracted from a given 2D image $u : I \to \mathbb{R}^c$ with c channels by taking partial derivatives channel-wise, e.g. $u_x = \frac{\partial u}{\partial x}$ and $u_{xy} = \frac{\partial^2 u}{\partial x \partial y}$. We compute these derivatives using binomial filters [HS87], which are a discrete approximation to the Gaussian derivative filters. A typical example for a feature map used in our experiments is

$$i \mapsto f^i = \Big(u, u_x, u_y, u_{xx}, \sqrt{2}u_{xy}, u_{yy}\Big)^\top \in \mathbb{R}^{6c} \tag{5.81}$$

where $(x, y)^\top$ denote the image coordinates at pixel $i \in I$. The corresponding covariance descriptor C_i with respect to a pixel neighborhood $\mathcal{N}_i \subset I$ is given by

$$C_i = \sum_{j \in \mathcal{N}_i} \omega_{ij}(f^j - \bar{f}^i)(f^j - \bar{f}^i)^\top + \varepsilon I_d \quad \text{with} \quad \bar{f}^i = \sum_{j \in \mathcal{N}_i} \omega_{ij} f^j, \quad 0 < \varepsilon \ll 1, \tag{5.82}$$

where $\omega_i = (\omega_{ij})_{j \in \mathcal{N}_i} \in \Delta_{|\mathcal{N}_i|}$ are weights. We add the identity matrix with a very small $\varepsilon > 0$ to ensure that all descriptors are positive definite, which otherwise may not hold in particular cases like homogeneous "flat" image regions.

HPD Manifold

Now we consider the task of clustering given covariance descriptors as points on the Riemannian manifold of Hermitian positive definite (HPD) matrices [Bha06]

$$\mathcal{P}_d \coloneqq \big\{X \in \mathbb{C}^{d \times d} : X^* = X,\ X \text{ is positive definite}\big\} \tag{5.83}$$

endowed with the Riemannian metric $g_X(U,V) = \operatorname{tr}(X^{-1}UX^{-1}V)$ on each tangent space $T_X\mathcal{P}_d = \{U \in \mathbb{C}^{d\times d} : U^* = U\}$ – cf. Section 2.1.4. The Riemannian gradient of a function $F : \mathcal{P}_d \to \mathbb{R}$ is given by $\operatorname{grad} F(X) = X\nabla F(X)X \in T_X\mathcal{P}_d$, where the Hermitian matrix $\nabla F(X)$ denotes the Euclidean gradient of F at X, and matrices X and $\nabla F(X)$ are multiplied as usual. Denoting the prototypes (labels) by $\{\Lambda_j\}_{j\in J} \subset \mathcal{P}_d$, the label flow of (5.42) reads

$$\dot{\Lambda}_j(t) = -\alpha \sum_{i\in I} \nu_{j|i}\big(W(t),\Lambda(t)\big)\Lambda_j(t)\partial_2 D\big(C_i,\Lambda_j(t)\big)\Lambda_j(t), \quad j \in J, \tag{5.84}$$

where $D(X,Y)$ is a proper divergence on $\mathcal{P}_d$ as discussed below and $\partial_2 D(X,Y)$ denotes its Euclidean gradient with respect to Y. In the following, we discuss possible choices of D. An obvious choice is the canonical divergence induced by the Riemannian distance

$$D_{\mathrm{R}}(X,Y) = \frac{1}{2} d_{\mathcal{P}_d}(X,Y)^2 = \frac{1}{2} \sum_{k\in[d]} \big(\log \lambda_k(X,Y)\big)^2, \tag{5.85}$$

which involves all generalized eigenvalues $\lambda_k(X,Y)$ of the matrix pencil (X,Y). Considering that $D(C_i,\Lambda_j)$ has to be computed for each pair of datum C_i and prototype Λ_j at *each point of time* when integrating the flow, the computation of the generalized eigenvalues would be very expensive computationally. As a more efficient alternative to D_{R}, we consider the Stein divergence [Sra13]

$$D_{\mathrm{S}}(X,Y) = \log\det\left(\tfrac{X+Y}{2}\right) - \tfrac{1}{2}\log\det(XY), \tag{5.86a}$$

$$\partial_2 D_{\mathrm{S}}(X,Y) = \frac{1}{2}\left(\left(\tfrac{X+Y}{2}\right)^{-1} - Y^{-1}\right). \tag{5.86b}$$

It involves the determinant and the inverse of a positive definite matrix, which both can be efficiently computed using the Cholesky decomposition. Moreover, it is shown in [Sra13] that D_{S} is a squared distance. Based on the choice $D = D_{\mathrm{S}}$, equation (5.84) takes the form

$$\dot{\Lambda}_j(t) = \frac{\alpha}{2}\Big(\Lambda_j(t) - \Lambda_j(t)Q_j(t)\Lambda_j(t)\Big) \quad \text{with} \quad Q_j(t) = \sum_{i\in I} \nu_{j|i}\big(W(t),\Lambda(t)\big)\left(\frac{C_i+\Lambda_j(t)}{2}\right)^{-1}. \tag{5.87}$$

Taking the exponential map

$$\operatorname{Exp}_X(U) = X^{\frac{1}{2}} \operatorname{expm}\big(X^{-\frac{1}{2}}UX^{-\frac{1}{2}}\big)X^{\frac{1}{2}} \tag{5.88}$$

into account, with expm denoting the matrix exponential, and discretizing the flow with the Riemannian explicit Euler scheme (5.44), gives the prototype update

$$\Lambda_j^{(t+1)} = \tilde{\Lambda}_j \operatorname{expm}\left(\frac{\alpha h}{2}\Big(I - \tilde{\Lambda}_j Q_j^{(t)} \tilde{\Lambda}_j\Big)\right)\tilde{\Lambda}_j \quad \text{with} \quad \tilde{\Lambda}_j = \left(\Lambda_j^{(t)}\right)^{\frac{1}{2}} \tag{5.89}$$

for the Stein divergence.

5.6.4 Rotation-Invariant Dissimilarity Measure on Covariance Descriptors

We additionally constructed a dissimilarity function on $\mathcal{P}_d$ that is invariant under rotations of the image domain. In contrast to the Stein divergence (5.86a), this dissimilarity function takes the special structure of covariance descriptors into account and hence depends on the underlying feature map.

Approach

Let $u, \tilde{u}\colon \mathbb{R}^2 \to \mathbb{R}$ denote two gray value images that are related by a Euclidean transformation

$$\begin{pmatrix}\tilde{x}\\ \tilde{y}\end{pmatrix} = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}\begin{pmatrix}x\\ y\end{pmatrix} + \begin{pmatrix}x_0\\ y_0\end{pmatrix}, \quad \theta \in [0, 2\pi) \tag{5.90}$$

of the image domain, i.e. $\tilde{u}(\tilde{x}, \tilde{y}) = u(x, y)$. Their derivatives transform as

$$\begin{pmatrix}\tilde{u}_x\\ \tilde{u}_y\end{pmatrix} = R_1(\theta)\begin{pmatrix}u_x\\ u_y\end{pmatrix}, \qquad \begin{pmatrix}\tilde{u}_{xx}\\ \sqrt{2}\tilde{u}_{xy}\\ \tilde{u}_{yy}\end{pmatrix} = R_2(\theta)\begin{pmatrix}u_{xx}\\ \sqrt{2}u_{xy}\\ u_{yy}\end{pmatrix}, \tag{5.91}$$

with rotation matrices $R_1(\theta) \in \mathrm{SO}(2)$ and $R_2(\theta) \in \mathrm{SO}(3)$ given by

$$R_1(\theta) = \begin{pmatrix}\cos\theta & -\sin\theta\\ \sin\theta & \cos\theta\end{pmatrix}, \qquad R_2(\theta) = \begin{pmatrix}\cos^2\theta & -\sqrt{2}\cos\theta\sin\theta & \sin^2\theta\\ \sqrt{2}\cos\theta\sin\theta & \cos^2\theta - \sin^2\theta & -\sqrt{2}\cos\theta\sin\theta\\ \sin^2\theta & \sqrt{2}\cos\theta\sin\theta & \cos^2\theta\end{pmatrix}. \tag{5.92}$$

It follows that covariance descriptors of an image $u\colon I \to \mathbb{R}^c$ with the feature map (5.81) transform as $\tilde{C} = R(\theta)CR(\theta)^\top$, with a rotation matrix $R(\theta) \in \mathrm{SO}(d)$. Setting

$$\mathcal{R} \coloneqq \big\{R(\theta)\colon \theta \in [0, 2\pi)\big\}, \tag{5.93}$$

it turns out that $\mathcal{R}$ is a one-dimensional subgroup of $SO(d)$, i.e. $R(\theta_1 + \theta_2) = R(\theta_1)R(\theta_2)$. Eventually, we construct the rotation-invariant dissimilarity function by minimizing over the Lie group action of $\mathcal{R}$, i.e.

$$D_{\mathrm{S},\mathcal{R}}(X, Y) \coloneqq \min_{R\in\mathcal{R}} D_{\mathrm{S}}\big(X, RYR^\top\big) = \min_{R\in\mathcal{R}} D_{\mathrm{S}}\big(R^\top XR, Y\big). \tag{5.94}$$

If $\partial_2 D_{\mathrm{S}}(\overline{R}^\top X\overline{R}, Y) = (\overline{R}^\top X\overline{R} + Y)^{-1} - \frac{1}{2}Y^{-1}$ is the same for all $\overline{R} \in \arg\min_{R\in\mathcal{R}} D_{\mathrm{S}}\big(R^\top XR, Y\big)$, then $D_{\mathrm{S},\mathcal{R}}(X, Y)$ is differentiable in Y and the derivative is given by [BS13, Theorem 4.13]

$$\partial_2 D_{\mathrm{S},\mathcal{R}}(X, Y) = \partial_2 D_{\mathrm{S}}(\overline{R}^\top X\overline{R}, Y), \qquad \overline{R} \in \operatorname*{arg\,min}_{R\in\mathcal{R}} D_{\mathrm{S}}\big(R^\top XR, Y\big). \tag{5.95}$$

This holds in particular if $\overline{R}$ is unique for a given pair (X, Y).
Using the divergence $D_{\mathrm{S},\mathcal{R}}$, equation (5.84) yields the same prototype update formulas (5.87) and (5.89) as for the Stein divergence, except for the modification

$$Q_j(t) = \sum_{i\in I} \nu_{j|i}\big(W(t), \Lambda(t)\big)\left(\frac{R_{ij}(t)^\top C_i R_{ij}(t) + \Lambda_j(t)}{2}\right)^{-1} \tag{5.96}$$

with $R_{ij}(t) \in \arg\min_{R\in\mathcal{R}} D_{\mathrm{S}}\big(R^\top C_i R, \Lambda_j(t)\big)$.

Numerical Aspects

Evaluating (5.94) amounts to solving a one-dimensional smooth but non-convex problem. We discuss a suitable optimization scheme for this problem. For given $X, Y \in \mathcal{P}_d$, we want to minimize

$$\delta(\theta) \coloneqq D_{\mathrm{S}}\big(X, R(\theta) Y R(\theta)^\top\big) \tag{5.97}$$

over the compact interval $[0, 2\pi]$. The naive approach would be sampling the interval $[0, 2\pi]$ and evaluating $\delta(\theta)$ at those sample points, but since the evaluation of

$$\delta(\theta) = \log\det\big(X + R(\theta) Y R(\theta)^\top\big) - d\log 2 - \tfrac{1}{2}\log\det X - \tfrac{1}{2}\log\det Y \tag{5.98}$$

involves computing the determinant of matrices in $\mathcal{P}_d$, this approach would be too costly, because this has to be done for each pair (C_i, Λ_j) of data points and labels. Therefore, the idea is to approximate (5.98) and evaluate the approximating function. Since $\delta(\theta)$ is an analytic periodic function, we consider a Hermite interpolation by a trigonometric polynomial at the points $\theta^k = \frac{\pi k}{N}$, $k = 0, \dots, 2N-1$. More precisely, we interpolate $\delta(\theta)$ by

$$\delta_N(\theta) = \sum_{n=0}^{3N} a_n \cos(n\theta) + \sum_{n=1}^{3N-1} b_n \sin(n\theta), \tag{5.99}$$

such that the function values and the first two derivatives at the points θ^k coincide:

$$\delta_N(\theta^k) = \delta(\theta^k), \quad \delta_N'(\theta^k) = \delta'(\theta^k), \quad \delta_N''(\theta^k) = \delta''(\theta^k) \qquad \text{for } k = 0, \dots, 2N-1. \tag{5.100}$$

Such a trigonometric polynomial always exists and is unique. Further, it converges exponentially fast with the number of sampling points, i. e. $\|\delta - \delta_N\|_\infty \le c_1 e^{-c_2 N}$. We refer to [Kre72] for further details on the Hermite interpolation. The reasoning for using the first two derivatives for the interpolation is that evaluating $\delta(\theta)$, $\delta'(\theta)$ and $\delta''(\theta)$ simultaneously at a point θ can be done efficiently as described next.

Since $\mathcal{R}$ is a one-dimensional subgroup of $\mathrm{SO}(d)$, we have the representation $R(\theta) = \exp\mathrm{m}(\theta\Omega)$ with a skew-symmetric matrix $\Omega \in \mathfrak{so}(d)$. For instance, the corresponding skew-symmetric matrices for the rotation matrices (5.92) are given by

$$\Omega_1 = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix}, \qquad \Omega_2 = \begin{pmatrix} 0 & -\sqrt{2} & 0 \\ \sqrt{2} & 0 & -\sqrt{2} \\ 0 & \sqrt{2} & 0 \end{pmatrix}. \tag{5.101}$$

The derivate of $R(\theta)$ is then given by $\frac{\mathrm{d}}{\mathrm{d}\theta} R(\theta) = R(\theta)\Omega = \Omega R(\theta)$. Setting $Y_R \coloneqq R(\theta) Y R(\theta)^\top$, we get

$$\delta(\theta) = \log\det(X + Y_R) - d\log 2 - \tfrac{1}{2}\log\det X - \tfrac{1}{2}\log\det Y_R, \tag{5.102a}$$

$$\delta'(\theta) = \operatorname{tr}\big((X + Y_R)^{-1}(\Omega^\top X + X\Omega)\big), \tag{5.102b}$$

$$\delta''(\theta) = -\operatorname{tr}\big((X + Y_R)^{-1}(\Omega Y_R + Y_R\Omega^\top)(X + Y_R)^{-1}(\Omega^\top X + X\Omega)\big). \tag{5.102c}$$

The Cholesky decomposition of $X + Y_R \in \mathcal{P}_d$ can be used to efficiently evaluate $\det(X + Y_R)$ as needed for $\delta(\theta)$ and to compute $(X + Y_R)^{-1}$ as needed for $\delta'(\theta)$ and $\delta''(\theta)$. Those are also the most expensive operations in the evaluation of (5.102).

We note that evaluating the polynomial (5.99) at sample points $\{\theta^k\}_k$ can be implemented as matrix-vector multiplication with a matrix containing the values of the basis functions, i.e. $\cos(n\theta^k)$ and $\sin(n\theta^k)$, and a vector containing the coefficients. Since only the coefficients vector change with different pairs (X, Y), the matrix can be reused.

Remark 5.6.1. We conclude this section with further comments on the invariant dissimilarity function (5.94) and its numerics.

1. We point out again that $\mathcal{R}$ (5.93) and its existence depends on the feature map f. For the specific case (5.81) considered above, a transformation of the form $\tilde{C} = R(\theta)CR(\theta)^\top$ exists since *all* derivatives up to a given order are involved. Furthermore, $\mathcal{R}$ is a subgroup of $\mathrm{SO}(d)$ due to the proper normalization of mixed derivatives. For the derivative $\left(\frac{\partial}{\partial x}\right)^{\alpha_1}\left(\frac{\partial}{\partial y}\right)^{\alpha_2} u$, this normalization factor is given by $\sqrt{\frac{(\alpha_1+\alpha_2)!}{\alpha_1!\alpha_2!}}$.

2. The dissimilarity function $D_{\mathrm{S},\mathcal{R}}$ is not a divergence function as introduced in Section 2.1.5, since $D_{\mathrm{S},\mathcal{R}}(X, Y) = 0$ does not imply $X = Y$, but only $[X]_\mathcal{R} = [Y]_\mathcal{R}$ with $[X]_\mathcal{R} = \{RXR^\top : R \in \mathcal{R}\}$. Unfortunately, this cannot be fixed by considering the quotient $\mathcal{P}_d / \sim$ with $X \sim Y$ if and only if $X \in [Y]_\mathcal{R}$, since $\mathcal{P}_d / \sim$ does not have a manifold structure (e.g. the equivalence class of the identity matrix is a singleton). Nevertheless, we can plug-in $D_{\mathrm{S},\mathcal{R}}$ into our approach that can be used with any differentiable dissimilarity function. The resulting prototypes are then representatives $\{\Lambda_j\}_{j\in J} \subset \mathcal{P}_d$ of classes $[\Lambda_j]_\mathcal{R}$.

3. The set of pairs $(X, Y) \in \mathcal{P}_d \times \mathcal{P}_d$, for which $D_{\mathrm{S},\mathcal{R}}(X, Y)$ is not differentiable, is negligible [RW09, Theorem 10.31]. But even for such pairs one can choose some optimal $R_{ij}(t)$ in (5.96), such that the prototype flow remains well-defined.

4. The rotation-invariant dissimilarity can also be extended to 3D images. The main difference to the 2D case is that the group $\mathcal{R}$ is then a three-dimensional non-Abelian Lie-group. This makes the evaluation of (5.94) significantly more difficult.

5. The minimization of (5.99) can be done by an eigenvalue computation of a complex non-Hermitian matrix [Boy06]. However, this approach would be too costly for our purpose.

6. Since the components of the matrix $R(\theta)$ are trigonometric polynomials (cf. (5.92)), $\delta(\theta)$ is the logarithm of a trigonometric polynomial due to equation (5.98), i.e.

$$\delta(\theta) = \log\Big(\tilde{a}_0 + \sum_{n=1}^{\tilde{N}} \tilde{a}_n \cos(n\theta) + \tilde{b}_n \cos(n\theta)\Big). \tag{5.103}$$

   Instead of interpolating $\delta(\theta)$, one might consider interpolating the trigonometric polynomial $\exp(\delta(\theta))$. Since the values of $\exp(\delta(\theta))$ can get quite large, we refrain from pursuing this approach due to possible numerical issues.

5.7 Experiments

In this section, we demonstrate and compare the proposed **unsupervised assignment flow (UAF)** using several synthetic and real-world images and different feature manifolds, as detailed

in Section 5.6. As described in Section 5.4.4, the geometric numerical integration of the **(UAF)** was carried out using the geometric implicit Euler scheme for the assignment component of the flow and a Riemannian explicit Euler scheme for the prototype component of the flow. For both schemes, we used the fixed step size $h = 0.1$ in all experiments. Additionally, we adopted in our implementation the renormalization step from [ÅPSS17] with $\varepsilon = 10^{-10}$ for the assignment component, to avoid numerical issues for assignments very close to the boundary of the simplex $\Delta_{|J|} = \overline{\mathcal{S}}$. Uniform weights (w_{ik}) were used for regularizing the assignments through geometric averaging (5.36). The integration process terminated when the average entropy of the assignment component dropped below 10^{-3} which indicates almost unique label desicions (probability vectors are close to unit vectors) and in turn that the weights $v_{j|i}(W, M)$ for the prototype evolution become stationary as well. We initialized the assignment component of the unsupervised assignment flow with the uninformative barycenter (all labels have the same probability). The initial prototypes were determined by greedy k-center metric clustering as discussed in Section 5.2.3, in order to obtain an almost uniformly sampled dictionary from the input data. The number of labels $|J|$ was chosen large enough to start with an overcomplete dictionary.

5.7.1 Parameter Influence

This experiment discusses the influence of the two model parameters σ and α of the **(UAF)** as defined by (5.42). Parameter σ determines the trade-off between the influence of the assignments (spatial regularization) and the influence of the distances in the feature space on the weights $v_{j|i}(W, M)$ which govern label evolution. $\sigma = \infty$ results in the coupled flow **(CFa)** where the weights $v_{j|i}(W, M)$ solely depend on the assignments, whereas $\sigma = \rho$ gives coupled flow **(CFb)** which incorporates both the spatially regularized assignment and the distances in the feature space into the dictionary update step. In general, the impact of spatial regularization on label evolution decreases with decreasing values of σ, and the influence of distances in the feature space on the evolution of labels is even stronger for $\sigma < \rho$.

Parameter α controls the relative speed of the evolution of labels vs. assignments. If α is set too small, i.e., if the evolution of labels is too slow, then hardly any label evolution occurs at all during the period of the assignment evolution so that the resulting assignment is effectively comparable to the *supervised* assignment flow (2.105) based on the initial set of labels. On the contrary, if α is set too large, labels adapt too fast to the current assignment, which may be undesirable especially if the assignment is still too close to the uninformative barycenter in the initial phase of its evolution.

In order to visualize clearly the role of σ and α, we consider in this section the RGB color space as feature space. The demonstrated effects carry over to other non-trivial feature manifolds, of course. We used a $|\mathcal{N}| = 3 \times 3$ neighborhood size for geometric spatial regularization and fixed the number of labels to $|J| = 8$.

Figure 5.4 illustrates the above discussion for an academic computer-generated color image with a smooth strong gradient, which was generated such that from left-to-right the red channel is increasing and the blue channel is decreasing, whereas the green channel is increasing from top to bottom. The boosted labels adaption for larger α, and the impact of spatial regularization is illustrated by the cell sizes of the final Voronoi diagram relative to the initial configuration.

Figure 5.5 demonstrates the same effects for a real image. The partitions corresponding to the unsupervised image labelings are additionally displayed using false colors in order to

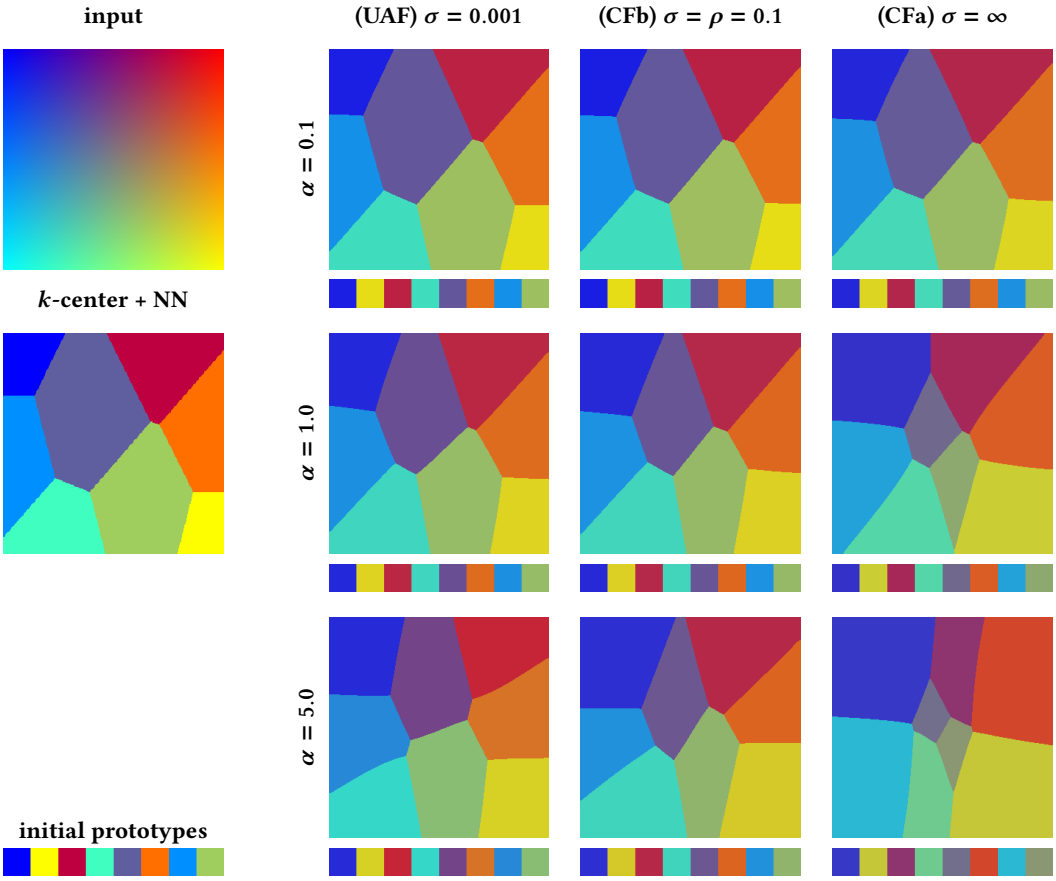


Figure 5.4: Influence of the parameters σ and α. The figure illustrates the influence of the parameters σ and α on the unsupervised assignment flow **(UAF)** in terms of the resulting labelings. From the smooth input image (left panel, top), the initial prototypes ($|J| = 8$) are extracted by greedy k-center clustering and assigned by the nearest neighbor (NN) rule. The right panel shows the labelings returned by the **(UAF)**, for different values of σ and α, after termination of the coupled evolution of labels and assignments. We observe for increasing values σ and α that regions are "attracted" toward the center of the image domain, since label colors are increasingly averaged through the spatially regularized assignments.

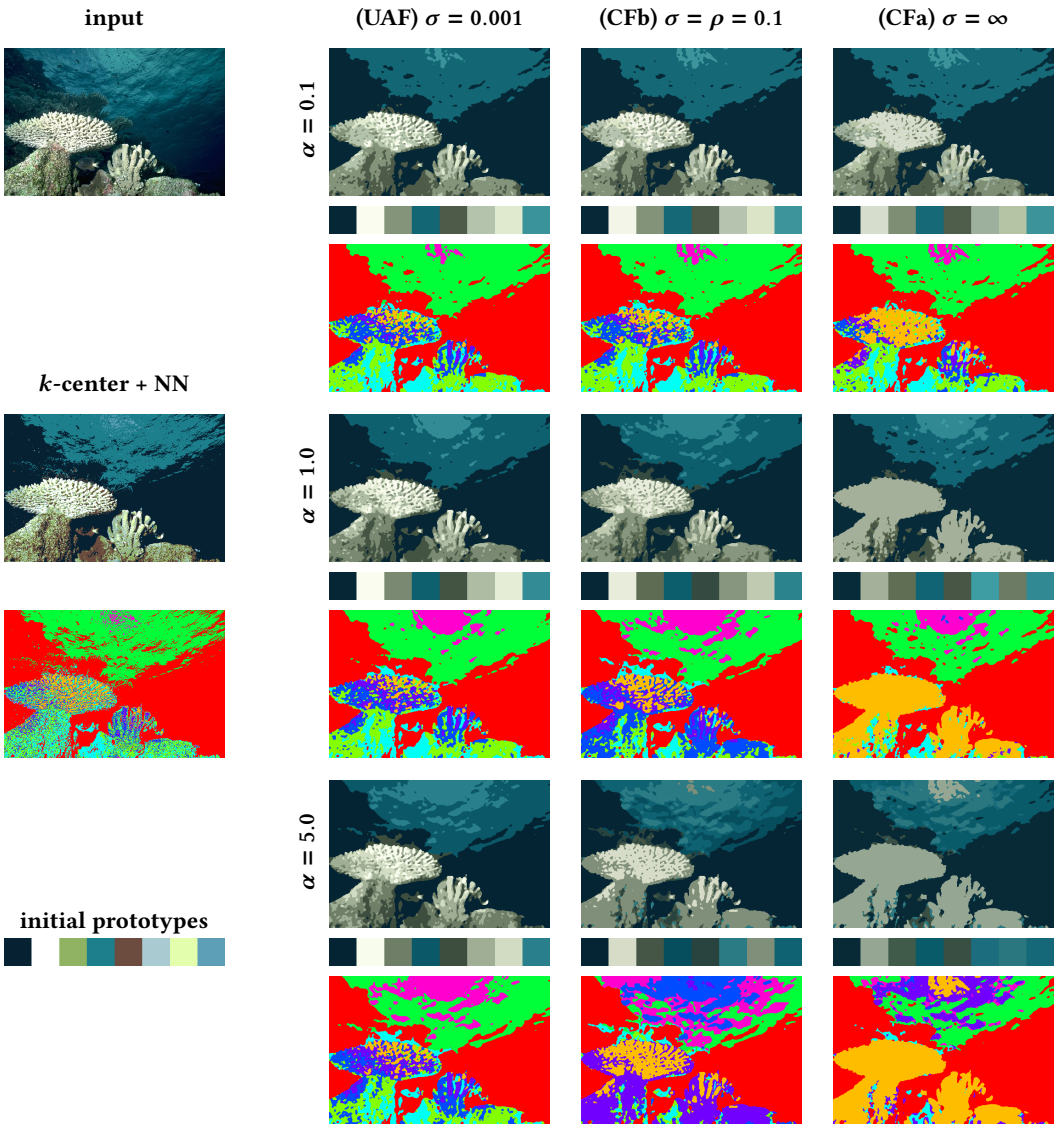


Figure 5.5: Influence of the parameters σ and α. Results of the **(UAF)** are shown that reproduce the effects illustrated by Figure 5.4 for a real image. Each labeling is additionally shown using false colors to ease the perception of differences. We observe for increasing σ an increasing impact of spatial regularization, whereas for increasing α labels adapt faster along with the size of the spatially regularized regions.

highlight the differences. The interpretation of the results for different values of σ and α is analogous to the effects shown by Figure 5.4.

Specifically, we observe that for a small value $\sigma = 0.001$ (column **(UAF)**), which increases the influence of the distances in the feature space, the resulting labeling preserves fine scales (e.g., see left coral in Figure 5.5) in comparison to the other extreme choice $\sigma = \infty$ (column **(CFa)**), where the influence of the spatial regularization through the assignments in the image domain is maximal and hence fine scales are removed from the resulting labeling. The intermediate parameter choice $\sigma = \rho = 0.1$ (column **(CFb)**) shows a good compromise between the effects caused by the two extreme values of σ.

The influence of parameter α controlling the relative speed of label and assignment evolution can be seen row-wise. For small $\alpha = 0.1$, the adaption of the prototypes is quite limited. For the choice $\alpha = 1.0$, we observe a good compromise between label evolution and spatial regularization through the assignment flow. Finally, a very large value $\alpha = 5.0$ results in strong spatial regularization, since the labels are adapting relatively quickly to the current assignments and consequently the regions assigned to labels grow faster.

5.7.2 Effect of Spatial Regularization

Figure 5.6 illustrates the effect of spatial regularization performed by the **(UAF)** on the evolution of both labels and label assignments, by comparing to basic k-means clustering and to greedy-based k-center clustering (Section 5.2.3), respectively, where no spatial regularization is involved at all. The parameter values $\alpha = 1.0$ and $\sigma = \infty$ were used.

Comparing k-means with k-center clustering shows that k-means clustering selects a more uniform quantization for the feature data, whereas the greedy k-center clustering rather picks more extremal points in the feature space which subsequently serve as initial prototypes for **(UAF)**. The remaining panels demonstrate that spatial regularization quickly sparsifies the label set as the scale (neighborhood size) of spatial regularization increases.

5.7.3 SO(3)-Valued Image Data: Orthogonal Frames in $\mathbb{R}^3$

Figure 5.7 depicts ground truth data in terms of orthogonal frames assigned to each pixel $i \in I$ and visualized with false colors. Each ground-truth label is also shown as trihedron by Figure 5.8.

The input data (Figure 5.7) were generated by independently sampling for each pixel $i \in I$ a vector $n_i \sim \mathcal{N}(0, \sqrt{0.5} I_3)$, determining a corresponding random skew-symmetric matrix $\Omega(n_i) \in \mathfrak{so}(3)$, and by replacing the ground-truth value R_i by $R_i \operatorname{expm}(\Omega(n_i))$.

We compare our method with hierarchical agglomerative clustering [Mül11]. As linkage criterion, we used the generalized Ward's criterion as presented in [Bat88], i.e., we replaced the squared Euclidean distance in the classical Ward's method by the Riemannian distance. This linkage criterion worked best in our experiments. We chose the threshold for this method such that we get the same number of clusters as in the ground truth. The labels were determined by computing the Riemannian mean within each cluster. The noisy clustering result (Figure 5.7) affects the computation of labels as can be seen in Figure 5.8.

As initialization for our method, we determined by greedy-based k-center clustering (Section 5.2.3) an overcomplete set of $|J| = 8$ prototypes as shown by Figure 5.8. The corresponding nearest neighbor (NN) assignments are shown by Figure 5.7. They clearly illustrate the need

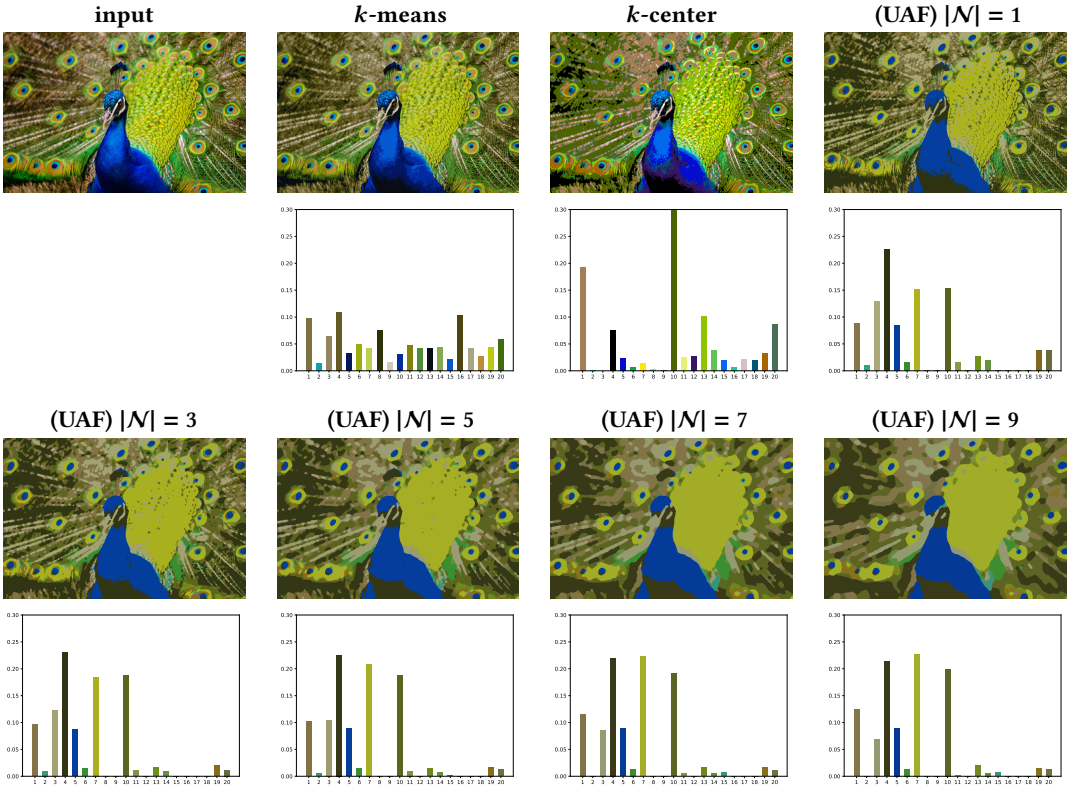


Figure 5.6: Effect of spatial regularization. We compare the proposed **(UAF)** to k-means clustering and k-center clustering, respectively, and demonstrate the effect of spatial regularization, parameterized by increasing neighborhood sizes used for geometric averaging, on the resulting label statistics and label assignments. The histogram bars are colored by the corresponding labels, and their heights indicate the relative amount of assigned pixels. We observe that as the scale (neighborhood size) of spatial regularization increases, the label set quickly becomes more sparse.

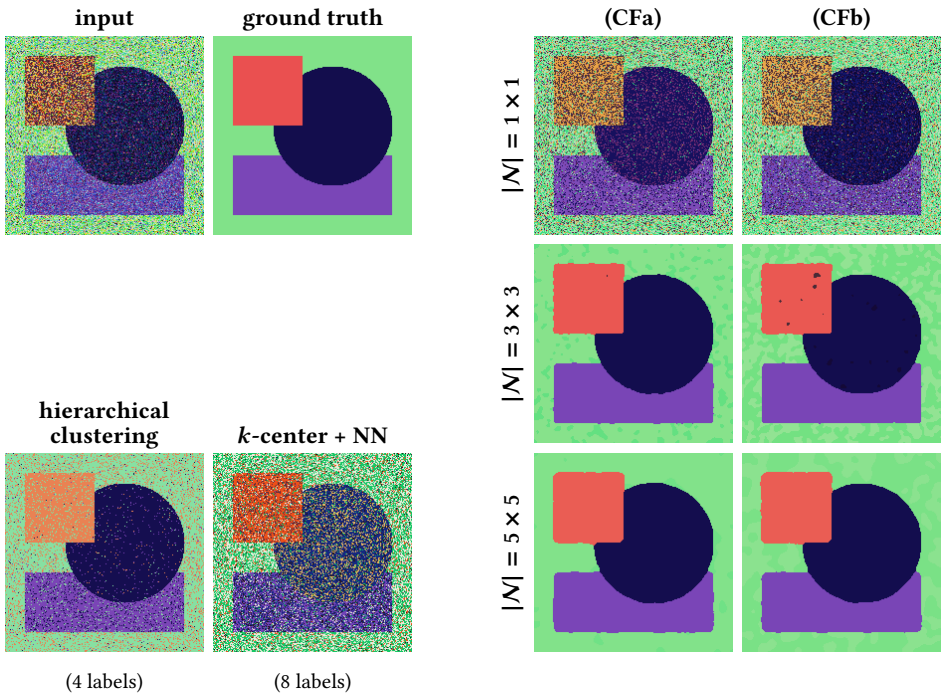


Figure 5.7: Unsupervised label learning for SO(3)-valued image data. Rotation matrices are color coded by the scheme adopted from [KMB15]. Each label (orthogonal frame, rotation matrix) is also depicted as trihedron by Figure 5.8 using as background the false color used here. The input data were generated from ground truth as described in the text. Hierarchical clustering with generalized Ward's linkage criterion produces a noisy labeling result. The panel 'k-center + NN' depicts the nearest neighbor assignments of the 8 labels which are selected from the input data by greedy k-center clustering (Section 5.2.3) and are used as initialization for **(UAF)**. Panels on the right depict both the labels and the assignment of these labels by the two versions **(CFa)** and **(CFb)** of the unsupervised assignment flow **(UAF)**. Spurious labels "die out" and, for a reasonably large neighborhood size used for spatial regularization, high-quality labelings are determined simultaneously. The resulting labels are visualized by Figure 5.8.

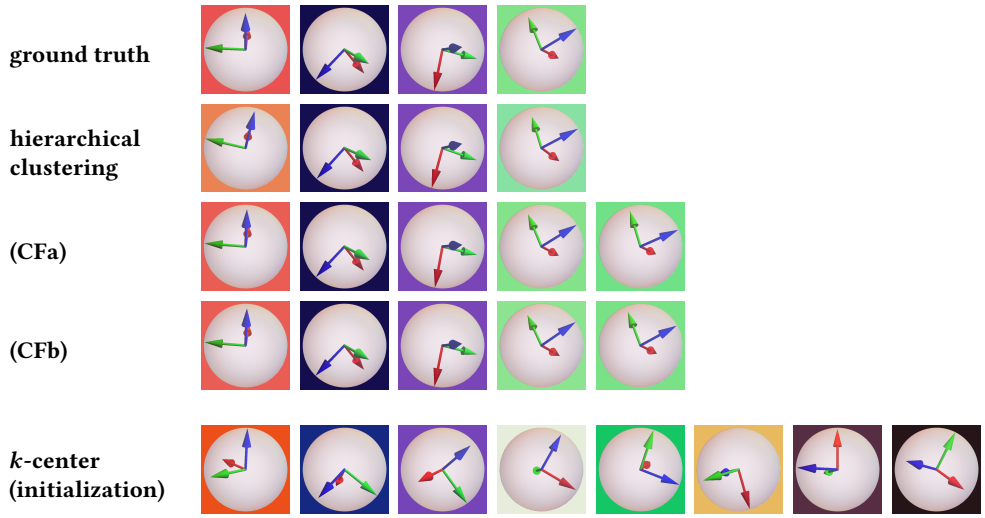


Figure 5.8: Label visualization for SO(3) data. Each label corresponding to the results depicted by Figure 5.7 is shown here as trihedron, using the false colors of Figure 5.7 as background colors. The labels obtained by hierarchical clustering are generally close to the ground truth labels but also deviate significantly in some cases, as is clearly visible in the first column. In addition, the label assignments in the spatial domain cannot cope with the noise of the input data. As for **(CFa)** and **(CFb)**, three labels of the initial label set (last row) "died out" during the unsupervised assignment flow evolution, whereas the remaining ones converged to values quite close to ground truth.

for spatially regularized assignments, *not only* for determining a reasonably coherent partition of the image domain *but also* for affecting label evolution, in order to determine proper labels enabling to find such a partition by assignment.

The labelings generated by unsupervised assignment flow **(UAF)** are shown by Figure 5.7, for the parameters $\sigma = \rho = 1.0$ and $\sigma = \infty$ corresponding to the specific versions **(CFa)** and **(CFb)** of the **(UAF)**, and using different neighborhood sizes $|\mathcal{N}| \in \{1\times 1, 3\times 3, 5\times 5\}$ for spatial regularization. The relative speed parameter α for the prototype evolution flow was set to the natural value $\alpha = 1$ (cf. Section 5.7.1). The results show that, for both flows **(CFa)** and **(CFb)**, spurious labels "die out" whereas the remaining labels converge to values quite close to ground truth (Figure 5.8). Specifically, for the large green background region, two labels close to the ground truth label are recovered due to the initial fluctuations within a large spatial region.

We point out that the only essential parameter value required for a reasonable result is the scale (neighborhood size) of spatial regularization.

5.7.4 Orientation Vector Fields

Given a grayscale image (Figure 5.9), we estimated orientations of local image structure from local gradient scatter matrices. Orientations are encoded at each pixel by the angle between the horizontal axis and the smallest eigenvector. The resulting data take values in $\mathbb{R}/\pi\mathbb{Z} \cong S^1$ after identifying antipodal points. Figure 5.9 shows the nearest neighbor assignments of the initial $|J| = 8$ prototypes determined by greedy k-center clustering from the noisy input data, together with labels and label assignments of the versions **(CFa)** and **(CFb)** of the unsupervised assignment flow **(UAF)** corresponding to the parameter choices $\sigma = \rho = 0.1$ and $\sigma = \infty$. The

relative speed parameter α for the prototype evolution was set to $\alpha = 0.5$ and $|\mathcal{N}| = 5 \times 5$ neighborhoods were used for spatial averaging.

Both flows managed to position a label correctly in the neighborhood of $0 \cong \pi$ (visualized in red) and only required seven labels to properly encode the data by labeling.

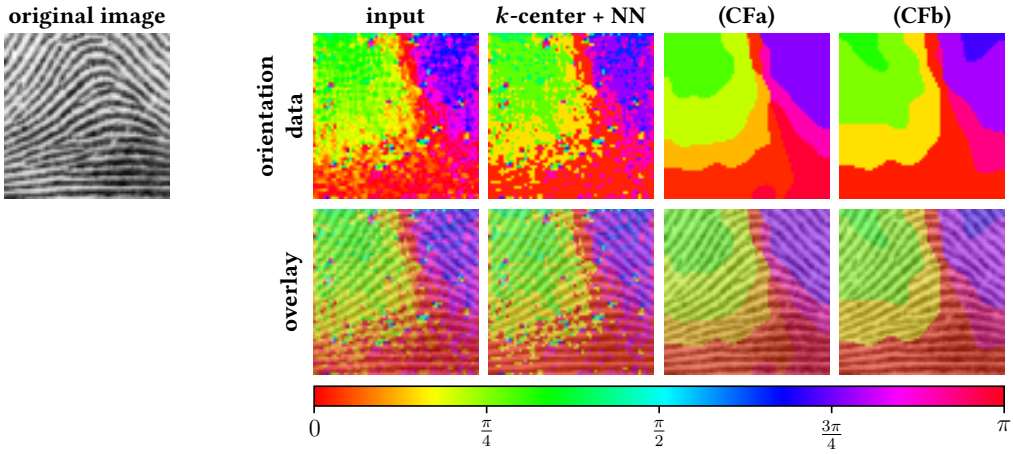


Figure 5.9: Unsupervised label learning from orientation vector fields. Orientations are extracted from the grayscale image, using the spectral decomposition of local scatter matrices of the image gradient, and represented as elements of $\mathbb{R}/_{\pi\mathbb{Z}} \cong S^1$ as described in the text. Using a corresponding distance function, the unsupervised assignment flow learns both proper labels, including their number, and label assignments for encoding the noisy input data.

5.7.5 Feature Covariance Descriptor Fields

We demonstrate the application of the unsupervised assignment flow to the manifold of positive definite matrices. For a given input image, we extracted the covariance descriptor using the feature map (5.81) and $\varepsilon = 10^{-7}$ in (5.82). We applied version **(CFa)** of the unsupervised assignment flow, i.e. setting $\sigma = \infty$ ensuring a strong effect of spatial regularization on label evolution. We compare this flow with the (supervised) assignment flow to demonstrate the effect of label adaption on the resulting labeling. Initial sets of labels were determined by metric clustering. Label assignments are visualized by false colors, i.e., different colors represent different labels. Due to the higher dimension of the feature space of this scenario, a larger value $\alpha = 10$ of the relative speed parameter controlling the prototype evolution turned out to be useful for both test instances.

Figure 5.10 shows an instance with $|J| = 8$ initial labels. For this example we used $|\mathcal{N}| = 7 \times 7$ for the covariance descriptor (5.82) and $|\mathcal{N}| = 5 \times 5$ for the spatial regularization. While the (supervised) assignment flow already suppresses noise in contrast to the nearest neighbor (NN) assignment, the unsupervised assignment flow **(CFa)** yields a more coherent result due to the label evolution that preserves visual features (eye region) or even enhances them (tip of the nose).

An instance with $|J| = 100$ initial labels is shown in Figure 5.11. Here, we used $|\mathcal{N}| = 5 \times 5$ for the covariance descriptor (5.82). The same observation applies as in the previous example, with a greater difference between the supervised and unsupervised assignment flow due to the

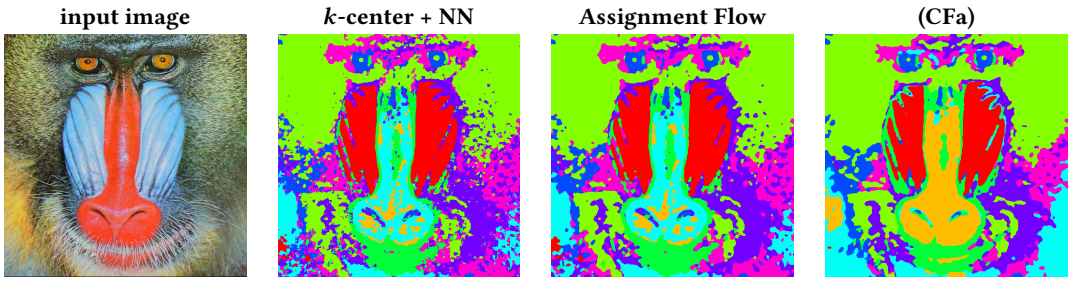


Figure 5.10: Impact of spatial regularization and label evolution (8 initial labels). Compared to the nearest neighbor (NN) assignment, the assignment flow consistently suppresses fine details. The additional label evolution of coupled flow **(CFa)** improves the spatially coherent assignment and preserves visual features (eye region, tip of the nose)

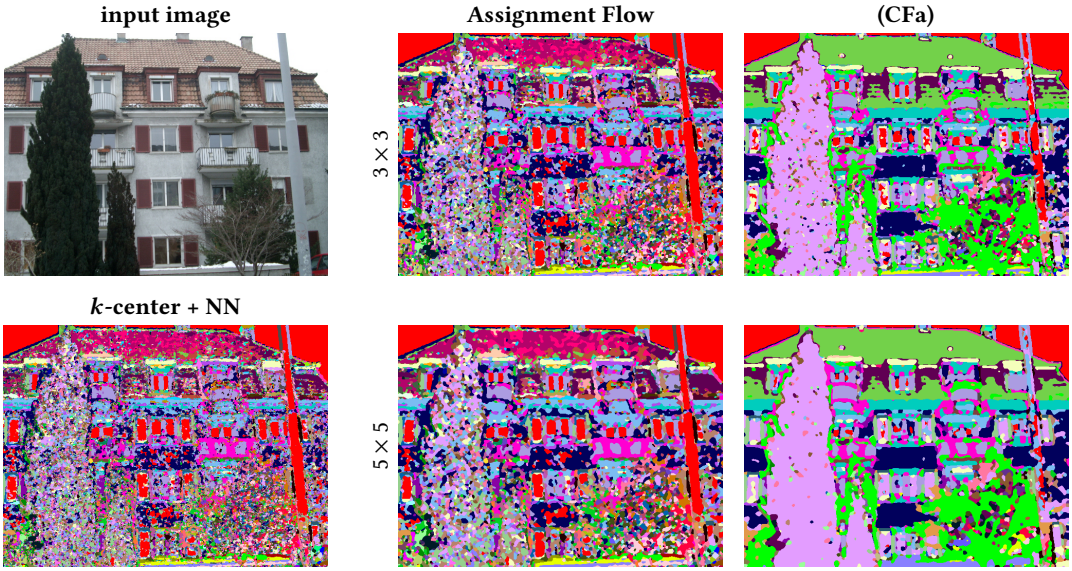


Figure 5.11: Impact of spatial regularization and label evolution (100 initial labels). The observations stated in the caption of Figure 5.10 hold here as well. We only point out that with sufficiently strong spatial regularization (bottom row), the coupled flow **(CFa)** managed to adapt labels on the feature manifold, and thus represents textured regions (tree, roof) more compactly in a spatially coherent way.

high number of labels. In particular, the simultaneous label evolution and spatial regularization of the unsupervised assignment flow **(CFa)** enable more compact representations of textured regions corresponding to the tree and the roof.

5.7.6 Rotation-Invariant Clustering of Covariance Descriptor Fields

We extend the experiments on covariance descriptors by using the rotation-invariant dissimilarity $D_{\mathrm{S},\mathcal{R}}$ (5.94). We continue using the feature map (5.81) and $|\mathcal{N}| = 5 \times 5$ for the covariance descriptor (5.82) as well as the version **(CFa)** of the unsupervised assignment flow with $\alpha = 10$. A synthetic and a real world image are considered, for which we use $|J| = 10$ initial labels in both cases, to ensure interpretation of the results visualized by false colors.

Before we proceed to the actual experiments, we first discuss the approximation of $\delta(\theta) = D_{\mathrm{S}}(X, R(\theta)YR(\theta)^{\top})$ by $\delta_N(\theta)$ (5.99), which is used for the numerical evaluation of $D_{\mathrm{S},\mathcal{R}}$. For this we have taken two covariance descriptors from the example in Figure 5.15 and depicted the approximation in Figure 5.12. As can be seen, the approximation for $N = 12$ is quite accurate, so that it can be used to locate the minimum of $\delta(\theta)$. Note that for $N = 8$ the approximation $\delta_N(\theta)$ would identify the wrong local minimum of $\delta(\theta)$ as the global minimum. Since we have observed this approximation accuracy of $\delta_N(\theta)$ for numerous other pairs (X, Y), we evaluate $D_{\mathrm{S},\mathcal{R}}$ in the following experiments using the approximation $\delta_N(\theta)$ with $N = 12$.

Figure 5.13 depicts a synthetic image with a texture rotated in steps of 15 degrees. $|\mathcal{N}| = 3 \times 3$ neighborhoods were used for spatial averaging and the constant of (5.82) was set to $\varepsilon = 10^{-5}$ to ensure strict positive definiteness even in completely homogeneous regions of this computer-generated image. Initial prototypes were extracted from the input data using the greedy k-center clustering with respect to the Stein divergence D_{S} and its rotation-invariant version $D_{\mathrm{S},\mathcal{R}}$, respectively. The experiments below should assess if numerical results display the rotational invariance of $D_{\mathrm{S},\mathcal{R}}$ that holds by construction mathematically (Section 5.6.4).

The six panels on the right of Figure 5.13 show columnwise the results of local label assignments (k-center + NN) and the assignments after label evolution performed by **(CFa)**, respectively, using either distance D_{S} or $D_{\mathrm{S},\mathcal{R}}$. Regarding the results depicted by the center column, greedy k-center clustering was performed using $D_{\mathrm{S},\mathcal{R}}$, while the nearest neighbor (NN) assignment and **(CFa)** were performed using D_{S}, in order to highlight the difference between D_{S} and $D_{\mathrm{S},\mathcal{R}}$ based on the same initial prototypes.

The result shows that using $D_{\mathrm{S},\mathcal{R}}$ leads to an unsupervised labeling of all textures with a single label only. Thus, depending on the application, using $D_{\mathrm{S},\mathcal{R}}$ instead of the basic Stein divergence D_{S} can lead to more compact label dictionaries determined by the proposed unsupervised assignment flow. Figure 5.14 underlines this finding from a different angle. The two panels on the left display the *pixelwise* distances D_{S} and $D_{\mathrm{S},\mathcal{R}}$ to some fixed (arbitrary) reference descriptor. The two images show quantitatively that D_{S} is highly non-uniform, unlike $D_{\mathrm{S},\mathcal{R}}$. The panel on the right of Figure 5.14 visualizes for each pixel the optimal angle minimizing (5.94) over (5.93) that has to be determined for the evaluation of $D_{\mathrm{S},\mathcal{R}}$. One can clearly see how the rotations of the textures of the input image of Figure 5.13 are recovered. This may be useful for some applications as well.

Figure 5.15 depicts a real world image. We used $|\mathcal{N}| = 5 \times 5$ neighborhoods for spatial averaging and $\varepsilon = 10^{-7}$ for the constant of (5.82) to ensure strict positive definiteness of the covariance descriptors. Analogous to Figure 5.13, we compared the nearest neighbor (NN) assignment and the result returned by **(CFa)** with respect to the Stein divergence D_{S} and its

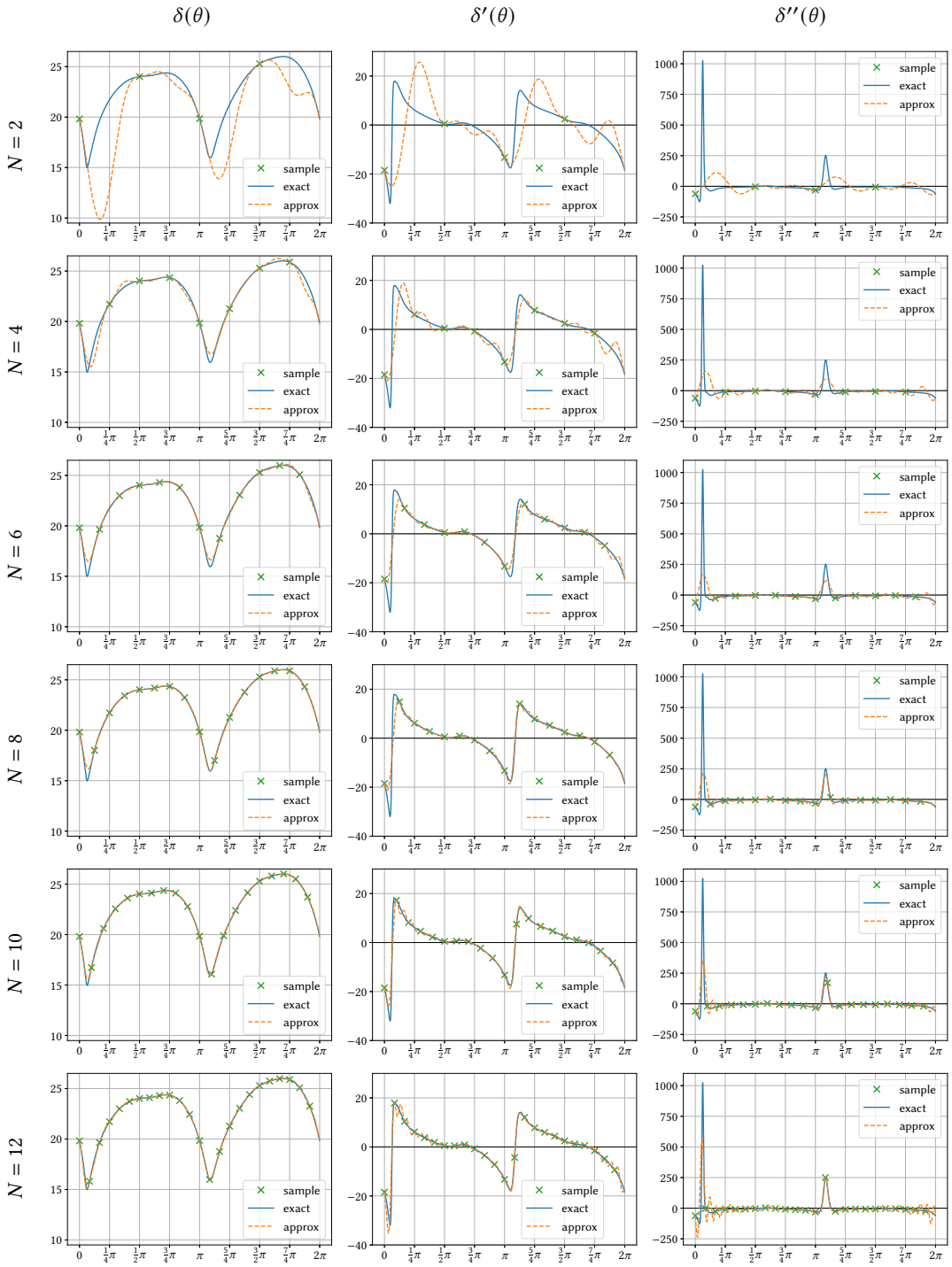


Figure 5.12: Approximation of $\delta(\theta)$ (5.98) **by trigonometric polynomial $\delta_N(\theta)$** (5.99)**.** The plots illustrate how well $\delta(\theta)$ (blue line) and its first two derivatives are approximated by the trigonometric polynomial $\delta_N(\theta)$ (orange dashed line) for different number of sampling points (marked with green crosses). Here, X and Y are covariance descriptors depicted from the input image shown in Figure 5.15. This example shows that for $N = 12$ the approximation is accurate enough to locate the minima of $\delta(\theta)$ using $\delta_N(\theta)$.

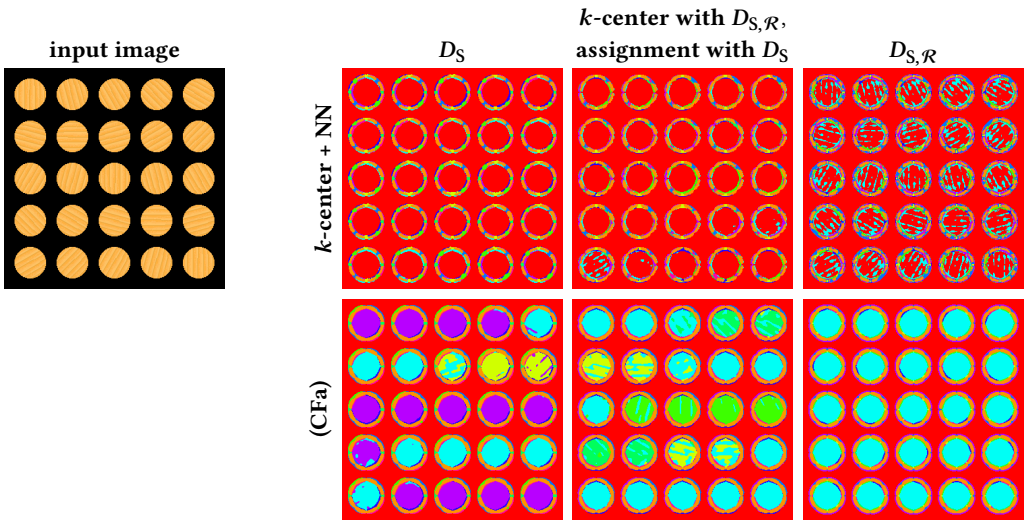


Figure 5.13: Unsupervised learning of rotationally invariant labels from covariance descriptors. The input data are covariance descriptors (5.82) extracted from the input image which comprises a texture rotated in steps of 15°. Both labels and label assignments are pixelwise visualized using false colors in the panels on the right (only color differences matter, rather than the colors themselves). The unsupervised assignment flow **(CFa)** together with the rotationally invariant Stein divergence $D_{S,\mathcal{R}}$ returns a small set of labels that encodes local image structure irrespective of its orientation. By contrast, using the Stein divergence D_S is less effective.

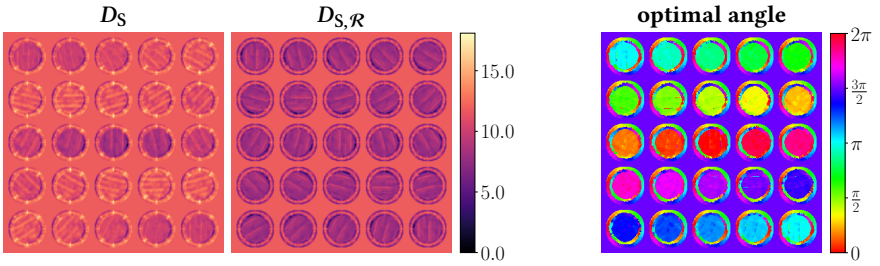


Figure 5.14: Comparing the Stein divergence D_S with its rotationally invariant variant $D_{S,\mathcal{R}}$. Using the covariance descriptors illustrated by Figure 5.13, the panels on the left show the pixelwise distances to some fixed (arbitrary) label. Contrary to the uniform distances $D_{S,\mathcal{R}}$, the distance D_S strongly depends on the orientation of the texture. On the right-hand side, the optimal rotation angles are shown corresponding to the evaluation of $D_{S,\mathcal{R}}$. These angles accurately recover rotations of the texture.

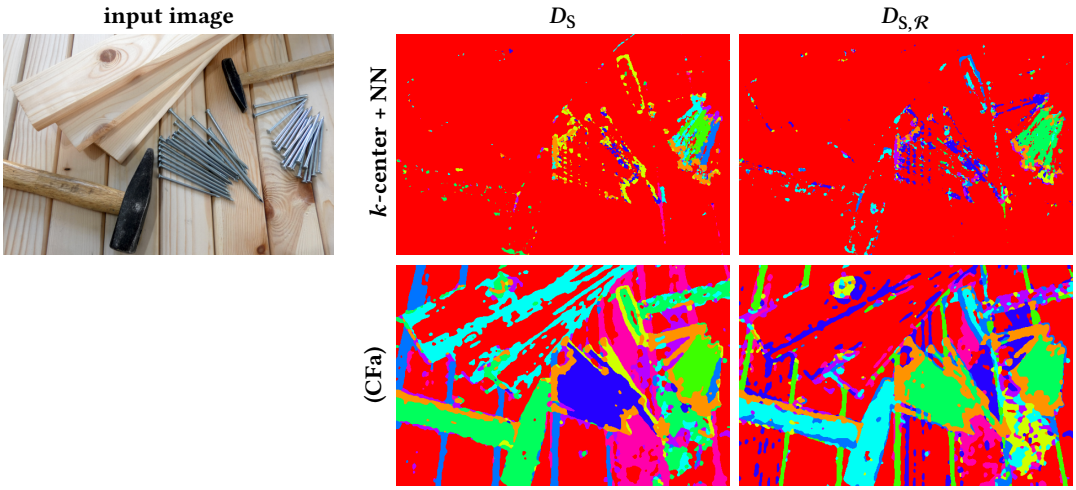


Figure 5.15: Unsupervised learning of covariance descriptor labels through label assignment. The depicted results were obtained for the real input image on the left and are analogous to the results of the synthetic scenario depicted by Figure 5.13. The local assignments of initial labels (top row on the right) highlight that metric clustering completely ignores the spatial structure of the input data. The results returned by the unsupervised assignment flow **(CFa)**, therefore, are impressive (bottom row): labels and label assignments jointly evolve so as to capture the spatial image structure. While the distance D_{S} is sensitive to orientations of texture, the distance $D_{\mathrm{S},\mathcal{R}}$ is not: the final labels and label assignments (bottom right) basically partition the image into wooden texture independent of the orientation of the wooden boards (encoded with red), nails and similar line structures in the background (encoded with green), the hammers (light-blue) and oriented wooden texture (blue), independent of the local orientation of these textures.

rotationally invariant version $D_{\mathrm{S},\mathcal{R}}$, respectively.

We observe that the rotationally invariant feature representation together with the unsupervised assignment flow ($D_{\mathrm{S},\mathcal{R}}$ / **(CFa)**; panel bottom-right) leads to an unsupervised label representation of the input data that basically partitions the image into wooden texture independent of the orientation of the wooden boards (encoded with red), nails and similar line structures in the background (encoded with green), the hammers (light-blue) and oriented wooden texture (blue).

Analogous to Figure 5.14, Figure 5.16 (first row) shows the pixelwise distances to a fixed label (located at the right pile of nails) for the distances D_{S} and $D_{\mathrm{S},\mathcal{R}}$, respectively. Comparing the distances to the two piles of nails illustrates once again and quantitatively the rotational invariance of $D_{\mathrm{S},\mathcal{R}}$. The bottom row of panels shows the corresponding optimal rotation angles corresponding to the evaluation of $D_{\mathrm{S},\mathcal{R}}$, as defined by (5.94). These angles recover the relative orientation of the textures which may be useful for some applications.

5.7.7 PolSAR Data

As the last application of our approach, we discuss segmentation of *PolSAR (Polarimetric synthetic aperture radar)* data, which is an effective tool to monitor ground surface and to perform terrain and land use classification [FD98, XJZ16].

PolSAR measures target reflectivity by using different polarization combinations. Follow-

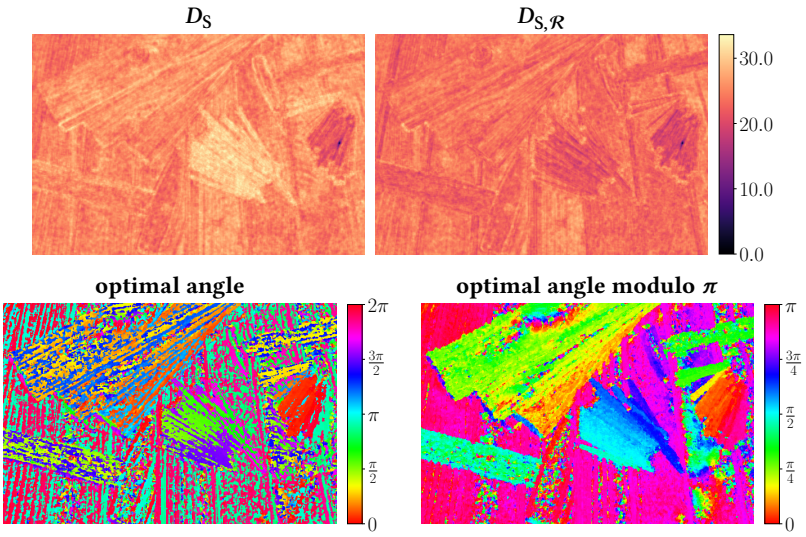


Figure 5.16: Comparing the Stein divergence D_S with its rotationally invariant variant $D_{S,\mathcal{R}}$. The depicted results correspond to the scenario of Figure 5.15 and are analogous to the results shown by Figure 5.14 for the synthetic scenario illustrated by Figure 5.13. The top row shows the pixelwise distances between each covariance descriptors extracted from the image of Figure 5.15 and a fixed prototype located at the pile of nails on the right. While the distance D_S considerably differs between two piles of nails due to the different orientations, the rotationally-invariant distance $D_{S,\mathcal{R}}$ is more uniform. The bottom row displays pixelwise the optimal rotation angle that determines $D_{S,\mathcal{R}}$. Up to unavoidable local errors of these locally computed estimates, the distance $D_{S,\mathcal{R}}$ recovers the local orientation of the real texture in the input image (bottom right).

ing the presentation in [NFC14], the polarimetric coherent information associates, at each frequency of operation, to each pixel a 2×2 complex matrix with entries S_{VV}, S_{VH}, S_{HV}, and S_{HH}, where S_{ij}, $i, j \in \{H, V\}$ is the backscattered signal for the i-th transmission and the j-th reception of linear polarization. Under mild conditions, $S_{HV} = S_{VH}$, and the scattering matrix can be simplified to the three-component complex vector $s = \left[S_{VV}, \sqrt{2}\, S_{VH}, S_{HH}\right]^\top$. This random vector can be modeled by the zero-mean multivariate complex Gaussian distribution [Goo63]. Different targets are characterized by different variances and, thus, are hard to discern visually. In order to reveal such differences and to improve the signal-to-noise ratio, L ideally independent measurements of the same target are obtained while processing the raw data. These observations are used to produce the *multilook sample covariance matrix format*

$$Z = \frac{1}{L} \sum_{i \in [L]} s^i (s^i)^*, \tag{5.104}$$

where the superscript * represents the complex conjugate transpose of a vector, and $\{s^i\}_{i \in [L]}$ are the L scattering vectors. Assuming that these vectors are independent, Z is a Hermitian positive definite matrix governed by a scaled complex Wishart distribution [ADE09]. Having Σ and L as parameters, the scaled complex Wishart distribution is represented by the probability density function

$$f_Z(z; \Sigma, L) = \frac{L^{3L} \det(z)^{L-3}}{\det(\Sigma)^L \Gamma_3(L)} \exp\big(-L \operatorname{tr}(\Sigma^{-1} z)\big), \tag{5.105}$$

where $\Gamma_3(L) = \pi^3 \prod_{i=0}^{2} \Gamma(L-i)$ for $L \geq 3$, $\Gamma(\cdot)$ is the gamma function, and Σ is the covariance matrix associated to s.

For the segmentation of real PolSAR data, we apply the unsupervised assignment approach worked out for the manifold of positive definite matrices in Section 5.6.3. For simplicity, we use the same parameters as for covariance descriptors, i. e., we use the version **(CFa)** with $\alpha = 10$ for the label evolution and $|\mathcal{N}| = 5 \times 5$ for the spatial regularization. We compare the performance of our approach with the performance of the following approaches:

1. *Supervised Wishart classification* **(SWC)**: each observation is assigned to the class whose density given by equation (5.105) is maximal.

2. *Supervised reaction diffusion equation* **(SRDE)**: Method proposed in [GAMF15, GAMF17]. Classification is obtained as the asymptotic state of a reaction diffusion equation system.

3. *Riemannian sparse representation with Stein divergence* **(RSR-S)**: a kernel embedding-based approach for dictionary learning and sparse coding for positive definite matrices, proposed in [HHLS16].

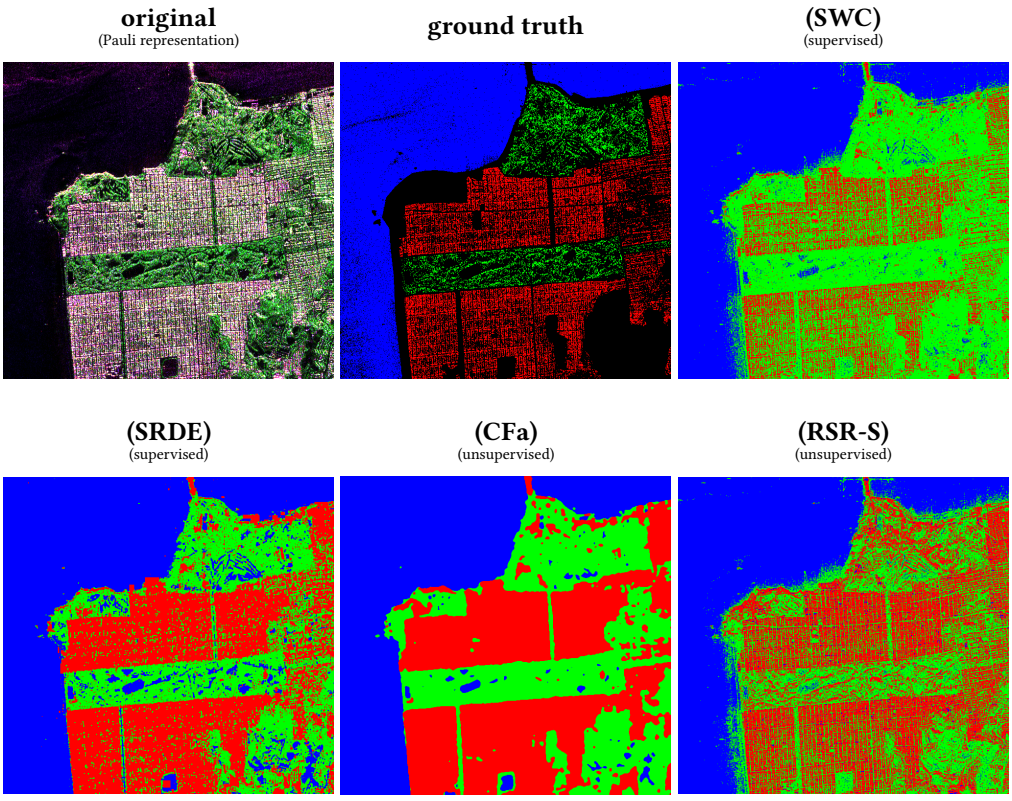


Figure 5.17: Experimental results for the San Francisco bay image. The input data is visualized by Pauli representation. The labels of the ground truth and classification results of the different approaches are encoded by the colors {■ = ocean, ■ = urban, ■ = forest}, black pixels in the ground truth image are not associated with any of the three classes. The unsupervised assignment flow **(CFa)** yields more homogeneous regions due to spatial geometric regularization of the assignments.

We use a PolSAR image obtained by the AIRSAR sensor over the San Francisco bay.[2] The image is shown in Figure 5.17 using the standard Pauli representation [LE09] which maps an Hermitian complex matrix into the RGB color space. The image size is 676×648 pixels. Three classes are readily identified: ocean, urban and forest areas.

Figure 5.17 shows the ground truth and the classification results of the different methods. Our approach yields more homogeneous regions due to spatial geometric regularization of the assignments. Table 5.1 lists the relative miss-classification errors and the accuracy for each individual class produced by each approach. Although our *unsupervised* approach **(CFa)** had no access to ground truth data, its performance is close to the performance of the leading *supervised* approach **(SRDE)**.

Method	**Type**	**Error**	**Accuracy by Class**		
		%	ocean	urban	forest
(SWC)	supervised	3.78	98.34	92.23	92.68
(SRDE)	supervised	1.34	99.91	97.47	92.56
(CFa)	unsupervised	1.61	99.95	97.31	89.38
(RSR-S)	unsupervised	9.14	96.66	97.96	18.53

Table 5.1: Classification scores for the San Francisco bay image. For each approach the error is shown as the percentage of miss-classified pixels relative to the ground truth data. Additionally, the accuracy is shown for each individual class as percentage. Our unsupervised approach **(CFa)** achieves results close to the leading *supervised* approach **(SRDE)**.

5.8 Conclusion

We proposed the unsupervised assignment flow for performing jointly label evolution on feature manifolds and spatially regularized label assignment to given feature input data. The approach alleviates the requirement for supervised image labeling to have proper labels at hand, because an initial set of labels can evolve and adapt to better values while being assigned to given data.

The derivation of our approach highlights that it encompasses related state-of-the-art approaches to unsupervised learning: soft-k-means clustering and EM-based estimation of mixture distributions with distributions of the exponential family as mixture components (class-conditional feature distributions). We generalized these approaches to manifold-valued data and defined the unsupervised assignment flow by coupling label evolution with the assignment flow from Section 2.2. We suggested greedy k-center clustering for determining an initial label set that works with linear complexity in any metric space and with fixed approximation error bounded from above.

The separation between feature evolution and spatial regularization through assignments enables the flexible application of our approach to various scenarios, provided some key operations (divergence function evaluation, exponential map) are computational feasible for

[2] Details about these freely available data can be accessed at ESA's web page `https://earth.esa.int/web/polsarpro/data-sources/sample-datasets`.

the particular feature manifold at hand. We demonstrated this property for three different feature manifolds and showed that coupling the evolution of labels and assignments has beneficial effects in either direction. The approach involved two parameters for the label evolution part whose role is well understood. The only essential parameter for the label assignment part is the neighborhood size used for spatial regularization.

Our unsupervised learning approach is consistent in that the very same approach that is used for supervised labeling is used for label learning, without need to resort to approximate inference due to the complexity of learning, as is the case, e. g., for learning with graphical models.

A key property of our approach is the sparsifying effect of spatial assignment regularization on unsupervised label learning. All experiments in this paper were conducted using uniform weights $\{\omega_{ik}\}_{k\in\mathcal{N}_i}$ for the spatial regularization of assignments – cf. (5.36). Learning these weights from data in order to represent the spatial context of typical feature occurrences as prior knowledge has been studied recently [HSPS19a]. Working out a mathematically consistent way to extend this approach to unsupervised scenarios, as studied in this chapter, defines an exciting modeling problem.

6 Conclusion

6.1 Summary

We focused on the image labeling problem and based our work on the assignment flow approach [ÅPSS17]. In addition to a mathematical investigation of the approach, we considered two important aspects of the image labeling problem. Firstly, we investigated how local label statistics can be included in the inference. Secondly, we addressed the question of how suitable labels for the labeling problem can be determined from the data.

In Chapter 3 we formulated and proved some theoretical statements concerning the asymptotic behavior of the assignment flow. This includes the statement postulated in [ÅPSS17] that the assignment flow for data in general position converges to integral assignment probabilities and thus unique label decisions. We showed this statement under mild conditions on the parameters. Furthermore, we examined the stability of limits and, for limits that fulfill a certain stability criterion, we derived a condition which makes it possible to verify convergence to these limits at an early stage. This condition provides a convergence guarantee, which we demonstrated on an academic example. The stability analysis was also carried out for the discretization [Sav+17] and we verified the convergence criterion for this discrete case as well. At the end, we considered a generalization of the S-flow [SS20], where the regularization also interacts between labels, and we discussed how the theoretical statements from this chapter apply to this flow.

Chapter 4 dealt with local label statistics induced by global convex constraints on filtering results of assignment probabilities. These constraints are incorporated into the labeling process by modifying the vector field of the assignment flow utilizing logarithmic barrier functions. The corresponding filters were learned from examples by solving an eigendecomposition and the effectiveness of the approach was evaluated in academic experiments.

In Chapter 5 we dealt with unsupervised label learning, i.e., no labels were given in advance for the labeling problem. As these labels should reflect the spatial regularization of the assignments, we took the approach of adapting initial labels over time to the assignment probabilities. For efficiency reasons, initial labels were determined with a greedy approximation of k-center clustering, which also provides a performance guarantee. The adaptation of the labels is done by a flow in the feature space, which is coupled to the assignment flow. This coupling is controlled by a parameter that balances the influence of data fidelity and spatial regularization on the label evolution. The simultaneous label learning and assignment not only improves the labels, but the improved labels also have a positive effect on the resulting image segmentation. The approach assumes that the data lie on a Riemannian manifold. We worked out the approach for the one-dimensional sphere, the Lie group of rotation matrices, and the set of positive definite matrices. Finally, we evaluated the approach in numerous experiments. These experiments demonstrated the influence of the parameters and the applicability for different manifolds.

6.2 Outlook

Finally, we would like to list some interesting research directions, in particular on label statistics, which have partly arisen from the results of this work:

- **Label statistics induced by weights.** The stability criteria in Corollary 3.6.5 show that label statistics can be induced by the weights that parameterize the similarity map (2.104). In this context, weights are considered which also interact between labels. Based on the flow in Section 3.6, a new approach for inducing label statistics could be developed. This would require further work on learning these weights on the basis of exemplary labelings. In some applications where label statistics are partially known, such as ordering constraints in OCT data [RSS14], the approach could also be tried with handcrafted weights. For the manual design of the weights one could rely on the stability criteria in Corollary 3.6.5.
- **Parameter learning using stability criteria.** Parameter learning is about determining parameters for given data and an optimality condition, e.g. given by a groundtruth, so that the limit of the assignment flow with these parameters meets the optimality condition. In [HSPS19b] a large-scale system of ODEs is considered for this problem. An interesting research question is whether the stability criteria in Corollary 3.4.5 and the estimates of the basin of attraction in Proposition 3.4.12 can be used to develop an efficient approach to this.
- **Prior label statistics for unsupervised label learning.** Label learning could be improved if additional information like prior label statistics are incorporated into the label learning process. This approach would be limited to applications where certain label statistics such as ordering constraints are known in advance. A challenge with this combination is that label statistics specify a certain arrangement of the labels, whereas (unsupervised) clustering, by its nature, does not have an arrangement of the clusters.

A Proofs of Propositions in Section 2.3

Proposition 2.3.9. *Let x^* be an equilibrium point of $\dot{x}(t) = F(x(t))$ with $F \in C^1(U, \mathbb{R}^n)$. Then*

(a) There exist a sufficiently small $\varepsilon_1 > 0$ and a (real) similarity transform

$$V^{-1} \tfrac{\partial F}{\partial x}(x^*) V = \begin{pmatrix} A_{sc} & 0 \\ 0 & A_u \end{pmatrix} = A, \tag{A.1}$$

such that

(i) $\mathrm{Re}(\lambda) \leq 0$ *for all eigenvalues λ of A_{sc},*

(ii) $\mathrm{Re}(\lambda) > 0$ *for all eigenvalues λ of A_u,*

(iii) $\langle y_{sc}, A_{sc} y_{sc} \rangle \leq \frac{\varepsilon_1}{4} \|y_{sc}\|^2$,

(iv) $\langle y_u, A_u y_u \rangle \geq \varepsilon_1 \|y_u\|^2$.

(b) Suppose $\frac{\partial F}{\partial x}(x^*)$ has at least one eigenvalue λ with $\mathrm{Re}(\lambda) > 0$. Considering an affine coordinate transform $y = V^{-1}(x - x^*)$ with $V \in \mathrm{GL}_n(\mathbb{R})$ from (a), the resulting flow

$$\dot{y} = G(y) = V^{-1} F(Vy + x^*), \tag{A.2}$$

which has the equilibrium $y^* = 0$ with $\frac{\partial G}{\partial y}(0) = V^{-1} \frac{\partial F}{\partial x}(x^*) V = A$, has the following property. There exist $\eta > 0$, $\delta > 0$ and $\varepsilon > 0$, such that if the flow starts at some point in the open truncated cone

$$U_{\eta,\delta} = \left\{ y = \left(\begin{smallmatrix} y_{sc} \\ y_u \end{smallmatrix} \right) \in \mathbb{R}^n \colon \|y_{sc}\|^2 < \eta \|y_u\|^2,\ \|y\| < \delta \right\} \subset B_\delta(0) = \left\{ y \in \mathbb{R}^n \colon \|y\| < \delta \right\}, \tag{A.3}$$

then the solution will not cross the conical portion of $\partial U_{\eta,\delta}$, i.e.

$$\left\{ y \in \mathbb{R}^n \colon \|y_{sc}\|^2 = \eta \|y_u\|^2,\ \|y\| < \delta \right\}, \tag{A.4}$$

and it fulfills $\|y(t)\| \geq \|y(0)\| e^{\varepsilon t}$ as long as $y(t) \in U_{\eta,\delta}$, i.e., $y(t)$ leaves the ball $B_\delta(0)$ at some time point. Especially, the equilibrium $y^* = 0$ is unstable. This property is accordingly transferred to the equilibrium x^* of $\dot{x}(t) = F(x(t))$ using $x(t) = Vy(t) + x^*$.

Proof. We work out the sketched proof in [SC16, Section 6.8.4].

(a) By [HJ12, Theorem 3.4.1.5], any real $n \times n$ matrix is similar to a real block diagonal matrix with blocks either of the form

$$J_k(\lambda) = \begin{pmatrix} \lambda & 1 & & \\ & \ddots & \ddots & \\ & & \ddots & 1 \\ & & & \lambda \end{pmatrix} \in \mathbb{R}^{k \times k} \tag{A.5}$$

in case of a real eigenvalue λ or of the form

$$C_k(a,b) = \begin{pmatrix} C_1(a,b) & I_2 & & \\ & \ddots & \ddots & \\ & & \ddots & I_2 \\ & & & C_1(a,b) \end{pmatrix} \in \mathbb{R}^{2k\times 2k} \qquad \text{with} \quad C_1(a,b) = \begin{pmatrix} a & -b \\ b & a \end{pmatrix} \tag{A.6}$$

in case of a pair of complex conjugate eigenvalues $\lambda = a \pm bi$. These blocks are unique up to their order. The size of these blocks, i.e. $k \in \mathbb{N}$, depends on the algebraic and geometric multiplicity of the corresponding eigenvalues. The nilpotent part in the blocks $J_k(\lambda)$ and $C_k(a,b)$ can be scaled by an $\varepsilon > 0$ using another similarity transform:

$$V^{-1} J_k(\lambda) V = \begin{pmatrix} \lambda & \varepsilon & & \\ & \ddots & \ddots & \\ & & \ddots & \varepsilon \\ & & & \lambda \end{pmatrix} \quad \text{for} \quad V = \operatorname{Diag}\left(1, \varepsilon, \varepsilon^2, \dots, \varepsilon^{k-1}\right), \tag{A.7}$$

$$V^{-1} C_k(a,b) V = \begin{pmatrix} C_1(a,b) & \varepsilon I_2 & & \\ & \ddots & \ddots & \\ & & \ddots & \varepsilon I_2 \\ & & & C_1(a,b) \end{pmatrix}$$
$$\text{for} \quad V = \operatorname{Diag}\left(1, 1, \varepsilon, \varepsilon, \varepsilon^2, \varepsilon^2, \dots, \varepsilon^{k-1}, \varepsilon^{k-1}\right). \tag{A.8}$$

We will denote the latter blocks by $J_{k,\varepsilon}(\lambda)$ and $C_{k,\varepsilon}(a,b)$. These blocks fulfill

$$\left(a-\varepsilon\right)\|x\|^2 \le \left\langle x, C_{k,\varepsilon}(a,b)x\right\rangle \le \left(a+\varepsilon\right)\|x\|^2 \qquad \forall x \in \mathbb{R}^{2k}, \tag{A.9a}$$
$$(\lambda-\varepsilon)\|x\|^2 \le \left\langle x, J_{k,\varepsilon}(\lambda)x\right\rangle \le (\lambda+\varepsilon)\|x\|^2 \qquad \forall x \in \mathbb{R}^{k}. \tag{A.9b}$$

Note that $\left\langle x, C_1(a,b)x\right\rangle = a\|x\|^2$. By ordering the blocks, we get a similarity transform

$$V^{-1}\frac{\partial F}{\partial x}(x^*)V = \begin{pmatrix} A_{\mathrm{sc}} & 0 \\ 0 & A_{\mathrm{u}} \end{pmatrix}, \tag{A.10}$$

where A_{sc} and A_{u} are block diagonal matrices with A_{sc} containing blocks $J_{k,\varepsilon}(\lambda)$ and $C_{k,\varepsilon}(a,b)$ for eigenvalues λ with $\operatorname{Re}(\lambda) \le 0$, and A_{u} containing the blocks for the remaining eigenvalues. By (A.9), we get

$$\langle y_{\mathrm{sc}}, A_{\mathrm{sc}} y_{\mathrm{sc}}\rangle \le \Big(\max_{\lambda\in\sigma(A_{\mathrm{sc}})} \operatorname{Re}(\lambda) + \varepsilon\Big)\|y_{\mathrm{sc}}\|^2 \le \varepsilon\|y_{\mathrm{sc}}\|^2, \tag{A.11a}$$
$$\langle y_{\mathrm{u}}, A_{\mathrm{u}} y_{\mathrm{u}}\rangle \ge \Big(\min_{\lambda\in\sigma(A_{\mathrm{u}})} \operatorname{Re}(\lambda) - \varepsilon\Big)\|y_{\mathrm{u}}\|^2 \tag{A.11b}$$

Hence, the choice

$$\varepsilon_1 = \frac{4}{5}\min_{\lambda\in\sigma(A_{\mathrm{u}})} \operatorname{Re}(\lambda) > 0, \qquad \varepsilon = \frac{\varepsilon_1}{4} \tag{A.12}$$

in the above transformations completes the proof.

(b) We have $G(y) = Ay + r(y)$ with $\frac{\|r(y)\|}{\|y\|} \to 0$ for $\|y\| \to 0$. We set

$$\eta = \frac{\varepsilon_1}{4\|A_{\mathrm{sc}}\|}, \qquad \varepsilon = \frac{\varepsilon_1}{2(1+\eta)}, \tag{A.13}$$

and $\delta > 0$ such that

$$\frac{\|r(y)\|}{\|y\|} < \min\left\{\frac{\varepsilon_1}{4(1+\eta)}, \frac{\varepsilon_1 \eta^{\frac{1}{2}}}{4(1+\eta)}\right\} \qquad \text{for} \quad \|y\| < \delta. \tag{A.14}$$

We bound

$$\langle y, G(y)\rangle = \langle y_{\mathrm{sc}}, A_{\mathrm{sc}} y_{\mathrm{sc}}\rangle + \langle y_{\mathrm{u}}, A_{\mathrm{u}} y_{\mathrm{u}}\rangle + \langle y, r(y)\rangle \tag{A.15}$$

by estimating each summand individually. We note that $y \in U_{\eta,\delta}$ implies

$$\|y_{\mathrm{sc}}\|^2 \le \eta \|y_{\mathrm{u}}\|^2, \tag{A.16a}$$

$$\|y\|^2 = \|y_{\mathrm{sc}}\|^2 + \|y_{\mathrm{u}}\|^2 \le (1+\eta)\|y_{\mathrm{u}}\|^2. \tag{A.16b}$$

Hence we get

$$\begin{aligned} \langle y_{\mathrm{sc}}, A_{\mathrm{sc}} y_{\mathrm{sc}}\rangle &\ge -\|A_{\mathrm{sc}}\|\|y_{\mathrm{sc}}\|^2 \overset{(A.16a)}{\ge} -\eta\|A_{\mathrm{sc}}\|\|y_{\mathrm{u}}\|^2 \overset{(A.13)}{=} -\tfrac{\varepsilon_1}{4}\|y_{\mathrm{u}}\|^2 \\ &\overset{(A.16b)}{\ge} -\tfrac{\varepsilon_1}{4(1+\eta)}\|y\|^2 \overset{(A.13)}{=} -\tfrac{\varepsilon}{2}\|y\|^2, \end{aligned} \tag{A.17a}$$

$$\langle y_{\mathrm{u}}, A_{\mathrm{u}} y_{\mathrm{u}}\rangle \overset{(a)}{\ge} \varepsilon_1 \|y_{\mathrm{u}}\|^2 \overset{(A.16b)}{\ge} \tfrac{\varepsilon_1}{1+\eta}\|y\|^2 \overset{(A.13)}{=} 2\varepsilon\|y\|^2, \tag{A.17b}$$

$$\langle y, r(y)\rangle \ge -\|y\|\|r(y)\| \overset{(A.14)}{\ge} -\tfrac{\varepsilon_1}{4(1+\eta)}\|y\|^2 \overset{(A.13)}{=} -\tfrac{\varepsilon}{2}\|y\|^2. \tag{A.17c}$$

Combining the estimates (A.17) with (A.15) gives

$$\langle y, G(y)\rangle \ge -\tfrac{\varepsilon}{2}\|y\|^2 + 2\varepsilon\|y\|^2 - \tfrac{\varepsilon}{2}\|y\|^2 = \varepsilon\|y\|^2. \tag{A.18}$$

This implies $\frac{\mathrm{d}}{\mathrm{d}t}\|y(t)\|^2 = 2\langle y, G(y)\rangle \ge 2\varepsilon\|y(t)\|^2$, i.e. $\|y(t)\|^2 \ge e^{2\varepsilon t}\|y(0)\|^2$ as long as $y(t) \in U_{\eta,\delta}$. It remains to show that $t \mapsto y(t)$ cannot leave $U_{\eta,\delta}$ through the conical portion of $\partial U_{\eta,\delta}$, i.e., $t \mapsto y(t)$ eventually leaves the ball $B_\delta(0)$ at a point in time. Let $y \in \partial U_{\eta,\delta}$ be a point on the conical portion, that means,

$$\|y_{\mathrm{sc}}\|^2 = \eta\|y_{\mathrm{u}}\|^2 \qquad \text{and} \qquad \|y\|^2 = \tfrac{1+\eta}{\eta}\|y_{\mathrm{sc}}\|^2. \tag{A.19}$$

The inward normal $N = N(y) = (-y_{\mathrm{sc}}, \eta y_{\mathrm{u}})$ at $y \in \partial U_{\eta,\delta}$ fulfills

$$\|N\|^2 = (1+\eta)\|y_{\mathrm{sc}}\|^2. \tag{A.20}$$

We consider each summand of

$$\langle N(y), G(y)\rangle = -\langle y_{\mathrm{sc}}, A_{\mathrm{sc}} y_{\mathrm{sc}}\rangle + \eta\langle y_{\mathrm{u}}, A_{\mathrm{u}} y_{\mathrm{u}}\rangle + \langle N, r(y)\rangle \tag{A.21}$$

separately and find

$$\langle y_{\mathrm{sc}}, A_{\mathrm{sc}} y_{\mathrm{sc}}\rangle \overset{(a)}{\le} \tfrac{\varepsilon_1}{4}\|y_{\mathrm{sc}}\|^2, \tag{A.22a}$$

$$\eta\langle y_{\mathrm{u}}, A_{\mathrm{u}} y_{\mathrm{u}}\rangle \overset{(a)}{\ge} \varepsilon_1 \eta \|y_{\mathrm{u}}\|^2 \overset{(A.19)}{=} \varepsilon_1 \|y_{\mathrm{sc}}\|^2, \tag{A.22b}$$

$$\begin{aligned} \langle N, r(y)\rangle &\ge -\|N\|\|r(y)\| \overset{(A.20)}{=} -(1+\eta)^{\frac{1}{2}}\|y_{\mathrm{sc}}\|\|r(y)\| \\ &\overset{(A.14)}{\ge} -(1+\eta)^{\frac{1}{2}} \tfrac{\varepsilon_1 \eta^{\frac{1}{2}}}{4(1+\eta)}\|y_{\mathrm{sc}}\|\|y\| \overset{(A.19)}{=} -\tfrac{\varepsilon_1}{4}\|y_{\mathrm{sc}}\|^2. \end{aligned} \tag{A.22c}$$

This implies

$$\langle N(y), G(y)\rangle \geq -\tfrac{\varepsilon_1}{4}\|y_{\mathrm{sc}}\|^2 + \varepsilon_1\|y_{\mathrm{sc}}\|^2 - \tfrac{\varepsilon_1}{4}\|y_{\mathrm{sc}}\|^2 = \tfrac{\varepsilon_1}{2}\|y_{\mathrm{sc}}\|^2 \geq 0 \tag{A.23}$$

with equality only for $y = 0$, which completes the proof.

□

Proposition 2.3.13. *Let x^* be an equilibrium point of $x^{(t+1)} = F(x^{(t)})$. Then*

(a) There exist a sufficiently small $\varepsilon_1 > 0$ and a (real) similarity transform

$$V^{-1}\tfrac{\partial F}{\partial x}(x^*)V = \begin{pmatrix} A_{sc} & 0 \\ 0 & A_u \end{pmatrix} = A, \tag{A.24}$$

such that

(i) $|\lambda| \leq 1$ for all eigenvalues λ of A_{sc},

(ii) $|\lambda| > 1$ for all eigenvalues λ of A_u,

(iii) $\|A_{sc}y_{sc}\|^2 \leq (1 + \frac{\varepsilon_1}{4})^2\|y_{sc}\|^2$,

(iv) $\|A_u y_u\|^2 \geq (1+\varepsilon_1)^2\|y_u\|^2$.

(b) Suppose $\frac{\partial F}{\partial x}(x^*)$ has at least one eigenvalue λ with $|\lambda| > 1$. Considering an affine coordinate transform $y = V^{-1}(x - x^*)$ with $V \in \mathrm{GL}_n(\mathbb{R})$ from (a), the resulting map

$$G(y) = V^{-1}\big(F(Vy + x^*) - x^*\big), \tag{A.25}$$

which has the fixed point $y^* = 0$ with $\frac{\partial G}{\partial y}(0) = V^{-1}\frac{\partial F}{\partial x}(x^*)V = A$, has the following property. There exist $\eta > 0$ and $\delta > 0$, such that if $y^{(0)}$ is a point in the open truncated cone $U_{\eta,\delta} = C_\eta \cap B_\delta(0)$ with

$$C_\eta = \Big\{y = \left(\begin{smallmatrix} y_{sc} \\ y_u \end{smallmatrix}\right) \in \mathbb{R}^n \colon \|y_{sc}\|^2 < \eta\|y_u\|^2\Big\}, \tag{A.26a}$$

$$B_\delta(0) = \big\{y \in \mathbb{R}^n \colon \|y\| < \delta\big\}, \tag{A.26b}$$

then we have

(i) $\big\|G(y^{(0)})\big\|^2 \geq (1+\varepsilon_1)\big\|y^{(0)}\big\|^2$,

(ii) $G(y^{(0)}) \in C_\eta$.

Hence, the sequence $y^{(t+1)} = G(y^{(t)})$ eventually leaves the ball $B_\delta(0)$ at some time point. Especially, the equilibrium $y^* = 0$ is unstable. This property is accordingly transferred to the equilibrium point x^* of $F(x)$ using $x^{(t)} = Vy^{(t)} + x^*$.

Proof. We adapt the proof of proposition 2.3.9.

(a) Analogously to the proof of proposition 2.3.9, we get a similarity transform

$$V^{-1}\tfrac{\partial F}{\partial x}(x^*)V = \begin{pmatrix} A_{\mathrm{sc}} & 0 \\ 0 & A_{\mathrm{u}} \end{pmatrix}, \tag{A.27}$$

where A_{sc} and A_{u} are block diagonal matrices with A_{sc} containing blocks $J_{k,\varepsilon}(\lambda)$ and $C_{k,\varepsilon}(a,b)$ for eigenvalues λ with $|\lambda| \leq 1$, and A_{u} containing the blocks for the remaining eigenvalues. The blocks $J_{k,\varepsilon}(\lambda)$ and $C_{k,\varepsilon}(a,b)$ fulfill $\left\|C_1(a,b)x\right\|^2 = (a^2+b^2)\|x\|^2$ and

$$\left(\sqrt{a^2+b^2}-\varepsilon\right)^2\|x\|^2 \leq \left\|C_{k,\varepsilon}(a,b)x\right\|^2 \leq \left(\sqrt{a^2+b^2}+\varepsilon\right)^2\|x\|^2 \qquad \forall x \in \mathbb{R}^{2k}, \tag{A.28a}$$

$$\left(|\lambda|-\varepsilon\right)^2\|x\|^2 \leq \left\|J_{k,\varepsilon}(\lambda)x\right\|^2 \leq \left(|\lambda|+\varepsilon\right)^2\|x\|^2 \qquad \forall x \in \mathbb{R}^{k}. \tag{A.28b}$$

This implies

$$\left\|A_{\mathrm{sc}}y_{\mathrm{sc}}\right\|^2 \leq \left(\max_{\lambda\in\sigma(A_{\mathrm{sc}})}|\lambda|+\varepsilon\right)^2\|y_{\mathrm{sc}}\|^2 \leq (1+\varepsilon)^2\|y_{\mathrm{sc}}\|^2, \tag{A.29a}$$

$$\left\|A_{\mathrm{u}}y_{\mathrm{u}}\right\|^2 \geq \left(\min_{\lambda\in\sigma(A_{\mathrm{u}})}|\lambda|-\varepsilon\right)^2\|y_{\mathrm{u}}\|^2. \tag{A.29b}$$

Hence, the choice

$$\varepsilon_1 = \frac{4}{5}\left(\min_{\lambda\in\sigma(A_{\mathrm{u}})}|\lambda|-1\right) > 0, \qquad \varepsilon = \frac{\varepsilon_1}{4} \tag{A.30}$$

completes the proof.

(b) We have $G(y) = Ay + r(y)$ with $\frac{\|r(y)\|}{\|y\|} \to 0$ for $\|y\| \to 0$. We set

$$\eta = \frac{\varepsilon_1}{2(1+2\varepsilon_1)}, \tag{A.31}$$

and $\delta > 0$ such that $\|y\| < \delta$ implies

$$\frac{\|r(y)\|}{\|y\|} < r_0 := \min\left\{\frac{\varepsilon_1}{4\|A\|}, \frac{\varepsilon_1}{4\|A_{\mathrm{u}}\|(1+\eta)^{\frac{1}{2}}}, \left(\frac{\eta\varepsilon_1}{2(1+\eta)}\right)^{\frac{1}{2}}, \frac{\eta^{\frac{1}{2}}\varepsilon_1}{4\|A_{\mathrm{sc}}\|(1+\eta)^{\frac{1}{2}}}\right\}. \tag{A.32}$$

We bound each summand of

$$|G(y)\|^2 = \|A_{\mathrm{sc}}y_{\mathrm{sc}}\|^2 + \|A_{\mathrm{u}}y_{\mathrm{u}}\|^2 + \|r(y)\|^2 + 2\langle Ay, r(y)\rangle \tag{A.33}$$

separately. For $y \in U_{\eta,\delta}$, we get

$$\|A_{\mathrm{sc}}y_{\mathrm{sc}}\|^2 \geq 0, \tag{A.34a}$$

$$\begin{aligned}\|A_{\mathrm{u}}y_{\mathrm{u}}\|^2 &\overset{(a)}{\geq} (1+\varepsilon_1)^2\|y_{\mathrm{u}}\|^2 \geq (1+2\varepsilon_1)\|y_{\mathrm{u}}\|^2 = (1+2\varepsilon_1)\|y\|^2 - (1+2\varepsilon_1)\|y_{\mathrm{sc}}\|^2 \\ &\geq (1+2\varepsilon_1)\|y\|^2 - (1+2\varepsilon_1)\eta\|y\|^2 \overset{(A.31)}{=} (1+\tfrac{3\varepsilon_1}{2})\|y\|^2,\end{aligned} \tag{A.34b}$$

$$\|r(y)\|^2 \geq 0, \tag{A.34c}$$

$$2\langle Ay, r(y)\rangle \geq -2\|A\|\|y\|\|r(y)\| \overset{(A.32)}{\geq} -\tfrac{\varepsilon_1}{2}\|y\|^2, \tag{A.34d}$$

and hence

$$\|G(y)\|^2 \geq (1+\varepsilon_1)\|y\|^2. \tag{A.35}$$

In order to verify $G(y) \in C_\eta$ for $y \in U_{\eta,\delta}$, we write

$$G(y) = \begin{pmatrix} G_{\mathrm{sc}}(y) \\ G_{\mathrm{u}}(y)\end{pmatrix} = \begin{pmatrix} A_{\mathrm{sc}} & \\ & A_{\mathrm{u}}\end{pmatrix}\begin{pmatrix} y_{\mathrm{sc}} \\ y_{\mathrm{u}}\end{pmatrix} + \begin{pmatrix} r_{\mathrm{sc}}(y) \\ r_{\mathrm{u}}(y)\end{pmatrix}. \tag{A.36}$$

We regard each summand of the right-hand side of

$$\begin{aligned}\eta\|G_{\mathrm{u}}(y)\|^2 - \|G_{\mathrm{sc}}(y)\|^2 = \eta\|A_{\mathrm{u}}y_{\mathrm{u}}\|^2 + \eta\|r_{\mathrm{u}}(y)\|^2 + 2\eta\langle A_{\mathrm{u}}y_{\mathrm{u}}, r_{\mathrm{u}}(y)\rangle \\ - \|A_{\mathrm{sc}}y_{\mathrm{sc}}\|^2 - \|r_{\mathrm{sc}}(y)\|^2 - 2\langle A_{\mathrm{sc}}y_{\mathrm{sc}}, r_{\mathrm{sc}}(y)\rangle\end{aligned} \tag{A.37}$$

separately. We have

$$\eta\|A_{\mathrm{u}}y_{\mathrm{u}}\|^2 \overset{\text{(a)}}{\geq} \eta(1+\varepsilon_1)^2\|y_{\mathrm{u}}\|^2, \tag{A.38a}$$

$$\eta\|r_{\mathrm{u}}(y)\|^2 \geq 0, \tag{A.38b}$$

$$\begin{aligned}2\eta\langle A_{\mathrm{u}}y_{\mathrm{u}}, r_{\mathrm{u}}(y)\rangle &\geq -2\eta\|A_{\mathrm{u}}\|\|y_{\mathrm{u}}\|\|r(y)\| \overset{(A.32)}{\geq} -2\eta(1+\eta)^{\frac{1}{2}}r_0\|A_{\mathrm{u}}\|\|y_{\mathrm{u}}\|^2 \\ &\overset{(A.32)}{\geq} -\tfrac{\eta\varepsilon_1}{2}\|y_{\mathrm{u}}\|^2,\end{aligned} \tag{A.38c}$$

$$\|A_{\mathrm{sc}}y_{\mathrm{sc}}\|^2 \overset{\text{(a)}}{\leq} (1+\tfrac{\varepsilon_1}{4})^2\|y_{\mathrm{sc}}\|^2 \overset{y\in U_{\eta,\delta}}{\leq} (1+\tfrac{\varepsilon_1}{4})^2\eta\|y_{\mathrm{u}}\|^2, \tag{A.38d}$$

$$\|r_{\mathrm{sc}}(y)\|^2 \leq \|r(y)\|^2 \overset{(A.32)}{\leq} r_0^2(1+\eta)\|y_{\mathrm{u}}\|^2 \overset{(A.32)}{\leq} \tfrac{\eta\varepsilon_1}{2}\|y_{\mathrm{u}}\|^2, \tag{A.38e}$$

$$\begin{aligned}2\langle A_{\mathrm{sc}}y_{\mathrm{sc}}, r_{\mathrm{sc}}(y)\rangle &\leq 2\|A_{\mathrm{sc}}\|\|y_{\mathrm{sc}}\|\|r(y)\| \overset{(A.32)}{\leq} 2\|A_{\mathrm{sc}}\|\eta^{\frac{1}{2}}(1+\eta)^{\frac{1}{2}}r_0\|y_{\mathrm{u}}\|^2 \\ &\overset{(A.32)}{\leq} \tfrac{\eta\varepsilon_1}{2}\|y_{\mathrm{u}}\|^2.\end{aligned} \tag{A.38f}$$

Combining the estimations (A.38) with (A.37) gives

$$\begin{aligned}\eta\|G_{\mathrm{u}}(y)\|^2 - \|G_{\mathrm{sc}}(y)\|^2 &\geq \eta(1+\varepsilon_1)^2\|y_{\mathrm{u}}\|^2 - \tfrac{\eta\varepsilon_1}{2}\|y_{\mathrm{u}}\|^2 \\ &\quad - (1+\tfrac{\varepsilon_1}{4})^2\eta\|y_{\mathrm{u}}\|^2 - \tfrac{\eta\varepsilon_1}{2}\|y_{\mathrm{u}}\|^2 - \tfrac{\eta\varepsilon_1}{2}\|y_{\mathrm{u}}\|^2 \qquad \text{(A.39a)} \\ &= \tfrac{15\eta\varepsilon_1^2}{16}\|y_{\mathrm{u}}\|^2 > 0. \qquad \text{(A.39b)}\end{aligned}$$

This shows $G(y) \in C_\eta$.

□

Bibliography

[ABB04] F. Alvarez, J. Bolte, and O. Brahic. "Hessian Riemannian Gradient Flows in Convex Programming". In: *SIAM Journal on Control and Optimization* 43.2 (2004), pp. 477–501.

[AC10] S.-I. Amari and A. Cichocki. "Information Geometry of Divergence Functions". In: *Bulletin of the Polish Academy of Sciences: Technical Sciences* 58.1 (2010), pp. 183–195.

[Ach+12] R. Achanta, A. Shaji, K. Smith, A. Lucchi, P. Fua, and S. Süsstrunk. "SLIC Superpixels Compared to State-of-the-art Superpixel Methods". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 34.11 (2012), pp. 2274–2282.

[ADE09] S. N. Anfinsen, A. P. Doulgeris, and T. Eltoft. "Estimation of the Equivalent Number of Looks in Polarimetric Synthetic Aperture Radar Imagery". In: *IEEE Transactions on Geoscience and Remote Sensing* 47.11 (2009), pp. 3795–3809.

[AJLS17] N. Ay, J. Jost, H. V. Lê, and L. Schwachhöfer. *Information Geometry*. Springer, 2017.

[AL10] F. Alvarez and J. López. "Convergence to the Optimal Value for Barrier Methods Combined with Hessian Riemannian Gradient Flows and Generalized Proximal Algorithms". In: *J. Convex Analysis* 17.3&4 (2010), pp. 701–720.

[Ama16] S.-I. Amari. *Information Geometry and its Applications*. Vol. 194. Springer, 2016.

[AN00] S.-I. Amari and H. Nagaoka. *Methods of Information Geometry*. Amer. Math. Soc. and Oxford Univ. Press, 2000.

[ÅPSS17] F. Åström, S. Petra, B. Schmitzer, and C. Schnörr. "Image Labeling by Assignment". In: *Journal of Mathematical Imaging and Vision* 58.2 (2017), pp. 211–238.

[Bar78] O. E. Barndorff-Nielsen. *Information and Exponential Families in Statistical Theory*. Chichester: Wiley, 1978.

[Bat88] V. Batagelj. "Generalized Ward and Related Clustering Problems". In: *Classification and Related Methods of Data Analysis* (1988), pp. 67–74.

[BB97] H. H. Bauschke and J. M. Borwein. "Legendre Functions and the Method of Random Bregman Projections". In: *J. Convex Analysis* 4.1 (1997), pp. 27–67.

[BFPS17] R. Bergmann, J. H. Fitschen, J. Persch, and G. Steidl. "Iterative Multiplicative Filters for Data Labeling". In: *International Journal of Computer Vision* 123.3 (2017), pp. 435–453.

[BGGS17] M. Benning, G. Gilboa, J. S. Grah, and C.-B. Schönlieb. "Learning Filter Functions in Regularisers by Minimising Quotients". In: *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer. 2017, pp. 511–523.

[Bha06] R. Bhatia. *Positive Definite Matrices*. Princeton Univ. Press, 2006.

[Bie+12] A. Biesdorf, K. Rohr, D. Feng, H. von Tengg-Kobligk, F. Rengier, D. Böckler, H.-U. Kauczor, and S. Wörz. "Segmentation and Quantification of the Aortic Arch Using Joint 3D Model-Based Segmentation and Elastic Image Registration". In: *Medical Image Analysis* 16.6 (2012), pp. 1187–1201.

[Bie+15] A. Biesdorf, S. Wörz, H. von Tengg-Kobligk, K. Rohr, and C. Schnörr. "3D Segmentation of Vessels by Incremental Implicit Polynomial Fitting and Convex Optimization". In: *Proc. ISBI*. 2015.

[BK04] Y. Boykov and V. Kolmogorov. "An Experimental Comparison of Min-Cut/Max-Flow Algorithms for Energy Minimization in Vision". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 26.9 (2004), pp. 1124–1137.

[BMDG05] A. Banerjee, S. Merugu, I. S. Dhillon, and J. Ghosh. "Clustering with Bregman Divergences". In: *J. Mach. Learn. Res.* 6 (2005), pp. 1705–1749.

[Boy06] J. P. Boyd. "Computing the Zeros, Maxima and Inflection Points of Chebyshev, Legendre and Fourier Series: Solving Transcendental Equations by Spectral Interpolation and Polynomial Rootfinding". In: *Journal of Engineering Mathematics* 56.3 (2006), pp. 203–219.

[BR09] G. Belitskii and V. Rayskin. "On the Grobman-Hartman Theorem in α-Holder Class for Banach Spaces". In: *preprint* (2009).

[BS02] N. P. Bhatia and G. P. Szegö. *Stability Theory of Dynamical Systems*. Springer Science & Business Media, 2002.

[BS13] J. F. Bonnans and A. Shapiro. *Perturbation Analysis of Optimization Problems*. Springer Science & Business Media, 2013.

[BS99] I. M. Bomze and V. Stix. "Genetic Engineering via Negative Fitness: Evolutionary Dynamics for Global Optimization". In: *Annals of Operations Research* 89 (1999), pp. 297–318.

[BV04] S. Boyd and L. Vandenberghe. *Convex Optimization*. Cambridge University Press, 2004.

[BVZ01] Y. Boykov, O. Veksler, and R. Zabih. "Fast Approximate Energy Minimization via Graph Cuts". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 23.11 (2001), pp. 1222–1239.

[CM02] D. Comaniciu and P. Meer. "Mean Shift: A Robust Approach Toward Feature Space Analysis". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24.5 (2002), pp. 603–619.

[CS16] A. Cherian and S. Sra. "Positive Definite Matrices: Data Representation and Applications to Computer Vision". In: *Algorithmic Advances in Riemannian Geometry and Applications*. Ed. by H. Minh and V. Murino. Springer, 2016, pp. 93–114.

[CSBP13] A. Cherian, S. Sra, A. Banerjee, and N. Papanikolopoulos. "Jensen-Bregman LogDet Divergence with Application to Efficient Similarity Search for Covariance Matrices". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 35.9 (2013), pp. 2161–2174.

[CU14] O. Calin and C. Udrişte. *Geometric Modeling in Probability and Statistics*. Springer, 2014.

[CZ97] Y. A. Censor and S. A. Zenios. *Parallel Optimization: Theory, Algorithms, and Applications*. New York: Oxford Univ. Press, 1997.

[Ela05] S. Elaydi. *An Introduction to Difference Equations*. Undergraduate Texts in Mathematics. Springer New York, 2005.

[FD98] A. Freeman and S. L. Durden. "A Three-Component Scattering Model for Polarimetric SAR Data". In: *IEEE Transactions on Geoscience and Remote Sensing* 36.3 (1998), pp. 963–973.

[Fen79] N. Fenichel. "Geometric Singular Perturbation Theory for Ordinary Differential Equations". In: *Journal of Differential Equations* 31.1 (1979), pp. 53–98.

[FH75] K. Fukunaga and L. Hostetler. "The Estimation of the Gradient of a Density Function, with Applications in Pattern Recognition". In: *IEEE Transactions on Information Theory* 21.1 (1975), pp. 32–40.

[Fri91] D. Friedman. "Evolutionary Games in Economics". In: *Econometrica: Journal of the Econometric Society* (1991), pp. 637–666.

[Gal07] O. Galor. *Discrete Dynamical Systems*. Springer Science & Business Media, 2007.

[GAMF15] L. Gomez, L. Alvarez, L. Mazorra, and A. C. Frery. "Classification of Complex Wishart Matrices with a Diffusion-Reaction System Guided by Stochastic Distances". In: *Philosophical Transactions of The Royal Society A* 373.2056 (2015), pp. 1–14.

[GAMF17] L. Gomez, L. Alvarez, L. Mazorra, and A. C. Frery. "Fully PolSAR Image Classification Using Machine Learning Techniques and Reaction-Diffusion Systems". In: *Neurocomputing* 255 (2017), pp. 52–60.

[GH02] J. Guckenheimer and P. Holmes. *Nonlinear Oscillations, Dynamical Systems, and Bifurcations of Vector Fields*. Applied Mathematical Sciences. Springer New York, 2002.

[Goo63] N. R. Goodman. "The Distribution of the Determinant of a Complex Wishart Distributed Matrix". In: *The Annals of Mathematical Statistics* 34.1 (1963), pp. 178–180.

[Har11] S. Har-Peled. *Geometric Approximation Algorithms*. AMS, 2011.

[HC08] W. M. Haddad and V. Chellaboina. *Nonlinear Dynamical Systems and Control: A Lyapunov-Based Approach*. Princeton University Press, 2008.

[HHLS16] M. T. Harandi, R. Hartley, B. Lovell, and C. Sanderson. "Sparse Coding on Symmetric Positive Definite Manifolds Using Bregman Divergences". In: *IEEE Transactions on Neural Networks and Learning Systems* 27.6 (2016), pp. 1294–1306.

[Hig08] N. J. Higham. *Functions of Matrices: Theory and Computation*. SIAM, 2008.

[HJ12] R. A. Horn and C. R Johnson. *Matrix Analysis*. Cambridge University Press, 2012.

[HS03] J. Hofbauer and K. Sigmund. "Evolutionary Game Dynamics". In: *Bull. Amer. Math. Soc.* 40.4 (2003), pp. 479–519.

[HS87] M. Hashimoto and J. Sklansky. "Multiple-Order Derivatives for Detecting Local Image Characteristics". In: *Computer Vision, Graphics, and Image Understanding* 39 (1987), pp. 28–55.

[HSPS19a] R. Hühnerbein, F. Savarino, S. Petra, and C. Schnörr. "Learning Adaptive Regularization for Image Labeling Using Geometric Assignment". In: *International Conference on Scale Space and Variational Methods in Computer Vision*. Springer. 2019, pp. 393–405.

[HSPS19b] R. Hühnerbein, F. Savarino, S. Petra, and C. Schnörr. "Learning Adaptive Regularization for Image Labeling Using Geometric Assignment". In: *arXiv preprint arXiv:1910.09976* (2019).

[HSS08] T. Hofmann, B. Schölkopf, and A. J. Smola. "Kernel Methods in Machine Learning". In: *Ann. Statistics* 36.3 (2008), pp. 1171–1220.

[Jos17] J. Jost. *Riemannian Geometry and Geometric Analysis*. 7th. Springer-Verlag Berlin Heidelberg, 2017.

[Kap+15] J. H. Kappes, B. Andres, F. A. Hamprecht, C. Schnörr, S. Nowozin, D. Batra, S. Kim, B. X. 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". In: *Int. J. Comp. Vision* 115.2 (2015), pp. 155–184.

[Kar77] H. Karcher. "Riemannian Center of Mass and Mollifier Smoothing". In: *Comm. Pure Appl. Math.* 30 (1977), pp. 509–541.

[Kel66] A. Kelley. "The Stable, Center-Stable, Center, Center-Unstable, Unstable Manifolds". In: *Journal of Differential Equations* (1966).

[Kel67] A. Kelley. "Stability of the Center-Stable Manifold". In: *Journal of Mathematical Analysis and Applications* 18.2 (1967), pp. 336–344.

[KMB15] A. Kleefeld, A. Meyer-Baese, and B. Burgeth. "Elementary Morphology for SO(2)- and SO(3)-Orientation Fields". In: *International Symposium on Mathematical Morphology and Its Applications to Signal and Image Processing*. Springer. 2015, pp. 458–469.

[Kre72] R. Kreß. "On General Hermite Trigonometric Interpolation". In: *Numerische Mathematik* 20.2 (1972), pp. 125–138.

[KSR19] L. Kostrykin, C. Schnörr, and K. Rohr. "Globally Optimal Segmentation of Cell Nuclei in Fluoroscence Microscopy Images using Shape and Intensity Information". In: *Medical Image Analysis* (2019).

[LA83] V. Losert and E. Akin. "Dynamics of Games and Genes: Discrete versus Continuous Time". In: *Journal of Mathematical Biology* 17.2 (1983), pp. 241–251.

[LE09] J.-S. Lee and E. Pottier. *Polarimetric Radar Imaging: From Basics to Applications*. CRC, Boca Raton, 2009.

[Lee06] J. M. Lee. *Riemannian Manifolds: An Introduction to Curvature*. Vol. 176. Springer Science & Business Media, 2006.

[Lee13] J. M. Lee. *Introduction to Smooth Manifolds*. Springer, 2013.

[Lee18] J. M. Lee. *Introduction to Riemannian Manifolds*. Vol. 2. Springer, 2018.

[LL05] J. Lafferty and G. Lebanon. "Diffusion Kernels on Statistical Manifolds". In: *Journal of Machine Learning Research* 6 (2005), pp. 129–163.

[Low04] D. G. Lowe. "Distinctive Image Features from Scale-Invariant Keypoints". In: *Int. J. Comp. Vision* 60.2 (2004), pp. 91–110.

[LS11] J. Lellmann and C. Schnörr. "Continuous Multiclass Labeling Approaches and Algorithms". In: *SIAM Journal on Imaging Sciences* 4.4 (2011), pp. 1049–1096.

[MP00] G. McLachlan and D. Peel. *Finite Mixture Models*. Wiley, 2000.

[MS18] P. Mertikopoulos and W. H. Sandholm. "Riemannian Game Dynamics". In: *Journal of Economic Theory* 177 (2018), pp. 315–364.

[Mül11] D. Müllner. "Modern Hierarchical, Agglomerative Clustering Algorithms". In: *arXiv preprint arXiv:1109.2378* (2011).

[MY09] J.-M. Morel and G. Yu. "ASIFT: A New Framework for Fully Affine Invariant Image Comparison". In: *SIAM Journal on Imaging Sciences* 2.2 (2009), pp. 438–469.

[NFC14] A. D. C. Nascimento, A. C. Frery, and R. J. Cintra. "Bias Correction and Modified Profile Likelihood Under the Wishart Complex Distribution". In: *IEEE Transactions on Geoscience and Remote Sensing* 52.8 (2014), pp. 4932–4941.

[NN94] Y. Nesterov and A. Nemirovskii. *Interior Point Polynomial Algorithms in Convex Programming*. SIAM, 1994.

[Per01] L. Perko. *Differential Equations and Dynamical Systems*. Vol. 7. Springer Science & Business Media, 2001.

[PFA06] X. Pennec, P. Fillard, and N. Ayache. "A Riemannian framework for Tensor Computing". In: *International Journal of Computer Vision* 66.1 (2006), pp. 41–66.

[PS00] J. Portilla and E. P. Simoncelli. "A Parametric Texture Model Based on Joint Statistics of Complex Wavelet Coefficients". In: *Int. J. Comp. Vision* 40.1 (2000), pp. 49–70.

[RPF97] F. Rocca, C. Prati, and A. Ferretti. "An Overview of ERS-SAR Interferometry". In: *ERS Symposium on Space at the Service of Our Environment, 3 rd, Florence, Italy*. 1997.

[RSS14] F. Rathke, S. Schmidt, and C. Schnörr. "Probabilistic Intra-Retinal Layer Segmentation in 3-D OCT Images Using Global Shape Regularization". In: *Medical Image Analysis* 18.5 (2014), pp. 781–794.

[RW09] R. T. Rockafellar and R. J.-B. Wets. *Variational Analysis*. 3rd. Springer, 2009.

[Sav+17] F. Savarino, R. Hühnerbein, F. Åström, J. Recknagel, and C. Schnörr. "Numerical Integration of Riemannian Gradient Flows for Image Labeling". In: *International Conference on Scale Space and Variational Methods in Computer Vision*. LNCS. Springer, 2017, pp. 361–372.

[SC16] D. G. Schaeffer and J. W. Cain. *Ordinary Differential Equations: Basics and Beyond*. Springer, 2016.

[Sch20] C. Schnörr. "Assignment Flows". In: *Variational Methods for Nonlinear Geometric Data and Applications*. Ed. by P. Grohs, M. Holler, and A. Weinmann. Springer, 2020, pp. 235–260.

[SG16] S. Schecter and H. Gintis. *Game Theory in Action: An Introduction to Classical and Evolutionary Models*. Princeton University Press, 2016.

[SM00] J. Shi and J. Malik. "Normalized Cuts and Image Segmentation". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 22.8 (2000), pp. 888–905.

[SM09] R. Subbarao and P. Meer. "Nonlinear Mean Shift over Riemannian Manifolds". In: *Int. J. Comp. Vision* 84.1 (2009), pp. 1–20.

[Sra13] S. Sra. "Positive Definite Matrices and the Symmetric Stein Divergence". In: *arXiv preprint arXiv:1110.1773* (2013).

[SS20] F. Savarino and C. Schnörr. "Continuous-Domain Assignment Flows". In: *Europ. J. Appl. Math.* (2020). (to appear).

[Teb07] M. Teboulle. "A Unified Continuous Optimization Framework for Center-Based Clustering Methods". In: *J. Mach. Learning Res.* 8 (2007), pp. 65–102.

[Tes12] G. Teschl. *Ordinary Differential Equations and Dynamical Systems*. Vol. 140. Grad. Studies Math. Amer. Math. Soc., 2012.

[TJ78] P. D. Taylor and L. B. Jonker. "Evolutionary Stable Strategies and Game Dynamics". In: *Mathematical Biosciences* 40.1-2 (1978), pp. 145–156.

[TPM06] O. Tuzel, F. Porikli, and P. Meer. "Region Covariance: A Fast Descriptor for Detection and Classification". In: *Proc. ECCV*. Springer. 2006, pp. 589–600.

[TPM08] O. Tuzel, F. Porikli, and P. Meer. "Pedestrian Detection via Classification on Riemannian Manifolds". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 30.10 (2008), pp. 1713–1727.

[TS16] P. K. Turaga and A. Srivastava, eds. *Riemannian Computing in Computer Vision*. Springer, 2016.

[Van00] C. F. Van Loan. "The Ubiquitous Kronecker Product". In: *J. Comput. Appl. Math.* 123 (2000), pp. 85–100.

[Wer07] T. Werner. "A Linear Programming Approach to Max-Sum Problem: A Review". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 29.7 (2007), pp. 1165–1179.

[WER18] T. Wollmann, C. S. Eijkman, and K. Rohr. "Adversarial Domain Adaptation to Improve Automatic Breast Cancer Grading in Lymph Nodes". In: *2018 IEEE 15th International Symposium on Biomedical Imaging (ISBI 2018)*. IEEE. 2018, pp. 582–585.

[Wig03] S. Wiggins. *Introduction to Applied Nonlinear Dynamical Systems and Chaos*. Vol. 2. Springer Science & Business Media, 2003.

[WJ08] M. J. Wainwright and M. I. Jordan. "Graphical Models, Exponential Families, and Variational Inference". In: *Foundations and Trends® in Machine Learning* 1.1–2 (2008), pp. 1–305.

[XJZ16] W. Xie, L. Jiao, and J. Zhao. "PolSAR Image Classification via D-KSVD and NSCT-Domain Features Extraction". In: *IEEE Geoscience and Remote Sensing Letters* 13.2 (2016), pp. 227–231.

[Zer+18] A. Zern, M. Zisler, F. Åström, S. Petra, and C. Schnörr. "Unsupervised Label Learning on Manifolds by Spatially Regularized Geometric Assignment". In: *German Conference on Pattern Recognition*. Springer. 2018, pp. 698–713.

[ZM97] S. C. Zhu and D. Mumford. "Prior Learning and Gibbs Reaction–Diffusion". In: *IEEE Transactions on Pattern Analysis and Machine Intelligence* 19.11 (1997), pp. 1236–1250.

[ZRS18] A. Zern, K. Rohr, and C. Schnörr. "Geometric Image Labeling with Global Convex Labeling Constraints". In: *International Workshop on Energy Minimization Methods in Computer Vision and Pattern Recognition*. Vol. 10746. LNCS. 2018, pp. 533–547.

[ZSPS20] A. Zeilmann, F. Savarino, S. Petra, and C. Schnörr. "Geometric Numerical Integration of the Assignment Flow". In: *Inverse Problems* 36 (2020), 034004 (33pp).

[ZZPS19a] A. Zern, M. Zisler, S. Petra, and C. Schnörr. "Spatially Regularized Geometric Assignment for Unsupervised Label Learning on Manifolds". In: *PAMM* 19.1 (2019), e201900258. DOI: 10.1002/pamm.201900258.

[ZZPS19b] M. Zisler, A. Zern, S. Petra, and C. Schnörr. "Unsupervised Labeling by Geometric and Spatially Regularized Self-Assignment". In: *International Conference on Scale Space and Variational Methods in Computer Vision*. 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". In: *Journal of Mathematical Imaging and Vision* (2020). DOI: 10.1007/s10851-019-00935-7. (in press).

[ZZPS20b] M. Zisler, A. Zern, S. Petra, and C. Schnörr. "Self-Assignment Flows for Unsupervised Data Labeling on Graphs". In: *SIAM Journal on Imaging Sciences* (2020). (in press).

[ZZS20] A. Zern, A. Zeilmann, and C. Schnörr. "Assignment Flows for Data Labeling on Graphs: Convergence and Stability". In: *arXiv preprint arXiv:2002.11571* (2020).